

Branching Brownian motion under selection

Contents

| | |
|---|-----------|
| 1 Effect of selection on population genealogies: Brunet Derrida conjecture | 1 |
| 2 Coalescent process | 3 |
| 2.1 Kingman's coalescent | 3 |
| 2.2 Multiplicative coalescent | 5 |
| 3 Continuous state-space branching process | 7 |
| 3.1 1-1 correspondence with Levy process | 8 |
| 3.2 Genealogy for a CSBP | 9 |
| 3.3 Stochastic flows associated to coalescent processes | 13 |
| 4 Branching Brownian motion with absorption | 16 |
| 4.1 Genealogy convergence to the Bolthausen-Sznitman coalescent for BBM with absorption | 20 |
| 4.1.1 Construction of the flow of bridges | 20 |
| 4.1.2 Convergence of bridges | 24 |

1 Effect of selection on population genealogies: Brunet Derrida conjecture

Branching Brownian motion (BBM) under selection is inspired concerning the effect of natural selection on the genealogy lines of a population. This connection can be viewed as follows: imagine that each individual in a population is represented by a position on the real line, which measures his or her fitness. Fitness of an individual is obviously influenced by fitness of his or her parent's fitness (here we are considering asexual haploid population), but due to mutation, fitness of an offspring evolves randomly. We describe here two basic models.

- (1) **Model 1: BRW with selection** We start with a population of size N individuals. Each individual $i \leq N$ is completely characterized by

a number $x_i \in \mathbb{R}$, which measures its selective advantage (i.e., fitness). Time is discrete and during each transition, each individual gives birth to $k (\geq 2)$ many individuals. Fitnesses of all offspring are given by their parent's fitness plus i.i.d. copies of a certain displacement law ρ . Now the natural selection works as follows : after each transition only the fittest N individuals remain (i.e., the rightmost N individuals).

- (2) **Model 2: BBM with selection** At time $t = 0$, N independent Branching Brownian motions with binary branching start from a cloud or *front* of N particles on \mathbb{R} . To study the effect of natural selection, at each branching event, we kill the leftmost particle. This ensures that you are always left with population of N fittest individuals (at that instant). Note that, if the leftmost particle tries to branch we ignore the branching event. In what follows we frequently refer this model as the *NBBM* model.

What can we say about the above two models of population under selection. Brunet, Derrida, Mueller and Munier (see [6], [7]) made three striking predictions:

- (1) **Speed conjecture:** Suppose at time t , the population is represented by $X_1(t) \geq X_2(t) \geq \dots \geq X_N(t)$ where $X_i(t)$ denotes the position of the i -th particle among the selected population of size N , ordered from right. As time goes to infinity, the typical diameter of the population of N particles will be of the order of $\log N$ and the population, taken as a whole moves ballistically with asymptotic speed depending on the size, i.e.,

$$\lim_{t \rightarrow \infty} \frac{X_1(t)}{t} = \dots = \lim_{t \rightarrow \infty} \frac{X_N(t)}{t} = v_N.$$

One remarkable prediction [6] is the following: as $N \rightarrow \infty$, the asymptotic speed v_N of the population converges to a limiting value v_∞ but at unexpectedly slow rate of $(\log N)^{-2}$. The precise conjecture is the following:

$$v_\infty - v_N \sim \frac{C}{(\log N)^2}.$$

- (2) **MRCA (most recent common ancestor) conjecture:** If two individuals are selected from the population at random in some generation, then the number of generations that we need to look back to find their most recent common ancestor is of the order of $(\log N)^3$.
- (3) **Genealogy conjecture:** If n individuals are sampled from the population at random (after large time), and their ancestral lines are traced backwards in time, the coalescence of these lineages can be described by the Bolthausen-Sznitman coalescent.

While the first conjecture is settled by Be ard and Gou er  [1] under some assumptions on the displacement distribution, the remaining two conjectures (for the above models) are still open. We should mention here that these two conjectures represent the striking effect of selection on the genealogy lines of population. Berestycki, Berestycki and Schweinsberg [2] able to prove these two conjectures for a closely related but modified model where they have more independence compared to the selection model. To discuss their results, we need to understand some notions. We first start with Bolthausen-Sznitman coalescent process.

2 Coalescent process

We first fix some vocabulary and notation. A partition π of \mathbb{N} is an equivalence relation on \mathbb{N} . The blocks of the partition are the equivalence classes of this relation. In other words, the elements i and j are equivalent w.r.t. π , they are in the same block of π . The blocks of π will be listed in the increasing order of their least elements: thus, B_1 is the block containing 1, B_2 is the block containing the smallest element not in B_1 , and so on. The space of partitions of \mathbb{N} is denoted by \mathcal{P} . There is a natural distance on \mathcal{P} , which is to take $d(\pi, \pi_1)$ to be equal to 1 over the largest n such that the restriction of π and π_1 to $[n] = \{1, \dots, n\}$ are identical.

Exercise Show that, equipped with this distance, \mathcal{P} is *Polish*.

This enables us to talk about convergence in distribution for random variables taking values over \mathcal{P} or \mathcal{P}_n , where \mathcal{P}_n denotes the space of partitions of $[n]$.

2.1 Kingman's coalescent

Next for each $n \geq 1$, we want to consider a process $\{\Pi_t^n : t \geq 0\}$ taking values in \mathcal{P}_n . We start with the simplest example of a (consistent and exchangeable) coalescent process.

Definition 1. *Kingman's coalescent is a process $\{\Pi_t^n : t \geq 0\}$ taking values in \mathcal{P}_n such that*

- (i) $\Pi_0^n := [n]$;
- (ii) $\{\Pi_t^n : t \geq 0\}$ is a continuous time strong Markov process, where the transition rates $q(\pi_1, \pi_2)$ are positive if and only if π_2 is obtained from merging exactly two blocks of π_1 , in which case the rate is given by $q(\pi_1, \pi_2) = 1$.

In words, this process starts with all singleton boxes and at each transition only binary merging is allowed. The evolution may be described as follows : every pair of blocks (irrespective of their sizes) has i.i.d. Poisson clocks with rate 1. Because of this, one may think of each block as

a particle. Each pair of particles merges at rate 1. Memoryless property of Poisson process ensures that $\{\Pi_t^n : t \geq 0\}$ is a continuous time Markov process. Sometime the role of the initial state $[n]$ is emphasized by referring it as Kingman's n -coalescent.

An important property of Kingman's n -coalescent is consistency: if we consider the natural restriction of \mathcal{P}_n to partitions in \mathcal{P}_m , where $m < n$, we obtain a new random process $\mathcal{P}_{m,n}$ whose distribution is same as the law of Kingman's m -coalescent (and is thus independent of n). In this case, it is not difficult to verify the claim. By Kokmogrov's extension theorem, one important consequence of this property is the following:

Proposition 1. *There exists a unique in law process $(\Pi_t : t \geq 0)$ with values in \mathcal{P} , such that the restriction of \mathcal{P} to \mathcal{P}_n is an n -coalescent. The process $(\Pi_t : t \geq 0)$ is called the Kingman's coalescent.*

For a variety of simple population models, the genealogy of a sample from that population can be approximated by Kingman's coalescent. This will usually be formalized by taking a scaling limit as the population size N tends to infinity, while the sample size n is fixed but arbitrarily large. A striking feature of these results is that the limiting process, Kingman's coalescent, is to some degree universal, i.e., the microscopic details of the underlying probability models are of little importance. However, this universality holds under a number of important assumptions. These assumptions can be informally described as follows:

- (1) Population of constant size, and individuals typically have few offspring.
- (2) Population is well-mixed (or mean-eld): everybody is liable to interact with anybody.
- (3) No selection acts on the population.

Wright-Fisher model and Moran model are the two basic models satisfying these assumptions and the limiting genealogy is given by Kingman's coalescent. In case of Wright-Fisher model time is discrete. Suppose at time $t \in \mathbb{Z}$ the population is given by $\{x_1, \dots, x_N\}$, then at time $t + 1$ the population is represented by $\{y_1, \dots, y_N\}$ where parent of each y_i is randomly chosen among $\{X_1, \dots, x_N\}$. In case of Moran model, the only difference is time is taken as continuous and each individual has i.i.d. Poisson clock of death. When an individual dies, a new individual/offspring appears whose parent is uniformly selected among the remaining $N - 1$ individuals present.

Theorem 2. *Fix $n \geq 1$ and let $\Pi_t^{N,n}$ denote the partition obtained from looking for the ancestors at time $-t$ of n randomly picked individuals at time 0. Then as $N \rightarrow \infty$, speeding up time by a factor N , we have*

$$(\Pi_{Nt}^{N,n} : t \geq 0) \Rightarrow (\Pi_t^n : t \geq 0),$$

where $(\Pi_t^n : t \geq 0)$ is the Kingman's n coalescent.

Next we describe coalescent processes which allow multiple merging.

2.2 Multiplicative coalescent

In the previous section, we saw that Kingman's coalescent is a suitable approximation for genealogy under certain conditions. Recall that Kingman's coalescent allows only pairwise merging and when these conditions are *not* satisfied, we may need different type of coalescent process to model. When the population size has excessive fluctuations, e.g., from time to time there are *bottlenecks* in which the population size becomes very small, then at these times large proportion of the ancestry lineages coalesce. Similarly, in the process of natural selection, individuals with strong favourable mutations will quickly occupy an important fraction of the population, which results into multiple coalescence of the ancestral lineages. These heuristics suggest that, for population under selection we may require a coalescent process which allows multiple merging.

With this motivation we first try to construct a coalescent process with multiple merging satisfying the consistency conditions. We consider an array of positive numbers $(\lambda_{b,k})_{2 \leq k \leq b}$ representing the rate at which any fixed k -tuple of blocks merges when there are b blocks in total and satisfying the recursion:

$$\lambda_{b,k} = \lambda_{b+1,k} + \lambda_{b+1,k+1}. \quad (1)$$

The interpretation is that a given group of b blocks may coalesce into k blocks in two ways, reveal an extra block $b + 1$ either these k coalesce by themselves without the extra block, or they coalesce together with it.

Definition 2. A Markovian \mathcal{P}_n valued n -coalescent, $(\Pi_t^n : t \geq 0)$, with Π_t^n exchangeable for any $t \geq 0$ and consistent in the sense that the law of Π^n restricted to $[m]$ is that of Π^m for every $1 \leq m \leq n$, is uniquely specified by an array of numbers $(\lambda_{b,k})_{2 \leq k \leq b}$ satisfying (1). The process $(\Pi_t : t \geq 0)$ (whose restriction to \mathcal{P}_n has law Π^n), is called Λ -coalescent.

The name of Λ -coalescent, comes from the following beautiful characterisation of coalescents with multiple merging, which is due to Pitman [14].

Theorem 3. Let $(\Pi_t : t \geq 0)$ be a multiple coalescent associated with an array of numbers $(\lambda_{b,k})_{2 \leq k \leq b}$ (satisfying (1)). Then there exists a finite measure Λ on the interval $[0, 1]$ such that

$$\lambda_{b,k} = \int_0^1 x^{k-2} (1-x)^{b-k} \Lambda(dx) \text{ for all } 2 \leq k \leq b. \quad (2)$$

The measure Λ uniquely characterizes the law of $(\Pi_t : t \geq 0)$, which is then called a Λ -coalescent.

Example Observe that for the choice Λ as the unit Dirac mass at 0, the corresponding Λ -coalescent is nothing but Kingman's coalescent (i.e., every pair of blocks is merging at rate 1).

Definition 3. For the choice Λ as uniform measure over $[0, 1]$, the corresponding Λ -coalescent is known as Bolthausen-Sznitman coalescent. In this case the transition rates $\lambda_{b,k}$ can be explicitly calculated as

$$\lambda_{b,k} = \frac{(k-2)!(b-k)!}{(b-1)!}.$$

The Bolthausen-Sznitman coalescent arose in connection with the physics of spin glass and believed to be a universal scaling limit for models which can be described by "random travelling waves". Now we take digression and explore a deeper property of Λ coalescent which will be useful later. Because we treat blocks as exchangeable particles, the coalescent process $(\Pi_t : t \geq 0)$ is actually a Levy process, in the sense that for every $0 \leq s \leq t$, given $\mathcal{F}_t = \sigma(\Pi_s : 0 \leq s \leq t)$,

$$\Pi_{t+s} \stackrel{d}{=} \Pi_t \star \Pi'_s,$$

where Π'_s is independent from \mathcal{F}_s and has the same distribution as Π_s . Here, \star operation is defined as follows : the partition $\pi_0 \star \pi_1$ is obtained by *coagulating all the blocks of π_0 whose labels are in the same block of π_1* .

Exercise

- (i) Show that the operation \star is *not* commutative.
- (ii) Show that (\mathcal{P}, \star) forms only a monoid but not a group.

It is useful to recall that a Levy process is a real-valued process $(X_t : t \geq 0)$ with independent and stationary increments. In other words, for every $t \geq 0$, the process $(X_{t+s} - X_t : s \geq 0)$ is independent of \mathcal{F}_t and has same distribution as the original process X . The simplest example of Levy processes are of course Brownian motion and the simple Poisson process.

Because of Theorem 2, if instead of picking N fittest or rightmost particles, at each generation if we choose N particles uniformly then the limiting coalescent process is given by the Kingman's coalescent, which allows only pairwise merging. So the third conjecture by Brunet and Derrida suggests that in presence of selection the limiting picture is completely different. In the next section, we introduce another stochastic process: continuous state-space branching process (CSBP).

3 Continuous state-space branching process

In this section we discuss continuous state-space branching process (CSBP). We first start with a continuous time branching process where each member has independent copies of Poisson death clock and once the clock rings, it produces offspring according to some common offspring distribution μ independent of everything else and dies. Clearly, the process $(Z_t : t \geq 0)$ is a continuous time Markov process with absorbing state 0, where Z_t denotes the size of the population at time t . It is easy to observe that for $Z_t^{m_1+m_2}$ denoting the size of the population at time t starting from $m_1 + m_2 \in \mathbb{N}$ individuals, we have

$$Z_t^{m_1+m_2} \stackrel{d}{=} Z_t^{m_1} + Z_t^{m_2} \text{ for all } t \geq 0.$$

This is referred as the fundamental *branching property*.

Definition 4. A continuous state-space branching process (CSBP) is a Markov process with RCLL paths with values in $[0, \infty]$ whose transition kernels $(P_t)_{t \geq 0}$ satisfy the branching property

$$P_t(x, \cdot) * P_t(y, \cdot) = P_t(x + y, \cdot) \text{ for all } t, x, y \geq 0. \quad (3)$$

In other words, if Z^x and Z^y are two independent copies of CSBP Z started respectively at x and y , then $Z^x + Z^y$ has the same law of Z started at $x + y$. CSBP was introduced by Jirina [?] and their connections with Galton-Watson processes have been studied in particular by Lamperti [12]. Continuous state-space branching process can also be observed as the limit of discrete state space branching processes (see [11]).

Consider a convex function $\psi : [0, \infty) \rightarrow \mathbb{R}$ of the type

$$\psi(u) = \alpha u + \beta u^2 + \int_0^\infty (e^{-xu} - 1 + xu \mathbf{1}_{x \leq 1}) \Lambda(dx),$$

where $\alpha \in \mathbb{R}$, $\beta \geq 0$, and Λ is a measure on $(0, \infty)$ such that $\int (1 \wedge x^2) \Lambda(dx) < \infty$. A CSBP with branching mechanism ψ is a time homogeneous Markov process $(Z_t : t \geq 0)$ with values in $[0, \infty]$, where the boundary points 0 and ∞ are absorbing states, whose semigroup can be characterized via its Laplace transform as follows. For every $\lambda > 0$ and $a \in [0, \infty)$,

$$\mathbb{E}(e^{-\lambda X_t} | X_0 = a) = \exp\{-a u_t(\lambda)\},$$

where the function $u_t(\lambda)$ solves

$$\frac{\partial u_t(\lambda)}{\partial t} = -\psi(u_t(\lambda)), u_0(\lambda) = \lambda.$$

So in short, for a CSBP, branching mechanism tells us everything about the process. One can decide whether a CSBP is subcritical, critical or supercritical depending upon its branching mechanism.

Next we mention that a CSBP can also be viewed as a time changed Levy process. This correspondence will be particularly useful for us.

3.1 1-1 correspondence with Levy process

Recall that a Levy process has stationary increments. We start with the continuous time discrete state space branching process $(Z_t : t \geq 0)$. We have i.i.d. offspring distributions here. But given $Z_t = m$, the next transition occurs at the minimum of m i.i.d. Poisson clock ticks. Hence the transition rate depends on the size of the population and the increments are not stationary. Since the minimum of m i.i.d. Poisson rate 1 arrivals has the same law as Poisson rate m arrival, if we appropriately slow down the clock, i.e., magnify the time scale accordingly (with a factor of the corresponding population size), then the increments become stationary. This heuristics suggest that a discrete state space branching process can be regarded as a time-changed Levy process. In what follows, we will make this correspondence more precise and show that a CSBP also enjoys this correspondence.

Let D denote the metric space of RCLL paths defined over $[0, \infty)$ with values in $[0, \infty]$ with the usual Skorohod metric. D_1 be the induced metric space obtained by restricting to the paths which get absorbed at either 0 or ∞ . This means that for any $f \in D_1$, if $f_1(s) = 0$ (or ∞) for some $s \geq 0$, then we must have $f_1(t) = 0$ (or ∞) for all $t \geq s$.

For $f \in D_1$, we define $\theta_t := \int_0^t f(s)ds \in [0, \infty]$ for all $t \geq 0$. Let κ denote the right inverse of θ on $[0, \infty]$ given by

$$\kappa_t := \inf\{u \geq 0 : \theta_u > t\} \in [0, \infty],$$

where we follow the convention that $\inf \emptyset = \infty$. Define the Lamperti transformation

$$L : D_1 \mapsto D_1 \text{ by } L(f) = f \circ \kappa, \tag{4}$$

where one remembers that $L(f)(t) = f(\infty)$ if $\kappa_t = \infty$.

Exercise

- (i) Show that the Lamperti transform defined above is actually a map from D_1 to D_1 , i.e., show that $L(f)$ is a RCLL path with absorbing state at 0 or ∞ .
- (ii) Show that the map $L : D_1 \mapsto D_1$ is bijective.
- (iii) Compute its inverse.

From the earlier discussions, it follows that for the continuous time discrete Markov chain $(Z_t : t \geq 0)$, Lamperti transform $(L(Z_t) : t \geq 0)$ gives a Levy process.

Similar argument suggests that it is possible to get back the branching process from the Levy process by adjusting the rate of the clocks. We first compute inverse of L (we have already commented that this is a bijective transformation). Setting $g = L(f)$, one rewrites κ as $\kappa_t := \int_0^t 1/g(s)ds \in$

$[0, \infty]$. Then it is not difficult to see that $f = g \circ \theta$ where θ is the right inverse of κ .

The above discussion shows that there is a one-one correspondence between continuous time discrete state space branching processes and Levy processes, more formally spectrally positive Levy processes, i.e., class of Levy processes with no negative jumps and which get absorbed at 0. Lamperti representation theorem states that this correspondence continues to hold for CSBP's as well.

Theorem 4. *The Lamperti transformation is a bijection between continuous-state branching processes and Levy processes with no negative jumps and stopped whenever reaching zero. Specifically, for any CSBP Z , the transformed process $L(Z)$ is a Levy process with no negative jumps and stopped whenever reaching zero. On the other hand for any Levy process X with no negative jumps and stopped whenever reaching zero, $L^1(X)$ is a CSBP.*

For a proof we first recall the fact that the CSBP's are limits of rescaled DSBP's (discrete state space branching process) (see [11]). For DSBP's we have already given a heuristic argument showing that the Lamperti transformed DSBP is a compound Poisson process. In general Lamperti transform is not continuous w.r.t. the Skorohod topology. But within the conservative and non-explosive scenario this continuity holds and proves the above theorem. Even in the explosive but non-conservative case Helland [9] was able to show that if the above convergence is *strong* enough then continuity of Lamperti transform holds and able to identify necessary and sufficient conditions for such type of convergence. In order to prove the above theorem in its full generality (i.e., including the non-conservative case also), [8] introduced a new topology on the Skorohod space and proved continuity of Lamperti transform w.r.t. new topology. We refer to [8] for details.

In the next subsection we are going to describe how to obtain the genealogical structure for a CSBP.

3.2 Genealogy for a CSBP

For a DSBP, it is straightforward to recover the genealogy information from the graphical tree representation. On the other hand, for a CSBP we don't have a tree representation as such and apparently it is not clear how to recover the genealogy information for a CSBP in a consistent manner. In a classic work, Bertoin and Le Gall [3] showed how to define precisely the notion of genealogy for a CSBP. Moreover they found that for a class of CSBP's (with certain type of branching mechanism) the genealogy is given by Bolthausen-Sznitman coalescent. In order to discuss their work we need to first introduce the notion of a *subordinators*.

Definition 5. *Subordinator is a non-decreasing Levy process taking values in \mathbb{R} . Here we will work with a modification which has RCLL paths almost surely.*

We first try to explain the heuristics using a discrete state-space branching process. Though in this case the genealogical information is evident from the graphical tree representation, but now we will try to present it in a mathematical way which does not use the graphical representation. Fix $0 \leq s \leq t$. For any individual at time t , the ancestor at time s can be found out in the following way: the population at time s is given by $\{1, 2, \dots, Z_s\}$. For each $1 \leq i \leq Z_s$, we consider the sub tree (or better to say sub population) at time t (or time $t - s$) obtained starting from $\{1, \dots, i\}$ at time s (or at time 0). We further assume that at time t the population members are ordered in a certain way which respects ancestry. One way to do this is the following: each member of $\{1, \dots, Z_t\}$ is appended with the label of its parent at time s which belongs to the set $\{1, \dots, Z_s\}$ and then follow lexicographic ordering. This construction gives a non-decreasing RCLL path and given this path it is evident how to read the ancestors from this path. At this point, it may appear confusing that at one hand we are saying that we will not use the graphical “tree”, on the other hand we are considering sub trees. Actually what we use is the fact that for all $1 \leq i \leq Z_s$, the DSBP’s starting from the initial set $\{1, \dots, i\}$ exist on the same probability space, which is evident in this set up. Now we will try to follow this intuition for a CSBP.

In order to highlight the role of an initial value, we write $Z_t = Z(t, a)$ to denote that the CSBP ($Z_t : t \geq 0$) starts from $Z_0 = a \in \mathbb{R}$. By the fundamental branching property, we have

$$Z(\cdot, a) + Z'(\cdot, b) \stackrel{d}{=} Z(\cdot, a + b),$$

where $Z'(\cdot, b)$ is independent of $Z(\cdot, a)$ and has the same law as $Z(\cdot, b)$. In other words there exists some probability space such that the CSBP’s, $Z(\cdot, a)$ and $Z(\cdot, a + b)$, are defined over it with the property that $Z(\cdot, a + b) - Z(\cdot, a)$ is independent of $Z(\cdot, a)$ and has the same law as $Z(\cdot, b)$. Next we fix any $0 \leq a_1 < a_2 < a_3 < \infty$ and consider three process defined over some probability space such that:

- (i) $Z(\cdot, a_1)$;
- (ii) $Z'(\cdot, a_2 - a_1)$ which is independent of $Z(\cdot, a_1)$ and has the same distribution as $Z(\cdot, a_2 - a_1)$;
- (ii) $Z''(\cdot, a_3 - a_2)$ which is independent of both $Z(\cdot, a_1)$, $Z'(\cdot, a_2 - a_1)$ and has the same distribution as $Z(\cdot, a_3 - a_2)$.

Using these three processes it is possible to construct copies of $Z(\cdot, a_i) : 1 \leq i \leq 3$ on the same probability space such that for $1 \leq j < i \leq 3$,

$Z(\cdot, a_i) - Z(\cdot, a_j)$ has the law $Z(\cdot, a_i - a_j)$ and independent of $Z(\cdot, a_l) : 1 \leq l \leq j$.

Clearly for any $k \in \mathbb{N}$, this construction can be carried over for k points $0 \leq a_1 < a_2 < \dots < a_k < \infty$. Moreover, this construction does not depend on in which order we pick these k points. Hence by Kolmogorov's theorem, we can construct a process $(Z(t, a) : t \geq 0 \text{ and } a \geq 0)$ such that $Z(\cdot, 0) = 0$ and for every $a, b \geq 0$, $Z(\cdot, a + b) - Z(\cdot, a)$ has law $Z(\cdot, b)$ and is independent of the family of the processes $\{Z(\cdot, c) : 0 \leq c \leq a\}$. Hence it follows that for any $t > 0$ the process $Z(t, \cdot)$ has *independent and stationary increments*. Further this process is non-decreasing a.s. This gives us that the process $Z(t, \cdot)$ is a 'subordinator' (defined as in (5)). Relating with our earlier discussions for DSBP's, we observe that the RCLL path obtained there is a subordinator. The next proposition (Proposition 1 of [3]) gives a deeper connection. We need to recall some notation.

Suppose the semigroup of the CSBP $(Z(t, a) : t \geq 0)$ can be expressed through Laplace transform characterized as follows:

$$\mathbb{E}(e^{-\lambda X_t} | X_0 = a) = e^{(-au_t(\lambda))}.$$

If the branching mechanism of the CSBP is given by $(\psi(u) : u \geq 0)$, then the function $u_t(\lambda)$ satisfies

$$\frac{\partial u_t(\lambda)}{\partial t} = -\psi(u_t(\lambda)), \quad u_0(\lambda) = \lambda.$$

Proposition 5. *On some probability space, there exists a process*

$$(S^{(s,t)}(a), 0 \leq s \leq t \text{ and } a \geq 0) \text{ such that:}$$

- (i) *For every $0 \leq s \leq t$, $(S^{(s,t)} = S^{(s,t)}(a), a \geq 0)$ is a subordinator with Laplace exponent $u_{t-s}(\cdot)$.*
- (ii) *For every integer $p \geq 2$ and $0 \leq t_1 \leq \dots \leq t_p$, the subordinators $S^{(t_1, t_2)}, \dots, S^{(t_{p-1}, t_p)}$ are independent and*

$$S^{(t_1, t_p)}(a) = S^{(t_{p-1}, t_p)} \circ \dots \circ S^{(t_1, t_2)}(a) \text{ for all } a \geq 0 \text{ a.s.}$$

Finally, the processes $(S^{(0,t)}(a), t \geq 0 \text{ and } a \geq 0)$ and $(Z(t, a) : t \geq 0 \text{ and } a \geq 0)$ have the same finite-dimensional marginals.

Proof: We present the basic steps here.

Step 1. Fix $0 \leq t_1 \leq \dots \leq t_p$ and consider $p - 1$ independent subordinators $S^{(t_1, t_2)}, \dots, S^{(t_{p-1}, t_p)}$ with respective Laplace exponents $u_{(t_2 - t_1)}(\cdot), \dots, u_{(t_p - t_{p-1})}(\cdot)$. For every $a \geq 0$, set $S^{(t_i, t_i)}(a) = a$ and

$$S^{(t_i, t_j)} = S^{(t_{j-1}, t_j)} \circ \dots \circ S^{(t_i, t_{i+1})} \text{ for all } 1 \leq i < j \leq p.$$

It follows from the construction that $S^{(t_i, t_j)}$ is a subordinator with Laplace exponent $u_{t_j - t_i}(\cdot)$ and for any $0 \leq s_1 \leq \dots \leq s_k \in \{t_1, \dots, t_p\}$ these subordinators $\{S^{(s_{i-1}, s_i)} : 1 \leq i \leq k\}$ are independent.

Step 2 (Consistency) Fix $p \in \mathbb{N}$ and consider p tuples (a_i, b_i) with $a_i \leq b_i$ for all $1 \leq i \leq p$.

- Observe that the earlier construction gives subordinators $S^{(a_i, b_i)}$ with respective Laplace exponents $u_{b_i - a_i}(\cdot)$ for $1 \leq i \leq p$.
- This construction does not depend on in which order these k tuples are picked.
- By Kolmogorov's theorem, the process $(S^{(s, t)}(a), 0 \leq s \leq t \text{ and } a \geq 0)$ exists.

Step 3 (Equality of fdd's)

- Fix $a \geq 0$ and Let \mathcal{F}_t be the σ -field generated by the subordinators $S^{(r, s)}(a)$ for $0 \leq r \leq s \leq t$. For any $s \geq 0$, $S^{(0, t+s)}(a)$ is defined as $S^{(t, t+s)} \circ S^{(0, t)}(a)$ where $S^{(t, t+s)}$ is an subordinator independent of \mathcal{F}_t with Laplace exponent $u_s(\cdot)$. Hence the process $S^{(0, \cdot)}(a)$ is a homogeneous Markov process starting from a with semigroup

$$\mathbb{E}(e^{-\lambda S^{(0, t+s)}(a)} | S^{(0, s)}(a) = x) = \mathbb{E}(e^{-\lambda S^{(s, t+s)}(x)}) = e^{-x u_t(\lambda)}.$$

The above shows that the processes $S^{(0, \cdot)}(a)$ and $Z(\cdot, a)$ have the same law.

- Finally in order to complete the proof we fix general $(a_i, t_i) : 1 \leq i \leq p$ with $t_i < t_{i+1}$ and show that there exists some probability space where we can have independent copies of these subordinators $S^{(0, t_i)}(a_i) : 1 \leq i \leq p$. Together with the earlier result, this completes the proof.

□

Once we have this construction of subordinators and CSBP's on the same probability space we are ready to recover the genealogies.

Definition 6. For every $b, c \geq 0$ and $0 \leq s < t$, we say that the individual c in the population at time t has ancestor (or is a descendant of) the individual b in the population at time s if b is a jump time of $S^{(s, t)}$ and

$$S^{(s, t)}(b) < c < S^{(s, t)}(b).$$

The set of individuals in the population at time t having an ancestor at time s is the random open subset of $[0, \infty)$ with canonical decomposition $\bigcup (S^{(s, t)}(b-), S^{(s, t)}(b))$ where the union is taken over the jump times of the

subordinator $S^{(s,t)}$. Its complement can be identified as the closed range of $S^{(s,t)}$. Under the assumption that $S^{(s,t)}$ has zero drift, the closed range has zero Lebesgue measure a.s. In other words, the individuals in the population at time t having no ancestor at time s are of Lebesgue measure zero a.s. Clearly an individual in the population at time t has at most one ancestor at time s .

Suppose $0 \leq r < s < t$. If the individual d in the population at time t has ancestor c in the population at time s , and if the latter has ancestor b in the population at time r , then by definition

$$S^{(s,t)}(c-) < d < S^{(s,t)}(c) \text{ and } S^{(s,t)}(b-) < c < S^{(s,t)}(b).$$

Since $S^{(r,t)} = S^{(s,t)} \circ S^{(r,s)}$, by monotonicity we have

$$S^{(r,t)}(b-) < d < S^{(r,t)}(b),$$

i.e., the individual d in the population at time t has ancestor b in the population at time r . This shows that the genealogies obtained through this construction is consistent.

The last thing we are going to discuss before discussing the results of [2] is how to study convergence of genealogies. This will be based on a pioneering work of Bertoin and Le Gall [4].

3.3 Stochastic flows associated to coalescent processes

We start with a haploid population of fixed size N denoted by $\{1, \dots, N\}$ and transitions due to generations happen at discrete time steps. For every generation $n \in \mathbb{Z}$ and $1 \leq i \leq N$, $\xi_{i,n}$ denote the number of children of the individual i at generation n . Since the total size of the population is fixed, we must have $\xi_{1,n} + \dots + \xi_{N,n} = N$. The individuals are labeled in a way which is compatible with the genealogy. Formally when $\xi_{i,n} \geq 1$, the labels of the individual(s) at generation $n+1$, who are the children of the individual i at the n -th generation, run from $\xi_{1,n} + \dots + \xi_{i-1,n} + 1$ to $\xi_{1,n} + \dots + \xi_{i,n}$. The ancestors of the population at the n -th generation can be obtained from the function $\Delta_n : \{0, \dots, N\} \mapsto \{0, \dots, N\}$ which is defined by

$$\Delta_n(i) = \begin{cases} 0 & \text{if } i = 0 \\ \xi_{1,n} + \dots + \xi_{i,n} & \text{if } 1 \leq i \leq n. \end{cases}$$

Note that, the increments of Δ_n are exactly the sizes of the blocks of the partition of the population at generation n who has the same ancestor in generation $n-1$.

For $m < n$, if we want to obtain the ancestors at generation m for the population at generation n , then we use the compositions

$$\Delta_{m,n} = \Delta_n \circ \Delta_{n-1} \circ \dots \circ \Delta_{m+1}.$$

Now when the sequence $\{(\xi_{1,n}, \dots, \xi_{N,n}) : n \in \mathbb{N}\}$ gives i.i.d. copies of exchangeable N -tuple of random variables $(\xi_{1,n}, \dots, \xi_{N,n})$, i.e., its distribution is invariant under any permutation of indices, then the sequence of *random maps* $\{\Delta_n : n \in \mathbb{Z}\}$ is i.i.d. with exchangeable increments. Further for any $m_1 < n_1 \leq m_2 < n_2$, the maps Δ_{m_1, n_1} and Δ_{m_2, n_2} are independent. This motivates the following definition.

Definition 7. We call $(B_{m,n} : m, n \in \mathbb{N} \text{ and } m \leq n)$ a *flow of discrete bridges (with exchangeable increments)* if the following conditions are satisfied :

- (a) for all $m \in \mathbb{N}$, $B_{(m,m)}$ denotes an identity map a.s.
- (b) (cocycle property) For every $0 \leq m \leq n$,

$$B_{0,n} = B_{m,n} \circ B_{0,m} .$$

- (c) The law of $B_{m,n}$ only depends on $n - m$, and for every $n_1 \leq \dots \leq n_k$, the bridges $B_{n_1, n_2}, \dots, B_{n_{k-1}, n_k}$ are independent.

In general, we observe that for all $0 \leq m \leq n$, the process $(B_{m,n}(s) : s \in [0, 1])$ has RCLL non-decreasing paths with exchangeable increments such that $B_{m,n}(0) = 0$ and $B_{m,n}(1) = 1$. With this discrete set up motivation, we proceed to study the continuous version.

Definition 8. A *bridge* $B = (B(r), r \in [0, 1])$ is a random process with non-decreasing RCLL paths and exchangeable increments, such that $B(0) = 0$ and $B(1) = 1$.

Let Id denote the identity mapping over $[0, 1]$. Following [4], we are going to define a continuous analog of Definition 7.

Definition 9. $B = (B_{s,t}(x), 0 \leq s \leq t, 0 \leq x \leq 1)$ is a *flow of bridges*, which is a collection $(B_{s,t}, 0 \leq s \leq t)$ of bridges such that:

- For every $s < t < u$, we have $B_{s,u} = B_{t,u} \circ B_{s,t}$.
- The law of $B_{s,t}$ only depends on $t - s$.
- If $s_1 < s_2 < \dots < s_n$, then the bridges $B_{s_1, s_2}, \dots, B_{s_{n-1}, s_n}$ are independent.
- $B_{0,0} = \text{Id}$ and $B_{0,t} \rightarrow \text{Id}$ as $t \rightarrow 0$ in probability, in the sense of Skorohod topology.

Definition 10. Consider $\{V_n : n \in \mathbb{N}\}$, an i.i.d. collection of $U(0, 1)$ random variables. Define a random partition $\Pi(B)$ of \mathbb{N} by the equivalence relation given by,

$$i \sim_{\Pi} j \Leftrightarrow B^{-1}(V_i) = B^{-1}(V_j) \text{ for } i, j \in \mathbb{N},$$

where B^{-1} denotes the usual right inverse given by

$$B^{-1}(r) = \inf\{s \in [0, 1] : B(s) > r\} \text{ for } r \in [0, 1] \text{ and } B^{-1}(1) = 1.$$

When B has exchangeable increments then $\Pi(B)$ is also an exchangeable partition.

We present a continuity lemma taken from [4], which is very useful for studying convergence of genealogies. Let $D([0, 1], \mathbb{R}_+)$ be the space of RCLL non-decreasing paths defined over $[0, 1]$ such that $f(0) = 0$ and $f(1) = 1$ for all $f \in D([0, 1], \mathbb{R}_+)$.

Lemma 6. *Consider a sequence of bridges $\{B_n : n \in \mathbb{N}\}$. The following are equivalent.*

- (i) *The exchangeable partition $\Pi(B_n) \Rightarrow \Pi(B_\infty)$ (w.r.t. the topology mentioned in the beginning of Section 2).*
- (ii) *$B_n \Rightarrow B_\infty$ in $D([0, 1], \mathbb{R}_+)$ endowed with the Skorohod topology.*

In a little more generally, this lemma can be stated as follows: if we start from a sequence $\{(B_n) : n \in \mathbb{N}\}$ of bridges and assume that the exchangeable partitions $\Pi(B_n)$ converge in distribution, then the limit has to be of the form $\Pi(B_\infty)$. We will not present the proof of this lemma here. In the next section we present a stronger version of this lemma and prove that.

Bertoin and Le Gall [3] proved that the genealogy for the CSBP with branching mechanism $\psi(u) = u \log(u)$, studied by Neveu, is given by Bolthausen-Sznitman coalescent process.

Theorem 7. *Fix $a, t > 0$. Let V_1, V_2, \dots be a sequence of random variables such that conditionally on \mathcal{F}_t , V_i 's are independently and uniformly distributed over $[0, Z(t, a)]$. For any $0 \leq s \leq t$ with probability 1 we can define an equivalence relation $\tilde{\Gamma}_s$ on \mathbb{N} by declaring $m \sim_{\tilde{\Gamma}_s} n$ (i.e., m, n belong to the same class of $\tilde{\Gamma}_s$ if V_m and V_n has the same ancestor at time $t - s$). Then*

$$(\tilde{\Gamma}_s : 0 \leq s \leq t) \text{ and } (\Gamma_s : 0 \leq s \leq t) \text{ has same f.d.d.'s.}$$

Here we just present how to construct a *flow of bridges* to represent genealogy when population size no longer remains constant. Then we have to normalize properly.

We fix an integer $p \geq 1$ and choose finitely many ordered time points $0 \leq t_0 < t_1 < \dots < t_p \leq t$. In order to simplify notation, for every $0 \leq k \leq p$, we take $a_k = Z(t_k, a) = S^{(0, t_k)}(a)$. For $0 \leq k \leq p - 1$, define a process $B_k = (B_k(s) : 0 \leq s \leq 1)$ by setting

$$B_k(s) = (S^{(t_k, t_{k+1})}(sa_k)) / (S^{(t_k, t_{k+1})}(a_k)) = (S^{(t_k, t_{k+1})}(sa_k)) / a_{k+1}.$$

Clearly, $B_k(0) = 0$ and $B_k(1) = 1$. By construction, B_k has non-decreasing RCLL paths.

Exercise Show that B_k 's are independent for different k 's.

For any $0 \leq k \leq p - 1$, we consider the compositions

$$B_{p1} \circ \dots \circ B_K(s) = a_p^{-1} S^{(t_k, t_p)}(sa_k).$$

4 Branching Brownian motion with absorption

In 1978, Kesten [10] introduced branching Brownian motion with absorption. In the beginning, one particle starts branching Brownian motion from a location $x > 0$ and as soon as a particle hits the line starting from the origin with slope μ it is killed. It is equivalent to say that this process follows the same dynamics as branching Brownian motion with drift μ , where $\mu > 0$, starting with one particle located at $x > 0$ in the beginning and particles are killed when they reach the origin. Kesten showed that there exists a critical value $\mu_c = \sqrt{2}$ such that if $\mu \geq \mu_c$, then the process dies out almost surely, while if $\mu < \mu_c$, the process survives with positive probability.

Branching Brownian motion with absorption can also be interpreted as a model for population with natural selection. Imagine that each individual in a population is represented by a position on the real line, representing his or her fitness. Due to mutations, fitness of an individual evolves according to Brownian motion, and initially the fitness of a child is identical to the fitness of the parent. Selection progressively eliminates all individuals whose fitness becomes too low and falls beyond a line with constant speed (or slope). Every individual whose fitness falls beyond the current threshold is instantly removed from the population. This prevents the population to grow too much. Clearly, the population size no longer remains the same, but we will see later if the threshold parameter μ (the drift of the threshold line) is chosen suitably, then the order of the population size remains the same. On the other hand, compared to the NBBM model where the number of Brownian particles are fixed due to natural selection (described in the beginning of Section 1), this model is relatively easier to deal with as the particles enjoy more independence. Berestycki, Berestycki and Schweinsberg [2] studied this model in detail. Before discussing their result, we describe some other predictions made by Brunet and Derrida ([6], [7]), which will help us to understand the heuristics of the choice of different parameters in [2].

- (1) Most of the time, the particles are in a meta-stable state. In this state, the cloud of particles (also called the front) has a diameter of $\log N + O(1)$. The empirical density of particles seen from the left-most is proportional to $e^{-x} \sin(\pi x / \log N)$, and the system moves at a linear speed $v_{\text{cutoff}} = 1 - \pi^2 / (2 \log^2 N)$. In particular, most of the particles are at $O(1)$ distance from the left-most particle.

- (2) They further observed that this meta-stable state is perturbed from time to time by particles moving far to the right. Fix a point in the bulk, for example the median, and say that a particle reaches a point y if it moves to distance y from that fixed point. It was predicted that a particle moving up to the point $\log N + x$ causes, after a relaxation time of order $\log^2 N$, a shift of the front by $\Delta = \log(1 + Ce^x/\log^3 N)$; for some constant $C > 0$. Hence, in order to have an effect on the position of the front, a particle has to reach a point near $\log N + 3 \log \log N$.
- (3) Assuming that such an event where a particle reaching to the point $\log N + x$ happens at rate proportional to e^x , the time taken by a particle to come close to $\log N + 3 \log \log N$ (and thus causing shifts of the front) is of the order of $\log^3 N$.

The first observation suggests that as long as the particles do not reach $\log N + 3 \log \log N$, the size of the population does not change too much and the particles hitting $\log N + 3 \log \log N$ require special attention. For the ease of notation we take $L_N = \log N + 3 \log \log N$. In order to study asymptotic results, the authors of [2] considered a sequence of branching Brownian motions $\{(X_N(t), t \geq 0) : N \in \mathbb{N}\}$ with absorption where for each $N \geq 2$ the corresponding drift $\mu = \mu_N$ is given by,

$$\mu_N := \sqrt{2 - \frac{2\pi^2 \log N}{(\log N + 3 \log \log N)}}. \quad (5)$$

Clearly this choice allows them to work with the near-critical set up and in order to keep the growth of the population in control, it is intuitive to take $\mu_n \uparrow \sqrt{2}$. But the specific choice of μ_N comes from the fact that, for a branching Brownian motion with drift μ starting from a single particle at x , and particles *killed* upon reaching 0 or L , the expected number of particles in a set B at a sufficiently large time t is approximately $\int_B p_t(x, y) dy$, where

$$p_t(x, y) = \frac{2}{L} e^{(1 - \mu^2/2 - \pi^2/(2L^2))t} \cdot e^{\mu x} \sin\left(\frac{\pi x}{L}\right) e^{-\mu y} \sin\left(\frac{\pi y}{L}\right).$$

In order to study the points hitting $L_N = \log N + 3 \log \log N$ separately (as these point produce a shift by producing large number of particles before getting absorbed), take $L = L_N$ as the right killing boundary. Note that the time parameter t appears only in the first exponential factor implying that the population size should be roughly constant provided μ is chosen such that $1 - \mu^2/2 - \pi^2/(2L^2) = 0$. This motivates the choice $\mu = \mu_N$ satisfying (5).

The above discussion suggests that for a branching Brownian motion with drift μ_N starting from a single particle, after a large time t , the number of particles whose ancestor at time s is within the interval $(0, L_N)$ for all $0 \leq s \leq t$ is roughly constant. On the other hand the fluctuation of the

population size is largely due to contributions of the points which go too far to the right and hit L_N . As predicted by the physicists, in [2] it was also proved that the time taken to hit L_N is of order $\log^3 N$ which is the proper time scale to look into. In order to obtain asymptotic results, they also ensure that the population size fluctuations (which are mainly due to the ‘far-away’ particles) are of the right *order* (i.e., of order N) on the time scale of interest.

It should be noted that in [2], actually the process started with many particles, rather than just one, satisfying some rather technical assumptions. In this expository article, we restrain to state the exact technical conditions at this moment and for interested reader we refer to Proposition 1 of [2]. We comment that these conditions ensure that the number of particles in the system stays of *order* of N on the time scale of interest, so that the process can be viewed as a model of a population of size *approximately* N . This again heuristically explains the connection to the NBBM model where the number of Brownian particles are fixed. Once again we comment that this model is relatively easier to deal with compared to the NBBM, viz., because the particles enjoy more independence. In what follows, we assume that these technical conditions hold.

With this brief heuristics, we present one of the main results of [2]. Recall that for each $N \in \mathbb{N}$, the process $(X_N(t) : t \geq 0)$ denotes a branching Brownian motion. Let $M_N(t)$ denote the number of the Brownian particles at time t and their positions are denoted by $X_{1,N}(t) \geq X_{2,N}(t) \geq \dots \geq X_{M_N,N}(t)$.

Theorem 8. *As $N \rightarrow \infty$, the finite-dimensional distributions of the process $\{M_N((\log^3 N)t) : t > 0\}$ converge to the finite-dimensional distributions of the continuous-state branching process (CSBP) with branching mechanism $\psi(u) = au + 2\pi u \log u$ started with distribution ν at time zero, where a and ν comes from the initial technical conditions.*

Remark 9. (i) *Our assumption about initial condition is not strong enough to guarantee the convergence for $t > 0$.*

(ii) *This result proves that the size of the population remains of order n at $\log^3 N$ time scaling.*

(ii) *Because of fluctuations, one can not have process convergence while working with the Skorohod topology. But it might be of interest to see whether it is possible to achieve process convergence with respect to some other topology (e.g., Skorohod M_1 topology).*

Here we present only the brief heuristics. As discussed earlier, by dividing the particles at time t into two parts; those that have stayed inside the interval $(0, L_N)$, and those that have hit the point L_N before hitting 0, roughly corresponds to the division of the process into a deterministic and

a stochastic part. For the particles that hit L_N , the authors of [2] further showed that

- (1) the number of descendants at a later time of such a particle is of the order of WN , where W is a random variable with tail $\mathbb{P}(W > x) \sim 1/x$ as $x \rightarrow \infty$ and
- (2) the rate at which particles hit the right barrier is of the order of $\log^3 N$.

These two findings together yield their result. Having said this, one should note that proving the above facts is quite involved and we refer to [2] for full details.

The next result proves the analog of the third conjecture of Brunet et al. ([6], [7]) for branching Brownian motion with absorption. Suppose n particles are picked randomly after a large time and we trace back their ancestral lines, the resulting coalescent process, with proper time scaling, converges to the Bolthausen-Sznitman coalescent. We recall that the Bolthausen-Sznitman coalescent allows multiple merging. Here multiple merging is due to the fact that occasionally particles go too far to the right and hit L_n and these particles before getting killed generate large number of offspring occupying a positive proportion of the population. This explains why multiple merging appears here. This is in sharp contrast with population models with no selections where Kingman's coalescent appear as the universal scaling limit (under some very general conditions).

For a precise formulation of the result, we choose n particles uniformly at random from the $M_N((\log^3 N)t)$ particles at time $(\log^3 N)t$, and label these particles at random by $1, \dots, n$. Fix $t > 0$. For $0 \leq s \leq 2\pi t$, let $\Pi_N(s)$ to be the partition of $\{1, \dots, n\}$ such that i and j are in the same block of $\Pi_N(s)$ if and only if i, j particles both are descended from the same ancestor at time $(t - s/(2\pi))(\log^3 N)$. On the other hand let $(\Pi(s) : 0 \leq s \leq 2\pi t)$ be the Bolthausen-Sznitman n coalescent running for time $2\pi t$ and restricted to $\{1, \dots, n\}$.

Theorem 10. *As $N \rightarrow \infty$, the finite-dimensional distributions of the process $(\Pi_N(s), 0 \leq s \leq 2\pi t)$ converge to those of $(\Pi(s), 0 \leq s \leq 2\pi t)$.*

We remark that the $(\log N)^3$ time scaling that appears here, as well as in the previous Theorem, shows that if two particles are chosen at random, then the time taken to reach to their most recent common ancestor is of the order $(\log N)^3$ and this justifies the second conjecture in the context of branching Brownian motion with absorption.

We recall that Bertoin and Le Gall [3] showed how to obtain the genealogy structure of a continuous-state branching process. They found that the genealogy of Neveu's continuous-state branching process is given by the Bolthausen-Sznitman coalescent. These results were further extended in [4]

and it was shown that the genealogy of any continuous-state branching process with branching mechanism, of the form $\psi(u) = au + bu \log u$, can still be described by the Bolthausen- Sznitman coalescent.

In order to prove Theorem 10 we first need to come up with a flow of discrete bridges $\{(B_{s,t}^N(x) : 0 \leq x \leq 1) : 0 \leq s \leq t\}$ such that for any $0 = t_0 < t_1 < \dots < t_k = t$, we have the equality of distribution

$$(\Pi(B_{t_{k-1},t_k}^N), \dots, \Pi(B_{t_0,t_k}^N)) \stackrel{d}{=} (\Pi_N(2\pi(t - t_{k-1})), \dots, \Pi_N(2\pi(t - t_0))),$$

where $\Pi_N(2\pi(t - t_i))$ is the coalescent process obtained by looking at the genealogy of n randomly picked individuals from the set of $M_N(\log^3 Nt)$ individuals over the time period $[\log^3 Nt_i, \log^3 Nt]$ at $(\log N)^3$ time scale. Then we show that, as $n \rightarrow \infty$, w.r.t. the Skorohod topology

$$\{(B_{t_{k-1},t_k}^N), \dots, \Pi(B_{t_0,t_k}^N) : N \in \mathbb{N}\} \Rightarrow \{(B_{t_{k-1},t}, \dots, B_{t_1,t})\}$$

where $\{(B_{s,t}(x) : 0 \leq x \leq 1) : 0 \leq s \leq t \leq 1\}$ represents the corresponding flow of bridge for a CSBP process with branching mechanism $\psi(u) = au + 2\pi^2 u \log u$ constructed as in the proof of Theorem 7. On the other hand, from [4] it follows that the finite dimensional distributions of $(\Pi(B_{s,t}) : 0 \leq s \leq t)$ are given by that of the Bolthausen-Sznitmann coalescent process. Hence because of Lemma 6 (strictly speaking we need a stronger lemma as the approximating bridges will no longer have exchangeable increments), we have that

$$\begin{aligned} & \{(\Pi_N(2\pi(t - t_{k-1})), \dots, \Pi_N(2\pi(t - t_0))) : N \in \mathbb{N}\} \\ & \Rightarrow (\Pi(2\pi(t - t_{k-1})), \dots, \Pi(2\pi(t - t_0))), \end{aligned}$$

as $n \rightarrow \infty$. This gives the result.

4.1 Genealogy convergence to the Bolthausen-Sznitman coalescent for BBM with absorption

In this subsection we will have a deeper look at Theorem 10 and discuss the steps of the proof in detail.

4.1.1 Construction of the flow of bridges

We first consider the CSBP $(Z(t) : t \geq 0)$ with branching mechanism $\psi(u) = au + 2\pi^2 u \log u$. Recall from Proposition 5 that there exists a flow of subordinators $(S^{(s,t)}(x) : 0 \leq s \leq t, x \geq 0)$ such that

- (i) for all $0 \leq s \leq t$, the process $S^{(s,t)} = (S^{(s,t)}(x) : x \geq 0)$ is a subordinator with Laplace exponent u_{t-s} where $u_{t-s}(\cdot)$ is as in (??);

- (ii) for $k \geq 2$ and for any $0 \leq t_1 \leq \dots \leq t_k$, the subordinators are independent, and

$$S^{(t_1, t_k)} = S^{(t_{k-1}, t_k)} \circ \dots \circ S^{(t_1, t_2)}, \text{ and}$$

- (iii) the processes $(Z(t, x) : t \geq 0, x \geq 0)$ and $(S^{(0, t)}(x) : t \geq 0, x \geq 0)$ have the same finite dimensional distributions.

Here $S^{(s, t)}(x)$ is understood as the size of the population at time t starting with a population x at time s .

Next we create a flow of bridges $(B_{s, t}(x) : 0 \leq x \leq 1, s \leq t)$ from the flow of subordinators as done in the ending of last section. Assuming that we are starting with initial population $Z(0) = z$, we set

$$B_{s, t} = S^{(s, t)}(xS^{(0, s)}(z))/S^{(0, t)}(z).$$

Clearly, $B_{s, t}(0) = 0$ and $B_{s, t}(1) = 1$.

Exercise Show that $B_{s, t}$ has non-decreasing RCLL paths with exchangeable increments.

Exercise Check that $(B_{s, t}(x) : 0 \leq s \leq t, 0 \leq x \leq 1)$ constitutes a *flow of bridges*, i.e., show that the following conditions hold:

- (i) for all $s < t < u$, $B_{s, u} = B_{s, t} \circ B_{t, u}$;
- (ii) $B_{s, t}$ only depends on $t - s$;
- (iii) for $s_1 < s_2 < \dots < s_n$, the bridges $B_{s_1, s_2}, B_{s_2, s_3}, \dots, B_{s_{n-1}, s_n}$ are independent;
- (iv) $B_{0, 0} = \text{Id}$ (identity mapping) and $B_{0, t} \rightarrow \text{Id}$ as $t \rightarrow 0$ in probability with respect to Skorohod topology.

Exercise Show that for $s < t < u$, we have $B_{s, u}^{-1} = B_{s, t}^{-1} \circ B_{t, u}^{-1}$ where $B^{-1}(u) = \inf\{v \in [0, 1] : B(v) \geq u\}$.

As explained in Subsection 3.3, given a flow of bridges $(B_{s, t} : 0 \leq s \leq t)$ we can consider the partition valued process $(\Pi_s = \Pi(B_{s, t}) : 0 \leq s \leq t)$. In fact Bertoin and Le Gall [4] established that there exists one-one correspondence between flow of bridges and exchangeable coalescent processes. In particular in this case, i.e., when the flow of bridges obtained from a CSBP process with branching mechanism $\psi(u) = au + 2\pi^2 u \log u$, the process $(\Pi(Bs/(2\pi), t) : 0 \leq s \leq 2\pi t)$ is the Bolthausen-Sznitman coalescent running for time $2\pi t$.

Now we need to create the flow of bridges representing the genealogy of branching Brownian motion with absorption and show that this flow of bridges converge to the flow $(B_{s, t} : 0 \leq s \leq t)$. In order to do that, we

need to introduce few notation and describe the two technical assumptions in detail. Let

$$\begin{aligned} Z_N(t) &= \sum_{i=1}^{M_N(t)} e^{\mu X_{i,N}(t)} \sin\left(\frac{\pi X_{i,N}(t)}{L}\right) \mathbf{1}_{X_{i,N}(t) \leq L} \text{ and} \\ Y_N(t) &= \sum_{i=1}^{M_N(t)} e^{\mu X_{i,N}(t)}, \end{aligned} \tag{6}$$

where $L_n := \frac{1}{\sqrt{2}}(\log N + 3 \log \log N)$. For a BBM with drift μ and two killing barriers at 0 and L , starting from $x \in (0, L)$ with a single particle, the expected number of particles in a set $B \subset [0, L]$ after large time, is approximately given by $\int_B p(x, y) dy$, where

$$p_t(x, y) = \frac{2}{L} e^{(1-\mu^2-\pi^2/(2L^2))t} e^{\mu x} \sin\left(\frac{\pi x}{L}\right) e^{-\mu y} \sin(\pi y/L).$$

The choice of μ_N and L_N ensures that the first exponential term is independent of t and if we sum over all initial positions, then this is proportional to $Z_N(0)$, defined as in (6). In this sense Z_N predicts the number of particles in a given set at later time and provides a useful measure of the size of the population. Now we are going to describe two technical assumptions about the initial configurations which are assumed to be hold.

- (1) Let $V_N(t) = \frac{1}{N(\log N)^2} Z_N((\log N)^3 t)$. As $N \rightarrow \infty$, the distribution of $V_N(0)$ converges to ν , a probability distribution on $[0, \infty)$ and
- (2) as $N \rightarrow \infty$, the quantity $Y_N(0)/N(\log N)^2$ converges to zero in probability.

It was proved in [2] that it is the value of $Z_N(t)$ rather than $M_N(t)$, that predicts the size of population a short time later. In fact Theorem 8 was proved first for the process $(Z_N((\log N)^3 t) : t \geq 0)$ (with a different scaling), and then using the fact that $Z_N(t)$ predicts the size of the population a short time later, Theorem 8 was proved. Precisely we have the following proposition (Proposition 1 of [2]):

Proposition 11. *(Under the assumptions on the initial configurations) there exists a constant $a \in \mathbb{R}$ such that as $N \rightarrow \infty$, the finite dimensional distributions of the process $(V_N(t) : t \geq 0)$ converge to the finite dimensional distributions of the CSBP with branching mechanism $\psi(u) = au + 2\pi^2 \log u$ started with initial distribution ν at time zero.*

This also explains why we do not have convergence for $M_N(0)$, the number of particles at time 0, as the conditions assumed about initial configurations do not necessarily give convergence for the initial *size* of the population.

Now we are ready to construct the bridges to study the genealogy of BBM with absorption. Fix $K \in \mathbb{N}$ and the time points $0 = t_0 < t_1 < \dots < t_K$. For $0 \leq i < j \leq k$, we will define a process $(B_{t_i, t_j}^N(s) : 0 \leq s \leq 1)$.

We first require an ordering of the population that respects ancestry, i.e., we want to assign label to the BBM particles at time $t_j \log^3 N$,

$$X_{1,N}(t_j \log^3 N) \geq X_{2,N}(t_j \log^3 N) \geq \dots \geq X_{M_N,N}(t_j \log^3 N).$$

Let $(v_{i,j} : i \geq 0, 0 \leq j \leq k)$ be a collection of i.i.d. uniform $(0, 1)$ random variables, independent of the BBM X_N . For Brownian particles at time 0, i.e., for $i \leq M_N(0)$, we take the label as $u_{i,0} = v_{i,0}$. For $j \geq 1$, $u_{i,j}$ are sequences of length $j + 1$ which are defined inductively by setting

$$u_{i,j} = (u_{p(i),j-1}, v_{i,j}),$$

where $u_{p(i),j-1}$ is the label of the parent particle at time $t_{j-1}(\log^3 N)$. In other words we concatenate $v_{i,j}$ with label of the parent particle to obtain the label of $X_{i,N}(t_j \log^3 N)$. The Brownian particles are then ordered using the lexicographic ordering w.r.t. their labels and this ordering respects ancestry. Let $x_{i,j}$ denote the position of the i -th individual (w.r.t. this ancestry ordering) at time $(t_j \log^3 N)$.

Next we assign weights to the individuals. For $0 \leq j \leq k$ and $1 \leq i \leq M_N(t_j \log^3 N)$ we take

$$w(i, j) = \begin{cases} \frac{1}{Z_N(t_j \log^3 N)} e^{\mu x_{i,j}} \sin(\pi x_{i,j}/L) \mathbf{1}_{x_{i,j} \leq L} & \text{if } 0 \leq j \leq k-1 \\ \frac{1}{M_N(t_k \log^3 N)} & \text{if } j = k. \end{cases} \quad (7)$$

i.e., the particles are weighted according to their contribution to (6). Let

$$A_i(j, k) = \{l : x_{l,k} \text{ (at time } (t_k \log^3 N)) \text{ is a descendant of } x_{i,j} \text{ at time } (t_j \log^3 N)\}.$$

For $0 \leq y \leq 1$ and $0 \leq j \leq k$ define

$$L_j(y) = \max\{l \in \mathbb{N} : \sum_{i=1}^l w(i, j) \leq y\}, \quad (8)$$

with the condition that maximum of the empty set is 0.

This allows us to construct the discrete bridges as follows: for $0 \leq y \leq 1$ and $0 \leq j < k \leq K$, let

$$B_{t_j, t_k}^N(y) = \sum_{i=1}^{L_j(y)} \sum_{m \in A_i(j, k)} w(m, k).$$

These processes are not exactly bridge in the sense of (8) as they do not have exchangeable increments.

Exercise For $i < j < k$ show the following:

- (i) $B_{t_i, t_k}^N = B_{t_j, t_k}^N \circ B_{t_i, t_j}^N$,
- (ii) $(B_{t_i, t_k}^N)^{-1} = (B_{t_i, t_j}^N)^{-1} \circ (B_{t_j, t_k}^N)^{-1}$,

4.1.2 Convergence of bridges

In order to prove the convergence of bridges we first show that, under some reasonable assumptions, if the population at time zero is divided into m random subgroups, then Proposition 11 holds jointly for branching Brownian motions starting from these m subgroups. In order to have this joint convergence the initial conditions must hold for each of these m random subgroups and the next result shows that these assumptions are enough to obtain joint convergence.

Lemma 12. *Suppose the initial population is divided into m random subgroups S_1, \dots, S_m , and given the position of the initial particles, they start BBM with drift μ killed at 0. Assume that $Y_N(0)/N(\log N)^2$ converges to 0 in probability. Let $Z_{i,N}(t)$ denote the contribution to the sum in (6) from particles descended from one of the particles that is in S_i at time 0. We further assume that the initial joint distribution of $(\frac{Z_{i,N}(0)}{N(\log N)^2})_{i=1}^m$ converges to some probability measure ρ on $[0, \infty)^m$. Then the finite dimensional distributions of the m -dimensional vector-valued processes*

$$\left\{ \left(\frac{Z_{i,N}(t(\log N)^3)}{N(\log N)^2} \right)_{i=1}^m : t \geq 0 \right\}$$

converge as $N \rightarrow \infty$ to the finite dimensional distributions of $\{(Z_i(t))_{i=1}^m : t \geq 0\}$ where $(Z_i(0))_{i=1}^m$ has distribution ρ , and conditional on $(Z_i(0))_{i=1}^m$, each Z_i evolves independently as a CSBP with branching mechanism $\psi(u) = au + 2\pi u \log u$.

Once one takes care of the fact that the m components of the joint process are not independent, but conditional on the given initial configuration they are independent, the proof follows from Proposition 11. It is not hard to see that using Theorem 8 it is possible to have similar joint convergence for population sizes of these m subgroups. Using the above lemma, we prove the following, which is our first step towards establishing convergence of bridges.

Lemma 13. *Fix $k \in \mathbb{N}$ and consider $0 = t_0 < t_1 < \dots < t_k$. Let $m \geq 1$ and let $0 = u_0 < u_1 < \dots < u_m = 1$. Then for each fixed i with $1 \leq i \leq k$, we have*

$$(B_{0, t_i}^N(u_j))_{j=1}^m \Rightarrow (B_{0, t_i}(u_j))_{j=1}^m.$$

Proof: It suffices to prove the joint convergence of the increments, $(B_{0, t_i}^N(u_j) - B_{0, t_i}^N(u_{j-1}))_{j=1}^m$. To do that we take $S_j = \{L_0(u_{j-1})+1, \dots, L_0(u_j)\}$

where L_0 is defined as in (8). In words denote the subset of population at time 0 associated with the quantiles in $[u_{j-1}, u_j)$, when the population is weighted according to $w_{\cdot,0}$ (7). By definition, S_j 's are disjoint and give a partition of the initial population into m random subgroups. From the initial assumption it is straightforward to see that $Y_{i,N}(0)/(N(\log N)^3) \rightarrow 0$ in probability as $N \rightarrow \infty$ for all $1 \leq i \leq m$. We also have assumed that the distribution of $V_N(0)/(N(\log N)^2)$ converges to a probability measure ν as $N \rightarrow \infty$. With some more work, it follows that joint distribution of $(V_{i,N}(0)/(N(\log N)^2))_{i=1}^m$ converges to $((u_j - u_{j-1})X)_{j=1}^m$ where X has distribution ν . Hence by Lemma 12 we have that

$$\left(\frac{Z_{j,N}(t_i(\log N)^3)}{N(\log N)^2}\right)_{j=1}^m \Rightarrow (Z_j(t_i))_{j=1}^m.$$

This together with some more work completes the proof. \square

Our goal is to first establish joint convergence of the bridges in terms of finite dimensional distributions. For any $t \geq 0$, let \mathcal{F}_t denote the σ -field generated by $(X_N(s) : 0 \leq s \leq t)$. We show that

Lemma 14. *Fix $k \in \mathbb{N}$ and consider $0 = t_0 < t_1 < \dots < t_k$. Let $m \geq 1$ and let $0 = u_0 < u_1 < \dots < u_m = 1$. Let $f : [0, 1]^{m+1} \rightarrow \mathbb{R}$ is bounded and continuous. For $0 \leq i \leq k-1$, we have that*

$$\mathbb{E}[f(B_{t_i, t_{i+1}}^N(u_0), \dots, B_{t_i, t_{i+1}}^N(u_m)) | \mathcal{F}_{t_i(\log N)^3}]$$

converges in probability to $\mathbb{E}[f(B_{t_i, t_{i+1}}(u_0), \dots, B_{t_i, t_{i+1}}(u_m))]$ as $N \rightarrow \infty$.

Proof: Note that, by Proposition 11 for all $1 \leq i \leq k$ we have

$$\frac{Z_N(t_i(\log N)^3)}{N(\log N)^2} \Rightarrow Z(t_i) \text{ as well as}$$

it was proved in [2] that $Y_N(t_i(\log N)^3)/(N(\log N)^3)$ converges to 0 in probability. Using the fact that given $\mathcal{F}_{t_i(\log N)^3}$, the branching Brownian motions evolve independently starting from the positions $X(t_i(\log N)^3)$, we apply Lemma 13 for time interval $[t_i(\log N)^3, t_{i+1}(\log N)^3]$ and complete the proof. \square

With this we are ready to prove the joint convergence of bridges in terms of finite dimensional distributions.

Lemma 15. *Fix $k \in \mathbb{N}$ and consider $0 = t_0 < t_1 < \dots < t_k$. Let $m \geq 1$ and let $0 = u_0 < u_1 < \dots < u_m = 1$. Then*

$$(B_{t_i, t_{i+1}}^N(u_j) : 1 \leq j \leq m)_{0 \leq i \leq k-1} \Rightarrow (B_{t_i, t_{i+1}}(u_j) : 1 \leq j \leq m)_{0 \leq i \leq k-1},$$

where the bridges $B_{t_i, t_{i+1}}, 0 \leq i \leq k-1$, are independent.

Proof: Use induction method. For $i = 1$, this follows from Lemma 13. Now use the fact that $B_{t_i, t_{i+1}}^N$ is conditionally independent of $B_{t_0, t_1}^N, B_{t_1, t_2}^N, \dots, B_{t_{i-1}, t_i}^N$ given $\mathcal{F}_{t_i(\log N)^3}$ together with Lemma 14 to complete the proof. \square

Now we are almost ready to prove Theorem 10. One important thing to observe here is that we can not apply Lemma 6 directly here as the discrete bridges B^N 's do not have exchangeable increments. We need a stronger version of that lemma. The next lemma which is in deterministic setting proves that exchangeable increments are not necessary for convergence.

Lemma 16. *Suppose b, b_1, b_2, \dots are functions from $[0, 1]$ to $[0, 1]$ that are non-decreasing and right continuous and have left limits at every point other than 0. Suppose $\lim_{n \rightarrow \infty} \rho(b_N, b) = 0$, where ρ denotes the Skorohod metric. Suppose $(x_n)_{n=1}^\infty$ and $(y_n)_{n=1}^\infty$ are sequences in $[0, 1]$ such that $x_n \rightarrow x$ and $y_n \rightarrow y$ as $n \rightarrow \infty$. Suppose x and y are not in the closure of the range of b . Then for sufficiently large n we have $b_n^{-1}(x_n) = b_n^{-1}(y_n)$ if and only if $b^{-1}(x) = b^{-1}(y)$. Furthermore,*

$$\lim_{n \rightarrow \infty} b_n^{-1}(x_n) = b^{-1}(x).$$

Exercise

- (1) Let $\Lambda : [0, 1] \mapsto [0, 1]$ denote the class of strictly increasing continuous onto functions. Show that for a collection of RCLL paths f, f_1, f_2, \dots , we have $\lim_{n \rightarrow \infty} \rho(f_n, f) = 0$ if and only if there exists a sequence of functions $(\lambda_n)_{n=1}^\infty$ in Λ such that

$$\begin{aligned} & \lim_{n \rightarrow \infty} \sup\{|f_n(\lambda_n(t))f(t)| : 0 \leq t \leq 1\} \\ &= \lim_{n \rightarrow \infty} \sup\{|\lambda_n(t)\lambda(t)| : 0 \leq t \leq 1\} = 0. \end{aligned}$$

- (2) Using the above interpretation of convergence with respect to Skorohod metric, prove Lemma 16.

Once we prove Lemma 16, the proof of Theorem 10 follows from backward method of induction. For the first step, from Lemmas 15 and 16 we have

$$\Pi(B_{t_{k-1}, t_k}^N) \Rightarrow \Pi(B_{t_{k-1}, t_k}) \text{ as } n \rightarrow \infty.$$

Assuming that the convergence holds for $(\Pi(B_{t_j, t_k}^N) : i + 1 \leq j \leq k - 1)$, in order to study $\Pi(B_{t_i, t_k}^N)$ we use the composition

$$B_{t_i, t_k}^N = B_{t_i, t_{i+1}}^N \circ B_{t_{i+1}, t_k}^N.$$

Further for a uniform $[0, 1]$ random variable U , because of Lemma 2 of [4] we have that the random variable $B_{t_{i+1}, t_k}^{-1}(U)$ also has the uniform distribution on $[0, 1]$. This fact together with Lemma 16 completes the backward induction and hence proves Theorem 10.

References

- [1] Bérard, J. and Guéré, J. B. BrunetDerrida behavior of branching-selection particle systems on the line. *Comm. Math. Phys.* **298** (2010) 323–342.
- [2] Berestycki, J., Berestycki, N. and Schweinsberg, J. The genealogy of branching Brownian motion with absorption. *Ann. Probab.* **41** (2013) 527–618.
- [3] Bertoin, J. and Gall, J. F. The BolthausenSznitman coalescent and the genealogy of continuous-state branching processes. *Probab. Theory Related Fields* **117** (2000) 249–266.
- [4] Bertoin, J. and Gall, J. F. Stochastic flows associated to coalescent processes. *Probab. Theory Related Fields* **126** (2003) 261–288.
- [5] Bolthausen, E. and Sznitman, A. S. On Ruelles probability cascades and an abstract cavity method. *Comm. Math. Phys.* **197** (1998) 247–276.
- [6] Brunet, E., Derrida, B., Mueller, A. H. and Munier, S. Noisy traveling waves: Effect of selection on genealogies. *Europhys. Lett.* **76** (2006) 1–7.
- [7] Brunet, E., Derrida, B., Mueller, A. H. and Munier, S. Effect of selection on ancestry: an exactly soluble case and its phenomenological generalization. *Phys. Rev. E* **76** 76 (2007) 041104.
- [8] Caballero, M. E., Lambert, A. and Uribe Bravo, G. Proof(s) of the Lamperti representation of the continuous-state branching processes. *Probab. Surv.* **6** (2009) 62–89.
- [9] Helland, I. S. Continuity of a class of random time transformations. *Stochastic Processes Appl.* **7** 1 (1978) 79–99.
- [10] Kesten, H. Branching Brownian motion with absorption. *Stochastic Process. Appl.* **7** (1978) 9–47.
- [11] Lamperti, J. The limit of a sequence of branching processes. *Z. Wahrsch. Verw. Gebiete* **7** (1967) 271–288.
- [12] Lamperti, J. Continuous state branching processes. *Bull. Amer. Math. Soc.* **73** (1967) 382–386.
- [13] Malliard, P. Speed and fluctuations of N -particle branching Brownian motion with spatial selection. 1304.0562.
- [14] Pitman, J. Coalescents with multiple collisions. *Ann. Probab.* **27** (1999) 1870–1902.