The Brownian net and selection in the spatial Λ-Fleming-Viot process

Alison Etheridge [*]   Nic Freeman [†]   Daniel Straulino[‡]

Abstract

We obtain the Brownian net of [24] as the scaling limit of the paths traced out by a system of continuous (one-dimensional) space and time branching and coalescing random walks. This demonstrates a certain universality of the net, which we have not seen explored elsewhere. The walks themselves arise in a natural way as the ancestral lineages relating individuals in a sample from a biological population evolving according to the spatial Lambda-Fleming-Viot process. Our scaling reveals the effect, in dimension one, of spatial structure on the spread of a selectively advantageous gene through such a population.

1 Introduction

The *Brownian net*, introduced in [24], arises as the scaling limit of a system of branching and coalescing random walk paths. It extends, in a natural way, the *Brownian web*, which originated in the work of [1]. In the Brownian web there is no branching. It can be thought of as the diffusive limit of a system of one-dimensional coalescing random walk paths, one started from each point of the space-time (diamond) lattice. Informally, the web is then a system of coalescing Brownian paths, one started from each space-time point. The Brownian web was formulated in [15] as a random variable taking its values in the space of compact sets of paths, equipped with a topology under which it is a Polish space. In this framework, the powerful techniques of weak convergence become available and as a result the Brownian web emerges as the limit of a wide variety of one-dimensional coalescing systems; e.g. [13], [20]. This points to a certain 'universality' of the Brownian web.

In the Brownian net, each path has a small probability (tending to zero in the scaling limit) of branching in each time step. The limiting object is (even) more difficult to

---

visualise than the Brownian web, as there will be a multitude of paths emanating from each space-time point. Nonetheless, [24] shows that it can be characterised through the systems of 'left-most' and 'right-most' paths from each point, each of which itself forms a Brownian web (with drift). Motivated by the study of perturbations of one-dimensional voter models, [19] shows that, by starting from systems of random walks that branch, coalesce and die on the diamond lattice, the Brownian net can be extended still further to include a killing term. However, we have not seen the 'universality' of the Brownian net explored. Our main result, Theorem 3.5, is a contribution in this direction. It establishes an appropriate scaling under which the paths traced out by a system of branching and coalescing *continuous time and space* random walks in one spatial dimension converges to the Brownian net.

The original motivation for our work was a question of interest in population genetics: when will the action of natural selection on a gene in a spatially structured population cause a detectable trace in the patterns of genetic variation observed in the contemporary population? We deal with the most biologically interesting case of a population evolving in a two-dimensional spatial continuum in [6]. Our work in this paper uncovers some of the rich mathematical structure underlying mathematical models for biological populations evolving in one-dimensional spatial continua. In particular, we study the systems of interacting random walks that, as dual processes (corresponding to ancestral lineages of the model), describe the relationships, across both time and space, between individuals sampled from those populations.

It is natural to ask whether the model of [19] has a biological interpretation. It does: killing corresponds to a mutation term. This was observed by [24] (c.f. [7]). However, in view of the technical challenges to be overcome to handle the additional killing term, even on a diamond lattice, we do not explore this further here.

Our starting point will be the Spatial Λ-Fleming-Viot process with selection (SΛFVS) which (along with its dual) was introduced and constructed in [10]. The dynamics of both the SΛFVS and its dual are driven by a Poisson Point Process of *events* (which model reproduction in the population) and will be described in detail in Section 2. Roughly, each event prescribes a region in which reproduction takes place. A proportion v of the population in the affected region is replaced by offspring of a single parent. We shall refer to v as the impact of the event. In the absence of selection, the dual process of ancestral lineages is a modification of the 'Poisson trees' of [13]. With selection, our dual follows 'potential' ancestral lineages, which introduces a branching mechanism, with the rate of branching determined by the presence of lineages in a region, but not increasing with their density. Our main result, Theorem 3.5, is that in one spatial dimension and when the impact $v = 1$ (which prevents ancestral lineages from jumping over one another), when suitably scaled the system of branching and coalescing ancestral lineages converges to the Brownian net.

Without selection, the corresponding objects converge (after scaling) to the Brownian web. In that setting, we believe (and [4] provides strong supporting evidence) that the random walks can even be allowed to jump over one another and the only effect on the limiting object is a simple scaling of time (given by one minus the crossing probability of 'nearby' paths). This would mirror the results of [20], in which systems of coalescing non-simple random walks with crossing paths are shown to converge to the Brownian web. When we try to include selection in this limit, allowing paths to cross has a more complicated effect, as we illustrate through simulations in Section 7. It is an intriguing open question to explain the pictures that we present there.

In [10], scaling limits of the SΛFVS were considered in which the local population density tends to infinity. In that case, the classical Fisher-KPP equation and its stochastic analogue can be recovered. The dual process of branching and coalescing lineages con-

verges to branching Brownian motion, with coalescence of lineages at a rate determined by the local time that they spend together. In this article we are interested in a very different regime, in which coalescence of lineages is instantaneous on meeting.

Although our result owes a lot to the existing literature, the continuum setting introduces some new features. In particular, some care is needed in extending the self-duality of the systems of branching and coalescing simple random walks that appear in [24] to an 'approximate' self-duality of the càdlàg walks in continuous time and space that are considered here.

In Section 2 we introduce the SΛFVS and its dual before providing a heuristic explanation for our scaling. In Section 3 we provide a self-contained account of the necessary background on the Brownian web and net. Our main result is then stated formally in Theorem 3.5, which is proved in Sections 4-6. Finally, Section 7 presents a brief numerical exploration of the effect of allowing the random walk paths of the dual process to cross on the positions at time one of the left-most and right-most paths emanating from a point.

Acknowledgement

2   The SΛFVS and its dual

In this section we introduce the set of branching and coalescing paths with which our main result is concerned. They arise as the dual to a special instance of the SΛFVS. The reader familiar with the SΛFVS can safely refer to Definition 2.2 (and the three lines preceding it) for notation, take note of Remark 2.3, and then skip to Section 3.

2.1   The SΛFVS

The Spatial Λ-Fleming-Viot process (SΛFV) without selection was introduced in [8, 2]. In fact the name does not refer to a single process, but rather to a framework for modelling the dynamics of frequencies of different genetic types found within a population that is evolving in a spatial continuum. It is distinguished from the classical models of population biology in that reproduction is based on 'events' rather than individuals. This introduces density dependence into reproduction in such a way that the clumping and extinction which plagues classical models is overcome, whilst the model remains analytically tractable. For a survey of the SΛFV we refer to [3].

There are very many different ways in which to introduce selection into the SΛFV. Here we adapt the approach typically adopted to introduce selection into the Moran model of classical population genetics. A full motivation of this approach can be found in [10], to which we refer the reader.

We suppose that the population is divided into two genetic types, which we denote a and A, and is evolving in a geographical space which is modelled by $\mathbb{R}$. It will be convenient to index time by the whole of $\mathbb{R}$. At each time t, the population will be represented by a random function $\{w_t(x),\, x \in \mathbb{R}\}$ defined, up to a Lebesgue null set of

$\mathbb{R}$, by

$$w_t(x) := \text{ proportion of type } a \text{ at spatial position } x \text{ at time } t.$$

A construction of an appropriate state space for $x \mapsto w_t(x)$ can be found in [25]. Using the identification

$$\int_{\mathbb{R}} \left\{ w(x)f(x,a) + (1 - w(x))f(x,A) \right\} dx = \int_{\mathbb{R}\times\{a,A\}} f(x,\kappa)\mathcal{M}(dx,d\kappa),$$

this state space is in one-to-one correspondence with the space $\mathcal{M}_\lambda$ of measures on $\mathbb{R} \times \{a,A\}$ with 'spatial marginal' Lebesgue measure, which we endow with the topology of vague convergence. By a slight abuse of notation, we also denote the state space of the process $(w_t)_{t\in\mathbb{R}}$ by $\mathcal{M}_\lambda$.

Definition 2.1 (One-dimensional SΛFV with selection (SΛFVS)). *Fix $\mathcal{R} \in (0,\infty)$ and $v \in (0,1]$ and let μ be a finite measure on $(0,\mathcal{R}]$. Further, let Π be a Poisson point process on $\mathbb{R} \times \mathbb{R} \times (0,\infty)$ with intensity measure*

$$dx \otimes dt \otimes \mu(dr). \tag{2.1}$$

The one-dimensional spatial Λ-Fleming-Viot process with selection *(SΛFVS) driven by (2.1) is the $\mathcal{M}_\lambda$-valued process $(w_t)_{t\in\mathbb{R}}$ with dynamics given as follows.*

If $(x,t,r) \in \Pi$, a reproduction event occurs at time t within the closed interval $[x-r, x+r]$. With probability $1 - \mathbf{s}$ the event (x,t,r) is neutral, *in which case:*

1. *Choose a parental location z uniformly at random within $(x-r, x+r)$, and a parental type, κ, according to $w_{t-}(z)$, that is $\kappa = a$ with probability $w_{t-}(z)$ and $\kappa = A$ with probability $1 - w_{t-}(z)$.*

2. *For every $y \in [x-r, x+r]$, set $w_t(y) = (1-v)w_{t-}(y) + v\mathbf{1}_{\{\kappa=a\}}$.*

With the complementary probability $\mathbf{s}$, (x,t,r) corresponds to a selective *event within $[x-r, x+r]$ at time t, in which case:*

1. *Choose two distinct 'potential' parental locations z, z' independently and uniformly at random within $(x-r, x+r)$, and at each of these locations 'potential' parental types κ, κ', according to $w_{t-}(z), w_{t-}(z')$ respectively.*

2. *For every $y \in [x-r, x+r]$ set $w_t(y) = (1-v)w_{t-}(y) + v\mathbf{1}_{\{\kappa=\kappa'=a\}}$. Declare the parental location to be z if $\kappa = \kappa' = a$ or $\kappa = \kappa' = A$ and to be z (resp. z') if $\kappa = A, \kappa' = a$ (resp. $\kappa = a, \kappa' = A$).*

In fact this is a very special case of the SΛFVS introduced in [10], and even more special than those constructed in [9], but it already provides a rich class of models. We use the assumption that μ has bounded support in Section 4, but it is far from necessary for the construction of the process. The assumption that parental locations are sampled uniformly from $(-r, r)$ has become standard in the literature, but at no point do we use it; our proofs work equally well for any symmetric distribution on $(-r, r)$. The parameter v is often refered to as the 'impact' of the event. It can be loosely thought of as inversely proportional to the local population density. For our rigorous results we shall take $v = 1$, meaning that during a reproduction event, *all* individuals in the affected region are replaced.

2.2 The dual process of branching and coalescing lineages

Our primary concern in this paper is the dual process of the SΛFVS, a system of branching and coalescing paths that encodes all the *potential ancestors* of individuals in a sample from the population.

When there is no selection, the dual contains only coalescing random walks, and each such walk corresponds to the ancestral lineage ℓ of some individual; meaning that ℓ traces out the locations in space-time occupied by the ancestors of that individual.

If selection is present, then, at a selective event, we cannot determine the genetic types of the potential ancestors of the event (and hence the type and location of the actual ancestor) without looking further into the past. To avoid intractable non-Markovian dynamics, in this case we define a dual which traces all the locations in space-time which *could* have contained ancestors of a sample S from the contemporary population. This leads to a system of branching and coalescing random walks, tracing all the *potential* ancestral lineages.

The dynamics of the dual are driven by the same Poisson point process of events, Π, that drove the forwards in time process. The distribution of this Poisson point process is invariant under time reversal and so we shall abuse notation by reversing the direction of time when discussing the dual.

We suppose that at time 0 (which we think of as 'the present') we sample k individuals from locations $x_1, \ldots, x_k$, and we write $\xi_t^1, \ldots, \xi_t^{N_t}$, for the locations of the N_t *potential ancestral lineages* that make up our dual at time t before the present.

Definition 2.2 (Branching and coalescing dual). *Fix $\mathcal{R} \in (0, \infty)$. Let Π be a Poisson point process on $\mathbb{R} \times \mathbb{R} \times (0, \infty)$ with intensity measure*

$$dx \otimes dt \otimes \mu(dr)$$

where μ is a finite measure on $(0, \mathcal{R}]$. The branching and coalescing dual process $(\Xi_t)_{t \geq 0}$ is the $\bigcup_{n \geq 1} \mathbb{R}^n$-valued Markov process with dynamics defined as follows: At each event $(x, t, r) \in \Pi$, with probability $1 - \mathbf{s}$, the event is neutral:

1. *for each i such that $\xi_{t-}^i \in [x - r, x + r]$, mark the ith lineage with probability υ, independently over i and of the past;*

2. *if at least one lineage is marked, all marked lineages disappear and are replaced by a single lineage (the 'parent' of the event), whose location at time t is drawn uniformly at random from within $(x - r, x + r)$.*

With the complementary probability $\mathbf{s}$, the event is selective:

1. *for each i such that $\xi_{t-}^i \in [x - r, x + r]$, mark the ith lineage with probability υ, independently over i and of the past;*

2. *if at least one lineage is marked, all marked lineages disappear and are replaced by two lineages (the 'potential parents' of the event), whose (almost surely distinct) locations are drawn independently and uniformly from within $(x - r, x + r)$.*

In both cases, if no lineage is marked, then nothing happens.

A potential ancestral lineage is any path obtained by following the locations of potential ancestors of an individual in the sample; whenever the potential ancestor corresponding to the current position of the lineage is marked in an event, the path jumps to the location of the potential parent of the event (if the event is neutral) or to the location of (either) one of the potential parents (if the event is selective). Of course there are now many such paths corresponding to each individual in the sample.

Remark 2.3. If we take the impact $\upsilon = 1$, then paths of the dual process cannot cross one another. We will impose this condition for our main result.

This dual process is the SΛFVS analogue of the Ancestral Selection Graph (ASG), introduced in the companion papers [16] and [18], which describes all the potential ancestors of a sample from a population evolving according to the Wright-Fisher diffusion with selection. Duality of this type can be expressed in several different ways, but perhaps the simplest is the statement that the ASG is the moment dual of the diffusion. To establish the analogous duality for the SΛFVS, we would need to be able to identify $\mathbb{E}[\prod_{i=1}^{n} w_t(x_i)]$ for any choice of points $x_1, \ldots, x_n \in \mathbb{R}$. The difficulty is that the SΛFVS $w_t(x)$ is only defined at Lebesgue almost every point x and so we have to be satisfied with a 'weak' moment duality.

Proposition 2.4 (Proposition 2.2 of [10]). *The spatial Λ-Fleming-Viot process with selection is dual to the process $(\Xi_t)_{t \geq 0}$ in the sense that for every $k \in \mathbb{N}$ and $\psi \in C(\mathbb{R}^k) \cap L^1(\mathbb{R}^k)$, we have*

$$\mathbb{E}_{w_0}\left[\int_{\mathbb{R}^k} \psi(x_1, \ldots, x_k) \left\{\prod_{j=1}^{k} w_t(x_j)\right\} dx_1 \ldots dx_k\right]$$

$$= \int_{\mathbb{R}^k} \psi(x_1, \ldots, x_k) \mathbb{E}_{\{x_1, \ldots, x_k\}}\left[\prod_{j=1}^{N_t} w_0(\xi_t^j)\right] dx_1 \ldots dx_k. \tag{2.2}$$

In fact, a stronger form of this duality holds, in which the forwards in time process of allele frequencies and the process of potential ancestors of a sample are realised on the same probability space through a lookdown construction; see [25] for the case without selection and [9] for the general case.

From now on

$$forwards\ in\ time\ refers\ to\ forwards\ for\ the\ dual\ process,$$

i.e. the reversal of that in Definition 2.1.

2.3 The scaling

We shall keep the impact of each reproduction event (i.e. the parameter υ) fixed, but we rescale the strength **s** of selection. In addition we perform a diffusive rescaling of time and space. For our main result we require $\upsilon = 1$, but the heuristic argument presented here and the numerical experiments of Section 7, suggest that there should be a non-trivial limit for any fixed $\upsilon \in (0, 1)$. Let us now describe the appropriate rescaling. The stages of our rescaling are indexed by $n \in \mathbb{N}$.

Recall that μ is a finite measure on $(0, \mathcal{R}]$. For each $n \in \mathbb{N}$, define the measure μ^n by $\mu^n(A) = \mu(n^{1/2}A)$, for all Borel subsets A of $\mathbb{R}$. At the nth stage of the rescaling, our rescaled dual is driven by the Poisson point process Π^n on $\mathbb{R} \times \mathbb{R} \times [0, \infty)$ with intensity

$$n^{1/2} dx \otimes n \, dt \otimes \mu^n(dr). \tag{2.3}$$

The $\sqrt{n}$ in front of dx arises since the rate at which centres of events fall in an interval of length l in the rescaled process is the rate at which they fall in an interval of length $\sqrt{n}l$ in the unscaled process. Each event of Π^n, independently, is neutral with probability $1 - \mathbf{s}_n$ and selective with probability $\mathbf{s}_n$, where $\mathbf{s}_n = \alpha/\sqrt{n}$ for some $\alpha \in (0, \infty)$. Thus, the nth rescaling of our dual process is precisely Definition 2.2 with (2.3) in place of (2.1).

Although not obvious for the SΛFVS itself, when considering the dual process it is not hard to understand why the scaling above should lead to a non-trivial limit. If we ignore the selective events, then each ancestral lineage follows a compound Poisson process and rescales to a (linear time change of) Brownian motion. Now, consider what happens

at a selective event. The two new lineages are born at a separation of order $1/\sqrt{n}$. If we are to 'see' both lineages in the limit then they must move apart to a separation of order 1 (after which they might, possibly, coalesce back together). Ignoring possible interactions with other lineages, the probability that a pair of lineages achieves such a separation is of order $1/\sqrt{n}$. Therefore, in order to obtain a non-trivial limit (which differs from that in the absence of selection) we need $\mathcal{O}(\sqrt{n})$ such branches per scaled unit of time, so we take $n\mathbf{s}_n = \alpha\sqrt{n}$ or $\mathbf{s}_n = \alpha/\sqrt{n}$. (This argument can also be used to identify the correct scaling of $\mathbf{s}_n$ in order to obtain a non-trivial limit in higher dimensions, see [6].)

Evidently we can extend the duality of Proposition 2.4 to lineages that are sampled at different times. For each point $p = (x, t) \in \mathbb{R}^2$, we think of an individual living at (x, t) and, at the nth stage of the rescaling, construct the set $\mathcal{P}_n^\uparrow(p)$ of the potential ancestral lineages of the individual at p. (The reason for the uparrow in the notation will become clear in Section 4.1.) Thus $P_n^\uparrow(p)$ is a set of branching and coalescing paths. Our main result will concern the limit when we consider the union of such sets of paths as p ranges over a countable dense set of space-time points.

3 The Brownian net

In order to state a precise result, we must introduce the Brownian net and, in particular, the state space in which convergence takes place. A short introduction to the Brownian web and net is provided in this section. For a detailed survey of the surrounding literature, see [22].

Once again, the reader familiar with this area can note our modification of the usual state space (detailed in Section 3.1) and Remark 3.1 for terminology, and then skip to the statement of our main result, which can be found in Section 3.3.

3.1 The state space

We now introduce the state space for our processes. Since our branching and coalescing paths are only càdlàg (not continuous), to capture the convergence of $\mathcal{P}_n^\uparrow(p)$ we will need a modification of the state space (introduced by [15]) that is commonly used for the Brownian web and net.

For $s \in [-\infty, \infty]$, we set

$$D[s] = \left\{ f : [s, \infty] \to [-\infty, \infty] \,;\, f \text{ is càdlàg on } [s, \infty] \cap (-\infty, \infty) \right\}.$$

For $f \in D[s]$, it will be convenient to define $\sigma(f) = \sigma_f = s$ to be the first time at which f is defined. We set

$$M = \bigcup_{t \in [-\infty, \infty]} D[t]. \tag{3.1}$$

For each $s \in [-\infty, \infty]$ and $f \in D[s]$ we define a function $\bar{f}$ as follows. Let $\kappa_t = \tanh^{-1}(t)$ and note that κ is an order preserving homeomorphism between $[-1, 1]$ and $[-\infty, \infty]$. (The specific choice of the function $\tanh$ is a convention in the literature. We use the symbol κ in place of $\tanh$ to denote a change of time rather than rescaling of space.) Then if $f \in M$ we define

$$\bar{f}(t) = \frac{\tanh(f(\kappa_t))}{1 + |\kappa_t|} \tag{3.2}$$

for $t \in [\kappa^{-1}(\sigma_f), 1]$. It follows immediately that $\bar{f}$ is càdlàg.

In Section 5.1 we define a generalization ρ of the Skorohod metric that acts on càdlàg paths with possibly different starting times. We show (in Section 5.2) that

$$d_M(f_1, f_2) = \rho(\bar{f}_1, \bar{f}_2) \vee |\tanh(\sigma_{f_1}) - \tanh(\sigma_{f_2})| \tag{3.3}$$
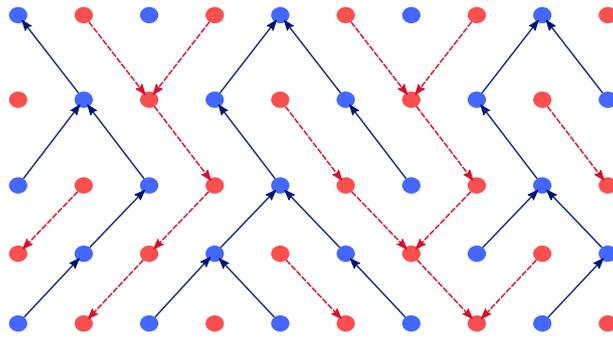
Figure 1: **Self duality of systems of coalescing random walks on the diamond lattice that converge to the Brownian web.** The blue arrows represent the forwards in time coalescing random walks, while the red arrows represent the backwards in time dual.

is a pseudo-metric on M. In standard fashion, from now on we implicitly work with equivalence classes of M and, with mild abuse of notation, treat (M, d_M) as a metric space. In view of (3.2), the intuition for (3.3) is that convergence in (M, d_M) can be described as local Skorohod convergence of the paths plus convergence of the starting times.

If we restrict to continuous paths and replace ρ with the usual L^∞ distance, then we recover the space $(\widetilde{M}, d_{\widetilde{M}})$ introduced by [15], see (5.21). Convergence in the corresponding metric on continuous paths can be described as locally uniform convergence of paths plus convergence of starting times.

We define the set $\mathcal{K}(M)$ of compact subsets of M, equipped with the Hausdorff metric, m, and including the empty set $\emptyset$ as an isolated point. We show (in Section 5.2) that (M, d_M) is complete and separable; the space $\mathcal{K}(M)$ inherits these properties. Similarly, we write $\mathcal{K}(\widetilde{M})$ for the space of all compact subsets of $\widetilde{M}$.

3.2 The Brownian web and net

Arratia [1] was the first to observe that the Brownian web exhibits a self-duality. It is most easily understood by first considering the prelimiting system of coalescing simple random walks, one started from each point of the diagonal space-time lattice. As illustrated in Figure 1, one can think of each path in the prelimiting system as the concatenation of a series of arrows, representing the jump made by the path out of each point of $\mathbb{Z}$ at each time $t \in \mathbb{Z}$, and there is then a natural dual system of arrows (on the dual lattice), pointing in the opposite direction of time, which 'fills out the gaps' between the walkers forwards in time. It is not hard to convince oneself that the law of the resultant system of backwards paths is equal to that of the forwards system, rotated by 180 degrees about the origin $(0,0)$. Under diffusive rescaling, the forwards and backwards systems converge jointly to a pair $(\mathcal{W}, \widehat{\mathcal{W}})$, known as the double Brownian web, in which $\mathcal{W}$ is the Brownian web and the dual web $\widehat{\mathcal{W}}$ has the same law as $\mathcal{W}$ rotated by 180 degrees.

[24] showed how to obtain an analogue of the Brownian web, which they dubbed the *Brownian net*, as the scaling limit of the paths traced out by a system of branching and coalescing simple random walks. If there is a branch at (x, t), then the random walker at point x at time t has two offspring which it places at $(x - 1)$ and $(x + 1)$, so that the space-time point (x, t) is connected by paths to each of $(x - 1, t + 1)$ and $(x + 1, t + 1)$. In order to obtain a non-trivial limit, the branching probability of each path in each time step is scaled to be $\mathcal{O}(1/\sqrt{n})$, corresponding exactly to the scaling in the dual to

the SΛFVS; that is at the nth stage of the rescaling the probability that two paths, one stepping left and one stepping right, emanate from a given point is $\zeta/\sqrt{n}$.

In contrast to the Brownian web, the Brownian net will have a multitude of paths coming out of each space-time point. The key to its characterisation is that it has a well-defined left-most and right-most path, which we denote l_z and r_z respectively, emanating from each point $z = (x, t) \in \mathbb{R}^2$ and these determine what is called a left-right Brownian web. Essentially, the set of left-most (resp. right-most) paths form a Brownian web with a leftwards (resp. rightwards) drift. Thus, for any deterministic pair of k-tuples of points $(z_1, \ldots, z_k)$, $(z_1', \ldots, z_{k'}')$, the left-most paths $l_{z_1}, \ldots, l_{z_k}$ are distributed as coalescing Brownian motions with drift ζ to the left, and the right-most paths $r_{z_1'}, \ldots r_{z_{k'}'}$ are distributed as coalescing Brownian motions with drift ζ to the right.

Before we can fully describe the Brownian net, we must explain how a left-most path $l_z = l_{(x,s)}$ and a right-most path $r_{z'} = r_{(x',s')}$ interact. Their joint evolution after time $s \vee s'$ is the unique weak solution to the left-right stochastic differential equation

$$
\begin{aligned}
dL_t &= \xi \mathbf{1}_{\{L_t \neq R_t\}} dB_t^l + \xi \mathbf{1}_{\{L_t = R_t\}} dB_t^c - \zeta dt, \\
dR_t &= \xi \mathbf{1}_{\{L_t \neq R_t\}} dB_t^r + \xi \mathbf{1}_{\{L_t = R_t\}} dB_t^c + \zeta dt,
\end{aligned}
\tag{3.4}
$$

where B_t^l, B_t^r and B_t^c are independent standard Brownian motions and if $s < t$ then $L_s \leq R_s \Rightarrow L_t \leq R_t$. [24] proved (weak) existence and uniqueness of the solution to this system.

A straightforward extension of (3.4) is sufficient to specify the joint distribution of any finite collection of left-right paths, which are known as left-right coalescing Brownian motions.

Remark 3.1. In [24] the drift parameter ζ of the left-right stochastic differential equation used to construct the Brownian net is allowed to vary but the diffusion constant, ξ^2, of the Brownian motions is always taken to be one. Applying a linear time change to their construction yields general ξ^2 and we will use such webs and nets (and results from elsewhere extended trivially to apply to them) without further comment. We shall refer to the Brownian net corresponding to the left-right system (3.4) as the net with drift ζ and diffusion constant ξ^2.

It remains to give a rigorous characterization of the Brownian net. One last ingredient is required.

Definition 3.2. Let $\alpha : [\sigma_\alpha, \infty) \to \mathbb{R}$ and $\alpha' : [\sigma_{\alpha'}, \infty) \to \mathbb{R}$ be paths. We say α crosses α' from left to right at time $t \in \mathbb{R}$ if there exists $t^- < t$ and $t^+ > t$ such that $\alpha(t^-) - \alpha'(t^-) < 0$ and $\alpha(t^+) - \alpha'(t^+) > 0$ and $t = \inf\{s \in (t^-, t^+) ; (\alpha(t^-) - \alpha'(t^-))(\alpha(s) - \alpha'(s)) < 0\}$.

We define a crossing of α' by α from right to left analogously. We say that α crosses α' if it does so from either left to right or right to left.

Given two paths α and α' which cross at say, time t, we can define a new path g by following α up until time t, and subsequently following α'. This procedure is known as hopping from α to α' at the (crossing) time t. Given a set P of paths, $\mathcal{H}_{cross}(P)$ is defined to be the set of paths obtained by hopping a finite number of times between paths within P.

Definition 3.3 ([24]). The Brownian net $\mathcal{N}$ is the $\mathcal{K}(\widetilde{M})$ valued random variable whose distribution is uniquely determined by the following properties:

1. For each deterministic $z \in \mathbb{R}^2$, almost surely $\mathcal{N}$ contains a unique left-most path l_z and a unique right-most path r_z.

2. For any finite deterministic set of points $z_1, \ldots, z_k, z_1', \ldots, z_{k'}' \in \mathbb{R}^2$, the collection of paths $l_{z_1}, \ldots, l_{z_k}, r_{z_1'}, \ldots, r_{z_{k'}'}$ has the distribution of a family of left-right coalescing Brownian motions.
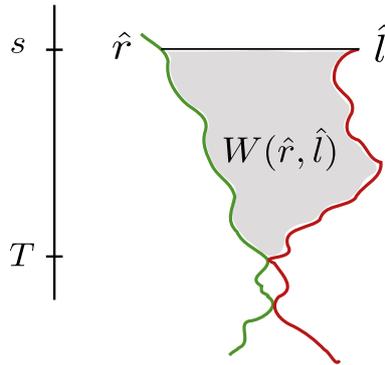
Figure 2: **A wedge.** In the notation of (3.5), the path $\hat{r}$ is in $\hat{X}_n^r$, and the path $\hat{l}$ is in $\hat{X}_n^l$. The last time at which both paths are defined is s, in this case given by $\sigma(\hat{l})$; $\hat{r}(s) < \hat{l}(s)$ and, tracing backwards in time, T is the first time at which $\hat{r} = \hat{l}$. The wedge is the shaded region.

3. For any deterministic dense countable sets $\mathcal{D}^l, \mathcal{D}^r \subseteq \mathbb{R}^2$,

$$\mathcal{N} = \overline{\mathcal{H}_{cross}(\{l_z \,;\, z \in \mathcal{D}^l\} \cup \{r_z \,;\, z \in \mathcal{D}^r\})}.$$

The proof of our main result rests on verifying the conditions of Theorem 3.4, which provides criteria under which a sequence of processes converges to the Brownian net. It is obtained by combining Theorem 6.11 and Remark 6.12 of [22]. To state it, we require the notion of a wedge. Let $(\hat{X}_n^l)$ and $(\hat{X}_l^r)$ be two random sets of paths such that their rotations by 180 degrees about $(0,0)$ are $\mathcal{K}(\widetilde{M})$ valued random variables. Take $\hat{l} \in \hat{X}_n^l$ and $\hat{r} \in \hat{X}_n^r$, defined on time intervals $(-\infty, \sigma(\hat{l})]$ and $(-\infty, \sigma(\hat{r})]$ respectively. We write $s = \sigma(\hat{l}) \wedge \sigma(\hat{r})$ for the largest time at which both paths are defined. Suppose that $\hat{r}(s) < \hat{l}(s)$ and define $T := \sup\{t < s : \hat{r}(t) = \hat{l}(t)\}$ to be the first time time the paths meet (as we trace backwards in time). We call the open set

$$W(\hat{r}, \hat{l}) := \{(x, u) \in \mathbb{R}^2 : T < u < s, \hat{r}(u) < x < \hat{l}(u)\} \tag{3.5}$$

a wedge. This set is illustrated in Figure 2. We say that a path π started at time σ_π enters W from the outside if there exists $\sigma_\pi \le u < t$ such that $(\pi(u), u) \notin \overline{W}$ and $(\pi(t), t) \in W$. Here, $\overline{W}$ denotes the closure of W.

Theorem 3.4 (Theorem 6.11, Remark 6.12 of [22])**.** *Let (X_n^l) and (X_n^r) be two sequences of $\mathcal{K}(\widetilde{M})$ valued random variables. Let $(\hat{X}_n^l)$ and $(\hat{X}_l^r)$ be two random sets of paths such that their rotations by 180 degrees about $(0,0)$ are $\mathcal{K}(\widetilde{M})$ valued random variables. Set $X_n = \mathcal{H}_{cross}(X_n^l \cup X_n^r)$ and $\hat{X}_n = \mathcal{H}_{cross}(\hat{X}_n^l \cup \hat{X}_n^r)$.*
    Suppose that:

($\mathscr{A}$) *Paths in X_n^l (resp. X_n^r) do not cross. No path in X_n crosses a path of X_n^l from right to left, and no path in X_n crosses a path of X_n^r from left to right. No path in X_n^l crosses a path of $\hat{X}_n^l$, and no path of X_n^r crosses a path of $\hat{X}_n^r$.*

($\mathscr{B}$) *For any $k \in \mathbb{N}$, and any $(z_1, \ldots, z_{2k}) \subseteq \mathbb{R} \times \mathbb{R}$ there exists a convergent sequence*

$$(l_{n,1}, \ldots, l_{n,k}, r_{n,1}, \ldots, r_{n,k}),$$

where $l_{n,i} \in X_n^l, r_{n,i} \in X_n^r$, whose limit (in distribution, in $\widetilde{M}^{2k}$, as $n \to \infty$) is a collection of left/right coalescing Brownian motions started at $(z_1, \ldots, z_{2k})$.

($\mathscr{C}$) *Whenever $k \in \mathbb{N}$ and $\hat{l}_n \in \hat{X}_n^l$ and $\hat{r}_n \in \hat{X}_n^r$ are such that $(\hat{l}_n, \hat{r}_n)$ converges (in $\widetilde{M}^2$, in distribution, as $n \to \infty$) to left/right Brownian motions $(\hat{l}, \hat{r})$, the first meeting time of $\hat{l}_n$ with $\hat{r}_n$ also converges in distribution to the first meeting time of $\hat{l}$ with $\hat{r}$.*

($\mathscr{D}$) *Paths of X_n do not enter wedges of $\hat{X}_n$ from the outside.*

Then, X_n converges (in $\mathcal{K}(\widetilde{M})$, in distribution) to the Brownian net.

3.3 Statement of the main result

We are finally in a position to give a formal statement of our result. Recall from Section 2.3, that $P_n^{\uparrow}(p)$ is the set of potential ancestral lineages of the individual at $p \in \mathbb{R}^2$ at the nth stage of our rescaling.

Let $(\mathcal{D}_n)_{n \in \mathbb{N}}$ be an increasing sequence of countable subsets of $\mathbb{R}^2$ such that, for each n, $\mathcal{D}_n$ is locally finite, and as $n \to \infty$ the set $\mathcal{D}_n$ becomes everywhere dense.

We define $\mathscr{A}(\mathcal{D}_n) = \bigcup_{p \in \mathcal{D}_n} P_n^{\uparrow}(p)$. The set $\mathscr{A}(\mathcal{D}_n)$ contains the potential ancestral lineages of all $p \in \mathcal{D}_n$. However, $\mathscr{A}(\mathcal{D}_n)$ is not an element of $\mathcal{K}(M)$, since it is not a closed subset of M, and so at the very least we should consider its closure. This requires that we augment $\mathscr{A}(\mathcal{D}_n)$ to also include ancestral lineages f that extend backwards in time until time $-\infty$, and we define $f(-\infty) = 0$ for such f. We include ∞ in the domain of each path f by defining $f(\infty) = 0$. Additionally, define the boundary paths

$$\mathcal{B} = \{f(\cdot) = -\infty \, ; \, \sigma_f \in [-\infty, \infty]\} \cup \{f(\cdot) = \infty \, ; \, \sigma_f \in [-\infty, \infty]\}. \tag{3.6}$$

We then set $P_n^{\uparrow}(\mathcal{D}_n) = \mathscr{A}(\mathcal{D}_n) \cup \mathcal{B}$. Lemma 6.4 shows that $P_n^{\uparrow}(\mathcal{D}_n)$ is an element of $\mathcal{K}(M)$.

Recall from Definition 2.2 that υ is the probability that an ancestral lineage that lies in $[x-r, x+r]$ at time $t-$ is affected by the event (x, t, r) and that $\mathbf{s}_n = \alpha/\sqrt{n}$ is the probability (at the nth stage of our rescaling) that an event is selective. Our main result is the following.

Theorem 3.5. *Let $\upsilon = 1$. As $n \to \infty$, $P_n^{\uparrow}(\mathcal{D}_n)$ converges weakly to $\mathcal{N}$ in $\mathcal{K}(M)$ where, in the terminology of Remark 3.1, $\mathcal{N}$ denotes the Brownian net with drift*

$$\zeta = \frac{2}{3}\alpha \int_0^{\mathcal{R}} r^2 \mu(dr), \tag{3.7}$$

and diffusion constant

$$\xi^2 = \frac{4}{9} \int_0^{\mathcal{R}} r^3 \mu(dr). \tag{3.8}$$

The proof of Theorem 3.5 can be found in Sections 4-6. It rests heavily on the theory of the Brownian web and net, in particular on Theorem 3.4. We will now place this result in the context of existing work and outline some of the additional difficulties that are encountered in our setting.

Consider, first, what would happen in the absence of selection. Our dual process reduces to a system of coalescing random walks and, as proved in [4], after a diffusive rescaling one recovers a system of (instantaneously) coalescing Brownian motions. If we set $\upsilon = 1$ and take the centre of the event, rather than a randomly chosen point, as the location of the parent, then this corresponds to the process of 'trajectoires d'exploration' of [17], who constructs a stochastic flow of maps by considering the dual started from every space-time point in the plane. If we specialise still further so that all events have radius 1 and we start 'exploration paths' only from the centre of each reproduction event then we recover a càdlàg version of the Poisson trees of [14]. By interpolation we recover the Poisson trees themselves. In [13] (see [12] for a more detailed account) it is shown that under a diffusive rescaling the Poisson trees converge to the Brownian web.
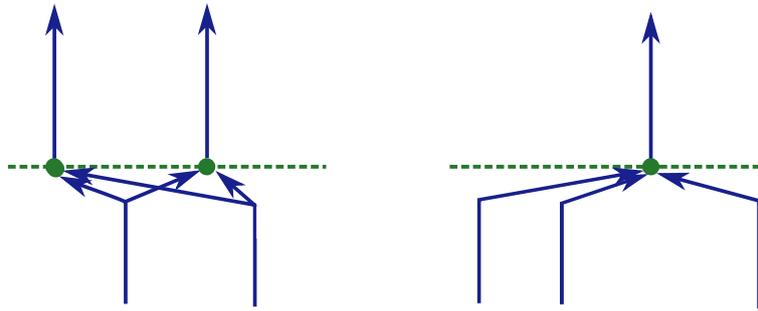
Figure 3: **Complications.** The diagram on the left illustrates the way in which paths can both coalesce and branch through the same event. The second diagram presents a case of multiple collisions in a 'neutral' event.

Even with the simplification $v = 1$, our prelimiting process is considerably more complex than that considered in [24]. When lineages are covered by the same neutral reproduction event, they coalesce. In particular, more than two lineages can coalesce in a single event. At selective events, when we must trace two potential parents, we can see either just branching or, if more than one lineage lies in the region affected by the event, a combination of branching and coalescence (see Figure 3).

Further complications compared to systems of branching and coalescing simple random walks arise since (a) our ancestral lineages jump at random times and the displacement caused by such jumps is random; and (b) the motion of distinct ancestral lineages becomes dependent when they are within distance $2\mathcal{R}$ of each other.

In spite of the additional complexity, it still makes sense to talk about left-most and right-most paths and this will be the key to our analysis. In fact (b) can be handled through elementary arguments; it turns out that the time periods during which ancestral lineages are 'nearby but not coalesced' are too brief to affect the limit.

In order to overcome (a), we must identify a dual system of (backwards in time) branching and coalescing lineages. At first sight, it is far from obvious that such a dual exists; in contrast to previous work, our pre-limiting systems will not be self-dual. We will construct a dual system with the property that, in contrast to Figure 1, after rotation by 180 degrees, although, separately, left-most and right-most paths in the dual have the same distribution as their forwards counterparts, the joint distributions of the forwards and backwards systems differ. The dual, which is defined in Section 4.1.1 is illustrated in Figure 4.

4 Convergence of left/right paths

We now turn to the proof of our main result. Recall that we take $v = 1$ so that if a lineage is in the interval covered by an event then it is necessarily affected by it.

4.1 Paths and arrows

In order to discuss the self-dual systems of branching and coalescing lineages that converge to the Brownian net, we must be precise about what we mean by 'branching-coalescing paths' and, in particular, have a notation for keeping track of the direction of time. We shall follow [15] in using segments of paths called arrows. Loosely speaking, paths are formed by concatenating arrows. A path (or an arrow) is an $\mathbb{R}$-valued function whose domain is a subinterval of $\mathbb{R}$. If a path/arrow is *forwards* (resp. *backwards*), then 'moving along it' means moving along the image of the path forwards (resp. backwards)

with respect to the usual (resp. reversed) order on the time domain. We shall use $\uparrow$ to denote forwards and $\downarrow$ to denote backwards paths.

For $a < b < c$, the concatenation of forwards paths $f : [a, b) \to \mathbb{R}$ and $g : [b, c) \to \mathbb{R}$ refers to the function $h : [a, c) \to \mathbb{R}$ which is equal to f on $[a, b)$ and equal to g on $[b, c)$. Concatenation of backwards paths is defined analogously.

When we are following a backwards path or arrow we interchange left and right, in the same way as left and right interchange if we reverse the direction in which we walk. For clarity, we reserve the terms *north, south, east and west* for global directions associated to the plane $\mathbb{R}^2$ and use the terms *right* and *left* for local directions whose frame of reference depends on the direction in which we are travelling.

4.1.1 Forwards and backwards paths

Recall from Section 2.3 that Π^n denotes the Poisson point process that drives the system of branching and coalescing paths at the nth stage of our rescaling. We refer to each $(x, t, r) \in \Pi^n$ as an event affecting the set $\{t\} \times [x - r, x + r]$ or, equivalently, affecting each point $y \in [x - r, x + r]$ at time t. The east- and west-most points of this event are $(x + r, t)$ and $(x - r, t)$ respectively. To each $(y, s) \in (-\infty, \infty) \times \mathbb{R}$ we associate a unique forwards arrow (pointing due north) and a unique backwards arrow (pointing due south), defined as follows. Let

$$T_{y,s}^{\uparrow} = \inf\{t \,;\, \exists (x, t, r) \in \Pi^n, y \in [x - r, x + r], t \geq s\},$$
$$T_{y,s}^{\downarrow} = \sup\{t \,;\, \exists (x, t, r) \in \Pi^n, y \in [x - r, x + r], t \leq s\},$$

be the times of the first event (non-strictly) north of (y, s) and the first event (non-strictly) south of (y, s), respectively, that affects the point y. Let $\star \in \{\uparrow, \downarrow\}$. An arrow starting at (y, s) is simply a path $\alpha_{y,s}^{\star} : [s, T_{y,s}^{\star}) \to \mathbb{R}$ defined to be the constant function $\alpha_{y,s}^{\star}(u) = y$. We shall call the event $(x, t, r) \in \Pi^n$ that defines $T_{y,s}^{\star}$ the *finishing event* of $\alpha_{y,s}$. It must be that $\lim_{u \uparrow T_{y,s}^{\star}} (\alpha(u), u) = (y, T_{y,s}^{\star}) \in [x - r, x + r] \times \{t\}$.

For each $\star \in \{\uparrow, \downarrow\}$, we can now associate two important sets of paths to each point (y, s). Let us first consider the forwards paths. The set $\mathcal{P}_n^{\uparrow}(y, s)$ is best described in words; it is the set of paths that are obtained by following the arrow $\alpha_{y,s}^{\uparrow}$ out of (y, s) and then, every time we finish an arrow, following a new arrow that starts from (one of) the (potential) parent(s) of the finishing event of $\alpha_{y,s}$. In other words, the forwards paths from (y, s) correspond precisely to the set of potential ancestral lineages of an individual who lived at the point y at time s, that we described in Section 2.3. We include time ∞ into the domain of each such forwards path, and set the location at time ∞ to be 0.

The set $P_n^{\downarrow}(y, s)$ of backwards paths is also best described in words. It is the set of paths obtained by first following the arrow $\alpha_{y,s}^{\downarrow}$ out of (y, s) and then, every time we finish an arrow $\alpha_{y',s'}^{\downarrow}$:

1. If the finishing event of $\alpha_{y',s'}$ is neutral with, parent at v, then

    (a) if $y' \leq v$, follow the arrow out of the west-most point of the finishing event of $\alpha_{y',s'}$,

    (b) if $y' > v$, follow the arrow out of the east-most point of the finishing event of $\alpha_{y',s'}$.

2. If the finishing event of $\alpha_{y',s'}$ is selective with potential parents at $v < v'$ then

    (a) if $y' < v$, follow the arrow out of the west-most point of the finishing event of $\alpha_{y',s'}$,

    (b) if $y' > v'$, follow the arrow out of the east-most point of the finishing event of $\alpha_{y',s'}$,

    (c) if $y' \in [v, v']$, a path can follow either one of the arrows out of the east-most/west-most points of the finishing event of $\alpha_{y',s'}$.
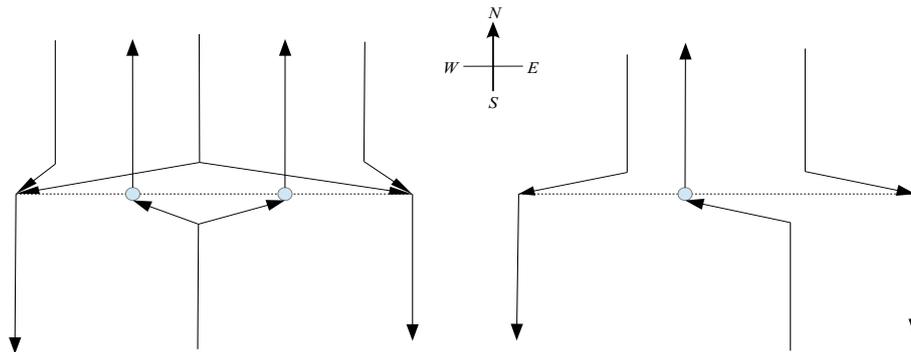
Figure 4: The movement of forwards and backwards paths (illustrated as interpolated arrows) about a selective event (left) and a neutral event (right). The events are shown as finely dotted horizontal lines and the (potential) parent(s) as small circles. Forwards paths travel northwards and backwards paths travel southwards, according to the compass shown between the two diagrams.

In analogy to forwards paths, we include time $-\infty$ into the domain of each such backwards path, and set the location at time $-\infty$ to be 0. See Figure 4 for an illustration of the forwards and backwards paths.

In keeping with our previous notation, for each forwards/backwards path f, we write $\sigma(f) = \sigma_f$, for the time at which it starts.

4.1.2 Interpolated paths and arrows

We wish to exploit the existing theory of Brownian webs and nets, which was developed in a setting restricted to continuous paths, and so we shall approximate the systems of (càdlàg) forwards and backwards paths of the last subsection by corresponding systems in which the jumps have been interpolated. This is achieved in [13] simply by taking paths that interpolate between the starting points of arrows. However, in our situation such interpolation would result in arrows which cross each other and, worse, would pass through reproduction events that did not previously affect them. Instead, we adopt a 'just in time' approach to our interpolation: we find small non-overlapping intervals of time and space about each event in which to interpolate.

Lemma 4.1. *Let $n \in \mathbb{N}$. For each $p = (x, t, r) \in \Pi^n$ and each $\epsilon > 0$ define the set*

$$\mathscr{B}_\epsilon(x, t, r) = \{(y, s) \in \mathbb{R}^2 \,;\, |x - y| \leq r, |t - s| \leq \epsilon\}.$$

Almost surely, there exists a map $\Upsilon : \Pi^n \to (0, \infty)$ such that the sets $(\mathscr{B}_{\Upsilon(p)}(p))_{p \in \Pi_n}$ are distinct.

Proof. This follows essentially immediately since Π^n has finite intensity; consequently the set of time coordinates of points of Π_n, restricted to any strip $[-K, K] \times \mathbb{R} \times [-\mathcal{R}, \mathcal{R}]$, where $K \in \mathbb{R}$, has (almost surely) no limit point. □

Let $f^\uparrow \in \mathcal{P}^\uparrow(y, s)$ and let $\alpha_{y,s}^\uparrow : [s, T_{y,s}^\uparrow) \to \mathbb{R}$ be one of the forwards arrows that make up $f^\uparrow$. Let $p = (x, t, r)$ denote the finishing event of $\alpha_{y,s}^\uparrow$ (so that, in particular, $T_{y,s}^\star = t$ a.s.). Suppose that α' is the next arrow in $f^\uparrow$ and write $z = \alpha'(T_{y,s}^\uparrow)$ for its starting point. We say that $\widetilde{\alpha}_{y,s}^\uparrow : [s, t) \to \mathbb{R}$ is the interpolated arrow of $\alpha_{y,s}^\uparrow$ if both

1. $\widetilde{\alpha}_{y,s}^\uparrow(u) = \alpha_{y,s}^\uparrow(u)$ for all $u \leq T_{y,s}^\uparrow - \Upsilon(p)$, and

2. $\widetilde{\alpha}_{y,s}^\uparrow(u)$ is linear on $[T_{y,s}^\uparrow - \Upsilon(p), t)$ and $\lim_{u \uparrow t} \widetilde{\alpha}_{y,s}^\uparrow(u) = z$.

Note that the interpolation of an arrow depends on the path f in which it is contained.

Given a forwards or backwards path $f \in \mathcal{P}_n^\uparrow(y,s)$, we define the continuous path $\tilde{f}$ to be the concatenation of the interpolations of the arrows within f, and additionally setting $f(\infty) = 0$ for forwards paths and $f(-\infty) = 0$ for backwards paths. We define

$$\widetilde{\mathcal{P}}_n^\uparrow(y,s) = \{\tilde{f}\,;\, f \in \mathcal{P}_n^\uparrow(y,s)\}$$

and define the set of interpolated backwards paths $\widetilde{\mathcal{P}}_n^\downarrow(y,s)$ in analogously. Of course, interpolated paths are close to their equivalent non-interpolated paths.

Lemma 4.2. *Let $(y,s) \in \mathbb{R}^2$ and let $f \in \mathcal{P}_n^\uparrow(y,s)$. Then $\sup_{t \in (\sigma_f,\infty)} |f(t) - \tilde{f}(t)| < 2\mathcal{R}n^{-1/2}$. The analogous estimate holds for backwards paths.*

Proof. Note that $\sigma_f = \sigma_{\tilde{f}}$. By definition, in the notation of Lemma 4.1, $f(u) = \tilde{f}(u)$ unless u is such that $(f(u), u) \in \mathscr{B}_{\Upsilon(x,t,r)}(x,t,r)$ for some $(x,t,r) \in \Pi^n$. When $(f(u),u) \in \mathscr{B}_{\Upsilon(x,t,r)}(x,t,r)$ we have $|f(u) - \tilde{f}(u)| \leq 2r$. Since, by definition of Π^n we have $r \leq \mathcal{R}n^{-1/2}$, this completes the proof. $\qquad\square$

4.1.3 Left-most and right-most paths

We now associate to each (y,s) four special paths.

Definition 4.3. *Left-most and right-most forward and backward paths are defined as follows.*

1. *The left-most forward path from (y,s) is the element of $\mathcal{P}_n^\uparrow(y,s)$ obtained by choosing the (forwards) arrow with the west-most potential parent, whenever a choice is available.*

2. *The right-most forward path from (y,s) is the element of $\mathcal{P}_n^\uparrow(y,s)$ obtained by choosing the (forwards) arrow with the east-most potential parent, whenever a choice is available.*

3. *The left-most backward path from (y,s) is the element of $\mathcal{P}_n^\downarrow(y,s)$ obtained by choosing the (backwards) arrow from the east-most point of the finishing event whenever a choice is available.*

4. *The right-most backward path from (y,s) is the element of $\mathcal{P}_n^\downarrow(y,s)$ obtained by choosing the (backwards) arrow from the west-most point of the finishing event, whenever a choice is available.*

We will sometimes shorten 'left-most' and 'right-most' to l-most and r-most.

For $D \subseteq \mathbb{R}^2$, $\dagger \in \{\uparrow, \downarrow\}$ and $\star \in \{l, r\}$ we define

$$\mathcal{Q}_n^{\star,\dagger}(D) = \{f\,;\, f = f_{y,s}^\star \text{ is the } \dagger\text{-most path of some } (y,s) \in D\}.$$

Recall from Section 3.3 that, in order to exploit the compactness properties of our state space, we must also include some extra paths, corresponding to ancestral lineages that extend backwards in time until $-\infty$. First, we say that a path $f : \mathbb{R} \to \mathbb{R}$ is an infinite extender of $\mathcal{Q}_n^{\uparrow,\dagger}(D)$ if there exists a sequence $(f_m)_{m=1}^\infty \subseteq \mathcal{Q}_n^{\uparrow,\dagger}(D)$ and a sequence (t_m) such that $t_m \downarrow -\infty$ and $f(t) = f_m(t)$ for all m and $t \geq t_m$. We make the corresponding definition for $\mathcal{Q}_n^{\downarrow,\dagger}(D)$ and, for $\star \in \{\uparrow, \downarrow\}$ and $\dagger \in \{l, r\}$ we define $\mathcal{Q}_n^{\star,\dagger,inf}(D)$ to be the set of infinite extenders of $\mathcal{Q}_n^{\star,\dagger}(D)$. Recall also the boundary paths $\mathcal{B}$ defined in (3.6). Then, define

$$\mathcal{P}_n^{\star,\dagger}(D) = \mathcal{Q}_n^{\star,\dagger}(D) \cup \mathcal{Q}_n^{\star,\dagger,inf}(D) \cup \mathcal{B}, \tag{4.1}$$

and $\widetilde{\mathcal{P}}_n^{\star,\dagger}(D) = \{\tilde{f}\,;\, f \in \mathcal{P}^{\star,\dagger}(D)\}$ to be the corresponding sets of interpolated paths.

We now verify Condition $(\mathscr{A})$ of Theorem 3.4. Recall, from Definition 3.2 what it means for two paths to cross.

Lemma 4.4. *Let D be any subset of $\mathbb{R}^2$ and let $n \in \mathbb{N}$. Let $\dagger \in \{l, r\}$ and $\star \in \{\uparrow, \downarrow\}$. Then, almost surely:*

1. *For all $f^\uparrow \in \mathcal{P}_n^{\uparrow,\dagger}(D)$ and $f^\downarrow \in \mathcal{P}_n^{\downarrow,\dagger}(D)$, the paths $f^\uparrow$ and $f^\downarrow$ do not cross.*

2. *For all $f^\star, g^\star \in \mathcal{P}^{\star,\dagger}(D)$, the paths $f^\star$ and $g^\star$ do not cross.*

Further, the same results hold for interpolated paths $f^\uparrow \in \widetilde{\mathcal{P}}_n^{\uparrow,\dagger}(D)$ and $f^\downarrow \in \widetilde{\mathcal{P}}_n^{\downarrow,\dagger}(D)$.

Proof. In the first case, note that two forwards paths can only cross if they are first coalesced and are then subsequently affected by the same selective reproduction event. In the second case, note that a forward path can only cross a backwards path if both are affected by a common event. In both cases, the fact that crossing cannot occur is then an easy consequence of the definitions (or see Figure 4). □

Remark 4.5. A forwards left-most path can cross a backwards right-most path, and a forwards right-most path can cross a backwards left-most path. Similarly, a forwards left-most path can cross a forwards right-most path (if they are both affected by the same selective event), and a backwards right-most path can cross a backwards left-most path.

Although not immediately obvious from the definition, the next lemma is a helpful feature of our construction.

Lemma 4.6. *A forwards left- (resp. right-) most path has the same distribution as a backwards left- (resp. right-) most path which has been rotated by 180 degrees.*

Proof. The proof is based on the movements of paths affected by reproduction events, which is depicted in Figure 4. It suffices to consider the case of left-most paths; the case of right-most paths then follows by symmetry.

First observe that the rate at which an event falls on (an arrow in) a path has the same distribution whether we look forwards or backwards in time and, when an event falls on a (forwards or backwards) path, the spatial position of the path will be uniformly distributed over the region affected by the event. Let us denote that position by V. Thus if the event corresponds to $p = (x, t, r)$, then V is uniformly distributed on $[-r, r]$.

Consider a left-most forwards path affected by a neutral event. The path jumps to the position of the parent, which we denote by U. Thus, on the event $V < U$ our path jumps a distance $U - V$ to the left, and on the event $U > V$ it jumps a distance $V - U$ to the right.

Now consider the left-most backwards path. Retaining the notation above, at a neutral event, on the event $V < U$ the path jumps to the west-most endpoint, which, once rotated by 180 degrees becomes a jump to the right of size $V - (-r)$. On the other hand, on the event $V > U$, the path jumps to the east-most endpoint, which upon rotation becomes a leftwards jump of magnitude $r - V$.

Conditional on $V < U$, V is uniform on $(-r, U)$, so $U - V \overset{d}{=} V - (-r)$. Similarly, conditional on $V > U$, V is uniform on (U, r) and $V - U \overset{d}{=} r - V$. Therefore, if we restrict to only neutral events, forwards left-most paths and backwards left-most paths rotated by 180 degrees have the same distribution.

Next, consider a selective event. We use a similar argument. The two potential parents are sampled uniformly from the event. We denote their positions by $U_1 < U_2$. Combined with V, we now have three independent uniformly distributed random variables on $[-r, r]$. Let us write them in ascending order as $U^{(1)}$, $U^{(2)}$, $U^{(3)}$. The following events may occur:

(a) $V = U^{(1)}$, in which case $U_1 = U^{(2)}$, so the path makes a rightwards jump of magnitude $U_1 - V$;

(b) $V \neq U^{(1)}$, in which case $U_1 = U^{(1)}$, so the path makes a leftwards jump of magnitude $V - U_1$.

Note that we are not concerned by the value of U_2, since we are interested in a left-most path. For a left-most backwards path, again at a selective event, the following events may occur:

(a) $V = U^{(1)}$, in which case the path jumps to the west end-point of the event, a jump which after rotation by 180 degrees becomes a rightwards jump of magnitude $V - (-r)$;

(b) $V \neq U^{(1)}$, in which case the path jumps to the east end-point of the event, a jump which after rotation by 180 degrees becomes a leftwards jump of magnitude $r - V$.

Again, because we consider a left-most path we are not concerned by the value of U_2.

We now compare the jumps in the (a) cases. Conditional on $V < U_1$, V is uniformly distributed on $(-r, U_1)$ and thus (as in the neutral case) $V - U_1 \stackrel{d}{=} V - (-r)$. Similarly, for the (b) cases, conditional on $U_1 < V$, V is uniformly distributed on (U_1, r) and thus (also as in the neutral case) $V - U_1 \stackrel{d}{=} r - V$. Thus the left-most forwards path and the rotated left-most backwards paths have the same distribution, which completes the proof. □

Remark 4.7. 1. Note that Lemma 4.6, with the same proof, remains true when the parent locations are sampled according to any symmetric distribution on $(-r, r)$.

2. In previous work on the Brownian web and net, there is a strict self-duality in the prelimiting systems. Here, we see a new feature. Although separately the left and right-most paths have the same distributions forwards and backwards in time, their joint distribution differs. As can be seen in Figure 4, our backwards paths branch less frequently than forwards ones, but when they do branch, they make larger jumps.

Recall the state space $\mathcal{K}(M)$ defined in Section 3.1. The space $\mathcal{K}(M)$ is an appropriate space in which to consider convergence of sets of (branching/coalescing) forwards paths, but it is not suitable for backwards paths. To remedy this, if P is a set of backwards paths then we define $-P = \{\hat{f} \,;\, f \in P\}$, where $\hat{f} : [-\sigma_f, \infty] \to [-\infty, \infty]$ given by $\hat{f}(t) = -f(-t)$ is the rotation of f by 180 degrees. Thus, $-P \in M$ is a set of forwards paths. With a slight abuse of notation, if f_n is a sequence of backwards paths and f is a backwards path, we will say $f_n \to f$ in M if $\hat{f}_n \to \hat{f}$ in M. Similarly, if P_n is a sequence of sets of backwards paths and P is a set of backwards paths we write $P_n \to P$ in $\mathcal{K}(M)$ to mean that $-P_n \to -P$ in $\mathcal{K}(M)$. We apply the same terminology to interpolated paths.

4.2 Convergence of a pair of left/right paths

We must ultimately verify that any limit point of our combined systems of left and right-most paths will satisfy condition $(\mathscr{B})$ of Theorem 3.4. As a first step, in this subsection we take the limit of a pair of paths, comprising one left-most path and one right-most path started at some time s (which, since Π^n is homogeneous in both space and time, we may, without loss of generality, take to be zero) and show that it satisfies the system (3.4). Our approach mirrors that in [24], and as far as possible we shall adhere to their notation. With this in mind, let L^n and R^n denote respectively the left-most and right-most forwards paths associated to the points $(y^{n,l}, 0)$ and $(y^{n,r}, 0)$. We assume that the sequences of starting points converge to $(y^l, 0)$ and $(y^r, 0)$ respectively.

Remark 4.8. A straightforward modification (in order to take into account the selective events and the resulting drift of the left and right-most paths) of Lemma 4.1 in [4] shows

that the pair (L^n, R^n) stopped when it first enters the 'coalesced' state converges in distribution to a pair of independent Brownian motions with drift $\pm\zeta$, stopped when they first meet. In particular, the first meeting times also (jointly) converge.

Following [24], using their Lemma 2.2, when $L_0 \leq R_0$ there is a one-to-one correspondence between weak solutions of (3.4) and solutions of the system

$$dL_s = \xi dB^l_{S_s} + \xi dB^c_{C_s} - \zeta ds, \tag{4.2}$$

$$dR_s = \xi dB^r_{S_s} + \xi dB^c_{C_s} + \zeta ds, \tag{4.3}$$

$$s = S_s + C_s, \tag{4.4}$$

$$0 = \int_0^s \mathbb{1}\{L_s < R_s\} dC_s, \tag{4.5}$$

where B^l, B^r and B^c are independent standard one dimensional Brownian motions. The infinitesimal variance ξ^2 of the Brownian motion and the drift ζ depend on α and μ and are given by (3.8) and (3.7) respectively. The solution (L_s, R_s) to this system is a $C_{\mathbb{R}^2}[0, \infty)$ valued process.

In the case $R_0 < L_0$, according to (3.4) both R_s and L_s evolve as independent Brownian motions, with drift $\pm\zeta$, until they meet. Thus, in view of Remark 4.8, it suffices to treat the case of $L_0 \leq R_0$, where $L_0 = y^l$ and $R_0 = y^r$.

The essence of (4.2)-(4.5) is that, once L_s and R_s meet, they will accumulate nontrivial time together as a result of a sticky interaction (see Proposition 2.1 in [24] for details). As part of the proof of their Lemma 2.2, [24] show that $S_s = \int_0^s \mathbb{1}\{L_u < R_u\} du$ and $C_s = \int_0^s \mathbb{1}\{L_u = R_u\} du$.

Proposition 4.9. *Let* $T \in (0, \infty)$. *As* $n \to \infty$, $(L^n_s, R^n_s)_{s \in [0,T]}$ *converges weakly to* $(L_s, R_s)_{s \in [0,T]})$ *in the sense of* $D_{\mathbb{R}^2}[0, T]$ *valued processes.*

The analogous result for interpolated paths, which we denote by $(\widetilde{L}^n, \widetilde{R}^n)$, follows easily:

Corollary 4.10. *Let* $T \in (0, \infty)$. *As* $n \to \infty$, $(\widetilde{L}^n_s, \widetilde{R}^n_s)_{s \in [0,T]}$ *converges weakly to* $(L_s, R_s)_{s \in [0,T]})$ *in the sense of* $C_{\mathbb{R}^2}[0, T]$ *valued processes.*

Proof. By Lemma 4.2, the weak convergence of Proposition 4.9 also holds (in $D_{\mathbb{R}^2}[0, T]$) when (L^n, R^n) is replaced by $(\widetilde{L}^n, \widetilde{R}^n)$. Since the space of continuous paths with the supremum topology is continuously embedded in the space of càdlàg paths with the Skorohod topology, it follows that the same convergence holds in $C_{\mathbb{R}^2}[0, T]$. □

The remainder of this subsection is devoted to the proof of Proposition 4.9. We begin by breaking down the evolution of the pair (L^n_s, R^n_s) into several different pieces. At time $s \geq 0$, we say L^n_s and R^n_s are

$$\text{coalesced if } L^n_s = R^n_s,$$

$$\text{nearby if } L^n_s \neq R^n_s \text{ and } |L^n_s - R^n_s| \leq \frac{2\mathcal{R}}{n^{1/2}},$$

$$\text{separated if } |L^n_s - R^n_s| > \frac{2\mathcal{R}}{n^{1/2}}.$$

For $s \geq 0$ we set

$$C^n_s = \int_0^s \mathbb{1}\{L^n_u, R^n_u \text{ are coalesced}\} du,$$

$$N^n_s = \int_0^s \mathbb{1}\{L^n_u, R^n_u \text{ are nearby}\} du, \tag{4.6}$$

$$S^n_s = \int_0^s \mathbb{1}\{L^n_u, R^n_u \text{ are separated}\} du,$$

and we note that $C_s^n + N_s^n + S_s^n = s$.

We define sequences of stopping times to track the changes of state of (L^n, R^n) during $[0, T]$. Firstly,

$$\tau_1^{n,C} = \inf\{s \geq 0\,;\, (L_s^n, R_s^n) \text{ are coalesced}\};$$
$$\tau_k^{n,C} = \inf\{s \geq \tau_{k-1}^{n,C}\,;\, (L_s^n, R_s^n) \text{ are coalesced}, (L_{s-}^n, R_{s-}^n) \text{ are not coalesced}\}.$$

Similarly, we define sequences $\tau_k^{n,N}$ and $\tau_k^{n,S}$ for 're-entrance' times of (L_r^n, R_s^n) to the states of 'nearby' and 'separated' respectively. It is easily seen that each $\tau_k^{n,C}, \tau_k^{n,N}, \tau_k^{n,S}$ is a stopping time and (L^n, R^n) is strong Markov.

Each jump of (L^n, R^n) is caused by one or both lineages being affected by a single event of Π^n. If (L^n, R^n) is coalesced immediately before this event then the event affects both L^n and R^n, whereas if they are separated the event affects only one of the two. The motion is more complicated when (L^n, R^n) is in the nearby state, when events can affect one or both lineages, but we shall see that the time spent in that state is negligible as we pass to the limit.

In order to identify the limiting objects, it is convenient to isolate the parts of the motion that contribute to the drift from those that contribute to the martingale terms in (4.2) and (4.3). The decomposition we make is not unique. Our particular choice highlights the fact that the martingale part of the motion of lineages is driven by neutral events, while the drift can be attributed to selection.

First we are going to define three random walks, from which we can build (L^n, R^n) when we are in the coalesced or separated states. To understand the origin of these, first suppose that a lineage is hit by a neutral event. When this happens, the position, y, of the lineage is uniformly distributed on the region affected by the event and it will jump to the position z of the parent, which is also uniformly distributed on the region. Neutral events fall according to a Poisson Point Process with intensity

$$n^{1/2}dx \otimes n(1 - \mathbf{s}_n)dt \otimes \mu^n(dr),$$

so they hit y at rate

$$K_n = n(1 - \mathbf{s}_n) \int_{-\infty}^{\infty} \int_0^{\infty} \int_{-r}^{r} \mathbb{1}\{y \in [x - r, x + r]\}\, dx\, \mu^n(dr)\, n^{1/2}dx$$
$$= 2n(1 - \mathbf{s}_n) \int_0^{\infty} r\mu(dr). \tag{4.7}$$

We define V^n to be a symmetric random walk driven by a Poisson Point Process with intensity

$$n(1 - \mathbf{s}_n)dt \otimes 2r\mu(dr).$$

At an event (t, r), the walk jumps with displacement $J_1/\sqrt{n}$ where

$$\mathbb{P}\left[J_1 \in A\right] = \mathbb{P}\left[Z_r - U_r \in A\right], \tag{4.8}$$

and U_r and Z_r are independent uniform random variables on $[0, 2r]$.

Now consider the motion due to selective events. If the pair is coalesced immediately before the event, then their position is uniformly distributed on the affected region and the left-most path will jump with displacement $z_1 - y$ and the right-most path jumps with displacement $z_2 - y$ where $z_1 < z_2$ are the (uniformly distributed) positions of the two potential parents of the event. If the pair (L^n, R^n) is separated, then only one of them will be affected by any given event. Selective events fall with intensity

$$n^{1/2}dx \otimes n\mathbf{s}_n dt \otimes \mu^n(dr).$$

We define a random walk $(D^{n,-}, D^{n,+})$, whose jumps are driven by a Poisson Point Process with intensity

$$n\mathbf{s}_n dt \otimes 2r\mu(dr).$$

At an event (t, r), $D^{n,-}$ jumps with displacement $(Z_1 - Y)/\sqrt{n}$ and $D^{n,+}$ jumps with displacement $(Z_2 - Y)/\sqrt{n}$ where $Z_1 = \min\{U_1, U_2\}$ and $Z_2 = \max\{U_1, U_2\}$ with U_1, U_2 and Y independent uniformly distributed random variables on $[0, 2r]$.

Lemma 4.11. *As $n \to \infty$, V^n converges weakly to ξB where B is a standard Brownian motion and*

$$\xi^2 = \frac{4}{9} \int_0^{\mathcal{R}} r^3 \mu(dr).$$

Proof. Evidently J_1 has mean zero and, conditional on r, its variance is $4r^2$ times the variance of the minimum of two independent uniform random variables on $[0, 1]$. Thus, conditional on r, the variance of J_1 is $2r^2/9$. The lemma now follows from the Functional Central Limit Theorem (see, for example, [11], Section 7.1). $\qquad\square$

Now consider $D^{n,\pm}$.

Lemma 4.12. *Let $T > 0$. As $n \to \infty$, $(D^{n,-}, D^{n,+})$ converges weakly to the deterministic process $s \mapsto (-\zeta s, \zeta s)$ where*

$$\zeta = \frac{2}{3}\alpha \int_0^{\mathcal{R}} r^2 \mu(dr).$$

Proof. Since these walks experience jumps of size $\mathcal{O}(1/\sqrt{n})$ at rate

$$2n\mathbf{s}_n \int_0^{\mathcal{R}} r\mu(dr) = 2\alpha\sqrt{n} \int_0^{\mathcal{R}} r\mu(dr), \tag{4.9}$$

which is proportional to $\sqrt{n}$, we see that we have a strong law rescaling. In the notation above, conditional on r, $\mathbb{E}[Z_2 - Y] = -\mathbb{E}[Z_1 - Y] = \frac{r}{3}$. By the law of large numbers as $n \to \infty$, $(D^{n,-}, D^{n,+})$ converges weakly to the deterministic process $s \mapsto (-\zeta s, \zeta s)$, with ζ as in the statement of the lemma. $\qquad\square$

When (L^n, R^n) is coalesced, its jumps have the same distribution as $(V^n + D^{n,-}, V^n + D^{n,+})$. When (L^n, R^n) is separated, its jumps have the same distribution as $(V^{n,l} + D^{n,l,-}, V^{n,r} + D^{n,r,+})$ where $V^{n,l}, V^{n,r}$ are independent copies of V^n and $D^{n,l,-}$, $D^{n,r,+}$ are independent and with the same distribution as $D^{n,-}$, $D^{n,+}$ respectively. When L^n and R^n are nearby the evolution is more complicated; in fact in this case the joint jump distribution depends on $|L^n - R^n|$. Happily, because of Lemma 4.15 (see below), we will not need to describe the evolution in this case explicitly and we will denote it simply by $(\mathcal{N}_s^{n,l}, \mathcal{N}_s^{n,r})$.

Since (L^n, R^n) is always in exactly one of the states 'coalesced', 'nearby', and 'separated', and using spatial and temporal homogeneity of Π^n, it follows from the above that we can represent the dynamics of (L^n, R^n) in terms of three independent copies of the triple $(V^n, D^{n\pm})$ which we denote $(V^{n,\alpha}, D^{n,\alpha,\pm})$ with $\alpha \in \{c, l, r\}$:

$$L_s^n = L_0^n + V_{S_s^n}^{n,l} + D_{S_s^n}^{n,l,-} + \mathcal{N}_{N_s^n}^{n,l} + V_{C_s^n}^{n,c} + D_{C_s^n}^{n,c,-}, \tag{4.10}$$

$$R_s^n = R_0^n + V_{S_s^n}^{n,r} + D_{S_s^n}^{n,r,+} + \mathcal{N}_{N_s^n}^{n,r} + V_{C_s^n}^{n,c} + D_{C_s^n}^{n,c,+}, \tag{4.11}$$

$$s = C_s^n + N_s^n + S_s^n, \tag{4.12}$$

$$0 = \int_0^s \mathbb{1}\{R_u^n > L_u^n\} dC_u^n. \tag{4.13}$$

Of course the 'clocks' (C^n, S^n, N^n) are coupled with the random walks $V^{n,\alpha}$ and $D^{n,\alpha,\pm}$.

Our next task is to prove that the time spent in the 'nearby' state is negligible. We require two preliminary estimates on the time between changes of state. The first says that each visit to the nearby state lasts at most $\mathcal{O}(1/n)$ units of time. The second estimate says that visits to the coalesced state last $\mathcal{O}(1/\sqrt{n})$ units of time. This will limit the possible number of such visits in the time interval $[0, T]$ to be $\mathcal{O}(\sqrt{n})$ and since, moreover, the number of visits to the nearby state before the pair visits the coalesced state is $\mathcal{O}(1)$, this in turn allows us to control the number of visits to the nearby state.

Lemma 4.13. *Let $k \in \mathbb{N}$ and let $\tau'_{k'}$ be the next state change after $\tau_k^{n,N}$. Then the random variables $(\tau'_{k'} - \tau_k^{n,N})_{k \in \mathbb{N}}$ are an independent sequence and there exists $A \in (0, \infty)$, not dependent on k, such that*

$$\mathbb{E}\left[\tau'_{k'} - \tau_k^{n,N}\right] \leq \frac{A}{n}.$$

Further, there exists $q > 0$, not dependent on k, such that the probability that (L^n, R^n) is coalesced at $\tau'_{k'}$ is greater than q.

Proof. Independence is clear since the jumps determining the distinct $\tau'_{k'} - \tau_k^{n,N}$ are driven by disjoint collections of events of Π^n. If (L^n, R^n) are nearby they must either be at a distance smaller than $\frac{3}{2}\mathcal{R}/\sqrt{n}$, or at a distance between $\frac{3}{2}\mathcal{R}/\sqrt{n}$ and $2\mathcal{R}/\sqrt{n}$. In the first scenario, the probability that they coalesce through the the next event that affects either of them is bounded away from 0. In the second scenario, the probability that the next event that affects either of them brings them closer than $\frac{3}{2}\mathcal{R}/\sqrt{n}$ is also bounded away from 0. This guarantees that the probability that (L^n, R^n) is coalesced at $\tau'_{k'}$ is bounded below by some $q > 0$. On the other hand, if the walkers are at a distance in $(\frac{3}{2}\mathcal{R}/\sqrt{n}, 2\mathcal{R}/\sqrt{n})$, the probability that they separate at the next step is strictly positive. Thus the number of jumps until they either coalesce of separate has finite mean and since events affect them at rate $\mathcal{O}(n)$ the result follows. $\square$

Lemma 4.14. *Let $k \in \mathbb{N}$ and let $\tau''_{k''}$ be the next state change after $\tau_k^{n,C}$. Then the random variables $(\tau''_{k''} - \tau_k^{n,C})_{k \in \mathbb{N}}$ are an i.i.d. sequence and there exists $A' \in (0, \infty)$, not dependent on k, such that*

$$\mathbb{E}\left[\tau''_{k''} - \tau_k^{n,C}\right] \geq \frac{A'}{\sqrt{n}}.$$

Proof. This is trivial, since any jump out of the coalesced state is due to a selective event and the rate at which these occur is given by (4.9). $\square$

Lemma 4.15. *Fix $T > 0$ and let N_T^n denote the total time spent in the nearby state up to time T. Then $N_T^n \to 0$ in probability as $n \to \infty$.*

Proof. The idea is simple. Since $C_T \leq T$, using Lemma 4.14, the number of visits to the coalesced state in $[0, T]$ has mean at most $T\sqrt{n}/A'$. But by Lemma 4.13, the expected number of visits to the coalesced state is at least q times the expected number of visits to the nearby state. Thus the expected number of visits to the nearby state is at most $T\sqrt{n}/(qA')$ and since, again by Lemma 4.13, each has expected duration at most A/n, $\mathbb{E}[N_T] \leq TA/(qA'\sqrt{n})$ and the result is proved. $\square$

Since L^n and R^n evolve as $V^n + D^{n,-}$ and $V^n + D^{n,+}$ respectively, both converge individually and so their joint law is tight. Moreover, since C^n, N^n and S^n are continuous increasing processes, with rate of increase bounded by one, their joint law is also tight. Evidently we now have that

$$(L_s^n, R_s^n, V_s^{n,l}, V_s^{n,r}, V_s^{n,c}, D_s^{n,l,-}, D_s^{n,r,+}, D_s^{n,c,-}, D_s^{n,c,+}, C_s^n, N_s^n, S_s^n)_{s \geq 0}$$

is tight and by passing to a subsequence we may assume that it converges weakly to some limiting process

$$(L_s, R_s, \xi B_s^l, \xi B_s^r, \xi B_s^c, -\zeta s, \zeta s, -\zeta s, \zeta s, C_s, 0, S_s)_{s \geq 0},$$

where B^l, B^r and B^c are independent (by construction). Here, Lemma 4.15 gives that $N_s^n \to 0$. By Skorohod's Representation Theorem, by passing to a further subsequence if necessary, we can assume that the convergence is almost sure. We claim that the limit $(L_s, R_s, C_s, S_s)_{s \geq 0}$ then satsifies (4.2-4.5). Indeed, letting $n \to \infty$ in (4.10), (4.11) and (4.12) we obtain precisely (4.2), (4.3) and (4.4). Note that here the term $\mathcal{N}_{N_s^n}^{n,\dagger}$ vanishes as an easy consequence of $N_s^n \to 0$.

Obtaining (4.5) from (4.13) requires a little more work (because the function $x \mapsto \mathbb{1}\{x > 0\}$ is not continuous), but we need only adapt the approach of [24]. For each $\delta > 0$ let ρ_δ be a continuous non-decreasing function such that $\rho_\delta(u) = 0$ for $u \in [0, \delta]$ and $\rho_\delta(u) = 1$ for $u \in [\delta, \infty)$. Using (4.13) we have

$$
\begin{aligned}
0 &= \int_0^T \mathbb{1}\{R_s^n > L_s^n\} dC_s^n \\
&= \int_0^T \mathbb{1}\left\{R_s^n - L_s^n > \frac{2\mathcal{R}}{n^{1/2}}\right\} dC_s^n + \int_0^T \mathbb{1}\left\{R_s^n - L_s^n \in \left(0, \frac{2\mathcal{R}}{n^{1/2}}\right]\right\} dC_s^n \\
&\geq \int_0^T \rho_\delta(R_s^n - L_s^n) dC_s^n \geq 0,
\end{aligned}
$$

provided that $\delta \geq 2\mathcal{R}/\sqrt{n}$. For such n we thus have $\int_0^T \rho_\delta(R_s^n - L_s^n) dC_s^n = 0$ and letting $n \to \infty$ we obtain $\int_0^T \rho_\delta(R_s - L_s) dC_s = 0$ for all $\delta > 0$. Letting $\delta \to 0$ we obtain $\int_0^T \mathbb{1}\{R_s > L_s\} dC_s = 0$ which is (4.5).

This completes the proof of Proposition 4.9.

Remark 4.16. An entirely analogous proof gives convergence of a pair of backwards right and left-most paths to left/right Brownian motions. In view of Remark 4.8, this convergence occurs jointly with convergence of their first meeting time. That the constants ξ and ζ are unchanged follows from Lemma 4.6.

5 Spaces of càdlàg paths

In this section, we construct the space $\mathcal{K}(M)$, which is the càdlàg path equivalent of the state space introduced by [15] for the Brownian web (and later used in [24] for the net).

5.1 Skorohod paths with different domains

We begin by studying the space

$$G = \{g : [\sigma_g, 2] \to [-1, 1] \, ; \, g \text{ is càdlàg, } \sigma_g \in [-1, 1], \, g \text{ is constant on } [1, 2]\}.$$

We wish to treat G as a space of paths with a Skorohod-like topology, but since paths in G can have different domains, we must extend the usual approach. We refer to Chapter 3, Section 12 of [5] and Chapter 3, Section 5 of [11], upon which our arguments are heavily based, for the standard theory of the Skorohod topology.

For $g, h \in G$, let $\Lambda'[g, h]$ denote the set of strictly increasing bijections from $[\sigma_g, 2] \to [\sigma_h, 2]$. We define $\Lambda[g, h]$ to be the subset of $\lambda \in \Lambda'[g, h]$ for which

$$\gamma_{g,h}(\lambda) = \sup_{\sigma_g \leq t < s \leq 2} \left| \log \frac{\lambda(s) - \lambda(t)}{s - t} \right| < \infty.$$

For such g, h, λ we define

$$d(g, h, \lambda) = \sup_{t \in [\sigma_g, 2]} |g(t) - h(\lambda(t))|.$$

and

$$\rho(g, h) = \inf_{\lambda \in \Lambda[g,h]} \Big(\gamma_{g,h}(\lambda) \vee d(g, h, \lambda) \Big). \tag{5.1}$$

Our main aim in this subsection is to show that G is a complete and separable metric space under the metric

$$d(g, h) = \rho(g, h) \vee |\sigma_g - \sigma_h|$$

Intuitively, this says that paths in G converge if their domains converge and, as the domains become close, the paths also become close (in the Skorohod sense). We take $\sigma_g \in [-1, 1]$ and the domain of $g \in G$ to be $[\sigma_g, 2]$ for technical reasons: if instead we took the domain $[\sigma_g, 1]$, $\Lambda'[g, h]$ would be empty whenever $\sigma_h < \sigma_g = 1$. For $s \in [-1, 1]$, we write $G[s] = \{g \in G : \sigma_g = s\}$.

Remark 5.1. For $s \in [-1, 1]$, $G[s]$ is precisely the space of càdlàg paths mapping $[s, 2] \to [-1, 1]$ that are constant on $[1, 2]$. Moreover, on $G[s]$, ρ coincides with the usual Skorohod metric.

Lemma 5.2. *The space (G, d) is a metric space.*

Proof. If $d(g, h) = 0$ then $\sigma_g = \sigma_h$, so by Remark 5.1 we have $g = h$. For any $\lambda \in \Lambda[g, h]$, we have $\lambda^{-1} \in \Lambda[h, g]$. Since

$$\gamma_{g,h}(\lambda) = \sup_{\sigma_g \leq t < s \leq 2} \left| \log \frac{\lambda(s) - \lambda(t)}{s - t} \right| = \sup_{\sigma_h \leq t < s \leq 2} \left| \log \frac{s - t}{\lambda^{-1}(s) - \lambda^{-1}(t)} \right| = \gamma_{h,g}(\lambda^{-1})$$

and, similarly, $d(g, h, \lambda) = d(h, g, \lambda^{-1})$, we have that d is symmetric.

It remains to prove that d satisfies the triangle inequality, for which it suffices to show that the triangle inequality holds for ρ. To see this, take $f, g, h \in G$. For $\lambda_1 \in \Lambda[f, g]$ and $\lambda_2 \in \Lambda[g, h]$ we have $\lambda_2 \circ \lambda_1 \in \Lambda[f, h]$ and

$$\begin{aligned}
\gamma_{f,h}(\lambda_2 \circ \lambda_1) &= \sup_{\sigma_f \leq t < s \leq 2} \left| \log \frac{(\lambda_2 \circ \lambda_1)(s) - (\lambda_2 \circ \lambda_1)(t)}{\lambda_1(s) - \lambda_1(t)} \frac{\lambda_1(s) - \lambda_1(t)}{s - t} \right| \\
&\leq \sup_{\sigma_g \leq t < s \leq 2} \left| \log \frac{\lambda_2(s) - \lambda_2(t)}{s - t} \right| + \sup_{\sigma_f \leq t < s \leq 2} \left| \log \frac{\lambda_1(s) - \lambda_1(t)}{s - t} \right| \\
&= \gamma_{g,h}(\lambda_2) + \gamma_{f,g}(\lambda_1). \tag{5.2}
\end{aligned}$$

Similarly,

$$\begin{aligned}
d(f, h, \lambda_1 \circ \lambda_2) &= \sup_{t \in [\sigma_f, 2]} |f(t) - h(\lambda_2(\lambda_1(t)))| \\
&\leq \sup_{t \in [\sigma_f, 2]} |f(t) - g(\lambda_1(t))| + \sup_{t \in [\sigma_g, 1]} |g(t) - h(\lambda_2(t))| \\
&= d(f, g, \lambda_1) + d(g, h, \lambda_2). \tag{5.3}
\end{aligned}$$

Combining (5.2) and (5.3) we have that $\rho(f, h) \leq \rho(f, g) + \rho(g, h)$, as required. $\quad\square$

Lemma 5.3. *The space (G, d) is separable.*

Proof. Let $g \in G$ and suppose $\sigma_g \in (-1, 1)$. Let (q_i) be an increasing sequence in $\mathbb{Q} \cap (-1, 1)$ such that $q_i \uparrow \sigma_g$ and, for each i, define $\lambda_i : [q_i, 2] \to [\sigma_g, 2]$ by setting

$\lambda_i(q_i) = \sigma_g$, $\lambda_i(1) = 1$, $\lambda_i(2) = 2$ and taking λ_i to be linear on $[q_i, 1]$ and on $[1, 2]$. Define $g_i \in G[q_i]$ by $g_i(t) = g(\lambda_i(t))$. Then $\lambda^{-1} \in \Lambda[g, g_i]$ and

$$\gamma_{g,g_i}(\lambda_i^{-1}) = \left| \log \frac{1 - q_i}{1 - \sigma_g} \right|, \quad d(g, g_i, \lambda_i^{-1}) = 0.$$

Hence $d(g, g_i) \to 0$ as $i \to \infty$. By Remark 5.1, for each $q \in \mathbb{Q} \cap [-1, 1]$ the space $(G[q], \rho)$ is separable, hence (G, d) is separable. $\qquad\square$

Before we address completeness, we recall that the Skorohod topology is often characterized using a metric with respect to which it is not complete; this characterization is useful primarily because it is easier to work with. The extension to G is as follows.

For $g, h \in G$ and $\lambda \in \Lambda'[g, h]$ define $\gamma'_{g,h}(\lambda) = \sup_{t \in [\sigma_g, 2]} |\lambda(t) - t|$. Then, let

$$\rho'(g, h) = \inf_{\lambda \in \Lambda'[g,h]} \left(\gamma'_{g,h}(\lambda) \vee d(g, h, \lambda) \right)$$

and define $d'(g, h) = |\sigma_g - \sigma_h| \vee \rho'(g, h)$. It can be checked, in similar style to the proof of Lemma 5.2, that (G, d') is a metric space.

Lemma 5.4. *The metrics d and d' generate the same topology on G.*

Proof. First note that $\Lambda[g, h] \subseteq \Lambda'[g, h]$ and, since $|x - 1| \leq e^{|\log x|} - 1$ for all $x > 0$, for $\lambda \in \Lambda[g, h]$ we have

$$\gamma'_{g,h}(\lambda) = \sup_{t \in (\sigma_g, 2]} |t - \sigma_g| \left| \frac{\lambda(t) - \sigma_g}{t - \sigma_g} - 1 \right| \leq \sup_{t \in (\sigma_g, 2]} |t - \sigma_g| \left\{ \left| \frac{\lambda(t) - \lambda(\sigma_g)}{t - \sigma_g} - 1 \right| + |\sigma_h - \sigma_g| \right\}$$

$$\leq 3 \left(e^{\gamma_{g,h}(\lambda)} - 1 + |\sigma_h - \sigma_g| \right). \quad (5.4)$$

(We have used that $\lambda(\sigma_g) = \sigma_h$ and the continuity of λ at σ_g.)

Let $(g_n) \subseteq G$ and $g \in G$. If $d(g_n, g) \to 0$ then it follows readily from (5.4) and the definitions that $d'(g_n, g) \to 0$. It remains to prove the converse; suppose instead that $d'(g_n, g) \to 0$.

Fix $N \in \mathbb{N}$. Since $d'(g_n, g) \to 0$ there exists a sequence $\lambda_n^N \in \Lambda[g_n, g]$ such that

$$\gamma'_{g_n, g}(\lambda_n^N) \vee d(g_n, g, \lambda_n^N) \vee |\sigma_{g_n} - \sigma_g| \to 0 \quad (5.5)$$

as $n \to \infty$. Define $\tau_0^N = \sigma_g$ and for $k = 1, 2, \ldots$ define

$$\tau_k^N = 2 \vee \inf \left\{ t > \tau_{k-1}^N \,;\, |g(t) - g(\tau_k^N)| > \frac{1}{N} \right\} \quad (5.6)$$

up until the first $k = k_N$ for which $\tau_k^N = 2$. Since g is càdlàg, $(\tau_k^N)_{k=0}^{k_N}$ is a finite, strictly increasing sequence, and $\tau_{k_N}^N = 2$.

For each $n \in \mathbb{N}$, define μ_n^N to be the unique piecewise linear function for which

$$\mu_n^N(\tau_k^N) = (\lambda_n^N)^{-1}(\tau_k^N) \quad (5.7)$$

for all $k = 0, \ldots, k_N$ (and is linear in between those points). Then, $\mu_n^N \in \Lambda'[g, g_n]$ and, moreover,

$$\gamma_{g,g_n}(\mu_n^N) = \sup_{k=1,\ldots,k_N} \left| \log \frac{(\lambda_n^N)^{-1}(\tau_k^N) - (\lambda_n^N)^{-1}(\tau_{k-1}^N)}{\tau_k^N - \tau_{k-1}^N} \right| < \infty$$

so that $\mu_n^N \in \Lambda[g, g_n]$. In fact, since (5.5) implies that $\lim_{n\to\infty} \lambda_n^N(\tau_k^N) = \tau_k^N$, we have $\gamma_{g,g_n}(\mu_n^N) \to 0$ as $n \to \infty$. Further,

$$\sup_{t\in[\sigma_{g_n},2]} \left|g_n(t) - g\left((\mu_n^N)^{-1}(t)\right)\right|$$

$$\leq \sup_{t\in[\sigma_{g_n},2]} \left|g_n(t) - g\left(\lambda_n^N(t)\right)\right| + \sup_{t\in[\sigma_{g_n},2]} \left|g\left(\lambda_n^N(t)\right) - g\left((\mu_n^N)^{-1}(t)\right)\right|$$

$$\leq d(g_n, g, \lambda_n^N) + \sup_{t\in[\sigma_g,2]} \left|g\left(\lambda_n^N \circ \mu_n^N(t)\right) - g(t)\right|$$

$$\leq d(g_n, g, \lambda_n^N) + \frac{2}{N}.$$

Here, the final line follows from (5.6) and (5.7). Hence, recalling that $d(g, g_n, \mu_n^N) = d(g, g_n, (\mu_n^N)^{-1})$, we have $d(g, g_n, \mu_n^N) \to 0$ as $n \to \infty$.

Combining the above with (5.5), we can choose a strictly increasing sequence $(n_N)_{N\in\mathbb{N}}$ of natural numbers such that, for all $n \geq n_N$,

$$\gamma_{g,g_n}(\mu_n^N) \leq \frac{1}{N}, \qquad d(g, g_n, \mu_n^N) \leq \frac{3}{N}, \qquad |\sigma_{g_n} - \sigma_g| \leq \frac{1}{N}.$$

Define $\kappa_n = \mu_n^N$ for all $n \in \mathbb{N}$ such that $n_N \leq n < n_{N+1}$. Then $d(g_n, g) \leq \gamma_{g,g_n}(\kappa_n) \vee d(g, g_n, \kappa_n) \vee |\sigma_{g_n} - \sigma_g|$ so $d(g_n, g) \to 0$ as $n \to \infty$. □

The space (G, d') is not complete (to see this, note first that by Remark 5.1 and Example 12.2 of [5], even the space $(G[s], d')$ is not complete). In order to prove completeness of (G, d), it will be useful to note that there exists $\epsilon^\star > 0$ such that for all $x \in [0, \epsilon^\star)$, we have

$$e^x - 1 \leq 2x. \tag{5.8}$$

Lemma 5.5. *The space (G, d) is complete.*

Proof. It suffices to show that any Cauchy sequence in (G, d) has a convergent subsequence. To this end, let (g_k) be a Cauchy sequence in (G, d). Thus σ_{g_k} is Cauchy, which implies that $\sigma_{g_k} \to \alpha$ for some $\alpha \in [-1, 1]$.

With mild abuse of notation, we pass to a subsequence (g_k) such that for all $j \geq k$ we have $d(g_k, g_j) \leq 2^{-k} e^{-k-1}$. Hence, there exists $\lambda_k \in \Lambda[g_k, g_{k+1}]$ such that, for all k,

$$\gamma_{g_k,g_{k+1}}(\lambda_k) \vee d(g_k, g_{k+1}, \lambda_k) \vee |\sigma_{g_k} - \alpha| \leq 2^{-k} \wedge \epsilon^\star. \tag{5.9}$$

For each k, define $\widetilde{\lambda}_k : [-1, 2] \to [-1, 2]$ to be the function that is equal to λ_k on $[\sigma_{g_k}, 2]$ and, if $\sigma_{g_k} > -1$, is linear in between $\widetilde{\lambda}_k(-1) = -1$ and $\widetilde{\lambda}_k(\sigma_{g_k}) = \sigma_{g_{k+1}}$. Note that this means $\widetilde{\lambda}_k$ has constant gradient on $[-1, \sigma_{g_k}]$, and that $\widetilde{\lambda}_k$ is a continuous bijection of $[-1, 2]$ to itself. Thus, we have

$$\sup_{-1\leq t\leq 2} \left|\widetilde{\lambda}_k(t) - t\right| \leq \sup_{\sigma_{g_k}\leq t\leq 2} |\lambda_k(t) - t|$$

$$\leq 3\left(e^{\gamma_{g_k}(\lambda_k)} - 1 + \left|\sigma_{g_{k+1}} - \sigma_{g_k}\right|\right)$$

$$\leq 3(2^{-k+1} + 2^{-k}) = 9 \cdot 2^{-k}. \tag{5.10}$$

Here, the second line follows from (5.4), and the final line from (5.8) and (5.9).

We now construct the limit of (g_k). Define $\mu_k^n : [\sigma_{g_k}, 2] \to [\sigma_{g_{n+k}}, 2]$ and $\widetilde{\mu}_k^n : [-1, 2] \to [-1, 2]$ by $\mu_k^n = \lambda_{k+n} \circ \ldots \circ \lambda_{k+1} \circ \lambda_k$, and $\widetilde{\mu}_k^n = \widetilde{\lambda}_{k+n} \circ \ldots \circ \widetilde{\lambda}_{k+1} \circ \widetilde{\lambda}_k$. By (5.10) we have

$$\sup_{t\in[-1,2]} |\widetilde{\mu}_k^{n+1}(t) - \widetilde{\mu}_k^n(t)| \leq \sup_{t\in[-1,2]} |\widetilde{\lambda}_{k+n+1}(t) - t| \leq 9 \cdot 2^{-k-n}.$$

It follows that $(\widetilde{\mu}_k^n)_{n=1}^\infty$ is a Cauchy sequence in $(C_{[-1,2]}[-1,1], ||\cdot||_\infty)$, and hence has a limit, which we denote by $\widetilde{\mu}_k$. Since the $\widetilde{\mu}_k^n$ are increasing, it is immediate that $\widetilde{\mu}_k(s) \geq \widetilde{\mu}_k(t)$ for $s \geq t$.

We define $\mu_k : [\sigma_{g_k}, 2] \to [-1, 1]$ by $\mu_k(t) = \widetilde{\mu}_k(t)$. Note that

$$\mu_k(\sigma_{g_k}) = \lim_{n\to\infty} \lambda_{k+n} \circ \ldots \circ \lambda_k(\sigma_{g_k}) = \lim_{n\to\infty} \sigma_{g_{k+n+1}} = \alpha, \tag{5.11}$$

$$\mu_k(2) = \lim_{n\to\infty} \lambda_{k+n} \circ \ldots \circ \lambda_k(2) = 2, \tag{5.12}$$

and also that, from (5.10),

$$|\mu_k(1) - 1| = \lim_{n\to\infty} |\mu_k^n(1) - 1| \leq \lim_{n\to\infty} \sum_{j=k}^{k+n} \sup_{t\in[\sigma_{g_k},2]} |\lambda_j(t) - t| \leq 9 \cdot 2^{-k+2}. \tag{5.13}$$

In similar style to (5.2),

$$\begin{aligned}
\sup_{\sigma_{g_k} \leq s < t \leq 2} \left| \log \frac{\mu_k^n(s) - \mu_k^n(t)}{s - t} \right| &\leq \sum_{j=k}^{k+n} \sup_{\sigma_{g_j} \leq s < t \leq 2} \left| \log \frac{\lambda_j(s) - \lambda_j(t)}{s - t} \right| \\
&= \sum_{j=k}^{k+n} \gamma_{g_j, g_{j+1}}(\lambda_j) \\
&\leq 2^{-k+1}.
\end{aligned} \tag{5.14}$$

Here, to deduce the final line we use (5.9). Letting $n \to \infty$ we have

$$\sup_{\sigma_{g_k} \leq s < t \leq 2} \left| \log \frac{\mu_k(s) - \mu_k(t)}{s - t} \right| \leq 2^{-k+1}. \tag{5.15}$$

Consequently, μ_k is strictly increasing. Thus from (5.11) and (5.12), we have that $\mu_k : [\sigma_{g_k}, 2] \to [\alpha, 2]$ is a strictly increasing bijection. In particular, it has an inverse $\mu_k^{-1} : [\alpha, 2] \to [\sigma_{g_k}, 2]$.

The proof now follows the usual strategy. For each k, we define $z_k : [-1, 2] \to [-1, 2]$ by

$$z_k(t) = \begin{cases} g_k \circ \mu_k^{-1}(t) & t \in [\alpha, 2] \\ 0 & t \in [-1, \alpha). \end{cases}$$

We have

$$\begin{aligned}
\sup_{t\in[-1,2]} |z_k(t) - z_{k+1}(t)| &= \sup_{t\in[-1,2]} \left| g_k\left(\mu_k^{-1}(t)\right) - g_{k+1}\left(\lambda_k(\mu_k^{-1}(t))\right) \right| \\
&= \sup_{\sigma_{g_k} \leq t \leq 2} \left| g_k(t) - g_{k+1}\left(\lambda_k(t)\right) \right| \\
&= d(g_k, g_{k+1}, \lambda_k) \\
&\leq 2^{-k}
\end{aligned}$$

Here, the second line follows by definition of μ_k and the final line follows by (5.9). Thus, by completeness of $\mathbb{R}$, there exists a function $z : [-1, 2] \to [-1, 1]$ such that

$$\sup_{t\in[-1,2]} |z_k(t) - z(t)| \to 0 \tag{5.16}$$

as $k \to \infty$. Since each z_k is càdlàg, z is càdlàg. By (5.13) and the fact that each g_k is constant on $[1, 2]$, it follows from (5.16) that z is constant on $(1, 2)$, hence by right continuity z is constant on $[1, 2]$.

We define g by setting $\sigma_g = \alpha$ and $g(t) = z(t)$ on $[\sigma_g, 2]$, and note that $g \in G$. We have shown above that μ_k is a strictly increasing bijection, hence $\mu_k \in \Lambda[g_k, g]$. Thus, from (5.16) we have

$$\sup_{t \in [\sigma_{g_k}, 2]} |g_k(t) - g(\mu_k(t))| = \sup_{t \in [\sigma_g, 2]} |g_k(\mu_k^{-1}(t)) - g(t)| \to 0 \quad \text{as} \ \ k \to \infty.$$

Combining the above equation with (5.15), $\gamma_{g_k, g}(\mu_k) \vee d(g_k, g, \mu_k) \to 0$ and since $\sigma_g = \lim_{k \to \infty} \sigma_{g_k}$ we have $g_k \to g$ in (G, d). $\qquad \square$

5.2 The space (M, d_M)

Recall the space M from (3.1) and the notation $\kappa_t = \tanh^{-1}(t)$. It will sometimes be useful to write $\kappa(t) = \kappa_t$. Each $f \in M$ corresponds to some $\bar{f} \in G$, essentially through the relation (3.2), that is

$$\bar{f}(t) = \frac{\tanh(f(\kappa_t))}{1 + |\kappa_t|}, \tag{5.17}$$

for $t \in [\kappa^{-1}(\sigma_f), 1]$. In order to treat $\bar{f}$ as an element of G we specify that additionally $\bar{f}(t) = 0$ for all $t \in [1, 2]$. Note that $\sigma_{\bar{f}} = \kappa^{-1}(\sigma_f)$. In this section will use notation from Section 5.1 without comment.

The map $f \mapsto \bar{f}$ naturally induces a pseudometric on M through the relation

$$d_M(f_1, f_2) = d(\bar{f}_1, \bar{f}_2). \tag{5.18}$$

It follows immediately from Lemmas 5.2 and 5.3 that the set of equivalence classes of M, under d_M, form a separable metric space. Note that it is necessary to use equivalence classes, since all $f \in D[\infty]$ map to the same $\bar{f} \in G$. From now on we abuse notation slightly and write (M, d_M) for the metric space of equivalence classes. This defines the metric d_M that appeared in (3.3).

Lemma 5.6. *The space* (M, d_M) *is complete.*

Proof. Let (f_k) be a Cauchy sequence in (M, d_M). Then $(\bar{f}_k)$ is a Cauchy sequence in (G, d) and by Lemma 5.5 there exists $g \in G$ such that $\bar{f}_k \to g$. It remains to show that there exists $f \in M$ such that $\bar{f} = g$, which will in turn follow immediately from (5.17) if we can show that

$$|g(t)| \leq \frac{1}{1 + |\kappa_t|} \tag{5.19}$$

for all $t \in [\sigma_g, 1]$.

Equation (5.19) is readily seen; note that, by Lemma 5.4, $\bar{f}_k \to g$ implies that there exists $\lambda_k \in \Lambda[g, \bar{f}_k]$ such that $\gamma'_{g, \bar{f}_k}(\lambda_k) \vee d(g, f_k, \lambda_k) \vee |\sigma_g - \sigma_{\bar{f}_k}| \to 0$. Therefore, $g(t) = \lim_{k \to \infty} \bar{f}_k(\lambda_k(t))$. By (5.17) we have $|\bar{f}_k(s)| \leq \frac{1}{1 + |\kappa_s|}$ for all s, hence

$$|g(t)| \leq \limsup_{k \to \infty} \frac{1}{1 + |\kappa(\lambda_k(t))|}. \tag{5.20}$$

We also have $\gamma'_{g, \bar{f}_k}(\lambda_k) = \sup_{t \in [\sigma_g, 1]} |\lambda_k(t) - t| \to 0$. Combining this with (5.20) proves (5.19). $\qquad \square$

Remark 5.7. Note that $d'_M(f_1, f_2) = d'(\bar{f}_1, \bar{f}_2)$ is a pseudo-metric on M, with the same equivalence classes as d_M. Hence, by Lemma 5.4, d'_M generates the same topology on M as d_M.

If we look at the subset $\widetilde{M}$ of M consisting only of continuous functions, with the metric

$$d_{\widetilde{M}}(f_1, f_2) = |\sigma_{\bar{f}_1} - \sigma_{\bar{f}_2}| \vee \sup_{t \in [-1,1]} |\bar{f}_1(t \vee \sigma_{f_1}) - \bar{f}_2(t \vee \sigma_{f_2})| \tag{5.21}$$

then we recover the space of continuous paths introduced by [15] (with a minor modification relating to the values of functions at $\{-\infty, \infty\}$, see the appendix of [24] for details).

We now establish the natural relationship between M and $\widetilde{M}$, which mirrors the 'usual' continuous embedding of spaces of continuous paths (with the $\|\cdot\|_\infty$ metric) into Skorohod spaces. Recall that $\mathcal{K}(M)$ (resp. $\mathcal{K}(\widetilde{M})$) denotes the space of all compact subsets of M (resp. $\widetilde{M}$).

Lemma 5.8. *The space $\widetilde{M}$ is continuously embedded in M. Moreover, $\mathcal{K}(\widetilde{M})$ is continuously embedded in $\mathcal{K}(M)$.*

Proof. Note that the first statement follows immediately from the second, so we will prove only the second statement. Recall that the topology generated by the Hausdorff metric (on $\mathcal{K}(M)$) depends only on the underlying topology (of M), and not on the underlying metric. In view of this fact and Remark 5.7, for the duration of this proof we take the Hausdorff metric on $\mathcal{K}(M)$ as that generated by (M, d'_M).

Let W_n, W be subsets of $\mathcal{K}(\widetilde{M})$ such that $W_n \to W$ in $\mathcal{K}(\widetilde{M})$. By Lemma A.1 the set $\mathscr{W} = W \cup \left(\bigcup_{n \in \mathbb{N}} W_n \right)$ is a compact subset of $\widetilde{M}$. A characterization of relative compactness in $\widetilde{M}$ is given in the proof of Lemma 4.6 of [24], based on the Ascoli-Arzela Theorem (or, for a more detailed treatment, see the appendix of [21]). It follows immediately from this characterization that the set $\bar{\mathscr{W}} = \{\bar{f} ; f \in \mathscr{W}\}$ is equicontinuous.

Let $\epsilon > 0$. By equicontinuity, there exists $\delta > 0$ such that $|s - t| \le \delta$ implies

$$\sup_{\bar{f} \in \bar{\mathscr{W}}} |\bar{f}(s) - \bar{f}(t)| \le \epsilon. \tag{5.22}$$

Without loss of generality we may choose $\delta \in (0, \epsilon)$. By definition of the Hausdorff metric, choose N such that for all $n \ge N$,

$$\sup_{g \in W_n} \inf_{h \in W} d_{\widetilde{M}}(g, h) \le \delta \quad \text{and} \quad \sup_{g \in W} \inf_{h \in W_n} d_{\widetilde{M}}(g, h) \le \delta. \tag{5.23}$$

By the first equality of (5.23), for any $g \in W_n$ and $n \ge N$, there exists $h \in W$ such that

$$|\sigma_{\bar{g}} - \sigma_{\bar{h}}| \le \delta \quad \text{and} \quad \sup_{t \in [\sigma_{\bar{g}} \vee \sigma_{\bar{h}}, 2]} |g(t) - h(t)| \le \delta. \tag{5.24}$$

Define $\lambda_{\bar{g}} : [\sigma_{\bar{g}}, 2] \to [\sigma_{\bar{h}}, 2]$ by setting $\lambda(\sigma_{\bar{g}}) = \sigma_{\bar{h}}$, $\lambda(1) = 1$, $\lambda(2) = 2$ and linear in between. Thus $\lambda_g \in \Lambda'[\bar{g}, \bar{h}]$. For $t \in [\sigma_{\bar{g}}, 2]$, we have $|t - \lambda_g(t)| \le |\sigma_{\bar{g}} - \sigma_{\bar{h}}| \le \delta$. This implies that $\gamma'_{\bar{g}}(\lambda_{\bar{g}}) \le \epsilon$ and, using (5.22) and (5.24), that

$$|\bar{g}(t) - \bar{h}(\lambda_{\bar{g}}(t))| \le |\bar{g}(t) - \bar{g}(\lambda_{\bar{g}}(t))| + |\bar{g}(\lambda_{\bar{g}}(t)) - \bar{h}(\lambda_{\bar{g}}(t))| \le 2\epsilon.$$

Thus, $d'(\bar{g}, \bar{h}) \le 2\epsilon$. Similarly, using the second equality of (5.23), for any $g \in W$ and $n \ge N$, there exists $h \in W_n$ such that $d'(\bar{g}, \bar{h}) \le 2\epsilon$. We thus have, for all $n \ge N$,

$$\max \left(\sup_{g \in W_n} \inf_{h \in W} d'(\bar{g}, \bar{h}), \sup_{g \in W} \inf_{h \in W_n} d'(\bar{g}, \bar{h}) \right) \le 2\epsilon. \tag{5.25}$$

Hence, $W_n \to W$ in $\mathcal{K}(M)$ as $n \to \infty$. $\qquad\square$

In the interests of brevity, we limit our further development of the space (M, d_M) to the following two results.

Lemma 5.9. *Let $f, g \in M$ with $\sigma_f = \sigma_g$ and $\sup_{t \in [\sigma_f, \infty]} |f(t) - g(t)| \le r$. Then $d'_M(f, g) \le r$.*

Proof. Since $\sigma_f = \sigma_g$, the identity function ι is an element of $\Lambda[\bar{f}, \bar{g}]$. Note that $\gamma_{\bar{f}}(\iota) = 0$. Hence $d'_M(f, g) \le \sup_{t \in [\sigma_{\bar{f}}, \infty]} |\bar{f}(t) - \bar{g}(t)| \le \sup_{t \in [\sigma_{\bar{f}}, 2]} |f(t) - g(t)| \le r$, as required. $\square$

Lemma 5.10. *Let $(f_m) \subseteq M$ and $f \in M$ with $\sigma_{f_m} = \sigma_f \in (-\infty, \infty)$. Then $d_M(f_m, f) \to 0$ if (the restrictions of) $f_m \to f$ in $D_{[\sigma_f, T]}(\mathbb{R})$ for all $T \in (\sigma_f, \infty)$.*

Proof. Note that $\tanh : [-\infty, \infty] \to [-1, 1]$ is a contraction. For $T \in (0, \infty)$, set $\bar{T} = \tanh(T)$, and note that $\tanh$ restricted to $[-T, T] \to [-\tanh(T), \tanh(T)]$ is bi-Lipschitz. Hence there are constants $C_T \in (0, \infty)$ such that

$$\sup_{-\bar{T} \le t < s \le \bar{T}} \left| \frac{\kappa_s - \kappa_t}{s - t} \right| \le C_T, \qquad \sup_{-\infty \le s < t \le \infty} \left| \frac{s - t}{\kappa_s - \kappa_t} \right| \le 1. \tag{5.26}$$

Let $\epsilon > 0$. Let $T \in (\sigma_f \vee 0, \infty)$ be such that $\frac{1}{1+T} \le \epsilon$. Thus,

$$\sup_{g \in M} \sup_{t \in [\bar{T} \vee \sigma_g, 2]} |\bar{g}(t)| \le \epsilon. \tag{5.27}$$

By Theorem 12.1 of [5] there exists $M \in \mathbb{N}$ such that for all $m \ge M$ there exists a continuous strictly increasing $\lambda_m : [\sigma_f, T] \to [\sigma_f, T]$ with

$$\sup_{t \in [\sigma_f, T]} |t - \lambda_m(t)| \le \frac{\epsilon}{C_T}, \qquad \sup_{t \in [\sigma_f, T]} |f(t) - f_m(\lambda_m(t))| \le \epsilon. \tag{5.28}$$

Define $\bar{\lambda}_m : [\sigma_{\bar{f}}, \bar{T}] \to [\sigma_{\bar{f}}, \bar{T}]$ by $\bar{\lambda}_m(t) = \kappa^{-1} \circ \lambda_m \circ \kappa(t)$, and note that by (5.28) for all $t \in [\sigma_{\bar{f}}, \bar{T}]$,

$$|t - \bar{\lambda}_m(t)| \le C_T |t - \lambda_m(t)| \le \epsilon, \tag{5.29}$$

and, by the right hand side of (5.26),

$$|\bar{f}(t) - \bar{f}_m(\bar{\lambda}_m(t))| = \frac{1}{1 + |\kappa_t|} |\tanh(f(\kappa_t)) - \tanh(f_m(\kappa(\bar{\lambda}_m(t))))|$$

$$\le |f(\kappa_t) - f_m(\lambda_m(\kappa_t))| \le \epsilon \tag{5.30}$$

Extend $\bar{\lambda}_m$ to $\eta_m : [\sigma_{\bar{f}}, 2] \to [\sigma_{\bar{f}}, 2]$ by setting $\bar{\lambda}_m(t) = t$ for $t \ge \bar{T}$. Then, combining (5.27), (5.29) and (5.30) we obtain $\gamma'_{\bar{f}}(\bar{\lambda}_m) \le \epsilon$ and $d'(\bar{f}, \bar{f}_m, \bar{\lambda}_m) \le 2\epsilon$. It follows that $d'(\bar{f}, \bar{f}_m) \to 0$, so the stated result now follows by Lemma 5.4. $\square$

6 Convergence to the Brownian net

6.1 Compactness

In order to use Theorem 3.4, we must verify that our various set of paths really are subsets of $\mathcal{K}(M)$. That is, we need to show that they are compact subsets of $\widetilde{M}$, which is the content of this subsection. We concentrate on forwards paths; analogous arguments apply to backwards paths.

We require three preparatory lemmas. The first two of these embody the key features of the argument; at any given time, within bounded intervals of space, the number of ancestral lines at distinct spatial locations is finite, and there do not exist ancestral lines that move arbitrarily fast across space.

Lemma 6.1. *Let $n \in \mathbb{N}$. Then, almost surely, for all (random) $a, b, t \in \mathbb{R}$ the set $E_{a,b,t} = [a, b] \cap \{f(t) \, ; \, f \in \widetilde{\mathcal{P}}_n^\uparrow(\mathcal{D}_n)\}$ is finite.*

Proof. For fixed deterministic a, b, t, it is easily seen that the set of (y, s) for which $s \leq t$, $y \in [a, b]$ and the line $\{y\} \times [s, t]$ is not affected by any reproduction events, is almost surely bounded (in $\mathbb{R}^2$). Thus, since $\mathcal{D}_n$ is locally finite, the set $E_{a,b,t}$ is almost surely finite. Take countably dense (deterministic) sequences $(a_m), (b_m)$ and (t_m) in $\mathbb{R}$; thus almost surely, for all $m_1, m_2, m_3 \in \mathbb{N}$ for which $a_{m_1} < b_{m_2}$, the set $E_{a_{m_1}, b_{m_2}, t_{m_3}}$ is finite. The stated result now follows from Lemma 4.1. □

Lemma 6.2. *Let* $n \in \mathbb{N}$. *Almost surely, there does not exist a (random) sequence* $(x_m, t_m, r_m)_{m=1}^\infty \subseteq \Pi_n$ *such that* $\sup_m |t_m| < \infty$, $\lim_{m \to \infty} x_m = \infty$ *and* $\sup_m |x_{m+1} - x_m| \leq 4\mathcal{R}_n$.

Proof. Let $K \in (0, \infty)$. Then, the probability that

$$\Pi_n \cap \Big([4k\mathcal{R}_n, 4(k+1)\mathcal{R}_n] \times [-K, K] \times [0, \mathcal{R}_n] \Big) = \emptyset$$

is positive and does not depend on k. Consequently, the probability that there exists a sequence $(x_m, t_m, r_m)_{m=1}^\infty \subseteq \Pi_n$ such that $\lim_{m \to \infty} x_m = \infty$ and $\sup_m |x_{m+1} - x_m| \leq 4\mathcal{R}_n$, with $\sup_m |t_m| \leq K$, is zero. Since K was arbitrary, the result follows. □

Recall that our ultimate goal is to prove Theorem 3.5, which claims convergence in distribution. In view of this, from now on we will (abuse notation slightly and) assume that the conclusions of Lemma 6.2 and 6.1 hold *surely*.

The next lemma asserts that any convergent sequence of paths that becomes close, in space, to touching ∞ within some bounded interval of time, must converge to a constant path at ∞.

Lemma 6.3. *Let* $n \in \mathbb{N}$. *Let* $(f_m)_{m=1}^\infty \subseteq \mathcal{P}_n^\uparrow(\mathcal{D}_n)$ *be a sequence of paths and suppose* $\sigma(f_m)$ *converges to* $v \in [-\infty, \infty]$. *Suppose also that there exists a bounded sequence* (t_m) *with* $t_m \geq \sigma(f_m)$ *for which* $f_m(t_m) \to \infty$ *as* $m \to \infty$.

Let $f_\infty \in M$ *be the path defined by* $\sigma(f_\infty) = v$ *and* $f_\infty(s) = \infty$ *for all* $s \in [v, \infty]$. *Then* $f_m \to f_\infty$ *in* M.

Moreover, if instead $(f_m) \subseteq \widetilde{\mathcal{P}}_n^\uparrow(\mathcal{D}_n)$, *then under the same hypothesis* $f_m \to f_\infty$ *in* $\widetilde{M}$.

Proof. Fix $K \in (0, \infty)$, large enough that $\sup_m |t_m| \leq K$, and define $x(m^*, K) = \inf\{f_m(s) \,;\, m \geq m^*, \sigma(f_m) \leq s, |s| \leq K\}$.

Suppose, for a contradiction, that $x(m^*, K)$ does not tend to ∞ as $m^* \to \infty$. Then, there exists $X \in (-\infty, \infty)$ and infinitely many m^* for which $x(m^*, K) \leq X$. For all such m^* we have some $m \geq m^*$ and $|s| \leq K$ such that $f_m(s) \leq X$, and (by our hypothesis) as $m^* \to \infty$ we have also $f_m(t_m) \to \infty$; since $f_m \in \mathcal{P}^\uparrow(\mathcal{D}_n)$ this is a contradiction to Lemma 6.2.

So, $x(m^*, K) \to \infty$ as $m^* \to \infty$. Thus, for any $K, X \in (-\infty, \infty)$ we can find $m^* \in \mathbb{N}$ such that, for all $m \geq m^*$ and $s \geq \sigma(f_m)$ such that $|s| \leq K$, we have $f_m(s) \geq X$. With this in hand, the stated results follow easily from (5.18) and (5.21). □

Recall that, by (4.1), the path f_∞ in the statement of Lemma 6.3 is an element of both $\mathcal{P}^\uparrow(\mathcal{D}_n)$ and $\widetilde{\mathcal{P}}^\uparrow(\mathcal{D}_n)$. Recall also that, in the notation of (4.1), both these sets also contain paths that are infinite extenders.

Lemma 6.4. *Let* $\star \in \{\uparrow, \downarrow\}$, $\dagger \in \{l, r\}$ *and* $n \in \mathbb{N}$. *Then,* $\mathcal{P}_n^\star(\mathcal{D}_n)$ *and* $\mathcal{P}_n^{\star,\dagger}(\mathcal{D}_n)$ *are compact subsets of* M, *also* $\widetilde{\mathcal{P}}_n^\star(\mathcal{D}_n)$ *and* $\widetilde{\mathcal{P}}_n^{\star,\dagger}(\mathcal{D}_n)$ *are compact subsets of* $\widetilde{M}$.

Proof. As usual, it suffices to consider the case of forward paths. Since Lemma 6.3, on which the following argument relies, holds in both M and $\widetilde{M}$, it will suffice to consider only cases in M. Moreover, the arguments required the case of $\mathcal{P}_n^{\uparrow,\dagger}(\mathcal{D}_n)$ are essentially

identical to those required for the case $\mathcal{P}_n^{\uparrow}(\mathcal{D}_n)$; thus, we aim to show that $\mathcal{P}_n^{\uparrow}(\mathcal{D}_n)$ is a sequentially compact subset of (the metric space) M.

Let $(f_m)_{m=1}^{\infty} \subseteq \mathcal{P}_n^{\uparrow}(\mathcal{D}_n)$ be a sequence of paths. We must show that (f_m) has a convergent subsequence, with limit in $\mathcal{P}_n^{\uparrow}(\mathcal{D}_n)$.

We now split into several cases.

1. If $(\sigma(f_m))_{m \geq 1}$ has a subsequence that converges to ∞ then, along this subsequence, f_m converges to the degenerate path f with $\sigma_f = \infty$ and $f(\infty) = 0$.

2. If $(\sigma(f_m))_{m \geq 1}$ has a bounded subsequence, then consider the sequence $x_m = f_m(\sigma_{f_m})$.

   (a) If $(x_m)_{m \geq 1}$ is bounded, then since $\mathcal{D}_n$ is locally finite there must be a subsequence along which $(\sigma(f_m), x_m)$ is eventually constant. Any given ancestral line moves to one of at most two locations in a reproduction event, thus $(f_m)_{m \geq 1}$ has a convergent subsequence; to construct the limit path we successively follow parent points that were followed by infinitely many of our f_m.

   (b) If $(x_m)_{m \geq 1}$ is not bounded, then without loss of generality we pass to a subsequence and assume that both $x_m \to \infty$ and $\sigma(f_m)$ converges. It then follows immediately from Lemma 6.3 that f_m converges along this subsequence.

3. If $(\sigma(f_m))_{m \geq 1}$ has a subsequence that converges to $-\infty$, then pass to that subsequence and set $t = \sup_m \sigma(f_m) < \infty$.

   (a) If $\{m \, ; \, |f_m(t)| \leq K\}$ is finite for all $K < \infty$, then essentially the same argument as in 2(b), reliant on Lemma 6.3, shows that f_m has a convergent subsequence.

   (b) If, for some $K < \infty$ the set $\{m \, ; \, |f_m(t)| \leq K\}$ is infinite, then pass to the subsequence of f_m such that $|f_m(t)| \leq K$. By Lemma 6.1, the set $[-K, K] \cap \{f_m(t) \, ; \, m \in \mathbb{N}\}$ is finite. Hence, there is some $|z| \leq K$ through which infinitely many f_m pass. Using the same method as in *2(a)*, we can construct a path $f : [t, \infty]$ with $f(t) = z$ that is followed by infinitely many f_m. So, pass to a further subsequence and assume $f_m(s) = f(s)$ for all $s \geq t$.

   We extend f backwards in time as follows. From location (z, t), look backwards in time until the most recent reproduction event (strictly) before t that affected z, say $p = (x, t', r)$. By Lemma 6.1, the set $\{f_m(t'-) \, ; \, p \text{ affects } f_m, m \in \mathbb{N}\}$ is finite. Pick some element z' of this set, and restrict to f_m for which $f_m(t'-) = z'$. Set $f(s) = z$ for $s \in [t', t)$ set $f(s) = z$. Then, look back from (z', t'), and repeat (in the language of (4.1), f is an 'infinite extender'). Thus, a subsequence of $(f_m)_{m \geq 1}$ converges to f.

Since $(\sigma(f_m))_{m \geq 1}$ must have a subsequence that converges in $[-\infty, \infty]$, at least one of the above cases occurs. This completes the proof. $\square$

6.2 Convergence of multiple left/right paths

We now extend Proposition 4.9 to larger collections of left and right-most paths. Let $N \in \mathbb{N}$. Given a finite set $D = \{(y_i, s_i) \in \mathbb{R}^2 \, ; \, i = 1, \dots, N\}$ of distinct points in $\mathbb{R}^2$ and a function $O : \{1, \dots, N\} \to \{l, r\}$, [24], Section 2.2, construct a system of left-right coalescing Brownian motions started from the points of D. (Recall that two left-most paths coalesce on meeting, as do two right-most paths, and that (3.4) describes the interaction between left-most and right-most paths.) We write

$$\mathcal{P}^{\uparrow}(D, O) = \{B^{\uparrow, (y_i, s_i)} \, ; \, i = 1, \dots, N\} \tag{6.1}$$

for this system, where $B^{\uparrow,(y_i,s_i)}$ denotes the path of a Brownian motion started from (y_i, s_i) with diffusion constant ξ^2 and drift ζ to the left if $O(i) = l$ and to the right if $O(i) = r$.

For each $i = 1, \ldots, N$ let $d_i^n = (y_i^n, s_i^n) \in \mathbb{R}^2$ be such that $d_i^n \to d_i = (y_i, s_i) \in \mathbb{R}^2$. Set $D^{(n)} = \{d_i^n \, ; \, i = 1, \ldots, N\}$. We define the set $\mathcal{P}_n^\uparrow(D^{(n)}, O) = \{f_1^\uparrow, \ldots, f_N^\uparrow\}$ where $f_i^\uparrow$ is the $O(i)$-most forward path from d_i^n driven by events in Π^ns and $\widetilde{\mathcal{P}}_n^\uparrow(D^{(n)}, O)$ for the corresponding space of interpolated paths. Note that both these sets are random elements of the product space M^N.

Lemma 6.5. *Let $N \in \mathbb{N}$ and let $D^{(n)}, D, O$ be as above. Then, as $n \to \infty$, $\mathcal{P}_n^\uparrow(D^{(n)}, O)$ converges weakly in M^N to $\mathcal{P}^\uparrow(D, O)$ and $\widetilde{\mathcal{P}}_n^\uparrow(D^{(n)}, O)$ converges weakly in $\widetilde{M}^N$ to $\mathcal{P}^\uparrow(D, O)$.*

Proof. The argument is essentially identical to that of the proof of Proposition 5.2 of [24]. The construction of $\mathcal{P}^\uparrow(D, O)$ in Section 2.2 of [24] is an inductive construction that views $\mathcal{P}^\uparrow(D, O)$ as made up of several independent pieces consisting of segments of either single left-most paths, single right-most paths or a pair of left/right paths. The same inductive construction breaks down $\mathcal{P}_n^\uparrow(D^{(n)}, O)$ into corresponding pieces. The stopping times used in this construction are continuous functionals on M^N with respect to the law of independent evolutions of paths within each such piece, so the first part of the lemma follows from Proposition 4.9 and Lemma 5.10. Similarly, $\widetilde{\mathcal{P}}_n^\uparrow(D^{(n)}, O)$ converges weakly to $\widetilde{\mathcal{P}}^\uparrow(D, O)$. □

6.3 Proof of Theorem 3.5

We complete the proof of Theorem 3.5 in three steps. Recall that the Brownian net is denoted by $\mathcal{N}$, and recall the function $\mathcal{H}_{cross}$ defined in Section 3.2.

Lemma 6.6. *As $n \to \infty$, we have that*

$$\mathcal{H}_{cross}\left(\widetilde{\mathcal{P}}_n^{\uparrow,l}(\mathcal{D}_n) \cup \widetilde{\mathcal{P}}_n^{\uparrow,r}(\mathcal{D}_n)\right) \to \mathcal{N},$$

in distribution in $\mathcal{K}(\widetilde{M})$.

Proof. We verify the conditions $(\mathscr{A})$-$(\mathscr{D})$ of Theorem 3.4. This theorem is applied with $X_n^\dagger = \widetilde{\mathcal{P}}_n^{\uparrow,\dagger}(\mathcal{D}_n)$ and $\hat{X}_n^\dagger = \widetilde{\mathcal{P}}_n^{\downarrow,\dagger}(\mathcal{D}_n)$, where $\dagger \in \{l, r\}$. By Lemma 6.4, all these sets of paths are (almost surely) elements of $\mathcal{K}(\widetilde{M})$ (after rotation by 180 degrees about $(0,0)$ for the backwards paths). We define $X_n = \mathcal{H}_{cross}(\widetilde{\mathcal{P}}_n^{\uparrow,l}(\mathcal{D}_n) \cup \widetilde{\mathcal{P}}_n^{\uparrow,r}(\mathcal{D}_n))$ and similarly for $\hat{X}_n$.

We now check the conditions in turn. For $(\mathscr{A})$, the required statements about non-crossing paths are precisely the content of Lemma 4.4. For $(\mathscr{B})$, the required convergence of multiple left/right paths to left/right Brownian motions is precisely the content of Lemma 6.5.

We now move on to $(\mathscr{C})$. If $(x_{n,1}, x_{n,2}) \to (x_1, x_2)$ and $\hat{l}_n, \hat{r}_n$ are respectively elements of $\widetilde{\mathcal{P}}_n^{\downarrow,l}(\mathcal{D}_n), \widetilde{\mathcal{P}}_n^{\downarrow,r}(\mathcal{D}_n)$ started at $(x_{n,1}, x_{n,2})$, then it follows by combining Lemmas 4.10 and Remark 4.16 that $(\hat{l}_n, \hat{r}_n) \to (\hat{l}, \hat{r})$ in distribution, where $(\hat{l}, \hat{r})$ are a pair of left/right Brownian motions. That the first meeting time of $\hat{l}_n$ with $\hat{r}_n$ also converges (jointly) in distribution to the first meeting time of $\hat{l}$ with $\hat{r}$ follows from Remark 4.8.

It remains to verify $(\mathscr{D})$. By Lemma 4.4, left-most fowards and left-most backwards paths cannot cross, and similarly for right-most paths. Therefore, a path of X_n that enters a wedge W of $\hat{X}_n$ from the outside must enter through the southern-most point of the wedge (i.e. precisely where the two paths $\hat{r}$ and $\hat{l}$ defining W meet, in Figure 2). Fix a wedge W of $\hat{X}_n$ and denote this event by E_W.

For the event E_W to occur, some reproduction event must have a potential parent situated at the spatial location of the southern-most point of W. The distribution of the

spatial location of a pre-parent (in a reproduction event occurring at given time) has no atoms, and thus the distribution of the spatial location of the meeting point of $\hat{r}$ and $\hat{l}$ also has no atoms; hence almost surely E_W does not occur.

The set $\hat{X}_n$ contains countably many paths and thus has at most countably many wedges. Hence, almost surely the event E_W does not occur for any wedge W of $\hat{X}_n$. Without loss of generality, we may assume this does not occur *surely*, so as ($\mathscr{D}$) holds. $\square$

Lemma 6.7. *As $n \to \infty$, we have that $\widetilde{\mathcal{P}}_n^\uparrow(\mathcal{D}_n)$ tends in distribution to $\mathcal{N}$*

Proof. If a left-most and a right-most path of $\widetilde{\mathcal{P}}_n^\uparrow(\mathcal{D}_n)$ cross, then the point at which they cross much be within one of the sets $\mathscr{B}_{\Upsilon(p)}(p)$ (defined in Lemma 4.1) associated to a selective event $p \in \Pi_n$. If an interpolated arrow finishes at p, then by definition there are interpolated arrows ending at both potential parents of p. Consequently, there is a one to one correspondence, $f \mapsto f'$ between paths $f \in \mathcal{H}_{cross}\left(\widetilde{\mathcal{P}}_n^{\uparrow,l}(\mathcal{D}_n) \cup \widetilde{\mathcal{P}}_n^{\uparrow,r}(\mathcal{D}_n)\right)$ and paths $f' \in \widetilde{\mathcal{P}}^\uparrow(\mathcal{D}_n)$ such that

$$|f(t) - f'(t)| \leq 2\mathcal{R}_n \tag{6.2}$$

for all $t \geq \sigma_f = \sigma_{f'}$, and $f(\sigma_f) = f'(\sigma_{f'})$.

By Lemma 6.4, we have that $\widetilde{\mathcal{P}}_n^\uparrow(\mathcal{D}_n)$ is an element of $\mathcal{K}(M)$. Combined with (6.2) this means that $\widetilde{\mathcal{P}}_n^\uparrow(\mathcal{D}_n)$ has the same limit (in distribution) as $\mathcal{H}_{cross}(\widetilde{\mathcal{P}}_n^{\uparrow,l}(\mathcal{D}_n) \cup \widetilde{\mathcal{P}}_n^{\uparrow,r}(\mathcal{D}_n))$ as $n \to \infty$. Thus, from Lemma 6.6, $\widetilde{\mathcal{P}}_n^\uparrow(\mathcal{D}_n)$ tends in distribution to $\mathcal{N}$. $\square$

To finish, we must upgrade the result of Lemma 6.7 from $\mathcal{K}(\widetilde{M})$ to $\mathcal{K}(M)$, and use càdlàg paths in place of interpolated paths.

By Lemma 6.4 we have that $\widetilde{\mathcal{P}}_n^\uparrow(\mathcal{D}_n) \in \mathcal{K}(M)$ for all n. Combining this fact with Lemma 4.2, and noting that there is a one to one correspondence between paths and interpolated paths, we obtain from Lemma 6.7 that also $\mathcal{P}_n^\uparrow(\mathcal{D}_n) \to \mathcal{N}$ in distribution in $\mathcal{K}(M)$. This completes the proof of Theorem 3.5.

7 Simulations

In all the work described in Section 3.3, lineages coalesce instantly on meeting and, in particular, they cannot 'jump over' one another. Our result also requires this property, which is achieved by setting $\upsilon = 1$. In the absence of selection, it is shown in [4] that if instead we fix $\upsilon \in (0, 1)$, the scaling limit of the paths relating a finite sample from the population is a system of coalescing Brownian motions, but with 'clock' rate υ (so that the whole process is slowed down). In particular, we obtain a simple time-change of the limit for $\upsilon = 1$.

It is natural to ask what happens when $\upsilon < 1$ in the presence of selection. Our method of proof certainly breaks down. In particular, we can no longer trace the left/right most paths starting from a single point in isolation: because paths can now cross, the current left-most path may not be affected by an event, whereas another path in the same region is, and if the parent (or one of the potential parents) of the event is to the left of the current left-most path, a new line of descent takes over as left-most.

We have been unable to find a rigorous result in this context and so in this brief section, instead, we present the results of a numerical experiment.

In order to approximate the limiting process, we simulated the system of branching-coalescing lineages (which was introduced in Section 2.3). In particular, we were interested in $\mathscr{P}(\upsilon)$, the expected position of the right-most ancestor at time 1 of a single particle, which starts at the origin at time 0. By symmetry the same analysis applies to the left-most particle.

As a result of the discussion above, and the result of [4], it is natural to ask if $\mathscr{P}(\upsilon)$ varies linearly with υ. It seems that this is not the case, as is shown in the left hand plot
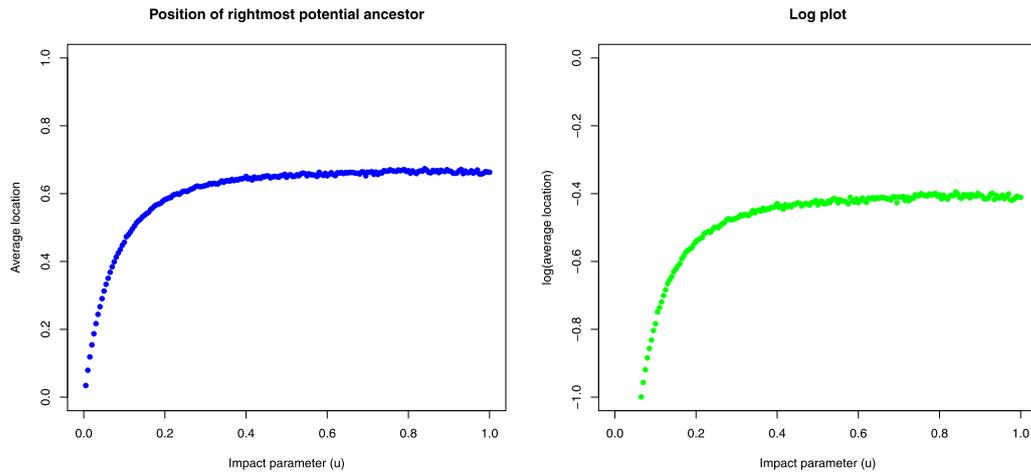
Figure 5: Right-most potential ancestor. The plot on the left shows $\mathscr{P}(u)$, the average location of the right most ancestor at time 1 of an individual located at the origin at time 0. Here we take $n = 1000$ and $\mu(dr) = \delta_1$. The plot contains 200 data points, each corresponding to a value of $v \in (0, 1]$, spaced evenly along the horizontal axis. Each data point is the mean of 2000 independent simulations with the corresponding value of v. The plot on the right is a logarithmic plot of the same data.

of Figure 5. Note that this does not remove the possibility that, as $n \to 0$, the SΛFVS rescales to a Brownian net, but it does imply that the speed of such a limiting net would not match the speed suggested by the simple time-change in the Brownian web limit of [4]. As can be seen from the right hand plot of Figure 5, $\mathscr{P}(v)$ is also not of the form $v^{-\alpha}$.

Simulating the SΛFVS when $v = 1$ is a quite different task from simulating it when v is close to 0. In the former case, it is more efficient to generate reproduction events by simulating the underlying Poisson Point Process, whereas in the latter case it is more efficient to track clusters of particles that are close enough to be affected by the same event and simulate their (correlated) motion directly. Our simulation employs both methods of event sampling and alternates between them based on which method is asymptotically more efficient for the given value of v.

In practice, for smaller values of v (i.e. closer to 0) larger values of n are needed to accurately simulate the SΛFVS. It seems very possible that, for v closer to 0, the numerics suggested by Figure 5 are not a reflection of the true behaviour as $n \to \infty$.

The C++ code which generated the data displayed in Figure 5 can be obtained from `http://www.github.com/nicfreeman1209`.

A On compactness

In the proof of Lemma 5.8 we used following result, which is almost certainly known but for which we were unable to find a reference.

Lemma A.1. *Let $(\mathscr{M}, d_{\mathscr{M}})$ be a metric space and let $\mathcal{K}(\mathscr{M})$ be the Hausdorff space of compact subsets of $\mathscr{M}$. Let $(W_n)_{n \in \mathbb{N}}$ be a sequence in $\mathcal{K}(\mathscr{M})$ such that $W_n \to W_\infty \in \mathcal{K}(\mathscr{M})$ as $n \to \infty$. Then $\mathscr{W} = W_\infty \cup \left(\bigcup_{n \in \mathbb{N}} W_n \right)$ is a compact subset of M.*

Proof. Since $\mathscr{M}$ is a metric space, a subset W of $\mathscr{M}$ is compact if and only if W is sequentially compact. Let $(w_n)_{n \in \mathbb{N}}$ be any sequence in $\mathscr{W}$ and define $m_n = \sup\{m \in$

$\mathbb{N} \cup \{\infty\} \, ; \, w_n \in W_m\}$. We will construct a convergent subsequence of $(w_n)_{n \in \mathbb{N}} \subseteq \mathscr{W}$.

If $\{m_n \, ; \, n \in \mathbb{N}\}$ is finite (as an unordered set) then there exists $m \in \mathbb{N} \cup \{\infty\}$ such that $w_n \in W_m$ eventually, in which case $(w_n)_{n \in \mathbb{N}}$ has a convergent subsequence by compactness of W_m. Alternatively, if $\{m_n \, ; \, n \in \mathbb{N}\}$ is infinite then, with slight abuse of notation, we may pass to a subsequence and assume that m_n is strictly increasing to ∞.

By definition of the Hausdorff metric, since $W_n \to W$ there exists $(h_n)_{n \in \mathbb{N}} \subseteq W_\infty$ such that $d_{\mathscr{M}}(w_n, h_n) \to 0$ as $n \to \infty$. Since W_∞ is compact, $(h_n)_{n \in \mathbb{N}}$ has a convergent subsequence, and with further slight abuse of notation we pass to this subsequence and assume that $h_n \to h \in W_\infty$ as $n \to \infty$. Then $d_{\mathscr{M}}(w_n, h) \to 0$ as $n \to \infty$, and the proof is complete. $\qquad\square$

References

[1] R A Arratia, *Coalescing Brownian motions on the line*, University of Wisconsin–Madison, 1979. MR-2630231

[2] N H Barton, A M Etheridge, and A Véber, *A new model for evolution in a spatial continuum*, Electron. J. Probab. **15** (2010), 162–216. MR-2594876

[3] N H Barton, A M Etheridge, and A Véber, *Modelling evolution in a spatial continuum*, Journal of Statistical Mechanics: Theory and Experiment **01** (2013), P01002. MR-3036210

[4] N Berestycki, A M Etheridge, and A Véber, *Large-scale behaviour of the spatial Λ-Fleming-Viot process*, Ann. Inst. H. Poincaré **49** (2013), no. 2, 374–401. MR-3088374

[5] P Billingsley, *Probability and Measure*, Wiley, 1995. MR-1324786

[6] A Etheridge, N Freeman, S Penington, and D Straulino, *Branching Brownian motion and Selection in the Spatial Lambda-Fleming-Viot process*, to appear in Ann. Appl. Probab., arXiv:1512.03766 (2016).

[7] A Etheridge, S Wang, and F Yu, *Conditioning the logistic branching process on non-extinction*, Preprint (2015).

[8] A M Etheridge, *Drift, draft and structure: some mathematical models of evolution*, Banach Center Publ. **80** (2008), 121–144. MR-2433141

[9] A M Etheridge and T G Kurtz, *Genealogical constructions of population models*, arXiv:1402.6724 (2014).

[10] A M Etheridge, A Véber, and F Yu, *Rescaling limits of the spatial Lambda-Fleming-Viot process with selection*, arXiv:1406.5884 (2014). MR-3024966

[11] S N Ethier and T G Kurtz, *Markov processes: characterization and convergence*, Wiley (1986). MR-0838085

[12] P A Ferrari, L R G Fontes, and X-Y Wu, *Two-dimensional poisson trees converge to the Brownian web*, arXiv:math/0304247 (2003).

[13] P A Ferrari, L R G Fontes, and X-Y Wu, *Two-dimensional poisson trees converge to the Brownian web*, Annales de l'Institut Henri Poincaré (B) Probability and Statistics **41** (2005), no. 5, 851–858. MR-2165253

[14] P A Ferrari, C Landim, and H Thorisson, *Poisson trees, succession lines and coalescing random walks*, Annales de l'Institut Henri Poincaré (B) Probability and Statistics **40** (2004), no. 2, 141–152. MR-2044812

[15] L R G Fontes, M Isopi, C M Newman, and K Ravishankar, *The Brownian web: characterization and convergence*, The Annals of Probability **32** (2004), no. 4, 2857–2883. MR-2094432

[16] S M Krone and C Neuhauser, *Ancestral processes with selection*, Theor. Pop. Biol. **51** (1997), 210–237.

[17] B Micaux, *Flots stochastiques d'opérateurs dirigés par des bruits gaussiens et poissonniens. 2007*, Thèse présentée pour obtenir le grade de docteur en sciences de l'université Paris XI, spécialité mathématiques (2007).

[18] C Neuhauser and S M Krone, *Genealogies of samples in models with selection*, Genetics **145** (1997), 519–534.

[19] C M Newman, K Ravishankar, and E Schertzer, *Brownian net with killing*, Stoch. Proc. Appl. **125** (2015), no. 3, 1148–1194. MR-3303971

[20] C M Newman, K Ravishankar, and R Sun, *Convergence of coalescing nonsimple random walks to the Brownian web*, Electron. J. Probab. **10** (2005), 21–60. MR-2120239

[21] E Schertzer, R Sun, and J M Swart, *Stochastic flows in the Brownian web and net*, Memoirs Amer. Math. Soc. **227** (2014). MR-3155782

[22] E Schertzer, R Sun, and J M Swart, *The Brownian web, the Brownian net, and their universality*, arXiv:1506.00724 (2015), 1–80. MR-2712322

[23] D Straulino, *Selection in a spatially structured population*, DPhil thesis, University of Oxford, 2014.

[24] R Sun and J M Swart, *The Brownian net*, The Annals of Probability (2008), 1153–1208. MR-2408586

[25] A Véber and A Wakolbinger, *The spatial Lambda-Fleming-Viot process: an event-based construction and a lookdown representation*, Ann. Inst. H. Poincaré Probab. Statist. (to appear), arXiv:1212.5909v2 (2013). MR-3335017