The Annals of Applied Probability
2022, Vol. 32, No. 3, 2211–2253
https://doi.org/10.1214/21-AAP1732

A TWO-PARAMETER FAMILY OF MEASURE-VALUED DIFFUSIONS WITH POISSON–DIRICHLET STATIONARY DISTRIBUTIONS

BY NOAH FORMAN[1,a], DOUGLAS RIZZOLO[2,b], QUAN SHI[3,c] AND MATTHIAS WINKEL[4,d]

[1]*Department of Mathematics & Statistics, McMaster University,* [a]*noah.forman@gmail.com*
[2]*Department of Mathematical Sciences, University of Delaware,* [b]*drizzolo@udel.edu*
[3]*Institut für Mathematik, Universität Mannheim,* [c]*quanshi.math@gmail.com*
[4]*Department of Statistics, University of Oxford,* [d]*winkel@stats.ox.ac.uk*

We give a pathwise construction of a two-parameter family of purely-atomic-measure-valued diffusions in which ranked masses of atoms are stationary with the Poisson–Dirichlet(α, θ) distributions, for $\alpha \in (0, 1)$ and $\theta \geq 0$. These processes resolve a conjecture of Feng and Sun (*Probab. Theory Related Fields* **148** (2010) 501–525). We build on our previous work on $(\alpha, 0)$- and (α, α)-interval partition evolutions. The extension to general $\theta \geq 0$ is achieved by the construction of a σ-finite excursion measure of a new measure-valued branching diffusion. Our measure-valued processes are Hunt processes on an incomplete subspace of the space of all probability measures and do not possess an extension to a Feller process. In a companion paper, we use generators to show that ranked masses evolve according to a two-parameter family of diffusions introduced by Petrov (*Funktsional. Anal. i Prilozhen.* **43** (2009) 45–66), extending work of Ethier and Kurtz (*Adv. in Appl. Probab.* **13** (1981) 429–452).

1. Introduction. The well-known *labeled infinitely-many-neutral-alleles model* introduced by Ethier and Kurtz [15, 16] is a one-parameter family of Fleming–Viot processes [19]. They showed that these measure-valued diffusions have stationary distributions that are now known [27, 39, 46] in the Bayesian nonparametrics community as the $\alpha = 0$ case of the Pitman–Yor process $\mathtt{PY}(\alpha, \theta, \mu)$, for a parameter $\theta > 0$ and a probability measure μ on an underlying space of alleles. For general $\alpha \in [0, 1)$, $\theta > -\alpha$, the ranked atom sizes of $\mathtt{PY}(\alpha, \theta, \mu)$ form a Poisson–Dirichlet sequence $\mathtt{PD}(\alpha, \theta)$ and their locations are independent with common distribution μ.

In this paper, we construct a family of measure-valued diffusions in the two-parameter setting $\alpha \in (0, 1)$ and $\theta \geq 0$. The existence of such processes has been conjectured by Feng and Sun [17]. Some progress has been made by analytic methods [8, 18, 34] to establish processes on spaces of decreasing sequences, on spaces allowing only finitely many atoms, and in the case $\alpha = 1/2$. Our approach is probabilistic, building on our previous work [21, 22] on interval-partition-valued diffusions. This enables us to study the evolution of individual atoms from their creation until reaching zero size again, corresponding to the creation and extinction of allelic types in the language of Ethier and Kurtz [13, 15].

Petrov [34] introduced $\mathtt{PD}(\alpha, \theta)$ diffusions, which generalize the process of ranked atom sizes [13] in the labeled infinitely-many-neutral-alleles diffusions [15]. In a companion paper [23], we use analytic arguments (generators) to show that projections of our construction onto the space ∇_∞ of ranked atom sizes evolve as Petrov's diffusions, hence fully resolving

Received October 2020; revised July 2021.
MSC2020 subject classifications. Primary 60J68, 60J60, 60J80; secondary 60G18, 60G52, 60G55.
Key words and phrases. Fleming–Viot process, Poisson–Dirichlet distribution, infinitely-many-neutral-alleles model, excursion theory, stable process.

Feng and Sun's conjecture. Our discussion here does not cover the case $\theta \in (-\alpha, 0)$, which is postponed to [43], as this requires considerable further refinement of our construction.

In the following, we model the space of alleles by the interval $[0, 1]$ and take as μ the uniform distribution $\texttt{Unif}(0,1)$, which may be thought of as the distribution of the allelic type of a mutant offspring of a gamete of any allelic type. This specialization is only for simplicity; the same results hold for any compact metric space endowed with any atom-free distribution μ. In this context we denote the Pitman–Yor distribution by $\texttt{PDRM}(\alpha, \theta)$, so that

$$\begin{aligned} &\sum_{i\geq 1} A_i \delta(U_i) \sim \texttt{PDRM}(\alpha, \theta), \\ &\text{with } \begin{matrix}(A_i, i \geq 1) \sim \texttt{PD}(\alpha, \theta), \\ U_i \sim \texttt{Unif} := \texttt{Unif}(0, 1), \quad i \geq 1,\end{matrix} \quad \text{independent.} \end{aligned} \tag{1.1}$$

These Poisson–Dirichlet random measures will be the stationary distributions of our measure-valued diffusions. Pitman–Yor processes are also called *three-parameter Dirichlet processes* in [4], as they are equivalent to what has also been referred to as the Dirichlet processes in Bayesian nonparametric statistics when $\alpha = 0$. We also refer to [28, 35] for more studies on this family.

1.1. *Statements of main results.* Let $\mathcal{M}^a \subset \mathcal{M}$ be the space of all purely atomic finite measures on $[0, 1]$, as a subspace of the Prokhorov space $(\mathcal{M}, d_{\mathcal{M}})$. For any measure $\pi \in \mathcal{M}$, write $\|\pi\| := \pi([0, 1])$ for its *total mass*. We denote by $\mathcal{M}_1^a = \{\pi \in \mathcal{M}^a : \|\pi\| = 1\}$ the subspace of atomic probability measures on $[0, 1]$. Our construction of $\mathcal{M}_1^a$-valued diffusions is in two steps. The first step is to construct self-similar $\mathcal{M}^a$-valued diffusions with a branching property, with fluctuating total mass. The second step is to time-change these diffusions and to renormalize by their total mass. In this measure-valued context, this is reminiscent—but not a special case—of the skew-product relationship between the measure-valued branching diffusions and probability-measure-valued Fleming–Viot process of [12, 31, 44].

The first step is most easily described by defining semigroups $(K_y^{\alpha,\theta}, y \geq 0)$ on $(\mathcal{M}^a, d_{\mathcal{M}})$ that possess the *branching property* under which the state at time y can be seen as the sum of a family of independent random measures indexed by the atoms of the initial measure (plus a further random measure—immigration). Indeed, we can think of a genealogy in which the atoms of each of the independent random measures are the descendants of the associated atom of the initial measure. Specifically, each atom $b\delta(x)$ of the initial measure will make:

- a trivial contribution (the zero measure $\mathbf{0}$), with some probability $\check{q}$,
- a contribution $\check{L}\delta(x) + \Pi$ that includes a "surviving" atom at x of some size $\check{L}$ and an atomic measure Π of further descendants, with probability $(1 - \check{q})\check{p}$,
- and a contribution of the form $\hat{L}\delta(U_0) + \Pi$ without x itself surviving, but with surviving descendants led by an atom of size $\hat{L}$ at some location U_0, with probability $(1 - \check{q})(1 - \check{p})$,

for some $\check{q}, \check{p} \in (0, 1)$, a random measure Π, some random atom sizes $\check{L}$ and $\hat{L}$, and a location $U_0 \sim \texttt{Unif}$, which we will assume to be jointly independent on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$. Setting $L = \check{L}$ with probability $\check{p}$ and $L = \hat{L}$ otherwise, we can disintegrate against the distribution of L and write the law of this contribution in the form

$$\begin{aligned} Q(\cdot) &= \check{q}\boldsymbol{\delta}_{\mathbf{0}}(\cdot) + (1 - \check{q})\check{p}\mathbf{P}\{\check{L}\delta(x) + \Pi \in \cdot\} + (1 - \check{q})(1 - \check{p})\mathbf{P}\{\hat{L}\delta(U_0) + \Pi \in \cdot\} \\ &= \check{q}\boldsymbol{\delta}_{\mathbf{0}}(\cdot) + (1 - \check{q}) \int_0^\infty \big(p(c)\mathbf{P}\{c\delta(x) + \Pi \in \cdot\} \\ &\quad + \big(1 - p(c)\big)\mathbf{P}\{c\delta(U_0) + \Pi \in \cdot\}\big)\mathbf{P}\{L \in dc\} \end{aligned} \tag{1.2}$$

for some measurable function $p\colon (0,\infty) \to [0,1]$, some random atom size L, location $U_0 \sim \texttt{Unif}$ and a random measure Π. Here, $\boldsymbol{\delta_0}$ denotes a Dirac mass on the zero measure $\mathbf{0}$, whereas $\delta(x)$ and $\delta(U_0)$ denote point masses on $[0,1]$. More precisely, we now define the kernels $K_y^{\alpha,\theta}$, where $y \geq 0$ is our time parameter, by a specific choice of these ingredients, which in the cases $\theta = 0$ and $\theta = \alpha$ are motivated by related semi-groups on spaces of interval partitions [22]. These in turn arose from level sets and jumps across level y in stable Lévy processes marked by squared Bessel excursions [20, 21], as we will recall in Section 1.2. These continuum constructions were originally motivated by down-up Chinese restaurant chains on integer compositions encoded by finite-activity Lévy processes marked by integer-valued paths; see [41] for a detailed discussion in the two-parameter model and [43] for stronger limit theorems in the self-similar regime, all in the setting of interval partitions.

DEFINITION 1.1 (Transition kernel $K_y^{\alpha,\theta}$). Fix $0 < \alpha < 1$. For $b > 0$ and $r > 0$, consider independent random variables $L_{b,r}$ with Laplace transform

$$\mathbf{E}\big[e^{-qL_{b,r}}\big] = \left(\frac{r+q}{r}\right)^{\alpha} \frac{e^{br^2/(r+q)} - 1}{e^{br} - 1}, \quad q \geq 0, \tag{1.3}$$

$U_0 \sim \texttt{Unif} := \texttt{Unif}[0,1]$, $\overline{\Pi} \sim \texttt{PDRM}(\alpha,\alpha)$ and $G_r \sim \texttt{Gamma}(\alpha, r)$. Scale the measure $\overline{\Pi}$ of unit mass to obtain a random measure $\Pi_r := G_r\overline{\Pi}$ of total mass G_r. Further consider extinction probabilities $\breve{q}_{b,r} = e^{-br}$ and mixing probabilities

$$p_{b,r}^{(\alpha)}(c) = \frac{I_{1+\alpha}(2r\sqrt{bc})}{I_{-1-\alpha}(2r\sqrt{bc}) + \alpha(2r\sqrt{bc})^{-1-\alpha}/\Gamma(1-\alpha)} \quad \text{and} \quad 1 - p_{b,r}^{(\alpha)}(c),$$

where for $v \in \mathbb{R}$, we denote by I_v the modified Bessel function of the first kind of index v. Then proceed as in (1.2) and define the distribution $Q_{b,x,r}^{(\alpha)}$ on $\mathcal{M}^a$ of a random measure as

$$\begin{aligned} Q_{b,x,r}^{(\alpha)}(\cdot) = e^{-br}\boldsymbol{\delta_0}(\cdot) + (1-e^{-br})\int_0^\infty \big(p_{b,r}^{(\alpha)}(c)\mathbf{P}\{c\delta(x) + \Pi_r \in \cdot\} \\ + (1 - p_{b,r}^{(\alpha)}(c))\mathbf{P}\{c\delta(U_0) + \Pi_r \in \cdot\}\big)\mathbf{P}\{L_{b,r} \in dc\}. \end{aligned} \tag{1.4}$$

Also fix $\theta \geq 0$. Let $\pi = \sum_{i\geq 1} b_i\delta(x_i) \in \mathcal{M}^a$. For a time-$y$ transition, $y > 0$, let $r = 1/2y$ and define $K_y^{\alpha,\theta}(\pi,\cdot)$ to be the distribution on $\mathcal{M}^a$ of the random measure $G^y\overline{\Pi}_0 + \sum_{i\geq 1}\Pi_i^y$ for independent $G^y \sim \texttt{Gamma}(\theta, 1/2y)$, $\overline{\Pi}_0 \sim \texttt{PDRM}(\alpha,\theta)$ and $\Pi_i^y \sim Q_{b_i,x_i,1/2y}^{(\alpha)}$, $i \geq 1$.

By [22], Proposition 2.2(iv), or [36], Corollary 4.8, and beta-gamma algebra, Π_r has the same distribution as $\sum_{i\geq 1} V_i\delta(U_i)$, where $(V_i, i \geq 1)$ is the ranked sequence of jump heights of a $\texttt{stable}(\alpha)$ subordinator stopped just before jumping across an independent $\texttt{Exponential}(r)$ threshold, and $U_i \sim \texttt{Unif}$, $i \geq 1$, are independent.

One can easily check that there are a.s. only finitely many $\Pi_i^y \neq 0$ (cf. [21], Lemma 6.1) and thus $K_y^{\alpha,\theta}(\pi,\cdot)$ is indeed well defined as a measure on the Borel sets of $(\mathcal{M}^a, d_{\mathcal{M}})$. Note, however, that it is not clear a priori that $(K_y^{\alpha,\theta}, y \geq 0)$ forms a transition semigroup nor that such a semigroup belongs to a nice Markov process. We do not attempt to settle in this paper the natural question of what other $L_{b,r}$, $p_{b,r}^{(\alpha)}(c)$, Π_r etc. lead to transition semigroups. See however Section 1.2 for a construction that sheds some light on this question and provides tools to study the processes. For now, we state our theorem for $(K_y^{\alpha,\theta}, y \geq 0)$.

THEOREM 1.2. *Fix $\alpha \in (0,1)$ and $\theta \geq 0$. The family $(K_y^{\alpha,\theta}, y \geq 0)$ forms the transition semigroup of a path-continuous Hunt process on $(\mathcal{M}^a, d_{\mathcal{M}})$.*

We shall refer to these measure-valued processes as *self-similar* (α, θ)*-superprocesses*, or $\mathtt{SSSP}(\alpha, \theta)$, in view of the following proposition. We write $\mathtt{SSSP}_\pi(\alpha, \theta)$ for the distribution of an $\mathtt{SSSP}(\alpha, \theta)$ starting from $\pi \in \mathcal{M}^a$. We typically denote an $\mathtt{SSSP}(\alpha, \theta)$ by $(\pi^y, y \geq 0)$, using y as its time parameter.

PROPOSITION 1.3 (Self-similarity). *Fix $\alpha \in (0,1)$, $\theta \geq 0$ and $\pi \in \mathcal{M}^a$. Then $(\pi^y, y \geq 0) \sim \mathtt{SSSP}_\pi(\alpha, \theta)$ implies $(c\pi^{y/c}, y \geq 0) \sim \mathtt{SSSP}_{c\pi}(\alpha, \theta)$ for all $c > 0$.*

PROPOSITION 1.4 (Additivity property). *Fix $\alpha \in (0,1)$, $\theta_1, \theta_2 \geq 0$, and mutually singular measures $\pi_1, \pi_2 \in \mathcal{M}^a$. If $(\pi_1^y, y \geq 0) \sim \mathtt{SSSP}_{\pi_1}(\alpha, \theta_1)$ and $(\pi_2^y, y \geq 0) \sim \mathtt{SSSP}_{\pi_2}(\alpha, \theta_2)$ are independent then $(\pi_1^y + \pi_2^y, y \geq 0)$ is an $\mathtt{SSSP}_{\pi_1+\pi_2}(\alpha, \theta_1 + \theta_2)$.*

For any $r \in \mathbb{R}$, $z \geq 0$ and B a standard one-dimensional Brownian motion, it is well known that there exists a unique strong solution to the equation

$$Z_t = z + rt + 2\int_0^t \sqrt{|Z_s|}\, dB_s,$$

which is called an *r-dimensional squared Bessel process* starting from z and denoted by $\mathtt{BESQ}_z(r)$. The Feller diffusion, which is a continuous-state branching process that arises as a scaling limit of critical Galton–Watson processes, is $\mathtt{BESQ}(0)$. For $r > 0$, $\mathtt{BESQ}(r)$ is a Feller diffusion with immigration. The case $r < 0$ can be interpreted as a Feller diffusion with emigration at rate $|r|$. In this case, as when $r = 0$, the boundary point 0 is not an entrance boundary for $(0, \infty)$, while exit at 0 (we will then force absorption) happens almost surely. For $r = d \in \mathbb{N}$, the squared norm of a d-dimensional Brownian motion is a $\mathtt{BESQ}(d)$. See [25, 33, 38].

THEOREM 1.5 (Total mass). *For $\alpha \in (0,1)$, $\theta \geq 0$ and $\pi \in \mathcal{M}^a$, the total mass of an $\mathtt{SSSP}_\pi(\alpha, \theta)$ evolves as a $\mathtt{BESQ}_{\|\pi\|}(2\theta)$.*

For the second step, to construct a probability-measure-valued diffusion, we consider $\boldsymbol{\pi} := (\pi^y, y \geq 0) \sim \mathtt{SSSP}_\pi(\alpha, \theta)$. We define a time-change function by

$$\rho_{\boldsymbol{\pi}}(u) := \inf\left\{ y \geq 0 \colon \int_0^y \|\pi^z\|^{-1}\, dz > u \right\}, \quad u \geq 0. \tag{1.5}$$

DEFINITION 1.6 ((α, θ)-Fleming–Viot process). Let $\boldsymbol{\pi} := (\pi^y, y \geq 0)$ be an $\mathtt{SSSP}_\pi(\alpha, \theta)$ and consider the time-change function $\rho_{\boldsymbol{\pi}}$ of (1.5). Then the $\mathcal{M}_1^a$-valued process $\overline{\boldsymbol{\pi}} := (\overline{\pi}^u, u \geq 0)$ defined by

$$\overline{\pi}^u := \|\pi^{\rho_{\boldsymbol{\pi}}(u)}\|^{-1} \pi^{\rho_{\boldsymbol{\pi}}(u)}, \quad u \geq 0,$$

is called an (α, θ)*-Fleming–Viot process*, or $\mathtt{FV}(\alpha, \theta)$.

Our main result is the following.

THEOREM 1.7. *For $\alpha \in (0,1)$ and $\theta \geq 0$, the $\mathtt{FV}(\alpha, \theta)$ is a path-continuous Hunt process on $(\mathcal{M}_1^a, d_{\mathcal{M}})$ and has $\mathtt{PDRM}(\alpha, \theta)$ as a stationary distribution.*

This is similar to the well-known normalization/time-change connection between the measure-valued branching diffusions and the Fleming–Viot processes of [12, 31, 44]. To extend the analogy, in our setting both $\mathtt{SSSP}(\alpha, \theta)$ and $\mathtt{FV}(\alpha, \theta)$ are time-homogeneous Markov

processes and SSSP(α,θ) enjoys the additivity property, which is a weaker version of the branching property arising in the above references.

Shiga ([44], (3.1)–(3.2)) gave a construction of a large class of measure-valued processes that include a process that, in the present setting, could be called SSSP$(0,\theta)$, based on a Poisson point process of BESQ(0) excursions. In the branching process view, this point process represents immigration over time. Our method for combining Shiga's $(0,\theta)$- with our $(\alpha,0)$- to construct (α,θ)-processes can be viewed as replacing each BESQ(0) excursion, which describes an individual immigrant, by a novel measure-valued excursion entering continuously from zero initial mass, which describes an immigrant family. In Section 7.1, we make this connection precise and use it to show that the FV(α,θ), along with some additional type labels, can be projected down to the infinitely-many-neutral-alleles model of Ethier and Kurtz [14, 15], which we refer to as FV$(0,\theta)$. This connection is a process-level analogue to the classic construction of PD(α,θ) by fragmenting each block of a PD$(0,\theta)$ into independent PD$(\alpha,0)$ proportions [36], (5.26), which we recall in (7.2). In our framework, we have not been able to identify a generator, which is a powerful tool in Shiga's framework. However, our approach does allow generator calculations, in principle, as we demonstrate in [23] to identify the ranked mass process. Beyond this, we leave the identification of a generator as an open problem.

A key innovation in this paper is the construction of the measure-valued excursions that we use to replace Shiga's BESQ(0).

1.2. *Construction from marked Lévy processes.* Although our construction can be viewed as a process-level fragmentation of Shiga's construction, at a technical level our approach is quite different. It is unclear how to adapt Shiga's martingale methods to the present context, and Feng and Sun's [17, 18] attempts to employ Dirichlet forms have encountered as yet insurmountable technical difficulties. In this paper we construct SSSP(α,θ) and FV(α,θ) by extending the methods of [21, 22], using marked stable Lévy processes to construct the immigrant families mentioned above. Specifically, following [21], denote by $\nu_{\mathrm{BESQ}}^{(-2\alpha)}$ the Pitman–Yor excursion measure of BESQ(-2α) on a space $\mathcal{E}$ of excursions away from zero [38], with normalization so that

$$\nu_{\mathrm{BESQ}}^{(-2\alpha)}\{\zeta > y\} = \frac{\alpha}{2^{\alpha}\Gamma(1-\alpha)\Gamma(1+\alpha)}\, y^{-1-\alpha},$$

where $\zeta(f) = \inf\{y > 0\colon f(y) = 0\}$ is the *length/lifetime* of the excursion $f \in \mathcal{E}$. We represent f as a function $f\colon \mathbb{R} \to [0,\infty)$ vanishing outside $[0,\zeta(f))$. Let $\mathbf{V}$ be a Poisson random

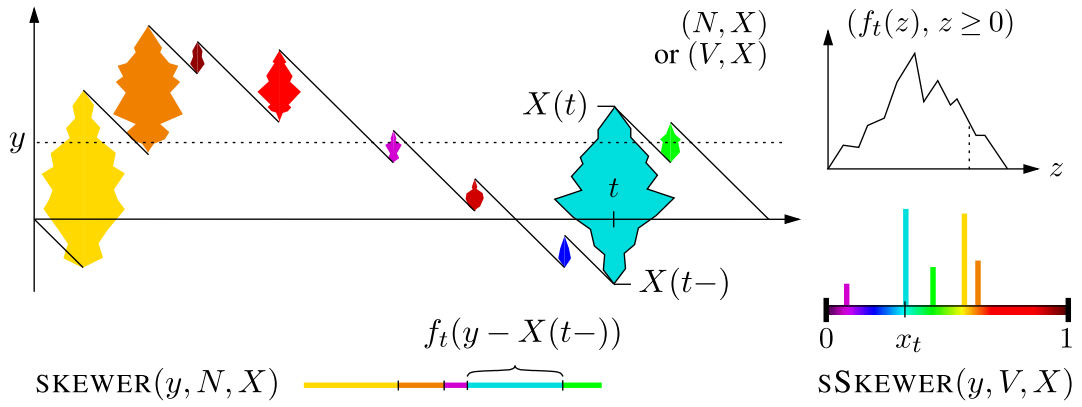

FIG. 1. *A scaffolding with marks (atom size evolutions as spindle-shapes and allelic types from a color spectrum coded by* $[0,1]$*) and the skewer and superskewer at level* y*, not to scale.*

measure (PRM) on $[0,\infty)\times\mathcal{E}\times[0,1]$ with intensity measure $\mathtt{Leb}\otimes\nu_{\mathtt{BESQ}}^{(-2\alpha)}\otimes\mathtt{Unif}$, where Leb denotes one-dimensional Lebesgue measure. Mapping all points (t, f_t, x_t) of $\mathbf{V}$ to times t and excursion lengths $\zeta(f_t)$, we obtain a PRM from which we can naturally define a zero-mean (stable) Lévy process $\mathbf{X}$ whose jumps are $\zeta(f_t)$ at each time t of a point (t, f_t, x_t) of $\mathbf{V}$. From the perspective of $\mathbf{X}$ we view (f_t, x_t) as an *atom size evolution* and *allelic type* marking the jump of size $\zeta(f_t)$ at time t. We refer to $\mathbf{X}$ as *scaffolding* to which the marks are attached. See Figure 1 for an illustration with finitely many jumps. We will refer to the excursions marking the jumps of $\mathbf{X}$ as *spindles*.

For $x\in[0,1]$ and $b>0$, denote by $\mathbf{Q}_{b,x}^{(\alpha)}$ the distribution of

$$\widehat{\mathbf{V}} := \delta(0,\mathbf{f},x) + \mathbf{V}|_{[0,T]\times\mathcal{E}\times[0,1]}, \quad \text{for } \begin{array}{l}\mathbf{f}\sim \mathtt{BESQ}_b(-2\alpha) \text{ independent,}\\ T := \inf\{t\ge 0\colon \zeta(\mathbf{f})+\mathbf{X}(t)=0\}.\end{array} \tag{1.6}$$

The Lévy process $\widehat{\mathbf{X}} := \zeta(\mathbf{f}) + \mathbf{X}|_{[0,T]}$ starting from $\zeta(\mathbf{f})>0$ and stopped at its first passage at 0 is fully determined by $\widehat{\mathbf{V}}$ and called the *scaffolding associated with* $\widehat{\mathbf{V}}$. We refer to $\widehat{\mathbf{V}}$ as a *clade starting from* $b\delta(x)$. Indeed, $\widehat{\mathbf{V}}$ provides the further size evolution of type x expressed by $\mathbf{f}$, and each atom (t, f_t, x_t) of $\widehat{\mathbf{V}}$, for $t>0$, is interpreted as a descendant allelic type x_t first created at level $\widehat{\mathbf{X}}(t-)$ with size evolving continuously according to f_t until its extinction at level $\widehat{\mathbf{X}}(t)=\widehat{\mathbf{X}}(t-)+\zeta(f_t)$; then the size of type x_t at level y is $f_t(y-\widehat{\mathbf{X}}(t-))$. See Figure 1.

DEFINITION 1.8. The *superskewer* at level y of a pair (V,X) of a point measure V of spindles and allelic types and a scaffolding X is defined as

$$\begin{aligned}\textsc{sSkewer}(y,V,X) &:= \int f\big(y-X(t-)\big)\delta(x)V(dt,df,dx)\\ &= \sum_{\text{points}(t,f_t,x_t)\text{of}V} f_t\big(y-X(t-)\big)\delta(x_t).\end{aligned}$$

We call this the superskewer as it constructs a superprocess by collecting from each level y of the marked scaffolding(s) and placing onto the type space $[0,1]$ the atoms of the superprocess. This is a variant of the skewer map introduced in [21], which similarly constructs an interval partition whose interval lengths (our atom sizes here) are placed onto $[0,\infty)$ by the skewer map in the left-to-right order of the jump times without leaving gaps as if on a skewer that pushes through the marked scaffolding from left to right. This skewer map does not record allelic types and therefore only depends on $N(dt,df) := V(dt,df,[0,1])$. See Figure 1 for an illustration and [21], Definition 1.2, or [22], Definition 1.7, for a precise definition of $\textsc{skewer}(y,N,X)$.

With $\widehat{\mathbf{V}}$ being a clade starting from $b\delta(x)$ and $\widehat{\mathbf{X}}$ being its associated scaffolding, $(\textsc{sSkewer}(y,\widehat{\mathbf{V}},\widehat{\mathbf{X}}),\, y\ge 0)$ gives a measure-valued process starting from $b\delta(x)$, whose total mass evolves according to a $\mathtt{BESQ}_b(0)$, as we will see later. To similarly construct a measure-valued excursion entering continuously from 0 initial mass, consider again $(\mathbf{V},\mathbf{X})$ as above. As we will demonstrate in Section 4, we can build scaffolding for immigration from excursions of $\mathbf{X}$ above its infimum process. Specifically, standard fluctuation theory for Lévy processes yields a σ-finite excursion measure $\nu_{\perp\mathrm{stb}}^{(\alpha)}$. In our setting we will define a σ-finite measure $\overline{\nu}_{\perp\mathrm{cld}}^{(\alpha)}$ on a suitable space of pairs (V,X) such that its pushforward onto the scaffolding X is $\nu_{\perp\mathrm{stb}}^{(\alpha)}$. The processes arising from $\overline{\nu}_{\perp\mathrm{cld}}^{(\alpha)}$ shall be interpreted as *clades starting from zero* and their associated scaffoldings. The superskewers of the processes arising from $\overline{\nu}_{\perp\mathrm{cld}}^{(\alpha)}$ are the measure-valued excursions, describing the evolutions of immigrant families and having $\mathtt{BESQ}(0)$ excursion total mass processes, alluded to at the end of Section 1.1.

We next construct $\mathtt{SSSP}(\alpha,\theta)$ and hence $\mathtt{FV}(\alpha,\theta)$ from these clades. For $\pi=\sum_{i\ge1} b_i\times\delta(x_i)\in\mathcal{M}^a$, consider independent $\mathbf{V}_i\sim\mathbf{Q}_{b_i,x_i}^{(\alpha)}$ and associated scaffolding $\mathbf{X}_i$, for all $i\ge1$

with $b_i > 0$. We write $\mathbf{Q}_\pi^{\alpha,0}$ for the distribution of the point measure $\mathbf{F} = \sum_{i\geq 1: b_i>0} \delta(\mathbf{V}_i, \mathbf{X}_i)$ on a suitable space of point measures on a suitable space of pairs (V, X) of point measures V of spindles and allelic types and scaffoldings X. See Section 2.2 for details. The superskewer at level $y \geq 0$ of $\mathbf{F} \sim \mathbf{Q}_\pi^{\alpha,0}$ for any $\pi \in \mathcal{M}^a$ is

$$\text{sSKEWER}(y, \mathbf{F}) := \sum_{\text{points } (V,X) \text{ of } \mathbf{F}} \text{sSKEWER}(y, V, X). \tag{1.7}$$

For $\mathbf{F} \sim \mathbf{Q}_0^{\alpha,\theta} := \text{PRM}(\frac{\theta}{\alpha}\text{Leb} \otimes \overline{\nu}_{\perp\text{cld}}^{(\alpha)})$, we interpret a point (z, V_z, X_z) as immigration at level $z \geq 0$ and define the superskewer as

$$\text{sSKEWER}(y, \mathbf{F}) := \sum_{\text{points } (z,V_z,X_z) \text{ of } \mathbf{F}} \text{sSKEWER}(y - z, V_z, X_z). \tag{1.8}$$

THEOREM 1.9. *Let $\alpha \in (0, 1)$ and $\pi \in \mathcal{M}^a$, and consider $\mathbf{F} \sim \mathbf{Q}_\pi^{\alpha,0}$.*

(i) *The process* $(\text{sSKEWER}(y, \mathbf{F}), y \geq 0)$ *is an* $\text{SSSP}_\pi(\alpha, 0)$ *as defined in/after Theorem 1.2.*

(ii) *Let $\theta > 0$ and $\mathbf{F} \sim \mathbf{Q}_0^{\alpha,\theta}$ independent of $\mathbf{F}$. Then*

$$\big(\text{sSKEWER}(y, \mathbf{F}) + \text{sSKEWER}(y, \mathbf{F}), y \geq 0\big) \tag{1.9}$$

is an $\text{SSSP}_\pi(\alpha, \theta)$ *as defined in/after Theorem 1.2.*

1.3. *Further properties of Fleming–Viot processes.* In this section, we state two further properties of $\text{FV}(\alpha, \theta)$ relating to the number of atoms at exceptional times and the α-diversity of the sequence of atom sizes as a proxy for the genetic diversity of types in the population modeled by a $\text{FV}(\alpha, \theta)$.

For $\pi \in \mathcal{M}^a$ denote by $N(\pi) \in \mathbb{N} \cup \{\infty\}$ the number of atoms of π. As $\text{PDRM}(\alpha, \theta)$ has infinitely many atoms a.s., for all $\theta \geq 0$, stationary $\text{FV}(\alpha, \theta)$ processes have infinitely many atoms a.s., at each time $u \geq 0$. However, the following result states that there are exceptional times where this is not so.

THEOREM 1.10. *Let $1 \leq n < \infty$. Then a* $\text{FV}(\alpha, \theta)$ *process visits $\{\pi \in \mathcal{M}^a : N(\pi) = n\}$ with positive probability if and only if $\theta + n\alpha < 1$.*

The corresponding result for Petrov's diffusions [34], which [23] identifies as the evolution of ranked atom sizes of a $\text{FV}(\alpha, \theta)$, was proved by Feng and Sun [17], Theorem 2.4, using Dirichlet form techniques. Our argument is based on the boundary behaviour of squared Bessel processes.

Let $\alpha \in (0, 1)$. For any $\pi \in \mathcal{M}^a$, if the following limit exists, then we say that the α-*diversity of* π is $\mathscr{D}_\alpha(\pi)$:

$$\mathscr{D}_\alpha(\pi) := \Gamma(1 - \alpha) \lim_{h\downarrow 0} h^\alpha \#\big\{x \in [0, 1] : \pi\big(\{x\}\big) > h\big\}. \tag{1.10}$$

It is well known that a $\text{PDRM}(\alpha, \theta)$ has an α-diversity almost surely, for all $\alpha \in (0, 1)$ and $\theta > -\alpha$. This really is a property of the ranked sequence of atom sizes, which is $\text{PD}(\alpha, \theta)$. We refer to Pitman [36], Theorem 3.13.

THEOREM 1.11. *If $\pi \in \mathcal{M}_1^a$ has an α-diversity, then* $\text{FV}_\pi(\alpha, \theta)$ *has a continuously evolving diversity process. If $\pi \in \mathcal{M}_1^a$ does not have an α-diversity, then* $\text{FV}_\pi(\alpha, \theta)$ *has α-diversities evolving continuously at all positive times.*

The corresponding result for the $(\alpha, 0)$- and (α, α)-IP evolutions of [22] is a consequence of [22], Theorems 1.2–1.3, in the special cases $\theta = 0$ and $\theta = \alpha$. Ruggiero et al. [42] study a modification of a Poisson–Dirichlet diffusion under which diversities evolve as a diffusion.

1.4. *Organization of the paper.* The structure of this paper is as follows. In Section 2, we recall from [20] and [21] relevant preliminaries about point measures of spindles and marked stable Lévy processes, which we enrich by further marking of jumps by independent allelic types. In Section 3, we study the special case $\theta = 0$ and prove Theorem 1.2 in this case, as well as Theorem 1.9(i). In Section 4, we study marked excursions of the stable process above its minimum and make precise the notion of the σ-finite measure $\overline{\nu}^{(\alpha)}_{\perp\mathrm{cld}}$. This is the key to completing the proof of Theorem 1.2 and Theorem 1.9 in Section 5, where we also establish Propositions 1.3–1.4 and Theorem 1.5. In Section 6 we turn to Fleming–Viot processes and prove Theorem 1.7. In Section 7, we further study Fleming–Viot processes, prove Theorems 1.10–1.11 and develop the connections to Shiga [44] described at the end of Section 1.1.

2. Preliminaries on marked Lévy processes and point measures. In this section, we recall from [20–22] material about point measures of spindles and scaffoldings used there for the construction of $\nu^{(-2\alpha)}_{\mathrm{BESQ}}$-interval-partition evolutions, but which we adapt here at the level of point measures to provide allelic types. We also show that the point measures and superskewers of Definition 1.8 are well defined.

2.1. *Spindles*: *Excursions to describe atom size evolutions.* Let $\mathcal{D}$ be the Skorokhod space of real-valued càdlàg functions $f : \mathbb{R} \to \mathbb{R}$ and $\mathcal{E}$ the subset of nonnegative excursions that are continuous except, possibly, at the beginning and at the end (to include incomplete excursions):

$$\mathcal{E} := \left\{ f \in \mathcal{D} \;\middle|\; \begin{array}{l} \exists\, z \in (0,\infty) \text{ s.t. } f|_{(-\infty,0)\cup[z,\infty)} = 0, \\ f \text{ positive and continuous on } (0,z) \end{array} \right\}. \tag{2.1}$$

We define the *lifetime* and *amplitude*, $\zeta, A : \mathcal{E} \to (0,\infty)$ via

$$\zeta(f) = \sup\{s \geq 0 : f(s) > 0\} \quad \text{and} \quad A(f) = \sup\{f(s), s \in [0, \zeta(f)]\}. \tag{2.2}$$

LEMMA 2.1 (Equation (13) in [25]). *Let* $B \sim \mathtt{BESQ}_z(-2\alpha)$ *for some* $z > 0$. *Then* $1/\zeta(B) \sim \mathtt{Gamma}[1+\alpha, z/2]$.

LEMMA 2.2 ([38], Section 3, [21], Section 2.3). *For the purpose of the following, for* $m > 0$ *let* $H^m : \mathcal{E} \to [0,\infty]$ *denote the first hitting time of level* m. *Then there exists a measure* $\nu^{(-2\alpha)}_{\mathrm{BESQ}}$ *on* $\mathcal{E}$ *such that, for every* $m > 0$,

$$\nu^{(-2\alpha)}_{\mathrm{BESQ}}\{f \in \mathcal{E} : A(f) > m\} = \frac{2\alpha(\alpha+1)}{\Gamma(1-\alpha)} m^{-1-\alpha},$$

and under the probability measure $\nu^{(-2\alpha)}_{\mathrm{BESQ}}(\cdot | A(f) > m)$, *the restricted canonical process* $(Z(y), y \in [0, H^m(Z)])$ *is a* $\mathtt{BESQ}_0(4+2\alpha)$ *stopped upon first hitting* m, *independent of* $(Z(H^m(Z)+s), s \geq 0) \sim \mathtt{BESQ}_m(-2\alpha)$. *Moreover,*

$$\nu^{(-2\alpha)}_{\mathrm{BESQ}}\{f \in \mathcal{E} : \zeta(f) > y\} = \frac{\alpha}{2^\alpha \Gamma(1-\alpha)\Gamma(1+\alpha)} y^{-1-\alpha}, \quad y > 0.$$

LEMMA 2.3 (Corollary 2.10 in [21]). *There is a unique* ζ*-disintegration* $\nu^{(-2\alpha)}_{\mathrm{BESQ}}(\cdot|\zeta)$ *of* $\nu^{(-2\alpha)}_{\mathrm{BESQ}}$ *with the scaling property that for all* $z \in (0,\infty)$ *and* $\mathbf{f} \sim \nu^{(-2\alpha)}_{\mathrm{BESQ}}(\cdot|\zeta = 1)$, *we have* $(z\mathbf{f}(s/z), s \geq 0) \sim \nu^{(-2\alpha)}_{\mathrm{BESQ}}(\cdot|\zeta = z)$.

2.2. *Scaffolding*: *Jumps describe births and deaths of atoms/allelic types*. As in Section 1.2, consider a Poisson random measure $\mathbf{V}$ on $[0,\infty) \times \mathcal{E} \times [0,1]$ with intensity $\mathtt{Leb} \otimes \nu_{\mathrm{BESQ}}^{(-2\alpha)} \otimes \mathtt{Unif}$, denoted by $\mathtt{PRM}[\mathtt{Leb} \otimes \nu_{\mathrm{BESQ}}^{(-2\alpha)} \otimes \mathtt{Unif}]$. Let $\mathbf{N} = \varphi(\mathbf{V})$, for the projection $\varphi(V)(dt, df) = V(dt, df, [0,1])$. Then $\mathbf{N} \sim \mathtt{PRM}[\mathtt{Leb} \otimes \nu_{\mathrm{BESQ}}^{(-2\alpha)}]$, as studied in [20, 21]. Let us make precise how $\mathbf{N}$, hence $\mathbf{V}$, induces the associated scaffolding $\mathbf{X}$. By standard mapping of PRMs,

$$\mathbf{M} := \sum_{\text{points } (t,f,x) \text{ of } \mathbf{V}} \delta\big(t, \zeta(f)\big) = \sum_{\text{points } (t,f) \text{ of } \mathbf{N}} \delta\big(t, \zeta(f)\big)$$

is a PRM on $[0,\infty) \times (0,\infty)$ with intensity $\mathtt{Leb} \otimes \Lambda$ given by Lemma 2.2 as

$$\Lambda(dz) = \nu_{\mathrm{BESQ}}^{(-2\alpha)}\{\zeta \in dz\} = \frac{\alpha(1+\alpha)}{2^{\alpha}\Gamma(1-\alpha)\Gamma(1+\alpha)} z^{-2-\alpha}\, dz.$$

Since Λ is a stable Lévy measure of index $1+\alpha \in (1,2)$, the following limit exists a.s. uniformly in t on compact subsets of $[0,\infty)$, for $N = \mathbf{N}$:

$$\xi_N(t) = \lim_{\varepsilon \downarrow 0} \Big(\int_{[0,t] \times \{g \in \mathcal{E}:\, \zeta(g) > \varepsilon\}} \zeta(f) N(du, df) - t \int_{(\varepsilon,\infty)} z \Lambda(dz) \Big). \tag{2.3}$$

We write $\mathbf{X} := \xi_{\mathbf{V}} := \xi_{\mathbf{N}} = (\xi_{\mathbf{N}}(t), t \geq 0)$ for this Lévy process. Then all jumps $\Delta\mathbf{X}(t) := \mathbf{X}(t) - \mathbf{X}(t-)$ of $\mathbf{X}$ are positive of size $\Delta\mathbf{X}(t) = \zeta(f)$, corresponding precisely to the points (t, f) of $\mathbf{N}$. We reserve the name "$\mathtt{Stable}(1+\alpha)$" to refer to spectrally positive stable Lévy processes with Lévy measure Λ. This process has Laplace exponent

$$\psi_{\alpha}(c) = 2^{-\alpha}\Gamma(1+\alpha)^{-1} c^{1+\alpha}. \tag{2.4}$$

For $(\mathcal{S}, d_{\mathcal{S}})$ a Borel subset of a complete and separable metric space, we denote by $\mathcal{N}(\mathcal{S})$ the space of boundedly finite point measures on that space. Let $\mathcal{V} := \mathcal{N}([0,\infty) \times \mathcal{E} \times [0,1])$. We equip $\mathcal{V}$ with the σ-algebra $\Sigma(\mathcal{V})$ generated by evaluation maps. The following result follows by standard marking properties of PRMs; cf. [21], Proposition 2.15, for the corresponding result with only spindle marks, but no allelic type marks.

LEMMA 2.4. *For $g \in \mathcal{D}$, denote by $\overline{\mu}_g$ the distribution of a point measure $\mathbf{V}_g = \sum_{t \in \mathbb{R}:\, \Delta g(t) > 0} \delta(t, f_t, x_t)$ that independently for each time t of a positive jump associates $(f_t, x_t) \sim \nu_{\mathrm{BESQ}}^{(-2\alpha)}(\cdot | \zeta = \Delta g(t)) \otimes \mathtt{Unif}$. Then $g \mapsto \overline{\mu}_g$ is a stochastic kernel from $\mathcal{D}$ to $\mathcal{V}$. Moreover, for $g = \mathbf{X} \sim \mathtt{Stable}(1+\alpha)$ we have $\mathbf{V}_{\mathbf{X}} \sim \mathtt{PRM}(\mathtt{Leb} \otimes \nu_{\mathrm{BESQ}}^{(-2\alpha)} \otimes \mathtt{Unif})$.*

Now consider $\widehat{\mathbf{V}} := \delta(0, \mathbf{f}, x) + \mathbf{V}|_{[0,T] \times \mathcal{E} \times [0,1]} \sim \mathbf{Q}_{b,x}^{(\alpha)}$ and $\widehat{\mathbf{X}} := \zeta(\mathbf{f}) + \mathbf{X}|_{[0,T]}$ for some $b > 0$, $x \in [0,1]$, an independent initial spindle $\mathbf{f} \sim \mathtt{BESQ}_b(-2\alpha)$ of lifetime $\zeta(\mathbf{f})$ and $T := \inf\{t \geq 0 \colon \zeta(\mathbf{f}) + \mathbf{X}(t) = 0\}$ as in Section 1.2. Then $T < \infty$ a.s. since $\mathbf{X}$ has zero mean and is spectrally positive, and the limit (2.3) exists for $N = \varphi(\widehat{\mathbf{V}})$, uniformly on $[0,T]$, with $\widehat{\mathbf{X}} = \xi_{\widehat{\mathbf{V}}}$, where we abuse notation and set $\xi_{\widehat{\mathbf{V}}}(t) = 0$ for $t \geq T$. Recall that we refer to $\widehat{\mathbf{V}}$ as a clade starting from $b\delta(x)$. The mass at type x evolves according to $\mathbf{f}$, and each atom (t, f_t, x_t) of $\widehat{\mathbf{V}}$, for $t > 0$, is interpreted as a descendant allelic type x_t, whose mass at level y is $f_t(y - \widehat{\mathbf{X}}(t-))$. Cf. Figure 1.

2.3. *Point measures of clades*. Recall from Section 1.2 our notation $\mathcal{M}$ for the set of finite Borel measures on $[0,1]$ and the Prokhorov distance

$$d_{\mathcal{M}}(\pi, \pi') = \inf\big\{\varepsilon > 0 \colon \pi(C) \leq \pi'(C^{\varepsilon}) + \varepsilon \text{ and } \pi'(C) \leq \pi(C^{\varepsilon}) + \varepsilon \ \forall C \in \mathcal{C}\big\},$$

where $\mathcal{C} = \{C \subseteq [0,1], C \text{ closed}\}$ and $C^{\varepsilon} = \{x \in [0,1] \colon \min_{y \in C} |x - y| \leq \varepsilon\}$ is the ε-thickening of C. The state space for our processes is the subspace $\mathcal{M}^a \subset \mathcal{M}$ of purely atomic

measures, which can be written in the form $\pi = \sum_{i\geq 1} b_i \delta(x_i)$ for some $b_i \geq 0$ and distinct $x_i \in [0,1]$, $i \geq 1$. Also recall $\mathcal{M}_1^a := \{\pi \in \mathcal{M}^a : \|\pi\| = 1\}$ for the subspace of atomic probability measures.

LEMMA 2.5 (2.13 in [11]). *The subsets $\mathcal{M}_1^a \subset \mathcal{M}^a \subset \mathcal{M}$ are Borel subsets of the complete and separable metric space $(\mathcal{M}, d_{\mathcal{M}})$.*

Denote by $(\mathcal{V}\times\mathcal{D})^\circ_{\text{fin}}$ the set of all pairs of a point measure $V \in \mathcal{V}$ and a scaffolding $X \in \mathcal{D}$ with the following additional finiteness properties:

(i) $V([0,\infty)\times\{f\in\mathcal{E}: \zeta(f) > z\}\times[0,1]) < \infty$ for all $z > 0$,
(ii) $\sum_{\text{points } (t,f,x) \text{ of } V} f(y - X(t-)) < \infty$ for all $y \in \mathbb{R}$.

Recall from (1.6) our notation $\mathbf{Q}^{(\alpha)}_{b,x}$ for the distribution on $\mathcal{V}$ of a clade $\widehat{\mathbf{V}}$ starting from $b\delta(x)$. Then the pair $(\widehat{\mathbf{V}}, \xi_{\widehat{\mathbf{V}}})$ takes values in $(\mathcal{V}\times\mathcal{D})^\circ_{\text{fin}}$ almost surely. On our spaces, σ-algebras generated by evaluation maps are Borel σ-algebras of complete separable metrics, so we may consider boundedly finite point measures on spaces such as $(\mathcal{V}\times\mathcal{D})^\circ_{\text{fin}}$. See [21], Supplement A, for measure-theoretic details, using the framework of [9, 10].

Now recall Definition 1.8 of the superskewer.

LEMMA 2.6. *For each $y > 0$, the map $(V, X) \mapsto \text{SSKEWER}(y, V, X)$ is measurable from the space $(\mathcal{V}\times\mathcal{D})^\circ_{\text{fin}}$ into the Prokhorov space $(\mathcal{M}^a, d_{\mathcal{M}})$.*

PROOF. Let $y > 0$. By property (ii), $\text{SSKEWER}(y, V, X)$ is a finite atomic measure on $[0,1]$, for all $(V, X) \in (\mathcal{V}\times\mathcal{D})^\circ_{\text{fin}}$. It therefore suffices to check that evaluation maps are measurable. Indeed, for any Borel set $A \subset [0,1]$,

$$\text{SSKEWER}(y, V, X)(A) = \int f\big(y - X(t-)\big)\mathbf{1}\{x \in A\}V(dt, df, dx)$$

is a measurable function of (V, X). □

PROPOSITION 2.7. *Consider $\pi = \sum_{i\geq 1} b_i\delta(x_i) \in \mathcal{M}^a$ and independent $\mathbf{V}_i \sim \mathbf{Q}^{(\alpha)}_{b_i,x_i}$, $i \geq 1$, with associated scaffolding $\mathbf{X}_i := \xi_{\mathbf{V}_i}$. Let $y > 0$. Then a.s.:*

(i) *at most finitely many $\mathbf{X}_i$ have height $\zeta_i^+ := \sup\{\mathbf{X}_i(t), t \in [0,\infty)\} > y$;*
(ii) *$\mathbf{F}_\pi = \sum_{i\geq 1:\, b_i>0} \delta(\mathbf{V}_i, \mathbf{X}_i)$ takes values in $\mathcal{N}((\mathcal{V}\times\mathcal{D})^\circ_{\text{fin}})$;*
(iii) *$\text{SSKEWER}(y, \mathbf{F}_\pi)$ is well defined and an $\mathcal{M}^a$-valued random variable;*
(iv) *$\pi \mapsto \mathbf{Q}^{\alpha,0}_\pi$, where $\mathbf{Q}^{\alpha,0}_\pi$ is the distribution of $\mathbf{F}_\pi$, is a stochastic kernel.*

PROOF. (i) This follows from [21], Lemmas 5.9 and 6.1, since the clades $\mathbf{V}_i$ for initial atom sizes b_i here yield projected clades $\mathbf{N}_i = \varphi(\mathbf{V}_i)$ for the construction of $\nu^{(-2\alpha)}_{\text{BESQ}}$-IP-evolutions starting from single intervals $(0, b_i)$ of length b_i in [21], with the same associated scaffolding $\mathbf{X}_i$.

(ii) Since $(\mathbf{V}_i, \xi_{\mathbf{V}_i})$ is $(\mathcal{V}\times\mathcal{D})^\circ_{\text{fin}}$-valued for all $i \geq 1$ almost surely and the number of points of $\mathbf{F}_\pi$ of height $\zeta_i^+ > y$ is finite for all $y > 0$, the point measure $\mathbf{F}_\pi$ is boundedly finite almost surely.

(iii) Lemma 2.6 yields that each $\text{SSKEWER}(y, \mathbf{V}_i, \mathbf{X}_i)$ is an $\mathcal{M}^a$-valued random variable. By (i), $\text{SSKEWER}(y, \mathbf{F}_\pi)$ is an a.s. finite sum of these.

(iv) First, there is a measurable enumeration of atoms $\pi \mapsto ((b_i, x_i), i \geq 1)$; see, for example, [30], Lemma 1.6. Second, $(b, x) \mapsto \mathbf{Q}^{(\alpha)}_{b,x}$ is measurable, by construction. Third, as the limit (2.3) exists a.s. uniformly for each $N = \mathbf{N}_i$, $i \geq 1$, the point measures $\mathbf{N}_i$ a.s. take

their values in a space on which $N \mapsto \xi_N$ is measurable, by [21], Proposition 2.14. The measurability of $\pi \mapsto \mathbf{Q}_\pi^{\alpha,0}$ follows as a composition and push-forward under measurable maps. □

COROLLARY 2.8. *Consider any $\pi_j \in \mathcal{M}^a$, $j \geq 1$, without common atom locations and $\pi := \sum_{j\geq 1} \pi_j \in \mathcal{M}^a$. Let $\mathbf{F}_j \sim \mathbf{Q}_{\pi_j}^{\alpha,0}$, $j \geq 1$, independent. Then $\sum_{j\geq 1} \mathbf{F}_j \sim \mathbf{Q}_\pi^{\alpha,0}$.*

2.4. *Continuity properties in point measures of spindles.* We record two results, one about uniform Hölder continuity of the spindles in a stopped PRM and one about the evolution of total mass of the superskewer.

LEMMA 2.9 (Proposition 6 of [20]). *Let $\mathbf{V} \sim \mathtt{PRM}(\mathtt{Leb} \otimes \nu_{\mathtt{BESQ}}^{(-2\alpha)} \otimes \mathtt{Unif})$ and $T \in (0,\infty)$ a random time. Set $\widetilde{\mathbf{V}} = \mathbf{V}|_{[0,T)}$, $\widetilde{\mathbf{N}} = \varphi(\widetilde{\mathbf{V}})$ and $\widetilde{\mathbf{X}} = \xi_{\widetilde{\mathbf{N}}}$. Let $\gamma \in (0, \min(1-\alpha, 1/2))$. Then a.s., the spindles of $\widetilde{\mathbf{N}}$ shifted to their birth levels according to $\widetilde{\mathbf{X}}$ can be partitioned into sequences $(g_j^n, j \geq 1)$, $n \geq 1$, such that for each level at most one spindle is alive from each sequence, and $|g_j^n(y) - g_j^n(x)| \leq D_n |y-x|^\gamma$ for all $x, y \in \mathbb{R}$, $j \geq 1$, $n \geq 1$, for a summable sequence $(D_n, n \geq 1)$ of (random) Hölder constants.*

PROPOSITION 2.10. *Let $\pi \in \mathcal{M}^a$. Then the total mass of the superskewer of $\mathbf{F} \sim \mathbf{Q}_\pi^{\alpha,0}$ evolves continuously a.s. Moreover, $(\|\mathrm{sSKEWER}(y, \mathbf{F})\|, y \geq 0)$ is a $\mathtt{BESQ}_{\|\pi\|}(0)$.*

PROOF. Let $\widehat{\mathbf{V}} \sim \mathbf{Q}_{b,x}^{(\alpha)}$ and $\widehat{\mathbf{X}} = \xi_{\widehat{\mathbf{V}}}$. Then [21], Proposition 3.8, shows that the skewer of $\widehat{\mathbf{N}} = \varphi(\widehat{\mathbf{V}})$ and hence its total mass evolve continuously a.s. By [22], Theorem 1.4, this total mass process is a $\mathtt{BESQ}_b(0)$. But since interval lengths of $\mathrm{SKEWER}(y, \widehat{\mathbf{N}}, \widehat{\mathbf{X}})$ are also the atom sizes $f_t(y - \widehat{\mathbf{X}}(t-))$ of $\mathrm{sSKEWER}(y, \widehat{\mathbf{V}}, \widehat{\mathbf{X}})$, the total masses coincide. Cf. Figure 1. For general $\pi \in \mathcal{M}^a$, the statements follow by additivity of total mass processes and of $\mathtt{BESQ}(0)$. □

2.5. *Clade statistics.* Let us introduce a scaling operator for $V \in \mathcal{V}$. For $c > 0$, we define

$$c \odot_{\mathrm{cld}}^{\alpha} V := \int \delta\big(c^{1+\alpha} t, c \odot_{\mathrm{spdl}} f, x\big) V(dt, df, dx), \tag{2.5}$$

where $c \odot_{\mathrm{spdl}} f := (cf(y/c), y \in \mathbb{R})$. Recall that for a pair $(V, X) \in \mathcal{V} \times \mathcal{D}$ with $X = \xi_V$, the heights of the scaffolding X correspond to times in the superskewer process $\mathrm{sSKEWER}(\cdot, V, X)$. In light of this, we define the *lifetime* of a clade $V \in \mathcal{V}$ for which ξ_V exists as

$$\zeta^+(V) := \zeta^+(\xi_V) := \sup_{t \in [0, \mathrm{len}(V)]} \xi_V(t), \tag{2.6}$$

where

$$\mathrm{len}(V) := \inf\big\{t > 0 \colon V\big([t,\infty) \times \mathcal{E} \times [0,1]\big) = 0\big\}. \tag{2.7}$$

Note that we have the following relations:

$$\zeta^+\big(c \odot_{\mathrm{cld}}^{\alpha} V\big) = c\zeta^+(V) \quad \text{and} \quad \mathrm{len}\big(c \odot_{\mathrm{cld}}^{\alpha} V\big) = c^{1+\alpha}\mathrm{len}(V). \tag{2.8}$$

LEMMA 2.11. *Let $\widehat{\mathbf{V}} \sim \mathbf{Q}_{b,x}^{(\alpha)}$. Then $c \odot_{\mathrm{cld}}^{\alpha} \widehat{\mathbf{V}} \sim \mathbf{Q}_{cb,x}^{(\alpha)}$ for all $c > 0$.*

PROOF. This follows directly from the scale invariance properties of $\mathtt{BESQ}(-2\alpha)$ and $\mathtt{Stable}(1+\alpha)$. Specifically, $\mathbf{f} \sim \mathtt{BESQ}_b(-2\alpha)$ yields $c \odot_{\mathrm{spdl}} \mathbf{f} = (c\mathbf{f}(y/c), y \geq 0) \sim \mathtt{BESQ}_{cb}(-2\alpha)$, and this is inherited by $\nu_{\mathtt{BESQ}}^{(-2\alpha)}$ in the form $\nu_{\mathtt{BESQ}}^{(-2\alpha)}(c \odot_{\mathrm{spdl}} \cdot) = c^{-1-\alpha} \nu_{\mathtt{BESQ}}^{(-2\alpha)}$. See, for example, [21], Lemma 2.9. □

PROPOSITION 2.12. *Consider a clade* $\widehat{\mathbf{V}} \sim \mathbf{Q}^{(\alpha)}_{b,x}$ *starting from* $b\delta(x)$ *with* $x \in [0,1]$, $b > 0$. *Set* $\widehat{\mathbf{X}} = \xi_{\widehat{\mathbf{V}}}$ *and* $\pi^y = \mathrm{SSKEWER}(y, \widehat{\mathbf{V}}, \widehat{\mathbf{X}})$, $y \geq 0$. *Then the lifetime* $\zeta^+(\widehat{\mathbf{V}})$ *of the clade has distribution* $\mathtt{InverseGamma}(1, b/2)$, *that is*,

$$\mathbf{P}\{\zeta^+(\widehat{\mathbf{V}}) > y\} = \mathbf{P}\{\pi^y \neq 0\} = 1 - e^{-b/2y}, \quad y \geq 0.$$

For any $y \geq 0$, *the distribution of* π^y *is* $Q^{(\alpha)}_{b,x,1/2y}$, *as defined in* (1.4).

PROOF. In the notation of [22], $\varphi(\widehat{\mathbf{V}}) \sim \nu_{\mathrm{cld}}(\cdot \mid m^0 = b)$. Then we obtain the law of its lifetime from [22], Proposition 3.4. Now fix $y \geq 0$. Recall that $\widehat{\mathbf{V}}$ has a point $(0, f_0, x)$ with $f_0 \sim \mathtt{BESQ}_b(-2\alpha)$. The mass of the leftmost spindle at level y is given by $m^y := f_r(y - \widehat{\mathbf{X}}(r-))$, where $r := \inf\{t \geq 0\colon (t, f_t, x_t)$ is a point of $\widehat{\mathbf{V}}$, $f_t(y - \widehat{\mathbf{X}}(t-)) > 0\}$. By [22], Lemma 3.5, the conditional distribution of m^y given $\{\pi^y \neq 0\}$ is the law of $L_{b,1/2y}$. The masses of the other blocks are conditionally independent of the leftmost block, given $\{\pi^y \neq 0\}$, and their distribution can be read from [22], Proof of Proposition 3.4: it is described by $(GA_i, i \geq 1)$, where $G \sim \mathtt{Gamma}(\alpha, 1/2y)$ and $(A_i, i \geq 1) \sim \mathtt{PD}(\alpha, \alpha)$ are independent.

Recall that the spindle f_0 has allelic type x and the others have i.i.d. $\mathtt{Unif}$ types. Let U_i, $i \geq 0$, be independent $\mathtt{Unif}$-distributed, independent of everything else. Then the conditional distribution of π^y given $\{\pi^y \neq 0\}$ is the law of $\lambda^y + G\overline{\Pi}$, where

$$\lambda^y := \mathbf{1}\{\zeta(f_0) > y\} m^y \delta(x) + \mathbf{1}\{\zeta(f_0) \leq y\} m^y \delta(U_0)$$

and $\overline{\Pi} := \sum_{i \geq 1} A_i \delta(U_i)$. Here λ^y and $G\overline{\Pi}$ represent respectively the contribution of the leftmost spindle at level y and the others. We have $\overline{\Pi} \sim \mathtt{PDRM}(\alpha, \alpha)$ by definition. It remains to study the law of λ^y. Using the [22], equations (A.6) and (A.7), we have the identity

$$\frac{\mathbf{P}(\zeta(f_0) > y, m^y \in dc \mid \zeta^+(\widehat{\mathbf{V}}) > y)}{\mathbf{P}(y \in [\zeta(f_0), \zeta^+(\widehat{\mathbf{V}})), m^y \in dc \mid \zeta^+(\widehat{\mathbf{V}}) > y)} = \frac{p^{(\alpha)}_{b,1/2y}(c)}{1 - p^{(\alpha)}_{b,1/2y}(c)}, \quad c > 0.$$

This implies that the conditional distribution of $\mathbf{1}\{\zeta(f_0) > y\}$ given $\{m^y = c\}$ is Bernoulli with parameter $p^{(\alpha)}_{b,1/2y}(c)$. Having already obtained the law of m^y, we can therefore identify the law of λ^y.

Summarizing, we deduce that the distribution of π^y is $Q^{(\alpha)}_{b,x,1/2y}$. $\square$

2.6. *Markov-like properties of point measures of spindles and clades.* Let $V \in \mathcal{V}$ and $X \in \mathcal{D}$ with $X = \xi_V$. Fix $y \geq 0$. If an atom (t, f_t, x_t) of V satisfies $y \in (X(t-), X(t))$, that is, the spindle f_t crosses level y, then we define $\hat{f}^y_t$ and $\check{f}^y_t$ to be its broken components split about that crossing. See Figure 2.

For a point measure $V \in \mathcal{V}$ and an interval $[a, b]$, the *shifted restriction* $V|^{\leftarrow}_{[a,b]} \in \mathcal{V}$ is defined as the restriction of V to $[a, b] \times \mathcal{E} \times [0, 1]$, translated so that it is supported on $[0, b-a] \times \mathcal{E} \times [0, 1]$. The shifted restriction of $X \in \mathcal{D}$, denoted by $X|^{\leftarrow}_{[a,b]} \in \mathcal{D}$, is defined correspondingly:

$$\begin{aligned} V|^{\leftarrow}_{[a,b]}([c,d] \times A) &= V\big(([c+a, d+a] \cap [a,b]) \times A\big), \quad A \in \Sigma(\mathcal{E} \times [0,1]), \\ X|^{\leftarrow}_{[a,b]}(t) &= \mathbf{1}\{t \in [0, b-a]\} X(t+a), \quad t \in \mathbb{R}. \end{aligned} \tag{2.9}$$

PROPOSITION 2.13 (Mid-spindle Markov property). *Fix* $y > 0$. *Let* $\mathbf{V} \sim \mathtt{PRM}(\mathtt{Leb} \otimes \nu^{(-2\alpha)}_{\mathtt{BESQ}} \otimes \mathtt{Unif})$, $\mathbf{X} = \xi_{\mathbf{V}}$ *and* $T = T^{\geq y} = \inf\{t \geq 0\colon \mathbf{X}(t) \geq y\}$. *Then* $(\mathbf{V}|^{\leftarrow}_{(T,\infty)}, \hat{f}^y_T, x_T)$ *is conditionally independent of* $(\mathbf{V}|^{\leftarrow}_{[0,T)}, \check{f}^y_T, x_T)$ *given the mass* $f_T(y - \mathbf{X}(T-)) = a$ *and type* $x_T = x$ *at level* y, *with regular conditional distribution* $\mathtt{PRM}(\mathtt{Leb} \otimes \nu^{(-2\alpha)}_{\mathtt{BESQ}} \otimes \mathtt{Unif}) \otimes \mathtt{BESQ}_a(-2\alpha) \otimes \delta_x$.

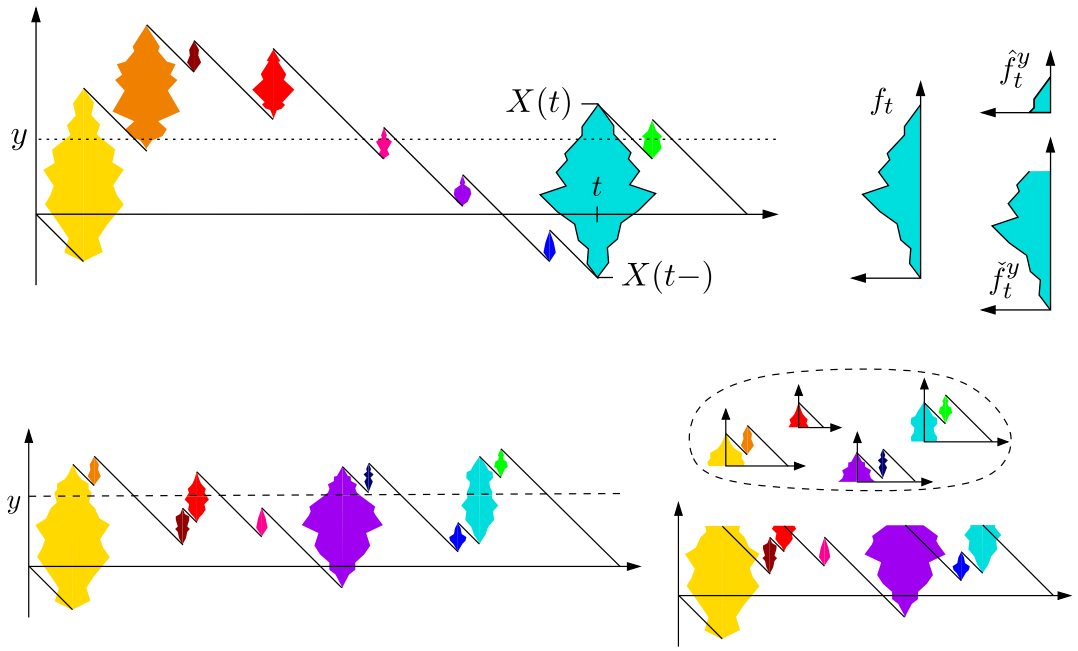

FIG. 2. *Spindle f_t of V at time t split about level y of the scaffolding X. Left: A scaffolding and spindles (V, X). Right bottom:* CUTOFF$^{\leq y}(V, X)$ *on its natural scaffolding; cf.* (3.3). *Right top:* $G_0^{\geq y}(V, X)$. *This is a point measure of clades, informally represented here as a "cloud" of clades, inside the dashed oval.*

PROOF. The Mid-spindle Markov property ([21], Lemma 4.13) of $\mathbf{N} := \varphi(\mathbf{V})$ at $T^{\geq y}$ states that $(\mathbf{N}|^{\leftarrow}_{(t,\infty)}, \hat{f}_T^y)$ is conditionally independent of $(\mathbf{N}|^{\leftarrow}_{[0,T)}, \check{f}_T^y)$ given $f_T(y - \mathbf{X}(T-)) = a$, with regular conditional distribution $\mathtt{PRM}(\mathtt{Leb} \otimes \nu_{\mathtt{BESQ}}^{(-2\alpha)}) \otimes \mathtt{BESQ}_a(-2\alpha)$. Hence, the proof is completed by independently marking all spindles by independent $\mathtt{Unif}$ allelic types, noting in particular that x_T is independent of everything else. □

For any level $y \geq 0$ and $(V, X) \in \mathcal{V} \times \mathcal{D}$, we define the *lower cutoff process*

$$\text{CUTOFF}^{\leq y}(V, X) := \sum_{\text{points } (t, f_t, x_t) \text{ of } V} \begin{pmatrix} \mathbf{1}\{y \in (X(t-), X(t))\}\delta(\Theta(t), \check{f}_t^y, x_t) \\ + \mathbf{1}\{X(t) \leq y\}\delta(\Theta(t), f_t, x_t) \end{pmatrix}, \tag{2.10}$$

where $\Theta(t) := \mathtt{Leb}\{u \leq t\colon X(u) \leq y\}$. This takes spindles f_t that X places below level y and the lower parts $\check{f}_t^y$ of spindles that X places across level y, and moves them (in time) so that intervals of excursions of X above level y are cut out. We also define the *upper point measure of clades*

$$G_0^{\geq y}(V, X) := \sum_{(a,b)} \delta\big(V|^{\leftarrow}_{(a,b)} + \mathbf{1}\{X(a) > y\}\delta(0, \hat{f}_a^y, x_a)\big), \tag{2.11}$$

where the sum is over all excursions of X above level y, recording for each excursion a clade consisting of the upper part of any initial spindle that X places across level y and the shifted point measure of spindles above level y. The cutoff process and upper point measure are illustrated in Figure 2.

PROPOSITION 2.14 (Markov-like property). *Let $\widehat{\mathbf{V}} \sim \mathbf{Q}_{b,x}^{(\alpha)}$, $\widehat{\mathbf{X}} = \xi_{\widehat{\mathbf{V}}}$ and $y \geq 0$. Then $G_0^{\geq y}(\widehat{\mathbf{V}}, \widehat{\mathbf{X}})$ is conditionally independent of* CUTOFF$^{\leq y}(\widehat{\mathbf{V}}, \widehat{\mathbf{X}})$ *given* SSKEWER$(y, \widehat{\mathbf{V}}, \widehat{\mathbf{X}}) = \pi$ *and has regular conditional distribution $\mathbf{Q}_\pi^{\alpha,0}$.*

PROOF. We define variants of $\text{CUTOFF}^{\leq y}(V, X)$ and $G_0^{\geq y}$ without allelic types,

$$\text{CUTOFF}^{\leq y}(N, X) := \sum_{\text{points } (t, f_t) \text{ of } N} \begin{pmatrix} \mathbf{1}\{y \in (X(t-), X(t))\}\delta(\Theta(t), \check{f}_t^y) \\ + \mathbf{1}\{X(t) \leq y\}\delta(\Theta(t), f_t) \end{pmatrix},$$

$$H_0^{\geq y}(N, X) := \sum_{(a,b)} \delta\big(N|_{(a,b)}^{\leftarrow} + \mathbf{1}\{X(a) > y\}\delta\big(0, \hat{f}_a^y\big)\big).$$

Then the Markov-like property of $\widehat{\mathbf{N}} := \varphi(\widehat{\mathbf{V}})$ at level y, [21], Proposition 5.6, yields that $\widehat{\mathbf{H}}_0^{\geq y} := H_0^{\geq y}(\widehat{\mathbf{N}}, \widehat{\mathbf{X}})$ is conditionally independent of the completion $\overline{\mathcal{F}}^y$ of the sigma-algebra $\sigma(\text{CUTOFF}^{\leq y}(\widehat{\mathbf{N}}, \widehat{\mathbf{X}}))$, given $\text{SKEWER}(y, \widehat{\mathbf{N}}, \widehat{\mathbf{X}})$, and a regular conditional distribution is provided by the distribution of

$$\sum_{\text{intervals } U \text{ of } \text{SKEWER}(y, \widehat{\mathbf{N}}, \widehat{\mathbf{X}})} \delta(\mathbf{N}_U),$$

where the $\mathbf{N}_U$ are conditionally independent. Moreover, comparing [21], Definition 5.1, with the definition of $\mathbf{Q}_{b,x}^{(\alpha)}$ above Definition 1.8 we see that $\mathbf{N}_U \sim \mathbf{Q}_{b,x}^{(\alpha)}(\varphi \in \cdot)$ for $b = \texttt{Leb}(U)$ and any $x \in [0, 1]$.

On the event $\{\text{SKEWER}(y, \widehat{\mathbf{N}}, \widehat{\mathbf{X}}) = \varnothing\} = \{\text{sSKEWER}(y, \widehat{\mathbf{V}}, \widehat{\mathbf{X}}) = 0\}$, the conditional independence and distribution trivially extend to the completion $\overline{\mathcal{G}}^y$ of the bigger σ-algebra $\sigma(\text{CUTOFF}^{\leq y}(\widehat{\mathbf{V}}, \widehat{\mathbf{X}}))$. See, for example, [21], equation (5.8).

On the event $\{\text{SKEWER}(y, \widehat{\mathbf{N}}, \widehat{\mathbf{X}}) \neq \varnothing\} = \{\text{sSKEWER}(y, \widehat{\mathbf{V}}, \widehat{\mathbf{X}}) \neq 0\}$, properties of $\texttt{Stable}(1 + \alpha)$ Lévy processes (see, e.g., [21], Proposition A.3) yield that there are, up to a null set, infinitely many points (t_i, f_i, x_i) of $\widehat{\mathbf{V}}$ that correspond to jumps of $\widehat{\mathbf{X}}$ across level y, with nontrivial lower and upper parts $\check{f}_i^y$ and $\hat{f}_i^y$, $i \geq 1$. Let $b_i := f_i(y - \widehat{\mathbf{X}}(t_i-))$, $i \geq 1$.

We recall that the conditional distribution of $\widehat{\mathbf{V}}$ given $\widehat{\mathbf{N}}$ is given by the marking kernel that independently marks each point (t, f_t) of $\widehat{\mathbf{N}}$ by an independent allelic type x_t. By elementary arguments based on the chain rule for conditional independence (e.g., [29], Proposition 6.8), this yields the conditional independence and distribution as statements conditionally given $\text{SKEWER}(y, \widehat{\mathbf{N}}, \widehat{\mathbf{X}})$ with intervals of lengths b_i and allelic types x_i, $i \geq 1$. But since this conditional distribution depends on $\text{SKEWER}(y, \widehat{\mathbf{N}}, \widehat{\mathbf{X}})$ and $(b_i, x_i)_{i \geq 1}$ only via $\pi = \sum_{i \geq 1} b_i \delta(x_i)$, the conditional independence and distribution also hold conditionally given $\text{sSKEWER}(y, \widehat{\mathbf{V}}, \widehat{\mathbf{X}}) = \pi$. □

3. Self-similar $(\alpha, 0)$-superprocesses.

The construction of $\texttt{SSSP}(\alpha, 0)$ and the proof of its properties follows the corresponding steps for $\nu_{\text{BESQ}}^{(-2\alpha)}$-interval partition evolutions in [21, 22]. For both processes, a substantial part of the technical work is carried out at the level of point measures in the framework of marked Lévy processes, as collected in Section 2.

In the present section, we use the superskewer to study the associated $\mathcal{M}^a$-valued processes. This also bears some similarities with the approach in [21, 22], but there are significant variations for several reasons. On the one hand, (natural) topologies on the state spaces involved are rather different, and our arguments crucially rely on both the topology of weak convergence and on total variation continuity. On the other hand, Lévy processes and interval partitions have an intrinsic left-to-right order, which is not captured in the state space $\mathcal{M}^a$.

3.1. *Path-continuity.* In addition to $d_{\mathcal{M}}$, we also consider the metric d_{TV} on $\mathcal{M}^a$ induced by the total variation norm: for any $\pi, \pi' \in \mathcal{M}^a$,

$$d_{\text{TV}}(\pi, \pi') := \sup_{B \in \mathcal{B}([0,1])} \big(|\pi(B) - \pi'(B)| + |\pi(B^c) - \pi'(B^c)|\big),$$

where $\mathcal{B}([0,1])$ denotes the Borel sets on $[0,1]$. Since the Prokhorov metric $d_{\mathcal{M}}$, unlike d_{TV}, induces a separable topology on $\mathcal{M}^a$, we use $d_{\mathcal{M}}$ where separability matters. But we know from [24], p.60–61, that the metric d_{TV} is Borel measurable on the metric space $(\mathcal{M}^a, d_{\mathcal{M}})$. It is therefore meaningful to study path-continuity for the stronger topology induced by d_{TV}.

PROPOSITION 3.1. *Let* $\mathbf{V} \sim \mathtt{PRM}(\mathtt{Leb} \otimes \nu_{\mathrm{BESQ}}^{(-2\alpha)} \otimes \mathtt{Unif})$ *and* $T \in (0,\infty)$ *be a random time. Set* $\widetilde{\mathbf{V}} = \mathbf{V}|_{[0,T)\times\mathcal{E}\times[0,1]}$. *Then the superskewer process* $(\mathrm{sSKEWER}(y, \widetilde{\mathbf{V}}, \xi_{\widetilde{\mathbf{V}}}), y \in \mathbb{R})$ *of the stopped* PRM *is a.s. Hölder-*γ *in* $(\mathcal{M}^a, d_{\mathrm{TV}})$ *for every* $\gamma \in (0, \min(1-\alpha, 1/2))$.

PROOF. Let $\widetilde{\mathbf{X}} := \xi_{\widetilde{\mathbf{V}}}$ and fix $\gamma \in (0, \min(1-\alpha, 1/2))$. Applying Lemma 2.9, we partition the marked spindles of $\widetilde{\mathbf{V}}$ into sequences $((g_j^n, x_j^n), j \geq 1)$ for $n \geq 1$. For every $y \in \mathbb{R}$, let $\pi^n(y) := \sum_{j\geq 1} g_j^n(y)\delta(x_j^n)$. Since $g_j^n(y) = 0$ for all but at most one $j \geq 1$, this captures a single type x_j^n for each $n \geq 1$. Furthermore, there is the identity

$$\mathrm{sSKEWER}(y, \widetilde{\mathbf{V}}, \widetilde{\mathbf{X}}) = \sum_{n\geq 1} \pi^n(y), \quad y \in \mathbb{R}.$$

Fix $y < z$. Let $A := \{n \geq 1 \colon \pi^n(r) \neq 0 \text{ for all } r \in [y,z]\}$. That is, A is the set of indices n for which $\widetilde{\mathbf{X}}$ places a single spindle in the sequence $(g_j^n, j \geq 1)$ across both levels y and z, that is, the type x_j^n corresponding to this spindle has positive mass during the entire interval $[y,z]$. Then we have

$$\begin{aligned} &d_{\mathrm{TV}}\big(\mathrm{sSKEWER}(y, \widetilde{\mathbf{V}}, \widetilde{\mathbf{X}}), \mathrm{sSKEWER}(z, \widetilde{\mathbf{V}}, \widetilde{\mathbf{X}})\big) \\ &\quad \leq \sum_{n\in A} \big|\|\pi^n(y)\| - \|\pi^n(z)\|\big| + \sum_{n\notin A} \|\pi^n(y)\| + \sum_{n\notin A} \|\pi^n(z)\|. \end{aligned} \tag{3.1}$$

By Lemma 2.9, a.s. each process $(\|\pi^n(r)\|, r \geq 0)$ is γ-Hölder continuous with Hölder constant D_n. Then we have $\big|\|\pi^n(y)\| - \|\pi^n(z)\|\big| \leq D_n|y-z|^\gamma$ for each $n \geq 1$. Moreover, for $n \notin A$, there exists some level $x \in [y,z]$ such that $\pi^n(x) = 0$. It follows that

$$\forall n \notin A, \ \max\big(\|\pi^n(y)\|, \|\pi^n(z)\|\big) \leq D_n|y-z|^\gamma.$$

So we deduce from (3.1) that $d_{\mathrm{TV}}(\pi^y, \pi^z) \leq 2\sum_{n\geq 1} D_n|y-z|^\gamma$. Since $(D_n, n \geq 1)$ is summable, we deduce the γ-Hölder continuity. □

COROLLARY 3.2. *For* $b > 0$ *and* $x \in [0,1]$, *let* $\widehat{\mathbf{V}} \sim \mathbf{Q}_{b,x}^{(\alpha)}$. *Set* $\widehat{\mathbf{X}} = \xi_{\widehat{\mathbf{V}}}$. *Then the superskewer process* $(\mathrm{sSKEWER}(y, \widehat{\mathbf{V}}, \widehat{\mathbf{X}}), y \geq 0)$ *is a.s. Hölder-*γ *in* $(\mathcal{M}^a, d_{\mathrm{TV}})$ *for every* $\gamma \in (0, \min(1-\alpha, 1/2))$.

PROOF. Let $\mathbf{V} \sim \mathtt{PRM}(\mathtt{Leb} \otimes \nu_{\mathrm{BESQ}}^{(-2\alpha)} \otimes \mathtt{Unif})$ and $\mathbf{f} \sim \mathtt{BESQ}_b(-2\alpha)$ be independent. Set $\mathbf{X} = \xi_{\mathbf{V}}$ and

$$\widetilde{\mathbf{V}} = \mathbf{V}|_{[0,T)\times\mathcal{E}\times[0,1]}, \quad \text{where } T := \inf\{t \geq 0 \colon \mathbf{X}(t) \leq -\zeta(\mathbf{f})\}.$$

By the definition of $\mathbf{Q}_{b,x}^{(\alpha)}$, we can write $\widehat{\mathbf{V}} = \widetilde{\mathbf{V}} + \delta(0, \mathbf{f}, x)$. Then we have the identity, for every $y \geq 0$,

$$\mathrm{sSKEWER}(y, \widehat{\mathbf{V}}, \widehat{\mathbf{X}}) = \mathrm{sSKEWER}\big(y - \zeta(\mathbf{f}), \widetilde{\mathbf{V}}, \xi_{\widetilde{\mathbf{V}}}\big) + \mathbf{1}\{y \leq \zeta(\mathbf{f})\}\mathbf{f}(y)\delta(x).$$

Applying Proposition 3.1 to $\widetilde{\mathbf{V}}$ and noting, for example, by [20], Corollary 34, that the $\mathtt{BESQ}(-2\alpha)$ process $\mathbf{f}$ is Hölder-γ for every $\gamma \in (0, 1/2)$, we deduce that $(\mathrm{sSKEWER}(\widehat{\mathbf{V}}, \widehat{\mathbf{X}}), y \geq 0)$ is the sum of two processes that are Hölder-γ in $(\mathcal{M}^a, d_{\mathrm{TV}})$, for any $\gamma < \min(1-\alpha, 1/2)$. This completes the proof. □

PROPOSITION 3.3. *Let $\pi \in \mathcal{M}^a$ and $\mathbf{F} \sim \mathbf{Q}_\pi^{\alpha,0}$. Then* $(\mathrm{sSKEWER}(y, \mathbf{F}),\ y \geq 0)$ *a.s. has continuous paths in* $(\mathcal{M}^a, d_{\mathrm{TV}})$ *starting from* $\mathrm{sSKEWER}(0, \mathbf{F}) = \pi$. *Moreover, it is a.s. Hölder-γ in* $(\mathcal{M}^a, d_{\mathrm{TV}})$ *for every* $\gamma \in (0, \min\{1-\alpha, 1/2\})$ *for all* $y \in (0, \infty)$.

Since the topology generated by d_{TV} is stronger than the topology generated by $d_{\mathcal{M}}$, Proposition 3.3 implies that the process also has continuous paths in $(\mathcal{M}^a, d_{\mathcal{M}})$.

PROOF. Our proof is an adaptation of arguments in the proof of [21], Theorem 1.3. For completeness, let us sketch it here.

For every $z > 0$, by Proposition 2.7, a.s. the post-z process is equal to the sum of a finite number of processes $(\mathrm{sSKEWER}(y, \widehat{\mathbf{V}}, \widehat{\mathbf{X}}),\ y \geq z)$, with $\widehat{\mathbf{V}} \sim \mathbf{Q}_{b,x}^{(\alpha)}$ and $\widehat{\mathbf{X}} = \xi_{\widehat{\mathbf{V}}}$. By Corollary 3.2, each of these superskewer processes is a.s. γ-Hölder continuous. Therefore, the sum is also a.s. γ-Hölder continuous. Since $z > 0$ is arbitrary, it only remains to prove the continuity at level 0.

Write $\pi = \sum_{i \geq 1} b_i \delta(x_i)$. Fix $\epsilon > 0$ and take $k \in \mathbb{N}$ large enough such that we have $b_0 := \|\pi\| - (\sum_{i=1}^k b_i) < \epsilon$. Recall the definition of $\mathbf{Q}_\pi^{\alpha,0}$, the law of $\mathbf{F}$, then we can write

$$\mathrm{sSKEWER}(y, \mathbf{F}) = \lambda^y + \sum_{i=1}^k f_i(y)\delta(x_i), \quad y \geq 0,$$

where $(f_i, i \leq k)$ are independent $\mathtt{BESQ}(-2\alpha)$ starting from b_i respectively.

By the continuity of the total mass process of Proposition 2.10, and by the continuity of $\mathtt{BESQ}(-2\alpha)$, there exists $h > 0$ such that for every $y < h$:

(i) $|\|\pi\| - (\|\lambda^y\| + \sum_{i=1}^k |f_i(y)|)| < \epsilon$;
(ii) for $i \leq k$, $|f_i(y) - b_i| < \epsilon/k$.

It follows from the first inequality and the choice of k that

$$\|\lambda^y\| \leq \sum_{i=1}^k |f_i(y) - b_i| + \epsilon + b_0 < 3\epsilon.$$

We conclude that

$$d_{\mathrm{TV}}\big(\pi, \mathrm{sSKEWER}(y, \mathbf{F})\big) \leq \sum_{i=1}^k |f_i(y) - b_i| + \|\lambda^y\| + b_0 < 5\epsilon.$$

This proves the continuity at level 0. □

3.2. *Markov property.* Let $\pi \in \mathcal{M}^a$. In Proposition 3.3, we saw that the superskewer process is $\mathbf{Q}_\pi^{\alpha,0}$-a.s. d_{TV}-path-continuous. In particular, it has a distribution on the space $\mathcal{C}([0, \infty), \mathcal{M}^a)$ of continuous paths in the separable metric space $(\mathcal{M}^a, d_{\mathcal{M}})$, since the total variation topology is stronger than the weak topology induced by the Prokhorov metric $d_{\mathcal{M}}$. We denote this distribution on $\mathcal{C}([0, \infty), \mathcal{M}^a)$ by $\mathbb{Q}_\pi^{\alpha,0}$.

LEMMA 3.4. *The map $\pi \mapsto \mathbb{Q}_\pi^{\alpha,0}$ is a stochastic kernel.*

PROOF. Consider the subset $(\mathcal{V} \times \mathcal{D})^* \subset (\mathcal{V} \times \mathcal{D})^\circ$ of pairs (V, X) with

$$\mathrm{sS\overline{KE}WER}(V, X) := \big(\mathrm{sSKEWER}(y, V, X),\ y \geq 0\big) \in \mathcal{C}\big([0, \infty), \mathcal{M}^a\big). \tag{3.2}$$

Recall that the Borel σ-algebra on $\mathcal{C}([0, \infty), \mathcal{M}^a)$ generated by the topology of uniform convergence on compact subsets of $[0, \infty)$ is also generated by the evaluation maps [29],

Theorem 12.5. Hence, Lemma 2.6 implies the measurability of $\overline{\text{SSKEWER}}$ as a function from $(\mathcal{V}\times\mathcal{D})^*$ to $\mathcal{C}([0,\infty),\mathcal{M}^a)$. By Proposition 3.3, $\mathbf{Q}_\pi^{\alpha,0}((\mathcal{V}\times\mathcal{D})^*)=1$. Therefore, this lemma follows from the kernel property of $\pi\mapsto\mathbf{Q}_\pi^{\alpha,0}$ established in Proposition 2.7(iv). $\square$

Let us now state the Markov property of the superskewer process in terms of the kernel $\pi\mapsto\mathbb{Q}_\pi^{\alpha,0}$. We also use notation $\mathbb{Q}_\mu^{\alpha,0}:=\int_{\mathcal{M}^a}\mathbb{Q}_\pi^{\alpha,0}\mu(d\pi)$ for Borel probability measures μ on $(\mathcal{M}^a,d_{\mathcal{M}})$.

PROPOSITION 3.5 (Markov property). *Let μ be a probability measure on $\mathcal{M}^a$ and $(\pi^z,z\ge0)\sim\mathbb{Q}_\mu^{\alpha,0}$. For any $y\ge0$, the process $(\pi^{y+r},r\ge0)$ is conditionally independent of $(\pi^z,z\le y)$ given $\pi^y=\pi$ and has regular conditional distribution $\mathbb{Q}_\pi^{\alpha,0}$.*

PROOF. Fix $b>0$, $x\in[0,1]$, let $\widehat{\mathbf{V}}\sim\mathbf{Q}_{b,x}^{(\alpha)}$ and $\widehat{\mathbf{X}}=\xi_{\widehat{\mathbf{V}}}$. By the Markov-like property of Proposition 2.14, the upper point measure of clades $G_0^{\ge y}(\widehat{\mathbf{V}},\widehat{\mathbf{X}})$ is conditionally independent of the lower cutoff process $V:=\text{CUTOFF}^{\le y}(\widehat{\mathbf{V}},\widehat{\mathbf{X}})$ given $\text{SSKEWER}(y,\widehat{\mathbf{V}},\widehat{\mathbf{X}})=\pi$ and has regular conditional distribution $\mathbf{Q}_\pi^{\alpha,0}$. Arguing as in the proof of Proposition 2.14, we can apply results from [21] in the present context. Specifically, [21], (5.3) and Lemma 5.5, yield

$$\text{SSKEWER}(z,V,\xi_V)=\text{SSKEWER}(z,\widehat{\mathbf{V}},\widehat{\mathbf{X}})\quad\text{for all } z\le y. \tag{3.3}$$

Now recall that $\mathbb{Q}_\pi^{\alpha,0}$ is the distribution of $(\text{SSKEWER}(r,\cdot),r\ge0)$ under $\mathbf{Q}_\pi^{\alpha,0}$. Applying $(\text{SSKEWER}(z,\cdot),z\le y)$ to (V,ξ_V) and $(\text{SSKEWER}(r,\cdot),r\ge0)$ to $G_0^{\ge y}(\widehat{\mathbf{V}},\widehat{\mathbf{X}})$, we find $(\text{SSKEWER}(y+r,\widehat{\mathbf{V}},\widehat{\mathbf{X}}),r\ge0)$ is conditionally independent of $(\text{SSKEWER}(z,\widehat{\mathbf{V}},\widehat{\mathbf{X}}),z\le y)$ given $\text{SSKEWER}(y,\widehat{\mathbf{V}},\widehat{\mathbf{X}})=\pi$, with conditional distribution $\mathbb{Q}_\pi^{\alpha,0}$.

For any $\pi'=\sum_{j\ge1}b_j\delta(x_j)\in\mathcal{M}^a$, we recall the construction of $\mathbf{F}_{\pi'}\sim\mathbf{Q}_{\pi'}^{\alpha,0}$ of Proposition 2.7 from independent $\mathbf{V}_j\sim\mathbf{Q}_{b_j,x_j}^{(\alpha)}$, $j\ge1$. By this independence, the conditional independence extends to $(\text{SSKEWER}(z,\mathbf{F}_{\pi'}),z\le y)$ and $(\text{SSKEWER}(y+r,\mathbf{F}_{\pi'}),r\ge0)$ given $(\text{SSKEWER}(y,\mathbf{V}_j,\xi_{\mathbf{V}_j}),j\ge1)=(\pi_j,j\ge1)$. The conditional distribution $\mathbb{Q}_\pi^{\alpha,0}$ of $(\text{SSKEWER}(y+r,\mathbf{F}_{\pi'}),r\ge0)$ is a consequence of Corollary 2.8; here $\pi:=\sum_{j\ge1}\pi_j$. Specifically, we note that $\pi\in\mathcal{M}^a$ a.s. by Proposition 2.7(iii); also the x_j, $j\ge1$, are distinct, and the independence of $(\mathbf{V}_j,\xi_{\mathbf{V}_j})$, $j\ge1$, each based on a countable family of independent `Unif` allelic types entails that the atom locations of π_j, $j\ge1$, are distinct a.s. Since this conditional distribution only depends on $(\pi_j,j\ge1)$ via $\pi=\sum_{j\ge1}\pi_j$, the conditional independence holds conditionally given just $\text{SSKEWER}(y,\mathbf{F}_{\pi'})=\pi$.

Since $\mathbb{Q}_{\pi'}^{\alpha,0}$ is the distribution of $(\text{SSKEWER}(z,\cdot),z\ge0)$ under $\mathbf{Q}_{\pi'}^{\alpha,0}$, this can be phrased purely in terms of $(\pi^z,z\ge0)\sim\mathbb{Q}_{\pi'}^{\alpha,0}$. The passage from fixed $\pi'\in\mathcal{M}^a$ to distributions μ on $\mathcal{M}^a$ is straightforward. $\square$

To strengthen the simple Markov property to a strong Markov property, we will use the standard approximation of a general finite stopping time by a decreasing sequence of stopping times taking values in refining discrete time grids. This argument requires some continuity of $\pi\mapsto\mathbb{Q}_\pi^{\alpha,0}$ along suitable sequences of states arising along such sequences of stopping times.

PROPOSITION 3.6 (Continuity in the initial state). *For a sequence $\pi_n\to\pi_\infty$ in $(\mathcal{M}^a,d_{\text{TV}})$, there is the weak convergence $\mathbb{Q}_{\pi_n}^{\alpha,0}\to\mathbb{Q}_{\pi_\infty}^{\alpha,0}$ in the sense of finite-dimensional distributions on $(\mathcal{M}^a,d_{\mathcal{M}})$.*

Note that we cannot expect a similar result under the weaker assumption that $\pi_n\to\pi_\infty$ under the Prokhorov metric. For example, with $\pi_n=\frac12\delta(2^{-1}-2^{-n-1})+\frac12\delta(2^{-1}+2^{-n-1})$,

we have $\pi_n \to \pi_\infty = \delta(2^{-1})$ weakly. Let $(\pi_n^y, y \geq 0) \sim \mathbb{Q}^{\alpha,0}_{\pi_n}$ and $(\pi_\infty^y, y \geq 0) \sim \mathbb{Q}^{\alpha,0}_{\pi_\infty}$, and consider $m_n^y := \pi_n^y((2^{-1} - \epsilon, 2^{-1} + \epsilon))$, $y \geq 0$, the mass evolution in an open interval with an arbitrarily small $\epsilon > 0$: for n large enough and $y_0 > 0$ small enough, the law of $(m_n^y, y \leq y_0)$ is "close" to the law of the sum of two independent $\mathrm{BESQ}_{1/2}(-2\alpha)$. But $(m_\infty^y, y \leq y_0)$ "nearly" has the law of a $\mathrm{BESQ}_1(-2\alpha)$.

PROOF. We fix $y > 0$ and shall establish the convergence of the one-dimensional distributions at level y; the multi-dimensional version can be proved inductively. See, for example, [21], Corollary 6.16, for an instance of this inductive argument that is easily adapted. Indeed, the convergence of one-dimensional distributions is also adapted from [21], Proposition 6.15, but we provide the details so as to be clear about the different topologies.

Suppose that $\pi_\infty = \sum_{i\geq 1} b_i \delta(x_i)$ for a nonincreasing sequence $(b_i)_{i\geq 1}$ of nonnegative numbers and distinct $x_i \in [0, 1]$, $i \geq 1$. Let us construct, on a suitable probability space, a family of coupled measure-valued processes $\boldsymbol{\lambda}_n \sim \mathbb{Q}^{\alpha,0}_{\pi_n}$, $n \in \mathbb{N} \cup \{\infty\}$. Specifically, let $(\mathbf{V}_{\infty,i}, i \geq 1)$ be a family of independent random point measures with $\mathbf{V}_{\infty,i} \sim \mathbf{Q}^{(\alpha)}_{b_i,x_i}$ and $\mathbf{X}_{\infty,i} := \xi_{\mathbf{V}_{\infty,i}}$ associated scaffolding. Set the process $(\lambda^z_{\infty,i}, z \geq 0)$ to be a d_{TV}-continuous version of $\mathrm{SSKEWER}(\mathbf{V}_{\infty,i}, \mathbf{X}_{\infty,i})$, whose existence is guaranteed by Corollary 3.2. By Proposition 3.3, we may further assume that $\lambda^z_\infty := \sum_{i\geq 1} \lambda^z_{\infty,i}$, $z \geq 0$, is d_{TV}-continuous. For an arbitrarily small $\epsilon > 0$, by Proposition 2.7 we can choose $\ell > 0$ large enough such that

$$\mathbf{P}\{\|\lambda^y_\infty\| > \ell\} < \epsilon. \tag{3.4}$$

We take the smallest $k \geq 1$ large enough so that $\sum_{i>k} b_i < \epsilon$. By Proposition 2.12, we have

$$\mathbf{P}\Big\{\sum_{i>k} \lambda^y_{\infty,i} \neq 0\Big\} \leq \frac{1}{2y} \sum_{i>k} b_i \leq \frac{\epsilon}{2y}. \tag{3.5}$$

Moreover, due to the path-continuity, there exists (random) $\Delta > 0$, such that for any $z \in (y - \Delta, y + \Delta)$, there is

$$d_{\mathrm{TV}}(\lambda^z_{\infty,i}, \lambda^y_{\infty,i}) < \frac{\epsilon}{k} \quad \text{for all } 1 \leq i \leq k. \tag{3.6}$$

By the convergence $d_{\mathrm{TV}}(\pi_n, \pi_\infty) \to 0$, there exists some $m_0 \geq 1$ large enough such that for all $n \geq m_0$, we can write $\pi_n = \sum_{i=1}^k b_{n,i}\delta(x_i) + \widetilde{\pi}_n$ with $\widetilde{\pi}_n$ having no mass at any x_i, $1 \leq i \leq k$, with $\|\widetilde{\pi}_n\| < \epsilon$, and we can choose (random) $M \geq m_0$ so that for all $1 \leq i \leq k$, $n > M$, we have $b_{n,i} > 0$,

$$|c_{n,i} - 1| < \frac{\epsilon}{\ell} \wedge 1, \quad \text{and} \quad \left|\frac{y}{c_{n,i}} - y\right| < \Delta, \quad \text{where } c_{n,i} := \frac{b_{n,i}}{b_i}. \tag{3.7}$$

Next, we set $\mathbf{V}_{n,i} := c_{n,i} \odot^\alpha_{\mathrm{cld}} \mathbf{V}_{\infty,i}$ and $\mathbf{X}_{n,i} := \xi_{\mathbf{V}_{n,i}}$ for $n \geq 1$, and we note the identity

$$\mathrm{SSKEWER}(y, \mathbf{V}_{n,i}, \mathbf{X}_{n,i}) = c_{n,i}\mathrm{SSKEWER}(y/c_{n,i}, \mathbf{V}_{\infty,i}, \mathbf{X}_{\infty,i}) = c_{n,i}\lambda^{y/c_{n,i}}_{\infty,i}.$$

Moreover, by Lemma 2.11, we have $\mathbf{V}_{n,i} \sim \mathbf{Q}^{(\alpha)}_{b_{n,i},x_i}$. For $n \geq 1$, let $\widetilde{\boldsymbol{\lambda}}_n \sim \mathbb{Q}^{\alpha,0}_{\widetilde{\pi}_n}$, independent of anything else. By Proposition 2.12, we have for all $n \geq m_0$

$$\mathbf{P}\{\widetilde{\lambda}^y_n \neq 0\} \leq \frac{\|\widetilde{\pi}_n\|}{2y} \leq \frac{\epsilon}{2y}. \tag{3.8}$$

Then $(\lambda^y_n, y \geq 0) := (\widetilde{\lambda}^y_n + \sum_{1\leq i\leq K} \mathrm{SSKEWER}(y, \mathbf{V}_{n,i}, \mathbf{X}_{n,i}), y \geq 0) \sim \mathbb{Q}^{\alpha,0}_{\pi_n}$.

For any $n \geq M$, we have by (3.6) and (3.7) that

$$\begin{aligned} d_{\mathrm{TV}}(c_{n,i}\lambda^{y/c_{n,i}}_{\infty,i}, \lambda^y_{\infty,i}) &\leq d_{\mathrm{TV}}(c_{n,i}\lambda^{y/c_{n,i}}_{\infty,i}, c_{n,i}\lambda^y_{\infty,i}) + d_{\mathrm{TV}}(c_{n,i}\lambda^y_{\infty,i}, \lambda^y_{\infty,i}) \\ &\leq 2\frac{\epsilon}{k} + \frac{\epsilon}{\ell}\|\lambda^y_{\infty,i}\|. \end{aligned}$$

Now let $m \geq 1$ so that $\mathbf{P}\{M > m\} < \epsilon$. Summarizing (3.4), (3.5) and (3.8), we deduce that, for any $n \geq m$,

$$\begin{aligned}
&\mathbf{P}\{d_{\mathrm{TV}}(\lambda_n^y, \lambda_\infty^y) > 3\epsilon\} \\
&\quad \leq \mathbf{P}\{\tilde{\lambda}_n^y \neq 0)\} + \mathbf{P}\Big\{\sum_{i>k} \lambda_{\infty,i}^y \neq 0\Big\} + \mathbf{P}\{\|\lambda_\infty^y\| > \ell\} \\
&\qquad + \mathbf{P}\Big\{d_{\mathrm{TV}}\Big(\sum_{i=1}^k c_{n,i}\lambda_{\infty,i}^{y/c_{n,i}}, \sum_{i=1}^k \lambda_{\infty,i}^y\Big) > 3\epsilon, \|\lambda_\infty^y\| \leq \ell\Big\} \\
&\quad \leq 2\frac{\epsilon}{2y} + 2\epsilon.
\end{aligned}$$

This completes the proof. □

PROPOSITION 3.7 (Strong Markov property). *For a probability measure μ on $\mathcal{M}^a$, let $(\pi^z, z \geq 0) \sim \mathbb{Q}_\mu^{\alpha,0}$. Denote its right-continuous natural filtration by $(\mathcal{F}^y, y \geq 0)$. Let Y be an a.s. finite $(\mathcal{F}^y, y \geq 0)$-stopping time. Then given $\mathcal{F}^Y$, the process $(\pi^{Y+y}, y \geq 0)$ has conditional distribution $\mathbb{Q}_{\pi^Y}^{\alpha,0}$.*

PROOF. This is now standard, so we only provide a sketch. Let $Y_n = 2^{-n}\lfloor 2^n Y + 1\rfloor \downarrow Y$. The strong Markov property at Y_n follows from the simple Markov property by elementary partitioning. Letting $n \to \infty$, we have $d_{\mathrm{TV}}(\pi^{Y_n}, \pi^Y) \to 0$ a.s., by Proposition 3.3. This implies $\mathbb{Q}_{\pi^{Y_n}}^{\alpha,0} \to \mathbb{Q}_{\pi^Y}^{\alpha,0}$ weakly in the sense of finite-dimensional distributions, by Proposition 3.6. □

3.3. *Proofs of the $\theta = 0$ cases of Theorems 1.2 and 1.9.* Let $\pi \in \mathcal{M}^a$ and $\mathbf{F} \sim \mathbf{Q}_\pi^{\alpha,0}$. Then for every $y \geq 0$, it follows from Proposition 2.12 that the superskewer $\mathrm{sSKEWER}(y, \mathbf{F})$ has distribution $K_y^{\alpha,0}(\pi, \cdot)$ defined in Definition 1.1. It remains to check the properties required of a path-continuous Hunt process on $(\mathcal{M}^a, d_{\mathcal{M}})$; see, for example, [32], Definition A.18.

Specifically, we noted in Lemma 2.5 that the state space $(\mathcal{M}^a, d_{\mathcal{M}})$ is a Lusin space (a Borel subset of a complete and separable metric space). In Proposition 3.3, we showed that the superskewer process is a.s. path-continuous under $\mathbf{Q}_\pi^{\alpha,0}$, with distribution on $\mathcal{C}([0,\infty), \mathcal{M}^a)$ denoted by $\mathbb{Q}_\pi^{\alpha,0}$ in Section 3.2. The map $\pi \mapsto K_y^{\alpha,0}(\pi, \cdot)$ is measurable by construction, it defines a semigroup on $(\mathcal{M}^a, d_{\mathcal{M}})$ by Proposition 3.5. Indeed, by Lemma 3.4, $\pi \mapsto \mathbb{Q}_\pi^{\alpha,0}$ is a kernel. The strong Markov property with respect to a right-continuous filtration was established in Proposition 3.7. □

4. Clades of the reflected process. In this section we study clades corresponding to excursions of the reflected scaffolding $\mathbf{X}(t) - \inf_{u \leq t} \mathbf{X}(u)$, $t \geq 0$, which form a key ingredient in the construction of the general two-parameter family of measure-valued diffusions.

4.1. *Preliminaries on reflected* $\mathtt{Stable}(1+\alpha)$ *processes.* For $\alpha \in (0,1)$, let $\mathbf{N}$ denote a $\mathtt{PRM}(\mathtt{Leb} \otimes \nu_{\mathrm{BESQ}}^{(-2\alpha)})$ on $[0,\infty) \times \mathcal{E}$ or $\mathbf{V}$ a $\mathtt{PRM}(\mathtt{Leb} \otimes \nu_{\mathrm{BESQ}}^{(-2\alpha)} \otimes \mathtt{Unif})$ on $[0,\infty) \times \mathcal{E} \times [0,1]$ and $\mathbf{N} = \varphi(\mathbf{V})$. Then $\mathbf{X} := \xi_{\mathbf{N}} \sim \mathtt{Stable}(1+\alpha)$. We call

$$\mathbf{X}_\perp(t) := \mathbf{X}(t) - \inf_{u \leq t} \mathbf{X}(u), \quad t \geq 0, \tag{4.1}$$

the *scaffolding process reflected at the infimum process*, or simply the *reflected scaffolding process*. See Figure 3. Since $\mathbf{X}_\perp$ is a strong Markov process [2], Proposition VI.1, Itô's theory of excursions applies to $\mathbf{X}_\perp$ [2], Chapters IV and VI, from which we record the following two results. Consider the *first passage process*

$$T^{-y} := \inf\{t \geq 0 \colon \mathbf{X}(t) < -y\} \quad \text{for } y \geq 0. \tag{4.2}$$

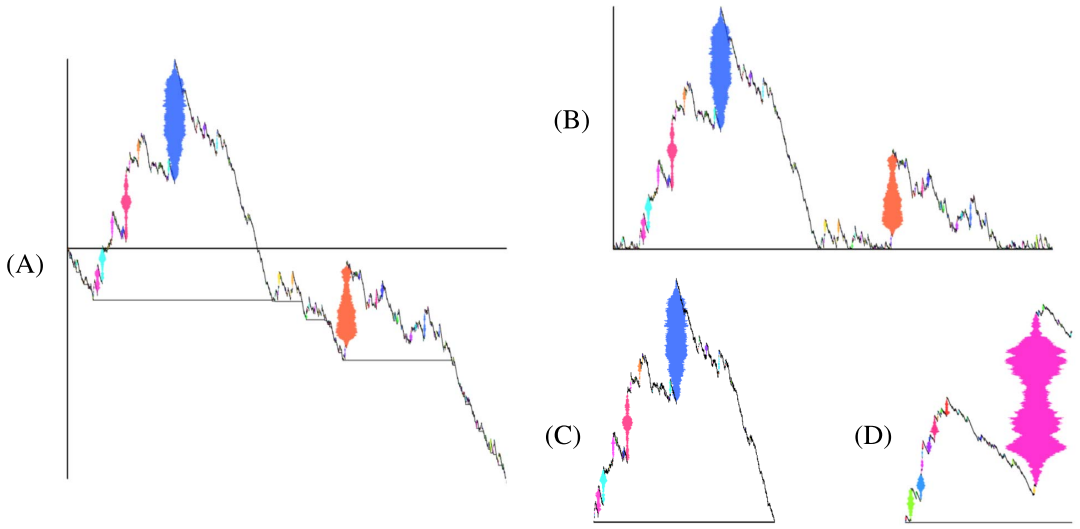

FIG. 3. *(A) Simulation of a* Stable(1.3) *scaffolding with spindles, run until hitting a negative level, superposed with the graph of its infimum process. (B) The associated reflected scaffolding with the same spindles, as in* (4.1). *(C) A single clade of the reflected process. (D) An enlarged* (6*-fold vertical,* 30*-fold horizontal) plot of the left end of this clade; note how there is no leftmost spindle but rather an accumulation of small spindles.*

PROPOSITION 4.1 ([2], Theorem VII.1). *The first passage process* $(T^{-y}, y \geq 0)$ *is a* Stable$(1/(1+\alpha))$ *subordinator. Its Laplace exponent is the inverse* $\phi_\alpha = \psi_\alpha^{-1}$ *of the Laplace exponent* ψ_α *of* $\mathbf{X}$:

$$\mathbf{E}[\exp(-qT^{-y})] = \exp(-y\phi_\alpha(q)), \quad \text{where} \\ \phi_\alpha(q) = \big(2^\alpha \Gamma(1+\alpha) q\big)^{1/(1+\alpha)}, \quad q \geq 0. \tag{4.3}$$

LEMMA 4.2 ([2], Theorem IV.10). *Define a point process on* $\mathcal{D}$ *by*

$$\mathbf{e}_\perp := \sum_{y \geq 0:\ T^{(-y)-} < T^{-y}} \delta\big(y, \mathbf{X}_\perp|^{\leftarrow}_{[T^{(-y)-}, T^{-y})}\big), \tag{4.4}$$

where $T^{(-y)-} := \sup_{z<y} T^{-z}$ *for* $y > 0$. *Then* $\mathbf{e}_\perp$ *is a* PRM$(\mathrm{Leb} \otimes \nu_{\perp\mathrm{stb}}^{1+\alpha})$, *where* $\nu_{\perp\mathrm{stb}}^{1+\alpha}$ *is known as the Itô measure of excursions of the reflected process* $\mathbf{X}_\perp$.

4.2. *The Itô measure of reflected clades.* Here, we extend the excursion theory, as we did for excursions away from fixed levels in [21], Section 4.3, to define an Itô measure $\overline{\nu}_{\perp\mathrm{cld}}^{(\alpha)}$ associated with clades of $(\mathbf{V}, \mathbf{X}_\perp)$: point measures of spindles corresponding to the excursions of $\mathbf{X}_\perp$.

Define a point process of clades

$$\mathbf{F}_\perp := \sum_{y \geq 0:\ T^{(-y)-} < T^{-y}} \delta\big(y, \mathbf{V}|^{\leftarrow}_{[T^{(-y)-}, T^{-y})}, \mathbf{X}_\perp|^{\leftarrow}_{[T^{(-y)-}, T^{-y})}\big). \tag{4.5}$$

The following statement follows readily from the marking property of PRMs and the existence of the limits (2.3) uniformly on all compact intervals.

PROPOSITION 4.3. *The point measure* $\mathbf{F}_\perp$ *is a* PRM$(\mathrm{Leb} \otimes \overline{\nu}_{\perp\mathrm{cld}}^{(\alpha)})$ *on* $[0, \infty) \times \mathcal{V} \times \mathcal{D}$, *where* $\overline{\nu}_{\perp\mathrm{cld}}^{(\alpha)}(dV, dg) = \overline{\mu}_g(dV)\nu_{\perp\mathrm{stb}}^{1+\alpha}(dg)$, *the kernel* $g \mapsto \overline{\mu}_g$ *is as in Lemma* 2.4 *and* $\nu_{\perp\mathrm{stb}}^{1+\alpha}$ *is the Itô measure of excursions of the reflected scaffolding* $\mathbf{X}_\perp$. *Moreover, a.s. for all points* (V, X) *of* $\mathbf{F}_\perp$, *we have* $X = \xi_V$.

See Figure 3 for a plot of the largest of a collection of clades sampled from a simulated approximation of $\overline{\nu}^{(0.3)}_{\perp\mathrm{cld}}$. In the rest of this section, we shall establish some relevant properties of $\overline{\nu}^{(\alpha)}_{\perp\mathrm{cld}}$. We begin with a scaling invariance property due to the $\mathtt{Stable}(1+\alpha)$-scaling property of $\mathbf{X}$, which we express in terms of the scaling operation $(c \odot^{1+\alpha}_{\mathrm{stb}} g)(t) = cg(t/c^{1+\alpha})$, for $c > 0$, $t \in \mathbb{R}$, $g \in \mathcal{D}$.

LEMMA 4.4 (Self-similarity of $\overline{\nu}^{(\alpha)}_{\perp\mathrm{cld}}$). *For $c > 0$, we have*

$$c\overline{\nu}^{(\alpha)}_{\perp\mathrm{cld}}\big((c \odot^{\alpha}_{\mathrm{cld}} A) \times (c \odot^{1+\alpha}_{\mathrm{stb}} B)\big) = \overline{\nu}^{(\alpha)}_{\perp\mathrm{cld}}(A \times B) \quad \text{for } A \in \Sigma(\mathcal{V}),\ B \in \Sigma(\mathcal{D}). \tag{4.6}$$

PROOF. Let us record the scaling invariance properties of $\nu^{(-2\alpha)}_{\mathrm{BESQ}}$ and $\mathtt{Stable}(1+\alpha)$; see, for example, [21], Lemmas 2.9 and 4.2:

$$c^{1+\alpha}\nu^{(-2\alpha)}_{\mathrm{BESQ}}(c \odot_{\mathrm{spdl}} B) = \nu^{(-2\alpha)}_{\mathrm{BESQ}}(B) \quad \text{and} \quad \big(\mathbf{X}(c^{1+\alpha}t), t \geq 0\big) \overset{d}{=} \big(c\mathbf{X}(t), t \geq 0\big).$$

The claim now follows from Proposition 4.3. □

Recall the notation ζ^+ and len from Section 2.5.

PROPOSITION 4.5.

(i) $\overline{\nu}^{(\alpha)}_{\perp\mathrm{cld}}\{\zeta^+ > z\} = \alpha z^{-1}, \quad z > 0.$

(ii) $\overline{\nu}^{(\alpha)}_{\perp\mathrm{cld}}\{\mathrm{len} > x\} = \frac{(2^\alpha\Gamma(1+\alpha))^{1/(1+\alpha)}}{\Gamma(\alpha/(1+\alpha))}x^{-1/(1+\alpha)}, \quad x > 0.$

(iii) $\overline{\nu}^{(\alpha)}_{\perp\mathrm{cld}}\{(V, g) \in \mathcal{V} \times \mathcal{D}\colon V(\{0\} \times \mathcal{E} \times [0,1]) > 0\} = 0.$

Less formally, (iii) states that clades of the reflected process $(\mathbf{V}, \mathbf{X}_\perp)$, corresponding to the atoms of $\mathbf{F}_\perp$, do not have spindles at time zero, and excursions of $\mathbf{X}_\perp$ do not begin with jumps. This proposition is essentially a consequence of known results on excursions of the reflected process $\mathbf{X}_\perp$. For completeness, we include a proof in the Appendix.

4.3. *Path-continuity and Markov-like properties under $\overline{\nu}^{(\alpha)}_{\perp\mathrm{cld}}$.*

PROPOSITION 4.6. *Let $(\mathbf{v}, \mathbf{x}) \sim \overline{\nu}^{(\alpha)}_{\perp\mathrm{cld}}(\cdot|\zeta^+ > y)$. Then $\mathbf{x} = \xi_{\mathbf{v}}$ a.s. Moreover, $\overline{\mathrm{SSKEWER}}(\mathbf{v}, \mathbf{x}) = (\mathrm{SSKEWER}(y, \mathbf{v}, \mathbf{x}), y \geq 0)$ defined as in Definition 1.8 is a well-defined $\mathcal{C}([0,\infty), (\mathcal{M}^a, d_{\mathcal{M}}))$-valued random variable. Furthermore, it is a d_{TV}-path-continuous excursion away from $0 \in \mathcal{M}^a$.*

PROOF. We may assume that $\mathbf{v}$ is the clade corresponding to the first excursion $\mathbf{x}$ above the minimum of $\mathbf{X}$ with height $\zeta^+(\mathbf{x}) > y$. Then we can write $\mathbf{v} = \mathbf{V}|^{\leftarrow}_{[T', T'')}$ for a pair of a.s. finite random times T', T''. We observe from Proposition 4.5(iii) that

$$\begin{aligned}&\overline{\mathrm{SSKEWER}}\big(\mathbf{V}|_{[0,T'')}, \mathbf{X}|_{[0,T'')} - \mathbf{X}(T')\big)\\ &\quad = \overline{\mathrm{SSKEWER}}\big(\mathbf{V}|_{[0,T')}, \mathbf{X}|_{[0,T')} - \mathbf{X}(T')\big) + \overline{\mathrm{SSKEWER}}(\mathbf{v}, \mathbf{x}).\end{aligned} \tag{4.7}$$

From Proposition 3.1, the first two superskewer processes in this formula are a.s. d_{TV}-continuous. The d_{TV}-continuity of $\overline{\mathrm{SSKEWER}}(\mathbf{v}, \mathbf{x})$ follows. Moreover, it follows from Proposition 4.5(iii) that $\mathrm{SSKEWER}(0, \mathbf{v}, \mathbf{x}) = 0$. □

The next statement, which is an analogue to Proposition 2.13, establishes the Markovian decoupling at the first passage above level y of $(\mathbf{v}, \mathbf{x}) \sim \overline{\nu}^{(\alpha)}_{\perp\mathrm{cld}}(\cdot|\zeta^+ > y)$.

LEMMA 4.7 (Mid-spindle Markov property (MSMP) for $\overline{\nu}^{(\alpha)}_{\perp\mathrm{cld}}$). *Fix $y > 0$. Consider $(\mathbf{v}, \mathbf{x})$ with law $\overline{\nu}^{(\alpha)}_{\perp\mathrm{cld}}(\cdot|\zeta^+ > y)$. Let (T, f_T, x_T) denote the point of the first spindle in $\mathbf{v}$ that crosses level y, and let $\hat{f}^y_T$ and $\check{f}^y_T$ denote its broken components. Finally, let $m^y(\mathbf{v}, \mathbf{x}) := \check{f}^y_T(y - \mathbf{x}(T-)) = \hat{f}^y_T(0)$. Given $m^y(\mathbf{v}, \mathbf{x})$ and x_T, the process $\mathbf{v}|^{\leftarrow}_{[0,T)} + \delta(T, \check{f}^y_T, x_T)$ is conditionally independent of $\mathbf{v}|^{\leftarrow}_{(T,\infty)} + \delta(0, \hat{f}^y_T, x_T)$. Moreover, under the conditional law, $(\mathbf{v}|^{\leftarrow}_{(T,\infty)}, \hat{f}^y_T)$ has the law of $(\mathbf{V}'|^{\leftarrow}_{[0,\tau]}, f')$, where $f' \sim \mathtt{BESQ}_{m^y(\mathbf{v},\mathbf{x})}(-2\alpha)$ is independent of $\mathbf{V}' \sim \mathtt{PRM}(\mathtt{Leb} \otimes \nu^{(-2\alpha)}_{\mathtt{BESQ}} \otimes \mathtt{Unif})$, and $\tau = \inf\{t \geq 0\colon \zeta(f') + \xi_{\mathbf{V}'}(t) = -y\}$.*

PROOF. Let $\mathbf{V} \sim \mathtt{PRM}\big(\mathtt{Leb} \otimes \nu^{(-2\alpha)}_{\mathtt{BESQ}} \otimes \mathtt{Unif}\big)$ and let $\mathbf{X}_\perp$ denote the associated scaffolding reflected at its minimum, as in (4.1). Let $T = T^{\geq y} := \inf\{t \geq 0\colon \mathbf{X}_\perp(t) \geq y\}$. The point measure $\mathbf{V}$ a.s. has a point (T, f_T, x_T); let $\check{f}^y_T$ and $\hat{f}^y_T$ denote the components of the spindle f_T broken around level y in $\mathbf{X}_\perp$, as in Figure 2. Let $m^y(\mathbf{V}, \mathbf{X}_\perp) := \hat{f}^y_T(0)$ denote the mass of the broken spindles at the break. Let

$$G := \sup\big\{t < T\colon \mathbf{X}_\perp(t) = 0\big\}, \qquad D := \inf\big\{t > T\colon \mathbf{X}_\perp(t) = 0\big\},$$

so $\mathbf{v}' := \mathbf{V}|^{\leftarrow}_{[G,D]}$ has the same distribution as $\mathbf{v}$ in the lemma statement.

The proof of the mid-spindle Markov property [21] of ($\mathbf{N}$, *Lemma* 4.13, $\mathbf{X}$) modified by (i) including type labels with the point measure $\mathbf{V}$ in place of $\mathbf{N} := \varphi(\mathbf{V}) \sim \mathtt{PRM}(\mathtt{Leb} \otimes \nu^{(-2\alpha)}_{\mathtt{BESQ}})$, as in the proof of Proposition 2.13, and (ii) substituting $\mathbf{X}_\perp$ in place of $\mathbf{X} := \xi(\mathbf{N})$, yields here that given $(m^y(\mathbf{V}, \mathbf{X}_\perp), x_T)$,

$$\mathbf{V}|^{\leftarrow}_{(T,\infty)} + \delta\big(0, \hat{f}^y_T, x_T\big) \text{ is conditionally independent of } \mathbf{V}|^{\leftarrow}_{[0,T)} + \delta\big(T, \check{f}^y_T, x_T\big), \tag{4.8}$$

with $\hat{f}^y_T$ being conditionally distributed as a $\mathtt{BESQ}(-2\alpha)$ started at $m^y(\mathbf{V}, \mathbf{X}_\perp)$ and $\mathbf{V}|^{\leftarrow}_{(T,\infty)} \overset{d}{=} \mathbf{V}$ being conditionally (and unconditionally) independent of $(\hat{f}^y_T, x_T)$.

Let $T' := T - G$. The random time G is a function of the measure in the second line of (4.8), and $D - T$ is the stopping time for the measure in the first line at which its associated scaffolding first hits $-y$. Thus, noting that $m^y(\mathbf{v}', \mathbf{x}') = m^y(\mathbf{V}, \mathbf{X}_\perp)$,

$$\begin{aligned}&\mathbf{v}'|^{\leftarrow}_{(T',\infty)} + \delta\big(0, \hat{f}^y_T, x_T\big) \text{ is conditionally independent of}\\ &\mathbf{v}'|^{\leftarrow}_{[0,T')} + \delta\big(T', \check{f}^y_T, x_T\big) \text{ given } \big(m^y(\mathbf{v}', \mathbf{x}'), x_T\big),\end{aligned}$$

with the desired regular conditional distribution. □

COROLLARY 4.8 (Markov-like property of $\overline{\nu}^{(\alpha)}_{\perp\mathrm{cld}}$). *Fix $y > 0$ and consider $(\mathbf{v}, \mathbf{x}) \sim \overline{\nu}^{(\alpha)}_{\perp\mathrm{cld}}(\cdot|\zeta^+ > y)$. Then the upper point measure $G_0^{\geq y}(\mathbf{v}, \mathbf{x})$ of clades is conditionally independent of the lower cutoff process $\mathrm{CUTOFF}^{\leq y}(\mathbf{v}, \mathbf{x})$ given $\mathrm{SSKEWER}(y, \mathbf{v}, \mathbf{x}) = \pi$ and has regular conditional distribution $\mathbf{Q}^{\alpha,0}_\pi$.*

PROOF. Using Lemma 4.7 and the notation in its statement, the conditional distribution that we wish to characterize is the same as the law of $G_0^{\geq 0}(V', \xi_{V'})$ given $\mathrm{SSKEWER}(0, V', \xi_{V'}) = \pi$, where $V' = \delta(0, f', x') + \mathbf{V}'|^{\leftarrow}_{[0,\tau]}$. By the Markov-like property of Proposition 2.14, we know that the latter conditional law is $\mathbf{Q}^{\alpha,0}_\pi$. This proves the corollary. □

4.4. *Entrance law of* $\overline{\mathrm{SSKEWER}}$ *under* $\overline{\nu}^{(\alpha)}_{\perp\mathrm{cld}}$. We first state a limit representation of the underlying excursions of the $\mathtt{Stable}(1+\alpha)$ scaffolding.

LEMMA 4.9. *Fix $y>0$. For $a\in(0,y)$, denote by $\mathbf{x}^a$ a* Stable$(1+\alpha)$ *process started at a, absorbed at 0 and conditioned to reach level y. Then we have the following weak convergence of measures on Skorokhod space*:

$$\mathbf{P}\{\mathbf{x}^a\in\cdot\}\to\nu_{\perp\mathrm{stb}}^{1+\alpha}(\cdot\mid\zeta^+>y)\quad\text{as } a\downarrow 0. \tag{4.9}$$

Similar results can be found in the literature, for example, [5], Corollaire 3, when conditioning to exceed an excursion length threshold rather than an excursion height threshold. Since we have been unable to find a reference for Lemma 4.9, we provide a proof in the Appendix.

PROPOSITION 4.10. *Fix $y>0$ and let $(\mathbf{v},\mathbf{x})\sim\overline{\nu}_{\perp\mathrm{cld}}^{(\alpha)}(\cdot|\zeta^+>y)$. Then*:

(i) $m^y(\mathbf{v},\mathbf{x})\sim$ Gamma$(1-\alpha,1/2y)$;

(ii) $\mathrm{sSKEWER}(y,\mathbf{v},\mathbf{x})\overset{d}{=}B^y\overline{\Pi}$, *where* $B^y\sim$ Exponential$(1/2y)$ *is independent of* $\overline{\Pi}\sim$ PDRM$(\alpha,0)$.

Together, Propositions 4.5(i) and 4.10, Corollary 4.8 and Theorem 1.9(i), yield an entrance law description of $\overline{\mathrm{sSKEWER}}$ under $\overline{\nu}_{\perp\mathrm{cld}}^{(\alpha)}$. Specifically, for each $y>0$, the point measure $\mathrm{sSKEWER}(y,\cdot)$ is nonzero at rate α/y, distributed as an Exponential$(1/2y)$ multiple of an independent PDRM$(\alpha,0)$, and $(\mathrm{sSKEWER}(z,\cdot),z\geq y)$ is an SSSP$(\alpha,0)$. Here is a consequence.

COROLLARY 4.11. *We have $\overline{\nu}_{\perp\mathrm{cld}}^{(\alpha)}\{\|\overline{\mathrm{sSKEWER}}\|\in\cdot\}=2\alpha\nu_{\mathrm{BESQ}}^{(0)}$ for all $\alpha\in(0,1)$, where $\nu_{\mathrm{BESQ}}^{(0)}$ is the Pitman–Yor excursion measure of* BESQ(0) *with the normalization $\nu_{\mathrm{BESQ}}^{(0)}\{\zeta>y\}=1/2y$, $y>0$.*

PROOF. This follows from Proposition 2.10 and the above entrance law discussion, together with the path-continuity noted in Proposition 4.6. □

PROOF OF PROPOSITION 4.10. Let $(\mathbf{v},\mathbf{x})\sim\overline{\nu}_{\perp\mathrm{cld}}^{(\alpha)}(\cdot\mid\zeta^+>y)$. For $k\geq 1$ let $\mathbf{x}_k$ denote a Stable$(1+\alpha)$ Lévy process started from $1/k$ and conditioned to exceed y before being absorbed upon hitting 0. Let $\mathbf{v}_k$ denote a point measure of spindles and allelic types associated with $\mathbf{x}_k$ in the natural manner described in Lemma 2.4, with each jump, including the initial jump up to level $1/k$, marked conditionally independently by a uniform type label and a BESQ spindle with lifetime equal to the jump height (the marking on the first jump is unimportant for this proof, except that $\xi_{\mathbf{v}_k}$ must equal $\mathbf{x}_k$). From (4.9), we may assume that these processes are coupled so that $\mathbf{x}_k$ converges to $\mathbf{x}$ in probability in the Skorokhod metric, as k increases.

(i) We write $\mathrm{len}(g):=\sup\{t\geq 0\colon g(t)\neq 0\}$ for $g\in\mathcal{D}$, and let $\mathcal{D}_{\mathrm{exc}}:=\{g\in\mathcal{D}\colon g|_{(-\infty,0)}=0,\ g|_{(0,\mathrm{len}(g))}\neq 0\}$ denote the set of càdlàg excursion paths. For an excursion in $g\in\mathcal{D}_{\mathrm{exc}}$ with $\zeta^+(g)>y$, let $J^{y-}(g)$ and $J^y(g)$ denote the values of the excursion immediately before and after the first time $T^{\geq y}(g)$ that g crosses level y, that is, $J^{y-}(g):=g(T^{\geq y}(g)-)$ and $J^y(g):=g(T^{\geq y}(g))$. For $\delta>0$, we define

$$S_\delta:=\big\{g\in\mathcal{D}_{\mathrm{exc}}\colon\zeta^+(g)>y,\ J^{y-}(g)<y-\delta<y+\delta<J^y(g)\big\}.$$

For every δ, S_δ is open in the Skorokhod topology on $\mathcal{D}_{\mathrm{exc}}$ and J^{y-} and J^y are Skorokhod-continuous on it. It follows from (4.9) that, conditional on $\{\mathbf{x}\in S_\delta\}$, we get

$$\big(J^{y-}(\mathbf{x}_k),J^y(\mathbf{x}_k)\big)\xrightarrow{P}\big(J^{y-}(\mathbf{x}),J^y(\mathbf{x})\big)\quad\text{as } k\to\infty.$$

For every $\epsilon > 0$ there is some $\delta > 0$ sufficiently small so that $\mathbf{P}\{\mathbf{x} \in S_\delta\} > 1 - \epsilon$; thus, this limit holds without conditioning.

Let $\mathbf{f}$ (resp. $\mathbf{f}_k$, $k \geq 1$) denote the leftmost spindle of $\mathbf{v}$ (resp. $\mathbf{v}_k$) that survives to level y. By Proposition 4.3, given $\mathbf{x}$, this spindle $\mathbf{f}$ has conditional law $\nu_{\mathrm{BESQ}}^{(-2\alpha)}(\cdot \mid \zeta = J^y(\mathbf{x}) - J^{y-}(\mathbf{x}))$. By [21], Proposition 4.9, which is stated for clades without type labels but adapts without modification to the present setting, for $k \geq \lceil 1/y \rceil$ the leftmost spindle of $\mathbf{v}_k$ to cross level y has conditional law $\mathbf{f}_k \sim \nu_{\mathrm{BESQ}}^{(-2\alpha)}(\cdot \mid \zeta = J^y(\mathbf{x}_k) - J^{y-}(\mathbf{x}_k))$ given $\mathbf{x}_k$.

The leftmost spindle masses of $\mathbf{v}$ and the $\mathbf{v}_k$ at level y are

$$m^y(\mathbf{v}, \mathbf{x}) = \mathbf{f}\big(y - J^{y-}(\mathbf{x})\big) \quad \text{and} \quad m^y(\mathbf{v}_k, \mathbf{x}_k) = \mathbf{f}_k\big(y - J^{y-}(\mathbf{x}_k)\big).$$

Since $y - J^{y-}(\mathbf{x}_k)$ is converging weakly to $y - J^{y-}(\mathbf{x})$ while $\mathbf{f}_k$ is converging weakly to $\mathbf{f}$, we conclude that these leftmost masses are converging weakly. From [22], Lemma A.4, for $c > 0$,

$$\mathbf{P}\big\{m^y(\mathbf{v}_k, \mathbf{x}_k) \in dc\big\} = \frac{\alpha 2^\alpha c^{-1-\alpha}}{\Gamma(1-\alpha)} \frac{e^{-c/2y} - e^{-c/2(y-k^{-1})}}{(y-k^{-1})^{-\alpha} - y^{-\alpha}}\, dc \to \frac{(2y)^{\alpha-1} e^{-c/2y}}{c^\alpha \Gamma(1-\alpha)}\, dc \tag{4.10}$$

as k increases, by L'Hôpital's rule. Thus, $m^y(\mathbf{v}, \mathbf{x}) \sim \mathtt{Gamma}(1-\alpha, 1/2y)$, as claimed.

(ii) We apply the MSMP of Lemma 4.7 to $(\mathbf{v}, \mathbf{x})$ at level y. Then, given $\mathbf{v}|_{[0,T^{\geq y}]}$, the point measure $(\mathbf{v}|^{\leftarrow}_{(T^{\geq y},\infty)})$ is conditionally a $\mathtt{PRM}\big(\mathrm{Leb} \otimes \nu_{\mathrm{BESQ}}^{(-2\alpha)} \otimes \mathtt{Unif}\big)$ killed when its scaffolding hits level $-J^y(\mathbf{x})$. This same situation has been addressed in Proposition 2.12 and previously in [22], proof of Proposition 3.4, with the conclusion that, if $\lambda := \textsc{sSkewer}(\mathbf{v}, \mathbf{x}, y) - m^y(\mathbf{v}, \mathbf{x})\delta_x$ denotes the superskewer minus the atom corresponding to the leftmost spindle at level y, then $G := \lambda([0,1]) \sim \mathtt{Gamma}(\alpha, 1/2y)$ and $\overline{\lambda} := \lambda/G \sim \mathtt{PDRM}(\alpha, \alpha)$. Moreover, $\overline{\lambda}$, G, and $m^y(\mathbf{v}, \mathbf{x})$ are independent of each other. Now, the claim follows from well-known descriptions of the $\mathtt{PD}(\alpha, \theta)$ distributions. In particular, if one takes a $\mathtt{PD}(\alpha, 0)$-distributed sequence, removes one of the entries chosen as a size-biased random pick, then renormalizes the remainder of the sequence, then the removed entry is independent of the renormalized remaining sequence, and the two have respective laws $\mathtt{Beta}(1-\alpha, \alpha)$ (the law of $m^y(\mathbf{v}, \mathbf{x})/(m^y(\mathbf{v}, \mathbf{x}) + G)$) and $\mathtt{PD}(\alpha, \alpha)$ (the law of ranked masses of λ/G) [36], Theorem 3.2, Definition 3.3. □

5. Self-similar (α, θ)-superprocesses. Recall that we outlined two approaches to self-similar (α, θ) superprocesses, $\mathtt{SSSP}(\alpha, \theta)$, in the introduction: the first is in terms of kernels that we claim, in Theorem 1.2, give rise to a Hunt process; the second is in terms of point measures that we claim, in Theorem 1.9(ii), have a Markovian superskewer process. As in the special case $\theta = 0$ that we established in Section 3, we will study the point measure construction and compute its semigroup to connect the two approaches.

5.1. *Marginal distributions of the superskewer process.* Fix $\theta > 0$. Recall the point measure construction in the statement of Theorem 1.9(ii): for any initial measure $\pi \in \mathcal{M}^a$, we construct a measure-valued process by adding two independent superskewer processes. The first is associated with $\mathbf{F} \sim \mathbf{Q}_\pi^{\alpha,0}$, which we know from Theorem 1.9(i), already proved, gives an $\mathtt{SSSP}_\pi(\alpha, 0)$. The second is associated with $\mathbf{F} \sim \mathbf{Q}_0^{\alpha,\theta} := \mathtt{PRM}(\frac{\theta}{\alpha}\mathrm{Leb} \otimes \overline{\nu}^{(\alpha)}_{\perp\mathrm{cld}})$, on $[0,\infty) \times \mathcal{V} \times \mathcal{D}$, as

$$\textsc{sSkewer}(y, \mathbf{F}) = \sum_{\text{points } (z, V_z, X_z) \text{ of } \mathbf{F}} \textsc{sSkewer}(y - z, V_z, X_z), \quad y \geq 0.$$

We interpret each atom (z, V_z, X_z) of $\mathbf{F}$ as a subpopulation—a clade, in the biological sense—that enters via immigration at level (time) z. The θ parameter is understood as the rate at

which new immigration enters. For any $y > 0$, the atomic measure $\mathrm{sSKEWER}(y, \mathbf{F})$ is the superskewer at level y of the population formed by clades that enter below level y. We stress that it comprises an infinite summation of superskewers of clades. The following proposition gives the marginal distribution of the process

$$\mathrm{sS\overline{KEWER}}(\mathbf{F}) := \big(\mathrm{sSKEWER}(y, \mathbf{F}),\, y \geq 0\big).$$

In particular, it confirms that the process takes values in the space $\mathcal{M}^a$ of finite atomic measures.

PROPOSITION 5.1. *Let* $\mathbf{F} \sim \mathtt{PRM}(\frac{\theta}{\alpha}\mathtt{Leb} \otimes \overline{\nu}^{(\alpha)}_{\perp\mathrm{cld}})$. *For any* $y \geq 0$, *consider* $G^y \sim \mathtt{Gamma}(\theta, 1/2y)$ *and an independent random measure* $\overline{\Pi}_0 \sim \mathtt{PDRM}(\alpha, \theta)$. *Then we have* $\mathrm{sSKEWER}(y, \mathbf{F}) \stackrel{d}{=} G^y \overline{\Pi}_0$.

COROLLARY 5.2 (Marginal distributions). *In the setting of Theorem* 1.9(*ii*),

$$\mathrm{sSKEWER}(y, \mathbf{F}) + \mathrm{sSKEWER}(y, \mathbf{F}) \sim K^{\alpha,\theta}_y(\pi, \cdot)$$

for each $y > 0$, *where* $K^{\alpha,\theta}_y(\pi, \cdot)$ *is the kernel of Theorem* 1.2.

PROOF. Given the definition of $K^{\alpha,\theta}_y$ in Definition 1.1, this follows straight from the proposition and the kernel $K^{\alpha,0}_y$ that we related to the superskewer construction in Theorem 1.9(i), as proved in Section 3.3. □

To prove the proposition, we require an intriguing multivariate distributional identity.

LEMMA 5.3. *Consider* $G \sim \mathtt{Gamma}(\theta, \rho)$ *and suppose that for* $n \geq 1$ *we have* $E_n \sim \mathtt{Exponential}(\rho)$ *and* $B_n \sim \mathtt{Beta}(\theta, 1)$, *with all of these variables jointly independent. Then*

$$\left(E_n \cdot \prod_{i=1}^{n} B_i,\ n \geq 1\right) \stackrel{d}{=} \left(G(1 - B_n) \cdot \prod_{i=1}^{n-1} B_i,\ n \geq 1\right). \tag{5.1}$$

Moreover, $\sum_{n\geq 1} E_n \cdot \prod_{i=1}^{n} B_i \sim \mathtt{Gamma}(\theta, \rho)$.

If we take each coordinate separately, then the claim follows from standard Beta-Gamma calculus. The joint distributional identity is subtler. It can be read from [45], Theorems 3–6, in the context of the limiting behaviour of birth processes with immigration, but it can also be quickly proved by more direct arguments, as we show here.

PROOF. The second conclusion follows from the first, as the terms in the sequence on the right in (5.1) form a telescoping series, adding up to G. Thus, we need only verify (5.1). Assume WLOG that $\rho = 1$. Let $J_n := E_n \cdot \prod_{i=1}^{n} B_i$ and $L_n := G(1 - B_n) \cdot \prod_{i=1}^{n-1} B_i$. It suffices to show that, for every $n \geq 1$ and $(r_1, \ldots, r_n) \in \mathbb{N}^n$, we have $\mathbf{E}[J_1^{r_1} \cdots J_n^{r_n}] = \mathbf{E}[L_1^{r_1} \cdots L_n^{r_n}]$.

For $j \in [n]$, let $r_{j:n}$ denote $r_j + r_{j+1} + \cdots + r_n$. Then

$$\begin{aligned}\mathbf{E}\big[J_1^{r_1} \cdots J_n^{r_n}\big] &= \mathbf{E}\big[B_1^{r_{1:n}} B_2^{r_{2:n}} \cdots B_n^{r_n} E_1^{r_1} \cdots E_n^{r_n}\big] \\ &= \left(\frac{\theta}{\theta + r_{1:n}}\right)\left(\frac{\theta}{\theta + r_{2:n}}\right) \cdots \left(\frac{\theta}{\theta + r_n}\right) r_1! \cdots r_n!,\end{aligned}$$

and

$$\begin{aligned}\mathbf{E}\big[L_1^{r_1}\cdots L_n^{r_n}\big] &= \mathbf{E}\big[(1-B_1)^{r_1}B_1^{r_{2:n}}(1-B_2)^{r_2}B_2^{r_{3:n}}\cdots(1-B_n)^{r_n}G^{r_{1:n}}\big]\\ &= \theta\bigg(\frac{\Gamma(\theta+r_{2:n})r_1!}{\Gamma(\theta+r_{1:n}+1)}\bigg)\theta\bigg(\frac{\Gamma(\theta+r_{3:n})r_2!}{\Gamma(\theta+r_{2:n}+1)}\bigg)\cdots\theta\bigg(\frac{\Gamma(\theta)r_n!}{\Gamma(\theta+r_n+1)}\bigg)\\ &\quad\times\frac{\Gamma(r_{1:n}+\theta)}{\Gamma(\theta)}.\end{aligned}$$

Clearly, these two products are equal, as desired. □

PROOF OF PROPOSITION 5.1. Consider the process

$$K^y(s) := \int_{[0,s]\times\mathcal{V}\times\mathcal{D}} \mathbf{1}\{\zeta^+(V)-r>y\}\mathbf{F}(dr,dV,dX), \quad s\in[0,y), \tag{5.2}$$

counting the clades that enter via immigration below level s and survive to level y. Using Proposition 4.5(i), we deduce that $(K^y(s), s\in[0,y))$ is an inhomogeneous Poisson process with intensity

$$\frac{\theta}{\alpha}\overline{\nu}^{(\alpha)}_{\perp\mathrm{cld}}\{\zeta^+>y-s\}\,ds = \theta(y-s)^{-1}\,ds \quad \text{for } s\in[0,y).$$

Let $S_0:=0$ and for $i\geq 1$ let S_i denote the time of the ith jump of K^y. Then almost surely, for $i\geq 1$,

$$\mathbf{P}(S_i>s\mid S_{i-1}) = \exp\bigg(-\int_{S_{i-1}}^{s}\theta(y-r)^{-1}\,dr\bigg) = \bigg(\frac{y-s}{y-S_{i-1}}\bigg)^{\theta}, \quad s\in[S_{i-1},y].$$

Thus, setting

$$B_i := \frac{y-S_i}{y-S_{i-1}} \quad \text{for each } i\geq 1,$$

the sequence $(B_i, i\geq 1)$ is i.i.d. $\mathtt{Beta}(\theta,1)$. Let (V_i,X_i), $i\geq 1$, denote the sequence of clades corresponding to the S_i, so $\mathbf{F}$ has a point at each (S_i,V_i,X_i). By the PRM definition of $\mathbf{F}$, the pairs $((V_i,X_i), i\geq 1)$ are conditionally independent given $(S_j, j\geq 1)$, with respective conditional distributions $\overline{\nu}^{(\alpha)}_{\perp\mathrm{cld}}\{\cdot\mid\zeta^+ + S_i>y\}$. Let

$$M_i := \big\|\textsc{sSkewer}(y-S_i,V_i,X_i)\big\|, \qquad \overline{\Pi}_i := \frac{1}{M_i}\textsc{sSkewer}(y-S_i,V_i,X_i).$$

By Proposition 4.10, conditionally given S_i, we get

$$M_i\sim\mathtt{Exponential}\big(1/2(y-S_i)\big) \quad\text{and}\quad \overline{\Pi}_i\sim\mathtt{PDRM}(\alpha,0),$$

with the two being independent. Since the distribution of $\overline{\Pi}_i$ does not depend on S_i, and since products of the (B_j) variables are telescoping, we get a nicer characterization by setting

$$E_i := M_i\bigg(\prod_{j=1}^{i}B_j\bigg)^{-1}\sim\mathtt{Exponential}(1/2y).$$

Indeed, the sequences (B_i), (E_i) and $(\overline{\Pi}_i)$ are each i.i.d. and are jointly independent of each other. Now,

$$\textsc{sSkewer}(y,\mathbf{F}) = \sum_{i\geq 1}\bigg(E_i\prod_{j=1}^{i}B_j\bigg)\overline{\Pi}_i. \tag{5.3}$$

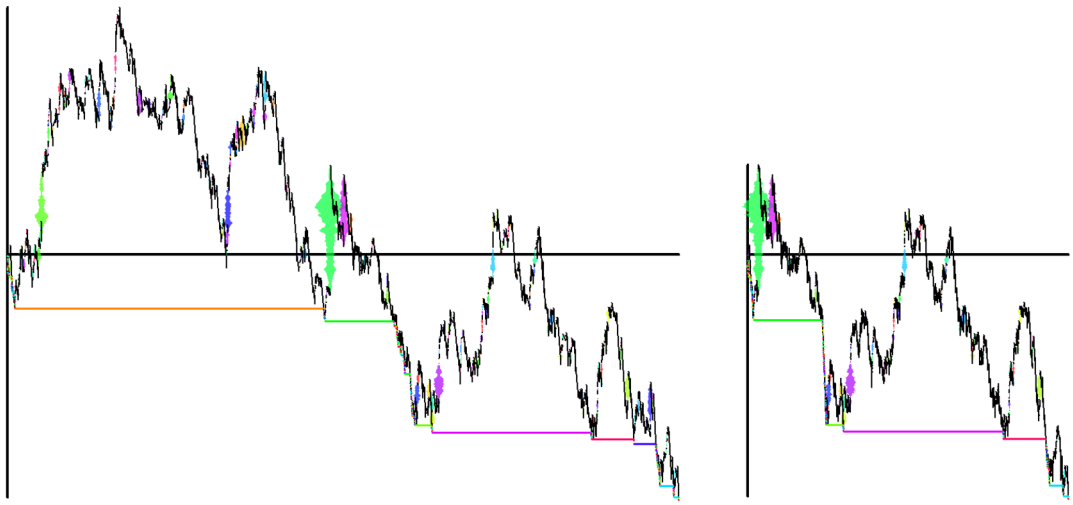

FIG. 4. *Simulated thinning of scaffolding and spindles as in the proof of Proposition* 5.4, *with* $(\alpha,\theta)=(0.6,0.4)$. *Left*: $(\mathbf{V},\mathbf{X})$. *Right*: $(\mathbf{V}_\theta,\mathbf{X}_\theta)$. *Corresponding clades of the reflected process are indicated by colored underlining* (*unrelated to spindle types*).

Note that the scaling factors in this concatenation are the terms on the left hand side in (5.1).

Consider $\overline{\Pi}' \sim \texttt{PDRM}(\alpha,\theta)$, $\overline{\Pi} \sim \texttt{PDRM}(\alpha,0)$ and $B \sim \texttt{Beta}(\theta,1)$ independent of each other. We appeal to a classical decomposition,

$$\overline{\Pi}' \stackrel{d}{=} B\overline{\Pi}' + (1-B)\overline{\Pi}, \tag{5.4}$$

for example, from [36], Proposition 3.16. Iterating (5.4) yields

$$\sum_{i\geq 1}\left((1-B_i)\prod_{j=1}^{i-1}B_j\right)\overline{\Pi}_i \sim \texttt{PDRM}(\alpha,\theta). \tag{5.5}$$

Scaling each term by a common factor of $G \sim \texttt{Gamma}(\theta, 1/2y)$, taken to be independent of the other variables, yields a concatenation of a sequence of independent $\texttt{PDRM}(\alpha,0)$, scaled by the terms of the sequence on the right hand side in (5.1). Thus, by (4.5) and Lemma 5.3,

$$\textsc{sSkewer}(y,\mathbf{F}) \stackrel{d}{=} \sum_{i\geq 1}\left(G(1-B_i)\prod_{j=1}^{i-1}B_j\right)\overline{\Pi}_i,$$

which has the claimed distribution. □

5.2. *Path-continuity and Markov property.*

PROPOSITION 5.4. *Let* $\theta \geq 0$, $\mathbf{F} \sim \texttt{PRM}(\frac{\theta}{\alpha}\texttt{Leb}\otimes\overline{\nu}^{(\alpha)}_{\perp\text{cld}})$ *and* $\pi^y := \textsc{sSkewer}(y,\mathbf{F})$, $y\geq 0$. *Then the process* $(\pi^y, y\geq 0)$ *is* d_{TV}*-path-continuous.*

We immediately deduce the following statement from Propositions 3.3 and 5.4.

COROLLARY 5.5. *In the setting of Theorem* 1.9(*ii*), *with* $\mathbf{F}\sim\mathbf{Q}^{\alpha,0}_\pi$, $\pi\in\mathcal{M}^a$, *the process* $(\textsc{sSkewer}(y,\mathbf{F})+\textsc{sSkewer}(y,\mathbf{F}), y\geq 0)$ *is* d_{TV}*-path-continuous starting from* π.

We denote the distribution of $(\textsc{sSkewer}(y,\mathbf{F})+\textsc{sSkewer}(y,\mathbf{F}), y\geq 0)$ on the space $\mathcal{C}([0,\infty),(\mathcal{M}^a,d_{\mathcal{M}}))$, by $\mathbb{Q}^{\alpha,\theta}_\pi$, for each $\pi\in\mathcal{M}^a$ and note the following consequence of the additivity of PRMs and other point measures.

COROLLARY 5.6 (Additivity property). *Let $\lambda_1, \lambda_2 \in \mathcal{M}^a$, with distinct atom locations. For independent $(\pi_1^y, y \geq 0) \sim \mathbb{Q}_{\lambda_1}^{\alpha,\theta_1}$ and $(\pi_2^y, y \geq 0) \sim \mathbb{Q}_{\lambda_2}^{\alpha,\theta_2}$, we have $(\pi_1^y + \pi_2^y, y \geq 0) \sim \mathbb{Q}_{\lambda_1+\lambda_2}^{\alpha,\theta_1+\theta_2}$.*

PROOF OF PROPOSITION 5.4. Fix $y_0 > 0$. To show d_{TV}-path-continuity on $[0, y_0]$, first consider the special case $\theta \in (0, \alpha]$. Our starting point is $\mathbf{V} \sim \mathtt{PRM}(\mathtt{Leb} \otimes \nu_{\mathtt{BESQ}}^{(-2\alpha)} \otimes \mathtt{Unif})$ with scaffolding $\mathbf{X} = \xi_{\mathbf{V}}$, and the associated $\mathtt{PRM}$ $\mathbf{F}_\perp$ of clades of the reflected process defined in (4.5), which we restrict to $[0, y_0] \times \mathcal{V} \times \mathcal{D}$. By thinning $\mathbf{F}_\perp$ retaining each clade with probability θ/α, we obtain a $\mathtt{PRM}(\frac{\theta}{\alpha}\mathtt{Leb} \otimes \overline{\nu}_{\perp\mathrm{cld}}^{(\alpha)})$, which we denote by $\mathbf{F}_\theta$. We concatenate the clades of $\mathbf{F}_\theta$ to a point measure $\mathbf{V}_\theta$ and a scaffolding $\mathbf{X}_\theta$ in which the excursions of the reflected process are still at the same levels as in $\mathbf{X}$; see Figure 4. By a space-time shift, we see that

$$\big(\mathrm{sSKEWER}(y, \mathbf{V}_\theta, y_0 + \mathbf{X}_\theta), y \in [0, y_0]\big) \stackrel{d}{=} \big(\pi^y, y \in [0, y_0]\big),$$

where $(\pi^y, y \geq 0)$ is as in the statement of the proposition.

Let $\gamma \in (0, 1/2)$. The $\mathtt{BESQ}(-2\alpha)$ excursions Z_t marking $\Delta\mathbf{X}_\theta(t)$ have Hölder constants $D_\gamma^*(t) = \sup_{0\leq x<y\leq \Delta X(t)} |Z_t(y) - Z_t(x)|/|y - x|^\gamma < \infty$ a.s., by [20], Corollary 36. If furthermore $\gamma < 1-\alpha$, then Lemma 2.9 yields that the set of jump times of $\mathbf{X}$ and hence $y_0 + \mathbf{X}_\theta$, by thinning, may a.s. be partitioned into a sequence of "piles" of jumps $\{J_j^n, j \geq 1\}$, $n \geq 1$, in such a way that the jump intervals $[\mathbf{X}(J_j^n-), \mathbf{X}(J_j^n)]$, $j \geq 1$, are disjoint for each $n \geq 1$ and the Hölder constants $D_n = \sup_{j\geq 1} D_\gamma^*(J_j^n)$ are summable in $n \geq 1$.

Furthermore, if $m\alpha < \theta \leq (m+1)\alpha$ for some $m \geq 1$, this conclusion applies with θ replaced by $\theta - m\alpha$, or with θ replaced by α. Taking the former and, independently, m copies of the latter, we can merge the $m+1$ sequences of piles into a single sequence of piles whose Hölder constants are still summable. Furthermore, we can superpose the $m+1$ Poisson random measures of clades of the reflected process to construct a point measure $\mathbf{V}_\theta$ and scaffolding $y_0 + \mathbf{X}_\theta$ such that the associated superskewer process is again distributed like $(\pi^y, y \in [0, y_0])$, as in the case $\theta \in (0, \alpha]$.

The proof of Proposition 3.1 is easily adapted to prove $(\pi^y, y \in [0, y_0])$ is a.s. Hölder-γ in $(\mathcal{M}^a, d_{\mathrm{TV}})$. □

To describe the Markov property we require additional notation. Recall the upper point measure of clades (2.11) in a pair $(V, X) \in \mathcal{V} \times \mathcal{D}$. For a point measure F on $[0, \infty) \times \mathcal{V} \times \mathcal{D}$ we similarly define

$$G_0^{\geq y}(F) := \sum_{\text{points } (s,V_s,X_s) \text{ of } F} \mathbf{1}\{s \in [0, y]\} G_0^{\geq y-s}(V_s, X_s), \tag{5.6}$$

and, extending the notation of (2.10), we define the lower cutoff process as

$$\mathrm{CUTOFF}^{\leq y}(F) := \sum_{\text{points } (s,V_s,X_s) \text{ of } F} \mathbf{1}\{s \in [0, y]\} \delta\big(s, \mathrm{CUTOFF}^{\leq y-s}(V_s, X_s)\big). \tag{5.7}$$

The following result is analogous to Proposition 2.14.

LEMMA 5.7 (Markov-like property). *Consider $\mathbf{F} \sim \mathtt{PRM}(\frac{\theta}{\alpha}\mathtt{Leb} \otimes \overline{\nu}_{\perp\mathrm{cld}}^{(\alpha)})$ on $[0, \infty) \times \mathcal{V} \times \mathcal{D}$ and $y > 0$. Then given $\mathrm{CUTOFF}^{\leq y}(\mathbf{F})$, we have the following conditional distribution:*

$$\big(\mathbf{F}|_{(y,\infty)\times\mathcal{V}\times\mathcal{D}}, G_0^{\geq y}(\mathbf{F})\big) \sim \mathtt{PRM}\Big(\frac{\theta}{\alpha}\mathtt{Leb}|_{(y,\infty)} \otimes \overline{\nu}_{\perp\mathrm{cld}}^{(\alpha)}\Big) \otimes \mathbf{Q}_{\pi^y}^{\alpha,0}, \tag{5.8}$$

where $\pi^y = \mathrm{sSKEWER}(y, \mathbf{F})$.

PROOF. We begin by decomposing $\mathbf{F}$ into a sum of three independent terms: the point process of clades that enter below level y and survive up to that level, those that enter below level y but do not survive to level y, and those that enter above level y;

$$\mathbf{F}_1 := \mathbf{F}|_{A^{\geq y}}, \qquad \mathbf{F}_2 := \mathbf{F}|_{A^{<y}}, \qquad \mathbf{F}_3 := \mathbf{F}|_{(y,\infty)\times\mathcal{V}\times\mathcal{D}}, \quad \text{where}$$
$$A^{\geq y} := \big\{(s, V, X) \in [0, y] \times \mathcal{V} \times \mathcal{D} \colon \zeta^+(V) + s \geq y\big\},$$
$$A^{<y} := \big\{(s, V, X) \in [0, y] \times \mathcal{V} \times \mathcal{D} \colon \zeta^+(V) + s < y\big\}.$$

By the Poisson property, these are three independent PRMs. The claimed conditional independence and distribution of the first coordinate in (5.8) follow immediately from three observations: this first coordinate equals $\mathbf{F}_3$, we are conditioning on a function of $\mathbf{F}_1 + \mathbf{F}_2$, and $G_0^{\geq y}(\mathbf{F}) = G_0^{\geq y}(\mathbf{F}_1)$.

Note that we are conditioning on $\text{CUTOFF}^{\leq y}(\mathbf{F})$, which equals $\text{CUTOFF}^{\leq y}(\mathbf{F}_1) + \text{CUTOFF}^{\leq y}(\mathbf{F}_2)$. Again, appealing to the independence given by the Poisson property, it remains only to show that given $\mathbf{G} := \text{CUTOFF}^{\leq y}(\mathbf{F}_1)$, we get $\mathbf{Q}^{\alpha,0}_{\pi^y}$ as the conditional distribution of $G_0^{\geq y}(\mathbf{F})$.

Recall from the proof of Proposition 5.1 the notation $((S_i, V_i, X_i), i \geq 1)$, for the sequence of points of $\mathbf{F}$ that give all the clades (V_i, X_i) surviving past level y in the increasing order of their immigration times S_i. In this notation,

$$G_0^{\geq y}(\mathbf{F}) = \sum_{i\geq 1} G_0^{\geq y - S_i}(V_i, X_i), \qquad \mathbf{G} = \sum_{i\geq 1} \delta\big(S_i, \text{CUTOFF}^{\leq y - S_i}(V_i, X_i)\big).$$

The Poisson property and the Markov-like property of Corollary 4.8 yield (i) $G_0^{\geq y - S_i}(V_i, X_i)$, $i \geq 1$, are conditionally independent given $\mathbf{G}$, and (ii) each $G_0^{\geq y - S_i}(V_i, X_i)$ then has conditional law $\mathbf{Q}^{\alpha,0}_{\pi_i^y}$, where $\pi_i^y = \text{SSKEWER}(y - S_i, V_i, X_i)$. Since $\pi^y = \sum_{i\geq 1} \pi_i^y$, and all types are distinct a.s., we deduce from Corollary 2.8 that $G_0^{\geq y}(\mathbf{F})$ has the conditional law $\mathbf{Q}^{\alpha,0}_{\pi^y}$, given $\mathbf{G}$. □

PROPOSITION 5.8 (Markov property). *Let μ be a probability measure on $\mathcal{M}^a$ and $(\pi^z, z \geq 0) \sim \mathbb{Q}^{\alpha,\theta}_\mu$. For any $y \geq 0$, the process $(\pi^{y+r}, r \geq 0)$ is conditionally independent of $(\pi^z, z \leq y)$ given $\pi^y = \pi$ and has regular conditional distribution $\mathbb{Q}^{\alpha,\theta}_\pi$.*

PROOF. Consider independent $\mathbf{F} \sim \mathbf{Q}^{\alpha,\theta}_0$ and $\mathbf{F} \sim \mathbf{Q}^{\alpha,0}_\pi$. Applying the superskewer process to the lower cutoff process and the upper point measure in Lemma 5.7 yields that $(\text{SSKEWER}(y+r, \mathbf{F}), r \geq 0) \sim \mathbb{Q}^{\alpha,\theta}_{\pi_1}$ is conditionally independent of $(\text{SSKEWER}(z, \mathbf{F}), z \leq y)$ given $\text{SSKEWER}(y, \mathbf{F}) = \pi_1$.

Independently, the Markov property of $\text{SS}\overline{\text{KEWER}}(\mathbf{F})$ in Proposition 3.5 yields that $(\text{SSKEWER}(y+r, \mathbf{F}), r \geq 0) \sim \mathbb{Q}^{\alpha,0}_{\pi_2}$ is conditionally independent of $(\text{SSKEWER}(z, \mathbf{F}), z \leq y)$ given $\text{SSKEWER}(y, \mathbf{F}) = \pi_2$.

By Corollary 5.6, as π_1 and π_2 a.s. have no shared types, the conditional distribution of the post-y superskewer process $(\text{SSKEWER}(y + r, \mathbf{F}) + \text{SSKEWER}(y + r, \mathbf{F}), r \geq 0)$ is $\mathbb{Q}^{\alpha,\theta}_\pi$, which only depends on $\text{SSKEWER}(y, \mathbf{F}) + \text{SSKEWER}(y, \mathbf{F}) = \pi_1 + \pi_2 =: \pi$. This completes the proof. □

PROPOSITION 5.9 (Continuity in the initial state). *For a sequence $\pi_n \to \pi_\infty$ in $(\mathcal{M}^a, d_{\text{TV}})$, there is the weak convergence $\mathbb{Q}^{\alpha,\theta}_{\pi_n} \to \mathbb{Q}^{\alpha,\theta}_{\pi_\infty}$ in the sense of finite-dimensional distributions on $(\mathcal{M}^a, d_{\mathcal{M}})$.*

PROOF. By Proposition 3.6 we have $\mathbb{Q}^{\alpha,0}_{\pi_n} \to \mathbb{Q}^{\alpha,0}_{\pi_\infty}$. Trivially, also $\mathbb{Q}^{\alpha,\theta}_0 \to \mathbb{Q}^{\alpha,\theta}_0$, and we conclude referring to Corollary 5.6 and by noting that addition is a continuous operation on $\mathcal{C}([0,\infty),(\mathcal{M}^a, d_{\mathcal{M}}))$. □

Finally, we have all the ingredients to apply the proof of Proposition 3.7 to establish the strong Markov property under $\mathbb{Q}^{\alpha,\theta}_\mu$.

PROPOSITION 5.10 (Strong Markov property). *For a probability measure μ on $\mathcal{M}^a$, let $(\pi^z, z \geq 0) \sim \mathbb{Q}^{\alpha,\theta}_\mu$. Denote its right-continuous natural filtration by $(\mathcal{F}^y, y \geq 0)$. Let Y be an a.s. finite $(\mathcal{F}^y, y \geq 0)$-stopping time. Then given $\mathcal{F}^Y$, the process $(\pi^{Y+y}, y \geq 0)$ has conditional distribution $\mathbb{Q}^{\alpha,\theta}_{\pi^Y}$.*

5.3. *Proofs of Theorems 1.2, 1.5, 1.9(ii) and Propositions 1.3–1.4.* We now pull the threads together and prove the remaining results about $\mathtt{SSSP}(\alpha,\theta)$ stated in the introduction, that is, the Hunt and path-continuity properties, transition kernels and the superskewer construction (Theorems 1.2 and 1.9, when $\theta > 0$), the self-similarity (Proposition 1.3) and additivity property (Proposition 1.4) and the $\mathtt{BESQ}(2\theta)$ total mass process (Theorem 1.5).

PROOF OF THE $\theta > 0$ CASES OF THEOREMS 1.2 AND 1.9. These results follow by the same arguments as those used to prove the $\theta = 0$ cases of Theorems 1.2 and 1.9 in Section 3.3, replacing the intermediate results with corresponding results in the $\theta > 0$ case, namely Proposition 5.2, Lemma 2.5, Corollary 5.5, Proposition 5.8 and Proposition 5.10. □

PROOF OF PROPOSITION 1.3. Let $\mathbf{F} \sim \mathtt{PRM}(\frac{\theta}{\alpha}\mathtt{Leb} \otimes \overline{\nu}^{(\alpha)}_{\perp\mathrm{cld}})$ and $\mathbf{F} \sim \mathbf{Q}^{\alpha,0}_\pi$ be independent. Then we can write

$$\pi^y = \mathrm{sSKEWER}(y,\mathbf{F}) + \mathrm{sSKEWER}(y,\mathbf{F}), \quad y \geq 0.$$

Recall the scaling operator defined in (2.5). We define

$$\mathbf{F}_c = \sum_{\text{points } (s,V,X) \text{ of } \mathbf{F}} \delta\big(cs, c \odot^\alpha_{\mathrm{cld}} V, c \odot^{1+\alpha}_{\mathrm{stb}} X\big),$$

where $c \odot^{1+\alpha}_{\mathrm{stb}} X = (cX(t/c^{1+\alpha}), t \geq 0)$. Similarly we define $\mathbf{F}_c$. Notice that, for any $V \in \mathcal{V}$ and $X \in \mathcal{D}$, we have the identity

$$\mathrm{sSKEWER}\big(y, c \odot^\alpha_{\mathrm{cld}} V, c \odot^{1+\alpha}_{\mathrm{stb}} X\big) = c \cdot \mathrm{sSKEWER}(y/c, V, X), \quad y \geq 0.$$

It follows that

$$c\pi^{y/c} = \mathrm{sSKEWER}(y,\mathbf{F}_c) + \mathrm{sSKEWER}(y,\mathbf{F}_c), \quad y \geq 0.$$

By Lemma 2.11 and Lemma 4.4 respectively, $\mathbf{F}_c \sim \mathbf{Q}^{\alpha,0}_{c\pi}$ and $\mathbf{F}_c \sim \mathtt{PRM}(\frac{\theta}{\alpha}\mathtt{Leb} \otimes \overline{\nu}^{(\alpha)}_{\perp\mathrm{cld}})$, which implies that $(c\pi^{y/c}, y \geq 0) \sim \mathbb{Q}^{\alpha,\theta}_{c\pi}$. □

PROOF OF PROPOSITION 1.4. This now follows from Corollary 5.6. □

PROOF OF THEOREM 1.5. Consider independent $(\pi^z_1, z \geq 0) \sim \mathbb{Q}^{\alpha,\theta}_0$ and $(\pi^z_2, z \geq 0) \sim \mathbb{Q}^{\alpha,0}_\pi$, as in Corollary 5.6, for any $\pi \in \mathcal{M}^a$. Then the process $(\pi^z = \pi^z_1 + \pi^z_2, z \geq 0)$ is an $\mathtt{SSSP}_\pi(\alpha,\theta)$. Denote its right-continuous natural filtration by $(\mathcal{F}^y, y \geq 0)$.

Proposition 2.10 shows that $(\|\pi^z_2\|, z \geq 0)$ is a $\mathtt{BESQ}_{\|\pi\|}(0)$. For any fixed $y \geq 0$, by Proposition 5.1 we deduce that $\|\pi^y_1\| \sim \mathtt{Gamma}(\theta, 1/2y)$, which is the marginal distribution of a $\mathtt{BESQ}_0(2\theta)$ at time y (see, e.g., [40], Corollary XI.(1.4)). Using the additivity property of

squared Bessel processes (see, e.g., [40], Theorem XI.(1.2)), we deduce that $\|\pi^y\|$ has the marginal distribution of a $\mathtt{BESQ}_{\|\pi\|}(2\theta)$ at time y.

Next, by the marginal distributions described above and the Markov property of the process $(\pi^z, z \geq 0)$, Proposition 5.8, we deduce that the total-mass process $(\|\pi^z\|, z \geq 0)$ possesses the Markov property with respect to the filtration $(\mathcal{F}^y, y \geq 0)$.

Finally, since the total-mass function is continuous on $(\mathcal{M}^a, d_{\mathcal{M}})$, we deduce by the path-continuity of an $\mathtt{SSSP}_\pi(\alpha,\theta)$ that the total-mass process has a.s. continuous paths. This completes the proof. □

5.4. *An emigration-immigration approach to* $\mathtt{SSSP}_{b\delta(x)}(\alpha, 0)$.

COROLLARY 5.11. *Fix* $x \in [0,1]$. *For* $\mathbf{f} \sim \mathtt{BESQ}_b(-2\alpha)$ *and* $(\pi^y, y \geq 0) \sim \mathtt{SSSP}_0(\alpha, \alpha)$ *independent, let* $\pi^y := \mathbf{f}(y)\delta(x) + \pi^y$, $0 \leq y \leq \zeta(\mathbf{f})$, *and conditionally given* $(\pi^y, y \in [0, \zeta(\mathbf{f})])$ *with* $\pi^{\zeta(\mathbf{f})} = \lambda$, *let* $(\pi^{\zeta(\mathbf{f})+z}, z \geq 0) \sim \mathtt{SSSP}_\lambda(\alpha, 0)$. *Then* $(\pi^y, y \geq 0) \sim \mathtt{SSSP}_{b\delta(x)}(\alpha, 0)$.

See [22], Proposition 3.15, for a related result for interval partition evolutions. Recalling from Theorem 1.5 that $\mathtt{SSSP}(\alpha,\alpha)$ has $\mathtt{BESQ}(2\alpha)$ total mass, we interpret $\mathtt{BESQ}(-2\alpha)$ here as a process with emigration of mass at rate -2α. This is compensated by immigration at rate 2α provided by $\mathtt{SSSP}(\alpha,\alpha)$ during the period $[0, \zeta(\mathbf{f})]$ of emigration. In fact, this corollary generalizes to a connection between $\mathtt{SSSP}(\alpha,\theta)$ and $\mathtt{SSSP}(\alpha,\theta-\alpha)$ for all $\theta \geq \alpha$; we leave the details to the reader.

PROOF. From independent $\mathbf{f} \sim \mathtt{BESQ}_b(-2\alpha)$ and $\mathbf{V} \sim \mathtt{PRM}(\mathtt{Leb} \otimes \nu_{\mathtt{BESQ}}^{(-2\alpha)} \otimes \mathtt{Unif})$, we can construct both $\widehat{\mathbf{V}} := \delta(0, \mathbf{f}, x) + \mathbf{V}|_{[0, T^{-\zeta(\mathbf{f})}]} \sim \mathbf{Q}_{b,x}^{(\alpha)}$ as in (1.6) and $\mathbf{F}_\perp \sim \mathtt{PRM}(\mathtt{Leb} \otimes \overline{\nu}_{\perp\mathrm{cld}}^{(\alpha)})$ as in (4.5) and Proposition 4.3. We also let

$$\mathbf{F}' = \sum_{\text{points } (z,V_z,X_z) \in [0,\zeta(\mathbf{f})] \times \mathcal{V} \times \mathcal{D} \text{ of } \mathbf{F}_\perp} \delta\big(\zeta(\mathbf{f}) - z, V_z, X_z\big).$$

By Theorem 1.9(ii), $(\mathtt{SSKEWER}(y, \mathbf{F}'), y \in [0, \zeta(\mathbf{f})])$ is distributed as an $\mathtt{SSSP}_0(\alpha,\alpha)$ stopped at an independent time $Y \overset{d}{=} \zeta(\mathbf{f})$. On the other hand, $\mathbf{f}\delta(x) + \overline{\mathtt{SSKEWER}}(\mathbf{F}') = \overline{\mathtt{SSKEWER}}(\widehat{\mathbf{V}}, \xi_{\widehat{\mathbf{V}}}) \sim \mathtt{SSSP}_{b\delta(x)}(\alpha, 0)$. Finally, by the strong Markov property of $(\pi^y, y \geq 0)$ $:= \overline{\mathtt{SSKEWER}}(\widehat{\mathbf{V}}, \xi_{\widehat{\mathbf{V}}})$ at the stopping time $\zeta(\mathbf{f})$, we get $(\mathtt{SSKEWER}(\zeta(\mathbf{f}) + y, \mathbf{F}'), y \geq 0) \sim \mathtt{SSSP}_\lambda(\alpha, 0)$ conditionally given $(\pi^y, y \in [0, \zeta(\mathbf{f})])$ with $\lambda = \pi^{\zeta(\mathbf{f})}$. □

6. (α,θ)-Fleming–Viot processes. Before we explicitly turn to the study of $\mathtt{FV}(\alpha,\theta)$ with $\mathtt{PDRM}(\alpha,\theta)$ stationary distribution in Section 6.2, by time-changing $\mathtt{SSSP}(\alpha,\theta)$ and normalising to unit total mass, let us, in Section 6.1, prepare the stationarity at the level of $\mathtt{SSSP}(\alpha,\theta)$, where we observe a certain decoupling of the nonstationary total mass evolution from stationary $\mathtt{PDRM}(\alpha,\theta)$ mass proportions. This was similarly observed in [22], Theorem 1.5, for the type-0 and type-1 interval partition evolutions that relate to the cases $\theta = \alpha$ and $\theta = 0$ here.

6.1. *Pseudo-stationarity of self-similar* (α,θ)*-superprocesses.*

THEOREM 6.1 (Pseudo-stationarity for $\mathtt{SSSP}(\alpha,\theta)$). *Let* $\theta \geq 0$. *Consider an independent pair* $(Z, \overline{\Pi})$, *where* $\overline{\Pi} \sim \mathtt{PDRM}(\alpha,\theta)$ *and* $Z = (Z(y), y \geq 0)$ *is a* $\mathtt{BESQ}(2\theta)$ *with an arbitrary initial distribution. Let* $(\pi^y, y \geq 0)$ *be an* $\mathtt{SSSP}(\alpha,\theta)$ *with* $\pi^0 \overset{d}{=} Z(0)\overline{\Pi}$. *Then for each fixed* $y \geq 0$ *we have* $\pi^y \overset{d}{=} Z(y)\overline{\Pi}$.

In Theorem 6.4 we will generalize this result from fixed times to certain stopping times. For now, to prove Theorem 6.1, we first consider a special case for the law of $Z(0)$.

PROPOSITION 6.2. *In the setting of Theorem 6.1 with $\theta > 0$, suppose that $Z(0) \sim \texttt{Gamma}(\theta, \rho)$ for some $\rho \in (0, \infty)$. Then $\pi^y \overset{d}{=} (2y\rho + 1)Z(0)\overline{\Pi}$, for every fixed $y \geq 0$.*

PROOF. Let $(\lambda^z, z \geq 0)$ be an $\texttt{SSSP}(\alpha, \theta)$ starting from 0. Fix $y \geq 0$. By Proposition 5.1, we have $\lambda^{1/2\rho} \overset{d}{=} Z(0)\overline{\Pi}$ and $\lambda^{y+1/2\rho} \overset{d}{=} (2y\rho+1)Z(0)\overline{\Pi}$, using the fact that $(2y\rho+1)Z(0) \sim \texttt{Gamma}(\theta, \rho/(2yp+1))$.

Since it follows from the simple Markov property that the process $(\widetilde{\lambda}^z := \lambda^{(1/2\rho)+z}, z \geq 0)$ is an $\texttt{SSSP}(\alpha, \theta)$ starting from $\lambda^{1/2\rho} \overset{d}{=} Z(0)\overline{\Pi}$, we have $\pi^y \overset{d}{=} \widetilde{\lambda}^y \overset{d}{=} (2y\rho + 1)Z(0)\overline{\Pi}$. This is the desired statement. □

PROOF OF THEOREM 6.1. First fix $\theta > 0$. For every $b \geq 0$, denote by $\mu_b^{\alpha,\theta}$ the distribution of $b\overline{\Pi}$ on $(\mathcal{M}^a, d_{\mathcal{M}})$. For every bounded continuous function f on $(\mathcal{M}^a, d_{\mathcal{M}})$ with $f(0) = 0$ and $y \geq 0$, Proposition 6.2 yields

$$\begin{aligned}
&\int_0^\infty e^{-\rho b}\Gamma(\theta)^{-1}\rho^\theta b^{\theta-1}\mathbb{Q}^{\alpha,\theta}_{\mu_b^{\alpha,\theta}}\big[f(\pi^y)\big]\,db \\
&\quad= \int_0^\infty \Gamma(\theta)^{-1}\rho^\theta b^{\theta-1}e^{-\rho b}\mathbf{E}\big[f\big((2y\rho+1)b\overline{\Pi}\big)\big]\,db \\
&\quad= \int_0^\infty \Gamma(\theta)^{-1}\frac{\rho^\theta}{(2y\rho+1)^\theta}c^{\theta-1}e^{-\rho c/(2y\rho+1)}\mathbf{E}\big[f(c\overline{\Pi})\big]\,dc.
\end{aligned}$$

For every $b \geq 0$, let $(Z_b(y), y \geq 0)$ be a $\texttt{BESQ}_b(2\theta)$, independent of $\overline{\Pi}$. It is known [40], Corollary XI.(1.4), that $Z_0(y) \sim \texttt{Gamma}(\theta, 1/2y)$ for every $y > 0$. Using this fact and the Markov property of $\texttt{BESQ}(2\theta)$, we deduce the identity

$$\begin{aligned}
&\int_0^\infty \Gamma(\theta)^{-1}\frac{\rho^\theta}{(2y\rho+1)^\theta}c^{\theta-1}e^{-\rho c/(2y\rho+1)}\mathbf{E}\big[f(c\overline{\Pi})\big]\,dc \\
&\quad= \mathbf{E}\big[\mathbf{E}\big[f\big(Z_0(y+1/2\rho)\overline{\Pi}\big)|Z_0(1/2\rho)\big]\big] \\
&\quad= \int_0^\infty e^{-\rho b}\Gamma(\theta)^{-1}\rho^\theta b^{\theta-1}\mathbf{E}\big[f\big(Z_b(y)\overline{\Pi}\big)\big]\,db.
\end{aligned}$$

By the uniqueness theorem for Laplace transforms, we find that

$$\mathbb{Q}^{\alpha,\theta}_{\mu_b^{\alpha,\theta}}\big[f(\pi^y)\big] = \mathbf{E}\big[f\big(Z_b(y)\overline{\Pi}\big)\big], \quad \text{for Lebesgue-a.e. } b > 0.$$

Since the map $b \mapsto b\overline{\Pi}$ is d_{TV}-continuous, we have by Proposition 5.9 that the map $b \mapsto \mathbb{Q}^{\alpha,\theta}_{\mu_b^{\alpha,\theta}}[f(\pi^y)]$ is also continuous. So we conclude that $\mathbb{Q}^{\alpha,\theta}_{\mu_b^{\alpha,\theta}}[f(\pi^y)] = \mathbf{E}[f(Z_b(y)\overline{\Pi})]$ for every $b \geq 0$. This identifies the distribution of π^y because bounded continuous functions separate points in $(\mathcal{M}^a, d_{\mathcal{M}})$, due to its Lusin property. This completes the proof for $\theta > 0$.

For any sequence $\theta_k \downarrow \theta_0 := 0$, consider independent $\overline{\Pi}'_k \sim \texttt{PDRM}(\alpha, \theta_k)$, $k \geq 0$, also independent of a sequence $B_k \sim \texttt{Beta}(\theta_k, 1)$, $k \geq 1$, with $B_k \downarrow 0 =: B_0$ a.s. By (5.4), we have $\overline{\Pi}_k := B_k\overline{\Pi}'_k + (1 - B_k)\overline{\Pi}'_0 \sim \texttt{PDRM}(\alpha, \theta_k)$, $k \geq 0$ and $d_{\mathrm{TV}}(\overline{\Pi}_k, \overline{\Pi}_0) \to 0$ a.s. For $Z_b^{(k)} \sim \texttt{BESQ}_b(2\theta_k)$, $k \geq 0$, we also have $Z_b^{(k)}(y) \to Z_b^{(0)}(y)$ in distribution, so the established time-y distribution for $\theta_k > 0$ converges to the claimed time-y distribution in the case

$\theta = 0$, weakly in $(\mathcal{M}^a, d_{\mathcal{M}})$. On the other hand, for all continuous $\phi\colon [0,1] \to [0,\infty)$

$$\begin{aligned}
&\mathbb{Q}^{\alpha,\theta_k}_{\mu_b^{\alpha,\theta_k}}\biggl[\exp\Bigl(-\int_{[0,1]}\phi(x)\pi^y(dx)\Bigr)\biggr]\\
&\quad= \mathbb{Q}^{\alpha,0}_{\mu_b^{\alpha,\theta_k}}\biggl[\exp\Bigl(-\int_{[0,1]}\phi(x)\pi^y(dx)\Bigr)\biggr]\mathbb{Q}^{\alpha,\theta_k}_{0}\biggl[\exp\Bigl(-\int_{[0,1]}\phi(x)\pi^y(dx)\Bigr)\biggr]\\
&\quad\to \mathbb{Q}^{\alpha,0}_{\mu_b^{\alpha,0}}\biggl[\exp\Bigl(-\int_{[0,1]}\phi(x)\pi^y(dx)\Bigr)\biggr],
\end{aligned}$$

by Proposition 3.6 for the first factor, using $d_{\mathrm{TV}}(b\overline{\Pi}_k, b\overline{\Pi}_0) \to 0$ a.s., and by the monotone convergence theorem (suitably coupled by thinning $\mathbf{F}_{\theta_k} \sim \mathtt{PRM}(\frac{\theta_k}{\alpha}\mathtt{Leb}\otimes\overline{\nu}^{(\alpha)}_{\perp\mathrm{cld}})$ as $\theta_k \downarrow 0$) for the second factor. By the uniqueness theorem for Laplace functionals, $\mathbb{Q}^{\alpha,0}_{\mu_b^{\alpha,0}}(\pi^y \in \cdot)$ is the claimed time-y distribution. □

We can further strengthen the pseudo-stationarity of Theorem 6.1. For the purpose of the following results, we write $\|\boldsymbol{\pi}\| := (\|\pi^y\|, y \ge 0)$.

PROPOSITION 6.3. *Let $\theta \ge 0$. Consider an independent pair $(Z, \overline{\Pi})$, where $\overline{\Pi} \sim \mathtt{PDRM}(\alpha,\theta)$ and $Z = (Z(y), y \ge 0)$ is a $\mathtt{BESQ}(2\theta)$ with an arbitrary initial distribution. Let $(\pi^y, y \ge 0)$ be an $\mathtt{SSSP}(\alpha,\theta)$ with $\pi^0 \overset{d}{=} Z(0)\overline{\Pi}$. Then for any $y \ge 0$ and nonnegative measurable functions $h\colon (\mathcal{M}_1^q, d_{\mathcal{M}}) \to [0,\infty)$ and $\eta\colon \mathcal{C}([0,\infty),[0,\infty)) \to [0,\infty)$, we have*

$$\mathbf{E}\bigl[\eta(\|\boldsymbol{\pi}\|)\mathbf{1}\{\pi^y \ne 0\}h(\|\pi^y\|^{-1}\pi^y)\bigr] = \mathbf{E}\bigl[\eta(Z)\mathbf{1}\{Z(y) \ne 0\}\bigr]\mathbf{E}\bigl[h(\overline{\Pi})\bigr].$$

PROOF. We adapt the proof of [22], Lemma 4.7. Denote by μ_b the distribution of $b\overline{\Pi}$ and consider $0 \le y_1 < \cdots < y_n = y$ and $f_1, \dots, f_n\colon [0,\infty) \to [0,\infty)$. Then the Markov property and pseudo-stationarity of $(\pi^y, y \ge 0)$ at y_1 yield

$$\begin{aligned}
&\mathbf{E}\biggl[\prod_{j=1}^n f_j(\|\pi^{y_j}\|)\mathbf{1}\{\pi^y \ne 0\}h(\|\pi^y\|^{-1}\pi^y)\biggr]\\
&\quad= \mathbf{E}\biggl[f_1(\|\pi^{y_1}\|)\mathbb{Q}^{\alpha,\theta}_{\pi^{y_1}}\biggl[\prod_{j=2}^n f_j(\|\pi^{y_j-y_1}\|)\mathbf{1}\{\pi^{y-y_1} \ne 0\}h(\|\pi^{y-y_1}\|^{-1}\pi^{y-y_1})\biggr]\biggr]\\
&\quad= \mathbf{E}\biggl[f_1(Z(y_1))\mathbb{Q}^{\alpha,\theta}_{\mu_{Z(y_1)}}\biggl[\prod_{j=2}^n f_j(\|\pi^{y_j-y_1}\|)\mathbf{1}\{\pi^{y-y_1} \ne 0\}h(\|\pi^{y-y_1}\|^{-1}\pi^{y-y_1})\biggr]\biggr].
\end{aligned}$$

For $n = 1$, we have $y = y_1$, so $\pi^{y-y_1} = \pi^0 \sim b\overline{\Pi}$ under $\mathbb{Q}^{\alpha,\theta}_{\mu_b}$, for all $b \ge 0$, so

$$\mathbb{Q}^{\alpha,\theta}_{\mu_b}\bigl[\mathbf{1}\{\pi^{y-y_1} \ne 0\}h(\|\pi^{y-y_1}\|^{-1}\pi^{y-y_1})\bigr] = \mathbf{1}\{b \ne 0\}\mathbf{E}\bigl[h(\overline{\Pi})\bigr],$$

as required, and for $n \ge 2$ we obtain inductively that

$$\begin{aligned}
&\mathbf{E}\biggl[\prod_{j=1}^n f_j(\|\pi^{y_j}\|)\mathbf{1}\{\pi^y \ne 0\}h(\|\pi^y\|^{-1}\pi^y)\biggr]\\
&\quad= \mathbf{E}\biggl[f_1(Z(y_1))\mathbb{Q}^{\alpha,\theta}_{\mu_{Z(y_1)}}\biggl[\prod_{j=2}^n f_j(\|\pi^{y_j-y_1}\|)\mathbf{1}\{\|\pi^{y-y_1}\| \ne 0\}\biggr]\mathbf{E}\bigl[h(\overline{\Pi})\bigr]\biggr]\\
&\quad= \mathbf{E}\biggl[\prod_{j=1}^n f_j(Z(y_j))\mathbf{1}\{Z(y) \ne 0\}\biggr]\mathbf{E}\bigl[h(\overline{\Pi})\bigr],
\end{aligned}$$

where the last step applied the Markov property of the $\mathtt{BESQ}(2\theta)$ total mass process at y_1. The argument is easily adapted by further applications of the Markov properties at $y = y_n$ to extend the product by further terms $f_j(\|\pi^{y_j}\|)$ for $y_j > y$, $n+1 \le j \le n+m$. The application of a monotone class theorem completes the proof. $\square$

THEOREM 6.4. *Let $(\pi^y, y \ge 0)$ be an $\mathtt{SSSP}(\alpha, \theta)$ with $\theta \ge 0$ and $\pi^0 \overset{d}{=} B\overline{\Pi}$, where B is an arbitrary nonnegative random variable and $\overline{\Pi} \sim \mathtt{PDRM}(\alpha, \theta)$ is independent of B. Let Y be an a.s. finite $\sigma(\|\pi^y\|, y \ge 0)$-measurable random time. Then for any measurable functions $h\colon (\mathcal{M}_1^a, d_{\mathcal{M}}) \to [0,\infty)$ and $\eta\colon \mathcal{C}([0,\infty),[0,\infty)) \to [0,\infty)$, we have*

$$\mathbf{E}\big[\eta(\|\boldsymbol{\pi}\|)\mathbf{1}\{\pi^Y \ne 0\}h(\|\pi^Y\|^{-1}\pi^Y)\big] = \mathbf{E}\big[\eta(\|\boldsymbol{\pi}\|)\mathbf{1}\{\pi^Y \ne 0\}\big]\mathbf{E}\big[h(\overline{\Pi})\big].$$

PROOF. We use the standard dyadic approximation. Let $Y_n = 2^{-n}\lceil 2^n Y\rceil$, which is a.s. finite and decreases to Y as $n \to \infty$. Then applying Proposition 6.3, for any bounded measurable function η, bounded continuous function h and $k \ge 1$, we have

$$\begin{aligned}
&\mathbf{E}\big[\eta(\|\boldsymbol{\pi}\|)\mathbf{1}\{\pi^{Y_n} \ne 0; Y_n = k2^{-n}\}h(\|\pi^{Y_n}\|^{-1}\pi^{Y_n})\big]\\
&\quad = \mathbf{E}\big[\eta(\|\boldsymbol{\pi}\|)\mathbf{1}\{Y_n = k2^{-n}\}\mathbf{1}\{\pi^{k2^{-n}} \ne 0\}h(\|\pi^{k2^{-n}}\|^{-1}\pi^{k2^{-n}})\big]\\
&\quad = \mathbf{E}\big[\eta(\|\boldsymbol{\pi}\|)\mathbf{1}\{Y_n = k2^{-n}\}\mathbf{1}\{\pi^{Y_n} \ne 0\}\big]\mathbf{E}\big[h(\overline{\Pi})\big].
\end{aligned}$$

Summing over $k \ge 0$ leads to

$$\mathbf{E}\big[\eta(\|\boldsymbol{\pi}\|)\mathbf{1}\{\pi^{Y_n} \ne 0\}h(\|\pi^{Y_n}\|^{-1}\pi^{Y_n})\big] = \mathbf{E}\big[\eta(\|\boldsymbol{\pi}\|)\mathbf{1}\{\pi^{Y_n} \ne 0\}\big]\mathbf{E}\big[h(\overline{\Pi})\big].$$

Letting $n \to \infty$, we deduce by the path-continuity of (π^y) and the dominated convergence theorem that the last identity still holds when Y_n is replaced by Y. This proves the assertion in the theorem for bounded measurable η and bounded continuous function h. We complete the proof by a monotone class theorem. $\square$

6.2. *De-Poissonization and (α,θ)-Fleming–Viot processes.* Fix $\alpha \in (0,1)$ and $\theta \ge 0$. Let $\boldsymbol{\pi} := (\pi^y, y \ge 0)$ be an $\mathtt{SSSP}_\pi(\alpha,\theta)$ for some $\pi \in \mathcal{M}^a \setminus \{0\}$. Recall the time-change function defined by (1.5):

$$\rho_{\boldsymbol{\pi}}(u) := \inf\left\{y \ge 0\colon \int_0^y \|\pi^z\|^{-1}\,dz > u\right\}, \quad u \ge 0.$$

Since by Theorem 1.5 the total mass of $\boldsymbol{\pi}$ evolves according to $\mathtt{BESQ}(2\theta)$, by rewriting known results on squared Bessel processes in our setting, we have the following lemma.

LEMMA 6.5 ([25], p. 314-5). *The time-change function $\rho_{\boldsymbol{\pi}}$ is continuous and strictly increasing. Moreover, with the usual convention* $\inf\varnothing = \infty$,

$$\lim_{u\uparrow\infty} \rho_{\boldsymbol{\pi}}(u) = \inf\{y > 0\colon \pi^y = 0\}.$$

The limit is a.s. finite if $\theta \in [0,1)$ and a.s. infinite if $\theta \ge 1$.

We define an $\mathcal{M}_1^a$-valued process $\overline{\boldsymbol{\pi}} := (\overline{\pi}^u, u \ge 0)$ via the following so-called *de-Poissonization*:

$$\overline{\pi}^u := \|\pi^{\rho_\pi(u)}\|^{-1}\pi^{\rho_\pi(u)}, \quad u \ge 0.$$

Recall from Definition 1.6 that we have referred to the de-Poissonized process $\overline{\boldsymbol{\pi}}$ as an (α,θ)-*Fleming–Viot process*, or $\mathtt{FV}(\alpha,\theta)$. We have by Lemma 6.5 and the path-continuity of an $\mathtt{SSSP}_\pi(\alpha,\theta)$ that the sample paths of an $\mathtt{FV}(\alpha,\theta)$ are still continuous.

For any $c>0$, let $\boldsymbol{\pi}_c=(\pi_c^y=c\pi^{y/c},\ y\geq 0)$. For every $u\geq 0$, since $\rho_{\boldsymbol{\pi}_c}(u)=c\rho_{\boldsymbol{\pi}}(u)$, we have the equality between the de-Poissonized processes

$$\|\pi^{\rho_{\boldsymbol{\pi}}(u)}\|^{-1}\pi^{\rho_{\boldsymbol{\pi}}(u)}=\|\pi_c^{\rho_{\boldsymbol{\pi}_c}(u)}\|^{-1}\pi_c^{\rho_{\boldsymbol{\pi}_c}(u)}.$$

Therefore, for any $\boldsymbol{\lambda}:=(\lambda^y,\ y\geq 0)\sim \mathtt{SSSP}(\alpha,\theta)$ with $\lambda^0\overset{d}{=}c\pi^0$ for some constant $c>0$, then their associated de-Poissonized processes have the same distribution. Indeed, this results from the following two facts: due to the self-similarity, Proposition 1.3, $\boldsymbol{\lambda}$ and the rescaled process $\boldsymbol{\pi}_c$ have the same law and so do their de-Poissonized processes; on the other hand, we have seen that $\boldsymbol{\pi}$ has the same de-Poissonized process as $\boldsymbol{\pi}_c$. As a result, the law of a $\mathtt{FV}(\alpha,\theta)$ is characterized by its initial value. For any probability measure $\overline{\mu}$ on $\mathcal{M}_1^a$, we can denote by $\overline{\mathbb{Q}}_{\overline{\mu}}^{\alpha,\theta}$ the law on $\mathcal{C}([0,\infty),(\mathcal{M}_1^a,d_{\mathcal{M}}))$ of the de-Poissonized process of $\boldsymbol{\pi}\sim\mathbb{Q}_{\overline{\mu}}^{\alpha,\theta}$.

PROPOSITION 6.6 (Strong Markov property). *For a probability measure $\overline{\mu}$ on $\mathcal{M}_1^a$, let $(\overline{\pi}^u,u\geq 0)\sim\overline{\mathbb{Q}}_{\overline{\mu}}^{\alpha,\theta}$. Denote its right-continuous natural filtration by $(\overline{\mathcal{F}}^u,u\geq 0)$. Let U be an a.s. finite $(\overline{\mathcal{F}}^u,u\geq 0)$-stopping time. Then given $\overline{\mathcal{F}}^U$, the process $(\overline{\pi}^{U+u},u\geq 0)$ has conditional distribution $\overline{\mathbb{Q}}_{\overline{\pi}^U}^{\alpha,\theta}$.*

PROOF. Let $\boldsymbol{\pi}=(\pi^y,\ y\geq 0)\sim\mathbb{Q}_{\overline{\mu}}^{\alpha,\theta}$ with its right-continuous natural filtration denoted by $(\mathcal{F}^y,\ y\geq 0)$. We may assume that $(\overline{\pi}^u,u\geq 0)$ is the de-Poissonized process of $\boldsymbol{\pi}$ and then we have $\overline{\mathcal{F}}^u\subseteq\mathcal{F}^{\rho_{\boldsymbol{\pi}}(u)}$. Since $\rho_{\boldsymbol{\pi}}$ is $(\overline{\mathcal{F}}^u)$-adapted, continuous and strictly increasing, we have by [29], Proposition 7.9, that $Y:=\rho_{\boldsymbol{\pi}}(U)$ is an $(\mathcal{F}^y)$-stopping time and $\overline{\mathcal{F}}^U\subseteq\mathcal{F}^Y$. Write $\boldsymbol{\pi}'=(\pi^{Y+y},\ y\geq 0)$ for the shifted process. Then for $u\geq 0$, we have

$$\overline{\pi}^{U+u}=\|\pi^{\rho_{\boldsymbol{\pi}}(U+u)}\|^{-1}\pi^{\rho_{\boldsymbol{\pi}}(U+u)}=\|\pi^{Y+\rho_{\boldsymbol{\pi}'}(u)}\|^{-1}\pi^{Y+\rho_{\boldsymbol{\pi}'}(u)}.$$

This is to say, $(\overline{\pi}^{U+u},u\geq 0)$ is the de-Poissonized process of $\boldsymbol{\pi}'$. Now the desired result follows from the strong Markov property of $\boldsymbol{\pi}$, Proposition 5.10. □

PROOF OF THEOREM 1.7. As in the proof of Theorem 1.2, we shall check the following properties to prove that it is a path-continuous Hunt process. The space $(\mathcal{M}_1^a,d_{\mathcal{M}})$ is Lusin by Lemma 2.5. By Lemma 6.5 and the path-continuity of an $\mathtt{SSSP}(\alpha,\theta)$, the sample paths of the time-changed process, the $\mathtt{FV}(\alpha,\theta)$, are still continuous. Moreover, by [26], Theorem 2.6, our time-change operations are continuous maps from Skorokhod space to itself. We deduce that the semigroup of the time-changed process is still measurable. Finally, Proposition 6.6 gives the required strong Markov property.

We finally verify the stationary distribution. For $\overline{\Pi}\sim\mathtt{PDRM}(\alpha,\theta)$, let $\boldsymbol{\pi}:=(\pi^y,\ y\geq 0)$ be an $\mathtt{SSSP}(\alpha,\theta)$ starting from $\overline{\Pi}$ and $(\overline{\pi}^u,u\geq 0)$ its de-Poissonized process. Applying Theorem 6.4 to $\eta=1$ and $Y=\rho_{\boldsymbol{\pi}}(u)$ with $u\geq 0$, we have, for every measurable function $h\colon\mathcal{M}_1^a\to[0,\infty)$,

$$\mathbf{E}[h(\overline{\pi}^u)]=\mathbf{E}[\mathbf{1}\{\pi^{\rho_{\boldsymbol{\pi}}(u)}\neq 0\}h(\|\pi^{\rho_{\boldsymbol{\pi}}(u)}\|^{-1}\pi^{\rho_{\boldsymbol{\pi}}(u)})]=\mathbf{E}[h(\overline{\Pi})],$$

as $\mathbf{P}(\pi^{\rho_{\boldsymbol{\pi}}(u)}\neq 0)=1$ by Lemma 6.5. This is the desired claim. □

7. Properties of $\mathtt{FV}(\alpha,\theta)$ and $\mathtt{SSSP}(\alpha,\theta)$. In Section 7.1 we establish the connection to Shiga [44] that was noted at the end of Section 1.1. Then, in Sections 7.2–7.3 we prove Theorems 1.10–1.11, respectively.

7.1. *Coupling* $\mathtt{FV}(\alpha,\theta)$ *and* $\mathtt{FV}(0,\theta)$. Shiga [44], (3.12) and Theorem 3.6, gave a Poissonian construction for a large class of measure-valued processes. Let us discuss it for a $\mathtt{BESQ}(0)$ excursion law $\nu^{(0)}_{\mathtt{BESQ}}$ in the sense of Pitman and Yor [38]: for $\pi\in\mathcal{M}$, consider independent $\mathbf{W}_\pi\sim\mathtt{PRM}(\nu^{(0)}_{\mathtt{BESQ}}\otimes\pi)$ and $\mathbf{W}\sim\mathtt{PRM}(2\theta\mathtt{Leb}\otimes\nu^{(0)}_{\mathtt{BESQ}}\otimes\mathtt{Unif})$. The process

$$\pi_0^y(dx)=\int_{(0,y]\times\mathcal{E}} f(y-s)\delta(x)\mathbf{W}(ds,df,dx)+\int_{\mathcal{E}} f(y)\delta(x)\mathbf{W}_\pi(df,dx), \tag{7.1}$$

$y>0$, uniquely solves the martingale problem associated with the generator

$$LF(\lambda)=2\int_{[0,1]}\frac{\delta^2F(\lambda)}{\delta\lambda(x)^2}\lambda(dx)+2\theta\int_{[0,1]}\frac{\delta F(\lambda)}{\delta\lambda(x)}\,dx,$$

on a domain of functions F of the form $F(\lambda)=g(\langle\phi_1,\lambda\rangle,\dots,\langle\phi_k,\lambda\rangle)$ for some bounded measurable $\phi_i\colon[0,1]\to\mathbb{R}$ and bounded twice continuously differentiable functions $g\colon\mathbb{R}^k\to\mathbb{R}$, $k\ge1$. Here, for such F

$$\frac{\delta F(\lambda)}{\delta\lambda(x)}:=\sum_{i=1}^k\phi_i(x)\partial_i g\big(\langle\phi_1,\lambda\rangle,\dots,\langle\phi_k,\lambda\rangle\big),\quad\lambda\in\mathcal{M},$$

and $\delta^2F(\lambda)/\delta\lambda(x)^2$ is defined similarly. Then $(\overline{\pi}^u_0,u\ge0)$ defined via de-Poissonization as Definition 1.6 is the Fleming–Viot process of Ethier and Kurtz [14], see also [15], Theorem 8.1, with stationary distribution $\mathtt{PDRM}(0,\theta)$, which we denote by $\mathtt{FV}(0,\theta)$. We also refer to $(\pi_0^y,y\ge0)$ as $\mathtt{SSSP}(0,\theta)$.

Recall from [38], Theorem (4.1), that $Z=(Z_y,y\ge0)\sim\mathtt{BESQ}_a(0)$ can be constructed from $\mathbf{M}\sim\mathtt{PRM}(a\nu^{(0)}_{\mathtt{BESQ}})$ as $Z_y=\sum_{\text{points }f\text{ of}\mathbf{M}}f(y)=\int_{\mathcal{E}}f(y)\mathbf{M}(df)$. We denote by $\kappa(g,dM)$ a regular conditional distribution of $\mathbf{P}(\mathbf{M}\in dM|Z=g)$. In (7.1), Shiga's construction builds $\mathtt{BESQ}_a(0)$ mass evolutions for atoms of initial mass a in this manner, via $\mathbf{W}_\pi$.

To obtain a more precise connection between Shiga's framework and our framework, consider the following maps on $\mathcal{V}\times\mathcal{D}$:

- $\varphi_{\mathrm{mass}}(V,X):=\|\overline{\mathrm{SSKEWER}}(V,X)\|$, which is $\mathcal{E}$-valued except on a subset that is $\overline{\nu}^{(\alpha)}_{\perp\mathrm{cld}}$-null and $\mathbf{Q}^{(\alpha)}_{b,x}$-null for all $b>0$, $x\in[0,1]$;
- $\varphi_{\mathrm{type}}(V,X):=x$ if V has a unique initial point of the form $b\delta(0,f,x)$; this is well defined $\mathbf{Q}^{(\alpha)}_{b,x}$-almost surely for all $b>0$, $x\in[0,1]$.

In the left panel of Figure 4 we see a descending scaffolding and spindles with clades of the reflected process underlined with distinct colors. One may think of those underline colors as representing independent $\mathtt{Unif}$ types with which the clades of $\mathbf{F}\sim\mathbf{Q}^{\alpha,\theta}_0$ are marked. We obtain a point measure $\mathbf{W}$ as in (7.1) by replacing each clade of $\mathbf{F}$ by its total mass evolution, marked by this independent type.

THEOREM 7.1. *Fix $\pi\in\mathcal{M}^a$. Let $\mathbf{F}_\pi\sim\mathbf{Q}^{\alpha,0}_\pi$ and $\mathbf{F}\sim\mathbf{Q}^{\alpha,\theta}_0$ be independent point measures. Define $\mathbf{W}$ by mapping $\mathbf{F}$ via $(y,V,X)\mapsto(y,\varphi_{\mathrm{mass}}(V,X))$ and marking by independent $\mathtt{Unif}$ types. Define $\mathbf{W}^\circ_\pi$ by mapping each point of $\mathbf{F}_\pi$ via $(V,X)\mapsto(g,x):=(\varphi_{\mathrm{mass}}(V,X),\varphi_{\mathrm{type}}(V,X))$, then marking (g,x) by the kernel $\kappa(g,dM)$. Finally superpose the points (g_i,x_i,M_i) of $\mathbf{W}^\circ_\pi$ to a measure $\mathbf{W}_\pi:=\sum_{\text{points }(g,x,M)\text{ of }\mathbf{W}^\circ_\pi}\sum_{\text{points }f\text{ of }M}\delta(f,x)$. Then $\mathbf{W}_\pi\sim\mathtt{PRM}(\nu^{(0)}_{\mathtt{BESQ}}\otimes\pi)$ and, independently, $\mathbf{W}\sim\mathtt{PRM}(2\theta\mathtt{Leb}\otimes\nu^{(0)}_{\mathtt{BESQ}}\otimes\mathtt{Unif})$.*

PROOF. By Corollary 4.11, $2\alpha\nu^{(0)}_{\mathtt{BESQ}}$ is the push forward of $\overline{\nu}^{(\alpha)}_{\perp\mathrm{cld}}$ via φ_{mass}. By Proposition 2.10, each point $(V_i,X_i)\sim\mathbf{Q}^{(\alpha)}_{b_i,x_i}$ of $\mathbf{F}_\pi$ is mapped to a $\mathtt{BESQ}_{b_i}(0)$, independently.

Therefore, this proof is completed by standard mapping, marking and superposition of Poisson random measures. □

In order to state the following corollary, it is helpful to bring outcomes $\omega \in \Omega$ into our notation for superprocesses, for example, $(\pi_\omega^y, y \geq 0)$ would be a sample path of an SSSP defined on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$.

COROLLARY 7.2. *Fix $\alpha \in (0,1)$ and $\pi \in \mathcal{M}^a$. There exists a coupling of $(\pi^y, y \geq 0) \sim \mathtt{SSSP}_\pi(\alpha,\theta)$ and $(\lambda^y, y \geq 0) \sim \mathtt{SSSP}_\pi(0,\theta)$ on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ and a measurable function $h\colon \Omega \times [0,1] \to [0,1]$ such that*

$$\lambda_\omega^y = \pi_\omega^y \circ h_\omega^{-1} \quad \textit{where } h_\omega(u) := h(\omega,u), \quad \textit{for all } \omega \in \Omega,\ y \geq 0,$$

that is, λ^y is the h-pushforward of π^y. In particular, $\|\lambda^y\| = \|\pi^y\|$ for all $y \geq 0$.

PROOF. We apply the map $\overline{\text{SSKEWER}}$ to the coupled point measures of Theorem 7.1 and take h_ω to be the function that maps all types in each clade of $\mathbf{F}(\omega)$ or $\mathbf{F}_\pi(\omega)$ to the type of the corresponding point in $\mathbf{W}(\omega)$ or $\mathbf{W}_\pi^\circ(\omega)$, respectively, while sending the remainder of $[0,1]$ arbitrarily to 0. □

We recall a Poisson–Dirichlet identity [36], (5.26). Let $(A_i', i \geq 1) \sim \mathtt{PD}(0,\theta)$ and independently for each i, let $(B_{ij}, j \geq 1) \sim \mathtt{PD}(\alpha,0)$. Independent of these variables, let $(U_i', i \geq 1)$ and, for each i, $(U_{ij}, j \geq 1)$, be mutually independent i.i.d. Unif sequences. Then, setting $A_{ij} := A_i' B_{ij}$,

$$\Pi' := \sum_{i\geq 1} A_i' \delta(U_i') \sim \mathtt{PDRM}(0,\theta) \quad \text{and} \quad \Pi := \sum_{i\geq 1}\sum_{j\geq 1} A_{ij}\delta(U_{ij}) \sim \mathtt{PDRM}(\alpha,\theta). \tag{7.2}$$

The random measure Π is called an $(\alpha,0)$-*fragmentation* of Π'.

COROLLARY 7.3. *For any $\alpha \in (0,1)$, there is a coupling of jointly stationary $\mathtt{FV}(\alpha,\theta)$ and $\mathtt{FV}(0,\theta)$, with the joint stationary law that the $\mathtt{PDRM}(\alpha,\theta)$ is an $(\alpha,0)$-fragmentation of the $\mathtt{PDRM}(0,\theta)$.*

PROOF. We consider initial measures Π' and Π as in (7.2). We will modify the coupling of Theorem 7.1 by noting that each point (V_{ij}, X_{ij}) of $\mathbf{F}_\Pi$ can be further associated with the corresponding type U_i' in Π'. Specifically, by mapping (V_{ij}, X_{ij}, U_i') to $(g,x) := (\varphi_{\text{mass}}(V_{ij}, X_{ij}), U_i')$, all initial mass $A_i' = \sum_{j\geq 1} A_{ij}$ will be positioned at U_i'. With $\mathbf{F}$ as in Theorem 7.1, the proof of that theorem yields $\mathbf{W}_{\Pi'}$ and $\mathbf{W}$ as required for $\mathtt{SSSP}(0,\theta)$ starting from Π'.

As in Corollary 7.2, the total mass processes and so the de-Poissonization time-changes of the associated $\mathtt{SSSP}(\alpha,\theta)$ and $\mathtt{SSSP}(0,\theta)$ coincide. This establishes a coupling of stationary $\mathtt{FV}(\alpha,\theta)$ and $\mathtt{FV}(0,\theta)$. The joint stationarity follows from (5.3) and the arguments in Section 6.1. □

Shiga's construction of [44], (3.12) and Theorem 3.6, requires that type mass evolutions—spindles, in our terminology—be continuous-state branching processes (CSBPs), that is, without emigration, while our $\mathtt{BESQ}(-2\alpha)$ spindle masses can be viewed as CSBPs with emigration [37]. Consequently, our $\mathtt{SSSP}(\alpha,\theta)$ are outside the class of processes considered by Shiga.

7.2. *Exceptional times with finitely many atoms.* Recall notation $N(\pi) = \#\{x \in [0,1]: \pi(\{x\}) > 0\}$ for the number of atoms of $\pi \in \mathcal{M}^a$ and let $1 \le n < \infty$. Theorem 1.10 claims that there are, with positive probability, exceptional times at which a $\mathtt{FV}(\alpha,\theta)$ has only n atoms if and only if $\theta + n\alpha < 1$. This may be surprising since there are infinitely many atoms with probability 1 under the stationary distribution $\mathtt{PDRM}(\alpha,\theta)$. Indeed, the relevant part of the transition kernel of the underlying $\mathtt{SSSP}(\alpha,\theta)$ adds in scaled $\mathtt{PDRM}(\alpha,\alpha)$ components, which each have infinitely many atoms with probability 1.

PROOF OF THEOREM 1.10. Since the desired property is not affected by either the continuous de-Poissonization time-change $u \mapsto \rho_{\boldsymbol{\pi}}(u)$ or by the normalization to unit mass, it suffices to establish that $\mathtt{SSSP}(\alpha,\theta)$ visits $A_n := \{\pi \in \mathcal{M}^a : N(\pi) = n\}$ before visiting zero with positive probability if and only if $\theta + n\alpha < 1$. Therefore, let $\pi \in \mathcal{M}^a \setminus \{0\}$, $\mathbf{F}_\pi \sim \mathbf{Q}_\pi^{\alpha,0}$ and $\mathbf{F}_\theta \sim \mathtt{PRM}(\frac{\theta}{\alpha}\mathtt{Leb} \otimes \overline{\nu}^{(\alpha)}_{\perp\mathrm{cld}})$ so that

$$\boldsymbol{\pi} = \big(\pi^y, y \ge 0\big) = \mathtt{SS}\overline{\mathtt{KE}}\mathtt{WER}(\mathbf{F}_\theta) + \mathtt{SS}\overline{\mathtt{KE}}\mathtt{WER}(\mathbf{F}_\pi) \sim \mathtt{SSSP}_\pi(\alpha,\theta).$$

Recall from Theorem 1.5 that $(\|\pi^y\|, y \ge 0) \sim \mathtt{BESQ}_{\|\pi\|}(2\theta)$. In particular, total mass visits 0 if and only if $\theta < 1$, by Lemma 6.5. To study times y when $\pi^y \in A_n$, we first suppose $\pi^0 = \pi \in A_n$, so that $\mathbf{F}_\pi = \sum_{1\le i\le n} \delta(\mathbf{V}_i, \mathbf{X}_i)$. Then the contribution of $\mathbf{F}_\theta$ to the total mass is a $\mathtt{BESQ}_0(2\theta)$, while each $(\mathbf{V}_i, \mathbf{X}_i)$ has a left-most spindle $\mathbf{f}_i$, and by Corollary 5.11, the remaining spindles contribute a $\mathtt{BESQ}_0(2\alpha)$ to total mass during $(0, \zeta(\mathbf{f}_i))$. In order for $\pi^y \in A_n$ at $y \in \bigcap_{1\le i\le n}(0,\zeta(\mathbf{f}_i))$, we need the sum of a $\mathtt{BESQ}_0(2\theta)$ and n independent $\mathtt{BESQ}_0(2\alpha)$, that is, a $\mathtt{BESQ}_0(2(\theta+n\alpha))$, to vanish, which happens with positive probability if and only if $\theta + n\alpha < 1$.

Now suppose we have more or fewer than n atoms initially. If more, there is positive probability that any choice of n atoms has lifetimes $\zeta(\mathbf{f}_i)$, $1 \le i \le n$, greater than all levels ζ_j^+ when the mass of $(\mathbf{V}_j, \mathbf{X}_j)$, $j > n$, vanishes. On this event, the previous argument applies. If there are fewer than n atoms initially, then $\mathbf{P}\{\pi^\epsilon \in A_\infty\} > 0$ for all $\epsilon > 0$, and by the Markov property, the previous argument applies. This completes the proof when $\theta + n\alpha < 1$.

For $\theta + n\alpha \ge 1$, it remains to show that A_n cannot be visited with positive probability even on events not yet considered. To this end, consider the set $A_n^{(1/m)} \subseteq A_n$ that requires all atoms to be of size strictly greater than $1/m$. If we do not have n atoms of this size, $A_n^{(1/m)}$ cannot be visited before the stopping time when our measure first includes n such large atoms, and we can apply the strong Markov property. Then the previous argument shows that the subset $A_n^{(1/m)}$ of A_n is not visited almost surely while none of these atoms have reached the end of their lifetime. At the stopping time when the first of them vanishes, we can keep repeating the argument. The differences between the relevant stopping times are stochastically bounded below by the law of the time of the first absorption among n independent $\mathtt{BESQ}_{1/m}(-2\alpha)$ diffusions. Thus, these stopping times tend to infinity a.s., so $A_n^{(1/m)}$ and hence $A_n = \bigcup_{m\ge1} A_n^{(1/m)}$ are not visited a.s., if $\theta + n\alpha \ge 1$. □

If there are times with precisely n atoms when $\theta + n\alpha < 1$, then by the argument above, locally, the set of such times has the structure of the zero set of $\mathtt{BESQ}(2(\theta+n\alpha))$, which is well known to be the range of a stable subordinator of index $1-(\theta+n\alpha)$ and therefore has Hausdorff dimension $1-(\theta+n\alpha)$. See, for example, [3].

7.3. *α-Diversity.* Theorem 1.11 claims that $(\overline{\pi}^u, u \ge 0) \sim \mathtt{FV}_\pi(\alpha,\theta)$ has α-diversities $\mathscr{D}(\overline{\pi}^u) := \Gamma(1-\alpha)\lim_{h\downarrow0} h^\alpha \#\{x \in [0,1]: \overline{\pi}^u(\{x\}) > h\}$ that evolve continuously. We develop this via a number of intermediate results about $\mathtt{SSSP}_\pi(\alpha,0)$ and $\mathtt{SSSP}_0(\alpha,\theta)$, beginning by strengthening the continuity of the superskewer process $\widetilde{\mathbf{V}}$, a $\mathtt{PRM}(\mathtt{Leb} \otimes \nu_{\mathtt{BESQ}}^{(-2\alpha)} \otimes \mathtt{Unif})$ stopped at a random time $T \in (0,\infty)$, as studied in Proposition 3.1.

PROPOSITION 7.4. *Let $\widetilde{\mathbf{V}}$ be as in Proposition 3.1 and $\widetilde{\mathbf{X}} = \xi_{\widetilde{\mathbf{V}}}$. For every $y \geq 0$, define $\pi^y = \mathrm{sSKEWER}(y, \widetilde{\mathbf{V}}, \widetilde{\mathbf{X}})$. Then a.s. the real-valued process $(\mathscr{D}(\pi^y), y \geq 0)$ is well defined and is Hölder-γ for every $\gamma \in (0, \alpha/2)$.*

PROOF. [21], Proposition 3.8, shows that the skewer of $\widetilde{\mathbf{N}} = \varphi(\widetilde{\mathbf{V}})$ a.s. has total α-diversity at every level and is Hölder-γ for every $\gamma \in (0, \alpha/2)$. But since interval lengths $f_t(y - \widetilde{\mathbf{X}}(t-))$ of $\mathrm{SKEWER}(y, \widetilde{\mathbf{N}}, \widetilde{\mathbf{X}})$ are also the atom sizes of $\mathrm{sSKEWER}(y, \widetilde{\mathbf{V}}, \widetilde{\mathbf{X}})$, their (total) diversities coincide. □

PROPOSITION 7.5. *Let $\pi \in \mathcal{M}^a$ and $\mathbf{F} \sim \mathbf{Q}_\pi^{\alpha,0}$. Then a.s. the real-valued process $(\mathscr{D}(\mathrm{sSKEWER}(y, \mathbf{F})), y > 0)$ is well defined and is Hölder-γ for every $\gamma \in (0, \alpha/2)$. If, in addition, the initial measure π has α-diversity $\mathscr{D}(\pi)$, then $\lim_{y \downarrow 0} \mathscr{D}(\mathrm{sSKEWER}(y, \mathbf{F})) = \mathscr{D}(\pi)$ a.s.*

PROOF. Like Proposition 7.4, this is really an assertion about the process of ranked atom sizes. The continuity of (total) α-diversity of interval partition valued processes with corresponding interval lengths was noted in [21], Corollary 1.5. More specifically, results including the Hölder continuity can be found in [21], Corollary 6.19 and Proposition 6.11, respectively. □

PROPOSITION 7.6. *The α-diversity process of an $\mathtt{SSSP}_0(\alpha, \theta)$ is well defined and path-continuous.*

PROOF. We use notation $(\pi^y, y \geq 0) \sim \mathtt{SSSP}_0(\alpha, \theta)$ of Proposition 5.4. More precisely, for every $0 \leq a < b \leq \infty$, we define for $\mathbf{F}_\theta \sim \mathtt{PRM}(\frac{\theta}{\alpha}\mathtt{Leb} \otimes \overline{\nu}^{(\alpha)}_{\perp\mathrm{cld}})$

$$\pi^y_{[a,b)} := \sum_{\text{points } (s,V_s,X_s) \text{ of } \mathbf{F}_\theta, s \in [a,b)} \mathrm{sSKEWER}\big(y - |s|, V_s, X_s\big), \quad y \geq 0.$$

That is, only those clades entering between levels a and b count for the process $\boldsymbol{\pi}_{[a,b)}$. In particular, $\pi^y_{[a,b)} = 0$ for all $y \leq a$. For every $z \geq 0$, let

$$\overline{D}^y_z := \Gamma(1-\alpha) \limsup_{h \downarrow 0} h^\alpha \#\big\{x \in [0,1] \colon \pi^y_{[z,\infty)}\big(\{x\}\big) > h\big\}, \quad y \geq 0,$$

$$\underline{D}^y_z := \Gamma(1-\alpha) \liminf_{h \downarrow 0} h^\alpha \#\big\{x \in [0,1] \colon \pi^y_{[z,\infty)}\big(\{x\}\big) > h\big\}, \quad y \geq 0.$$

Then the α-diversity $\mathscr{D}(\pi^y)$ exists at level y if and only if $\overline{D}^y_0 = \underline{D}^y_0$. Moreover, we then have $\mathscr{D}(\pi^y) = \overline{D}^y_0 = \underline{D}^y_0$.

We now fix $y_0 > 0$ and control uniformly in level $[0, y_0]$ the contributions of newly entered clades. Specifically, the following holds almost surely.

CLAIM 1. *Almost surely, for all $\epsilon > 0$, there exists $\delta' > 0$ (that depends on ϵ and the realization) such that*

$$\sup_{y \in (z, z+\delta']} \overline{D}^y_z < \epsilon/3, \quad \textit{for every } z \in [0, y_0]. \tag{7.3}$$

To prove this, it suffices to prove that for any fixed $\epsilon > 0$ there exists $\delta' > 0$ a.s. such that (7.3) holds. Then we deduce the desired property by taking the intersection of the almost sure events for each $\epsilon \in \{1/n, n \in \mathbb{N}\}$.

We first consider the case $\theta = \alpha$. In this case, we are in the setting of Proposition 7.4, so the existence and continuity of α-diversity process $(\mathscr{D}(\pi^y), y \geq 0)$ has been justified. Then

for each $z \geq 0$, the same conclusion holds for the process $\boldsymbol{\pi}_z := (\pi^{z+y}_{[z,\infty)}, y \geq 0)$. Fix $\epsilon > 0$. Let $r_0 = 0$. Inductively, by continuity of $y \mapsto \mathscr{D}(\pi^y_{r_n})$, we have for all $n \geq 0$

$$r_{n+1} := r_n + \inf\Big\{x \geq 0 \colon \sup_{y \in (0,2x)} \mathscr{D}\big(\pi^{r_n+y}_{[r_n, r_n+x)}\big) \geq \epsilon/9\Big\}$$
$$\geq r_n + \inf\Big\{x \geq 0 \colon \sup_{y \in (0,2x)} \mathscr{D}\big(\pi^{y}_{r_n}\big) \geq \epsilon/9\Big\}.$$

Since also $\boldsymbol{\pi}_{r_n}$ is independent of $(r_0, \ldots, r_n)$, the process $(r_n, n \geq 0)$ is a random walk with i.i.d. increments $\delta_n = r_{n+1} - r_n > 0$ a.s. In particular, $k = \inf\{n \geq 1 \colon r_n > y_0\} < \infty$ and $\delta = \min\{\delta_0, \ldots, \delta_k\} > 0$ a.s. Now for all $z \in [0, y_0]$ and n such that $r_n \leq z < r_{n+1}$, we have $z + \delta < r_{n+1} + \delta \leq r_{n+2}$ and

$$\sup_{y \in (z, z+\delta]} \overline{D}^y_z \leq \sup_{y \in (r_n, r_{n+2})} \overline{D}^y_{r_n} = \sup_{y \in (r_n, r_{n+2})} \big(\mathscr{D}(\pi^y_{[r_n, r_{n+1})}) + \mathscr{D}(\pi^y_{[r_{n+1}, r_{n+2})})\big),$$

is at most $2\epsilon/9 < \epsilon/3$. This ends the proof of Claim 1 if $\theta = \alpha$. The general case $\theta \neq \alpha$ follows by thinning and superposition as in the proof of Proposition 5.4, since $\overline{D}^y_z$ decreases when thinning and is (sub)additive in superpositions.

Next, we control each clade and observe:

CLAIM 2. *Almost surely, for all points (s, V_s, X_s) of $\mathbf{F}_\theta$ with $s \in [0, y_0)$, the process $y \mapsto \mathscr{D}(\mathrm{SSKEWER}(y, V_s, X_s))$ is continuous.*

This follows straight from (4.7), Proposition 7.4, and standard properties of Poisson random measures, which also yield the following.

CLAIM 3. *Almost surely, for all $\delta' > 0$, there are at most finitely many points (s, V_s, X_s) of $\mathbf{F}_\theta$ with $s \in [0, y_0)$ and $\zeta^+(V_s) > \delta'$.*

For the remainder of this proof, we will argue on the intersection of the three almost sure events, on which Claims 1–3 hold. Fix any $\epsilon > 0$ and take $\delta' > 0$ so that (7.3) holds. By Claims 2 and 3, the evolution of "long-living" clades

$$\lambda^y = \sum_{\text{points } (s,V_s,X_s) \text{ of } \mathbf{F}_\theta, s \in [0,y_0) \zeta^+(V_s) > \delta'} \mathrm{SSKEWER}(y - s, V_s, X_s), \quad y \in [0, y_0],$$

has continuously evolving diversity. In particular, there is $\delta \in (0, \delta']$ such that for all $x, z \in [0, y_0]$ with $|x - z| < \delta$, we have $|\mathscr{D}(\lambda^x) - \mathscr{D}(\lambda^z)| < \varepsilon/3$, and hence, by Claim 1,

$$|\overline{D}^x_0 - \underline{D}^z_0| \leq \overline{D}^x_{\max\{x-\delta',0\}} + \overline{D}^z_{\max\{z-\delta',0\}} + |\mathscr{D}(\lambda^x) - \mathscr{D}(\lambda^z)| < \varepsilon. \tag{7.4}$$

Since ε was arbitrary, the choice $x = z$ yields $\overline{D}^z_0 = \underline{D}^z_0$ and hence the existence of $\mathscr{D}(\pi^z) = \overline{D}^z_0 = \underline{D}^z_0$ for all $z \in [0, y_0]$. In particular, the LHS in (7.4) is $|\mathscr{D}(\pi^x) - \mathscr{D}(\pi^z)|$ and we conclude the continuity of $z \mapsto \mathscr{D}(\pi^z)$. □

PROOF OF THEOREM 1.11. By the construction of $\mathrm{FV}_\pi(\alpha, \theta)$ processes from $\mathrm{SSSP}_\pi(\alpha, \theta)$ processes via a continuous time-change and normalization by the continuous total mass process, $\mathrm{FV}_\pi(\alpha, \theta)$ inherits the continuity of the diversity process from $\mathrm{SSSP}_\pi(\alpha, \theta)$. Since, by Proposition 1.4, an $\mathrm{SSSP}_\pi(\alpha, \theta)$ can be constructed by adding independent $\mathrm{SSSP}_\pi(\alpha, 0)$ and $\mathrm{SSSP}_0(\alpha, \theta)$, the continuity of its α-diversity process follows from Propositions 7.5 and 7.6, with the subtleties about the α-diversity at time 0 obtained in Proposition 7.5 □

APPENDIX: PROOFS OF PROPERTIES OF STABLE PROCESS

In this appendix, we complete the proofs of Proposition 4.5 and Lemma 4.9.

PROOF OF PROPOSITION 4.5. (ii) The measure $\overline{\nu}^{(\alpha)}_{\perp\mathrm{cld}}\{\mathrm{len}\in\cdot\}$ is just the Lévy measure of the subordinator in Proposition 4.1, therefore, we find its density by the following identity:

$$\int_0^\infty (1-e^{-qx})\overline{\nu}^{(\alpha)}_{\perp\mathrm{cld}}\{\mathrm{len}\in dx\} = \big(2^\alpha\Gamma(1+\alpha)q\big)^{1/(1+\alpha)}$$
$$= \int_0^\infty (1-e^{-qx})\frac{(2^\alpha\Gamma(1+\alpha))^{1/(1+\alpha)}}{\Gamma(\frac{\alpha}{1+\alpha})(1+\alpha)}x^{-1-1/(1+\alpha)}\,dx.$$

(i) It is known that [1], Page 222,

$$\overline{\nu}^{(\alpha)}_{\perp\mathrm{cld}}\{\zeta^+>z\} = \nu^{1+\alpha}_{\perp\mathrm{stb}}\Big\{g\in\mathcal{D}\colon \sup_{t\in\mathbb{R}} g(t)>z\Big\} = W'(z)/W(z),$$

where $W(z)$ is the scale function of the stable process with Laplace exponent ψ_α of (2.4), satisfying $\int_0^\infty e^{-cz}W(z)\,dz = (\psi_\alpha(c))^{-1} = 2^\alpha\Gamma(1+\alpha)c^{-(1+\alpha)}$. Then

$$W(z) = 2^\alpha z^\alpha, \tag{A.1}$$

and we deduce the desired formula.

(iii) Similar arguments appear around [21], Propositions A.2, A.3. It suffices to show that our $\mathtt{Stable}(1+\alpha)$ scaffolding, $\mathbf{X}$, does not jump away from a historic minimum, that is, there is a.s. no $t>0$ for which $\mathbf{X}(t-)=\inf\{\mathbf{X}(s)\colon s<t\}$ and $\mathbf{X}(t-)<\mathbf{X}(t)$. And by the Lévy process properties of $\mathbf{X}$, it suffices to prove this for $t\in(0,1)$.

By the strong Markov property, there is a.s. no $t\in(0,1)$ at which $\mathbf{X}(t-)<\mathbf{X}(t)$ and $\mathbf{X}(t)=\sup\{\mathbf{X}(s)\colon s\in(t,1]\}$, since each of the countably many jumps of $\mathbf{X}$ can be captured by a stopping time. Now, the desired property follows by invariance under increment reversal. □

PROOF OF LEMMA 4.9. Let X be $\mathtt{Stable}(1+\alpha)$ and $\mathbb{P}_a$ the law on $\mathcal{D}_{\mathrm{exc}}$ of X starting from a and absorbed when first reaching 0. Recall notation $\mathrm{len}(X)=\sup\{t\ge 0\colon X(t)\neq 0\}$ for the excursion length. Let $t>0$. By [5], Corollary 3, and the Markov property,

$$\mathbb{E}_a[H|\mathrm{len}>t]\to\nu^{1+\alpha}_{\perp\mathrm{stb}}[H|\mathrm{len}>t]\quad\text{as } a\downarrow 0,$$

for all bounded continuous $H\colon\mathcal{D}\to\mathbb{R}$. Also, by [6], Lemma 1, and [7], Example in Section 4,

$$\mathbb{P}_a\{\mathrm{len}>t\}/\mathbb{P}_a\{\zeta^+>y\}\to\nu^{1+\alpha}_{\perp\mathrm{stb}}\{\mathrm{len}>t\}/\nu^{1+\alpha}_{\perp\mathrm{stb}}\{\zeta^+>y\}\quad\text{as } a\downarrow 0.$$

Now consider any open $A\subseteq\mathcal{D}_{\mathrm{exc}}$. Then for every $t>0$, since ζ^+ is continuous, we have by the Portmanteau theorem

$$\begin{aligned}\liminf_{a\downarrow0}\mathbb{P}_a(A|\zeta^+>y) &\ge \liminf_{a\downarrow0}\mathbb{P}_a(A\cap\{\mathrm{len}>t\}|\zeta^+>y)\\ &= \liminf_{a\downarrow0}\frac{\mathbb{P}_a\{\mathrm{len}>t\}}{\mathbb{P}_a\{\zeta^+>y\}}\mathbb{P}_a(A\cap\{\zeta^+>y\}|\mathrm{len}>t)\\ &\ge \frac{\nu^{1+\alpha}_{\perp\mathrm{stb}}\{\mathrm{len}>t\}}{\nu^{1+\alpha}_{\perp\mathrm{stb}}\{\zeta^+>y\}}\nu^{1+\alpha}_{\perp\mathrm{stb}}(A\cap\{\zeta^+>y\}|\mathrm{len}>t)\\ &= \nu^{1+\alpha}_{\perp\mathrm{stb}}(A\cap\{\mathrm{len}>t\}|\zeta^+>y).\end{aligned}$$

Letting $t\downarrow0$, we find that $\liminf_{a\downarrow0}\mathbb{P}_a(A|\zeta^+>y)\ge\nu^{1+\alpha}_{\perp\mathrm{stb}}(A|\zeta^+>y)$, and the conclusion follows from the Portmanteau theorem. □

Funding. This research is partially supported by NSF Grants DMS-1204840, DMS-1308340, DMS-1612483, DMS-1855568, UW-RRF grant A112251, EPSRC Grant EP/K029797/1.

REFERENCES

[1] Avram, F., Kyprianou, A. E. and Pistorius, M. R. (2004). Exit problems for spectrally negative Lévy processes and applications to (Canadized) Russian options. *Ann. Appl. Probab.* **14** 215–238. MR2023021 https://doi.org/10.1214/aoap/1075828052

[2] Bertoin, J. (1996). *Lévy Processes*. *Cambridge Tracts in Mathematics* **121**. Cambridge Univ. Press, Cambridge. MR1406564

[3] Bertoin, J. (1999). Subordinators: Examples and applications. In *Lectures on Probability Theory and Statistics* (*Saint-Flour*, 1997). *Lecture Notes in Math.* **1717** 1–91. Springer, Berlin. MR1746300 https://doi.org/10.1007/978-3-540-48115-7_1

[4] Carlton, M. A. (2002). A family of densities derived from the three-parameter Dirichlet process. *J. Appl. Probab.* **39** 764–774. MR1938169 https://doi.org/10.1017/s0021900200022026

[5] Chaumont, L. (1994). Sur certains processus de Lévy conditionnés à rester positifs. *Stoch. Stoch. Rep.* **47** 1–20. MR1787140 https://doi.org/10.1080/17442509408833880

[6] Chaumont, L. (1997). Excursion normalisée, méandre et pont pour les processus de Lévy stables. *Bull. Sci. Math.* **121** 377–403. MR1465814

[7] Chaumont, L. and Doney, R. A. (2005). On Lévy processes conditioned to stay positive. *Electron. J. Probab.* **10** 948–961. MR2164035 https://doi.org/10.1214/EJP.v10-261

[8] Costantini, C., De Blasi, P., Ethier, S. N., Ruggiero, M. and Spanò, D. (2017). Wright-Fisher construction of the two-parameter Poisson-Dirichlet diffusion. *Ann. Appl. Probab.* **27** 1923–1950. MR3678488 https://doi.org/10.1214/16-AAP1252

[9] Daley, D. J. and Vere-Jones, D. (2003). *An Introduction to the Theory of Point Processes. Vol. I: Elementary Theory and Methods*, 2nd ed. *Probability and Its Applications* (*New York*). Springer, New York. MR1950431

[10] Daley, D. J. and Vere-Jones, D. (2008). *An Introduction to the Theory of Point Processes. Vol. II: General Theory and Structure*, 2nd ed. *Probability and Its Applications* (*New York*). Springer, New York. MR2371524 https://doi.org/10.1007/978-0-387-49835-5

[11] Dubins, L. and Freedman, D. (1964). Measurable sets of measures. *Pacific J. Math.* **14** 1211–1222. MR0174687

[12] Etheridge, A. and March, P. (1991). A note on superprocesses. *Probab. Theory Related Fields* **89** 141–147. MR1110534 https://doi.org/10.1007/BF01366902

[13] Ethier, S. N. and Kurtz, T. G. (1981). The infinitely-many-neutral-alleles diffusion model. *Adv. in Appl. Probab.* **13** 429–452. MR0615945 https://doi.org/10.2307/1426779

[14] Ethier, S. N. and Kurtz, T. G. (1987). The infinitely-many-alleles model with selection as a measure-valued diffusion. In *Stochastic Methods in Biology* (*Nagoya*, 1985). *Lecture Notes in Biomathematics* **70** 72–86. Springer, Berlin. MR0893637 https://doi.org/10.1007/978-3-642-46599-4_6

[15] Ethier, S. N. and Kurtz, T. G. (1993). Fleming-Viot processes in population genetics. *SIAM J. Control Optim.* **31** 345–386. MR1205982 https://doi.org/10.1137/0331019

[16] Ethier, S. N. and Kurtz, T. G. (2005). *Markov Processes*: *Characterization and Convergence*. *Wiley Series in Probability and Mathematical Statistics*: *Probability and Mathematical Statistics*. Wiley, New York.

[17] Feng, S. and Sun, W. (2010). Some diffusion processes associated with two parameter Poisson-Dirichlet distribution and Dirichlet process. *Probab. Theory Related Fields* **148** 501–525. MR2678897 https://doi.org/10.1007/s00440-009-0238-2

[18] Feng, S. and Sun, W. (2019). A dynamic model for the two-parameter Dirichlet process. *Potential Anal.* **51** 147–164. MR3983502 https://doi.org/10.1007/s11118-018-9706-6

[19] Fleming, W. H. and Viot, M. (1979). Some measure-valued Markov processes in population genetics theory. *Indiana Univ. Math. J.* **28** 817–843. MR0542340 https://doi.org/10.1512/iumj.1979.28.28058

[20] Forman, N., Pal, S., Rizzolo, D. and Winkel, M. (2018). Uniform control of local times of spectrally positive stable processes. *Ann. Appl. Probab.* **28** 2592–2634. MR3843837 https://doi.org/10.1214/17-AAP1370

[21] Forman, N., Pal, S., Rizzolo, D. and Winkel, M. (2020). Diffusions on a space of interval partitions: Construction from marked Lévy processes. *Electron. J. Probab.* **25** Paper No. 133, 46. MR4169174 https://doi.org/10.1214/20-ejp521

[22] FORMAN, N., PAL, S., RIZZOLO, D. and WINKEL, M. (2021). Diffusions on a space of interval partitions: Poisson–Dirichlet stationary distributions. *Ann. Probab.* **49** 793–831. MR4255131 https://doi.org/10.1214/20-aop1460

[23] FORMAN, N., PAL, S., RIZZOLO, D. and WINKEL, M. (2021). Ranked masses in two-parameter Fleming-Viot diffusions. arXiv:2101.09307 [math.PR].

[24] GHOSH, J. K. and RAMAMOORTHI, R. V. (2003). *Bayesian Nonparametrics*. *Springer Series in Statistics*. Springer, New York. MR1992245

[25] GÖING-JAESCHKE, A. and YOR, M. (2003). A survey and some generalizations of Bessel processes. *Bernoulli* **9** 313–349. MR1997032 https://doi.org/10.3150/bj/1068128980

[26] HELLAND, I. S. (1978). Continuity of a class of random time transformations. *Stochastic Process. Appl.* **7** 79–99. MR0488203 https://doi.org/10.1016/0304-4149(78)90040-6

[27] ISHWARAN, H. and JAMES, L. F. (2003). Generalized weighted Chinese restaurant processes for species sampling mixture models. *Statist. Sinica* **13** 1211–1235. MR2026070

[28] JAMES, L. F., LIJOI, A. and PRÜNSTER, I. (2008). Distributions of linear functionals of two parameter Poisson-Dirichlet random measures. *Ann. Appl. Probab.* **18** 521–551. MR2398765 https://doi.org/10.1214/07-AAP462

[29] KALLENBERG, O. (2002). *Foundations of Modern Probability*, 2nd ed. *Probability and Its Applications* (*New York*). Springer, New York. MR1876169 https://doi.org/10.1007/978-1-4757-4015-8

[30] KALLENBERG, O. (2017). *Random Measures, Theory and Applications*. *Probability Theory and Stochastic Modelling* **77**. Springer, Cham. MR3642325 https://doi.org/10.1007/978-3-319-41598-7

[31] KONNO, N. and SHIGA, T. (1988). Stochastic partial differential equations for some measure-valued diffusions. *Probab. Theory Related Fields* **79** 201–225. MR0958288 https://doi.org/10.1007/BF00320919

[32] LI, Z. (2011). *Measure-Valued Branching Markov Processes*. *Probability and Its Applications* (*New York*). Springer, Heidelberg. MR2760602 https://doi.org/10.1007/978-3-642-15004-3

[33] PAL, S. (2013). Wright-Fisher diffusion with negative mutation rates. *Ann. Probab.* **41** 503–526. MR3077518 https://doi.org/10.1214/11-AOP704

[34] PETROV, L. A. (2009). A two-parameter family of infinite-dimensional diffusions on the Kingman simplex. *Funktsional. Anal. i Prilozhen.* **43** 45–66. MR2596654 https://doi.org/10.1007/s10688-009-0036-8

[35] PITMAN, J. (1996). Some developments of the Blackwell-MacQueen urn scheme. In *Statistics, Probability and Game Theory*. *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **30** 245–267. IMS, Hayward, CA. MR1481784 https://doi.org/10.1214/lnms/1215453576

[36] PITMAN, J. (2006). *Combinatorial Stochastic Processes*. *Lecture Notes in Math.* **1875**. Springer, Berlin. Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002, With a foreword by Jean Picard. MR2245368

[37] PITMAN, J. and WINKEL, M. (2018). Squared Bessel processes of positive and negative dimension embedded in Brownian local times. *Electron. Commun. Probab.* **23** Paper No. 74, 13. MR3866047 https://doi.org/10.1214/18-ECP174

[38] PITMAN, J. and YOR, M. (1982). A decomposition of Bessel bridges. *Z. Wahrsch. Verw. Gebiete* **59** 425–457. MR0656509 https://doi.org/10.1007/BF00532802

[39] PITMAN, J. and YOR, M. (1997). The two-parameter Poisson-Dirichlet distribution derived from a stable subordinator. *Ann. Probab.* **25** 855–900. MR1434129 https://doi.org/10.1214/aop/1024404422

[40] REVUZ, D. and YOR, M. (1999). *Continuous Martingales and Brownian Motion*, 3rd ed. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **293**. Springer, Berlin. MR1725357 https://doi.org/10.1007/978-3-662-06400-9

[41] ROGERS, D. and WINKEL, M. (2020). A Ray–Knight representation of up-down Chinese restaurants. Bernoulli. To appear. arXiv:2006.06334 [math.PR].

[42] RUGGIERO, M., WALKER, S. G. and FAVARO, S. (2013). Alpha-diversity processes and normalized inverse-Gaussian diffusions. *Ann. Appl. Probab.* **23** 386–425. MR3059239 https://doi.org/10.1214/12-AAP846

[43] SHI, Q. and WINKEL, M. (2020). Up-down ordered Chinese restaurant processes with two-sided immigration, emigration and diffusion limits. arXiv:2012.15758 [math.PR].

[44] SHIGA, T. (1990). A stochastic equation based on a Poisson system for a class of measure-valued diffusion processes. *J. Math. Kyoto Univ.* **30** 245–279. MR1068791 https://doi.org/10.1215/kjm/1250520071

[45] TAVARÉ, S. (1987). The birth process with immigration, and the genealogical structure of large populations. *J. Math. Biol.* **25** 161–168. MR0896431 https://doi.org/10.1007/BF00276387

[46] TEH, Y. W. (2006). A hierarchical Bayesian language model based on Pitman–Yor processes. In *Proceedings of the* 21*st International Conference on Computational Linguistics and the* 44*th Annual Meeting of the Association for Computational Linguistics* 985–992. Association for Computational Linguistics, Sydney.