Genealogies in Expanding Populations

Rick Durrett and Wai-Tong (Louis) Fan

July 2, 2015

Abstract

Mueller and Tribe [32] have shown that rescaled long-range voter models in one-dimension converge to a Wright-Fisher SPDE, also known as a stochastic Fisher-Kolmogorov-Petrovskii-Piscounov (FKPP) wave. Recently Hallatschek and Nelson [20] have described the asymptotic behavior of genealogies in a closely related model. Their answer is expressed in terms of a diffusion in a very singular random environment. Here we prove rigorous results that partially confirm their analysis. Brunet et al. [7] have conjectured that genealogies of all models in the FKPP universality class are described by the Bolthausen-Sznitman coalescent, see [33]. However, in the model we study there are no simultaneous coalescences.

1 Introduction

The goal of this work is to understand the shape of genealogies in growing tumors but similar issues arise in the study of populations expanding into new territory, [9, 15, 26]. The last two papers consider a 25×100 grid of demes of carrying capacity 50. Here, we will study the biased voter model on $\mathbb{Z}^d$ with each site occupied by one cell that is a wild type (0) or a cancer cell (1). The fitness of cancer cells is $\lambda = 1 + s$ times that of wild type cells. Let f_i be the fraction of neighbors in state i. A site changes

$$\begin{aligned} &\text{from state 0 to 1 at rate } \lambda f_1, \\ &\text{from state 1 to 0 at rate } f_0. \end{aligned}$$

As we will describe in more detail in Section 1.2, the biased voter model, ξ_t, is dual to the branching coalescing random walk: $\zeta_r^{x,t}$, which can be defined on the same space so that

$$\{\xi_t(x) = 1\} = \{\xi_{t-r}(y) = 1 \text{ for some } y \in \zeta_r^{x,t}\}. \tag{1}$$

In $\zeta_r^{x,t}$ particles jump at rate 1 to a randomly chosen nearest neighbor, and give birth at rate s to a new particle sent to a randomly chosen nearest neighbor. If a particle lands on a site occupied by another particle they coalesce to become a single particle.

Let $A_0^{x,t} = \{y \in \zeta_t^{x,t} : \xi_0(y) = 1\}$. If this set is empty, then $\xi_t(x) = 0$; if not, then $A_0^{x,t}$ gives the potential ancestors of the cell at x at time t. To find the actual ancestor, the traditional method is to run the biased voter model forward in time from this initial set.

This procedure, invented by Krone and Neuhauser, [28, 35], is called the ancestral selection graph. Here, as we show in Section 1.2, one can instead impose an ordering on points in the dual so that the ancestor is the first occupied site in the dual.

The shape theorem of Bramson and Griffeath [5, 4], shows that if we start the biased voter model from a single type 1 and the process does not die out, then at time t, the set of 1's is roughly $tD \cap Z^d$, where D is the limiting shape. D is a convex set that has the same symmetries as those of $\mathbb{Z}^d$ that leave the origin invariant, i.e., rotations and reflections. Suppose we sample a site x at random from tD. Let $C_t = \cup_{s=0}^{t} (sD \times \{s\})$ denote the space time cone swept out by the expanding ball. As we work backwards in time in the dual process, there will be no successful branching events until our genealogy exits C_t, so this part of the genealogy will be a random walk. Random walks move by an amount $O(\sqrt{t})$ in time t while the cone is growing linearly, so to a first approximation the genealogy will go straight down until it hits the boundary of the cone. What happens when the genealogy hits the boundary is hard to visualize, but there must be enough successful branching events to allow the genealogy to move at the same speed as which the biased voter model advances.

Some insight into the nature of the genealogies can be found in experimental papers. Sottoriva et al. [39] did genomic profiling of 349 individual glands sampled from the opposite sides of 15 colorectal cancer tumors. Consistent with the hypothesis that the cancer came from a single aberrant colon crypt, some mutations were present in all glands. However, there were also mutations that were only present in small regions of the tumor. This is consistent with the mental picture of genealogies working backwards in time along the edge of the space time cone. Genealogies of cells sampled from opposite edges of the tumor at time t will coalesce near the time when the tumor began growing. See Figure 1 in [39].

Hallatschek, Hersen, Ramanathan, and Nelson [19] inserted two initially well mixed populations of fluorescent *E. coli* into the center of a Petri dish. Over time the system developed well-defined sector-like regions with "fractal" boundaries. The authors concluded that the formation of these regions was driven by random fluctuations that originate in a thin band of pioneers at the expanding frontier. This system and similar experiments involving yeast have been studied by approximate calculations and heuristic arguments leading to a description of the genealogies, see [20, 27, 30, 36].

Our goal is to obtain rigorous results. One related model that has been analyzed mathematically is the multi-color version of first passage percolation, which was introduced by Haggström and Pemantle [18]. In this system, occupied sites are red or blue and vacant sites become occupied by either a red or a blue particle at a rate proportional to the number of occupied sites of that color. In [18] it was shown that if we start with one red and one blue then with positive probability the red and blue populations become infinitely large, and will each occupy a linearly growing cone. For more recent results, see [17] and references therein.

1.1 Long range voter models in one dimension

To try to find a mathematically tractable model, we will leave biological reality and consider a one-dimensional system. Perhaps the most famous of these is the model introduced by Brunet, Derrida, Mueller, and Munier [7]. In one of the many variants, at each time step, each of the N particles dies and gives birth to two offspring, displaced from its current location by an independent amount. Then out of the $2N$ particles we choose the right-most

N of them. In a second version, which is analytically more tractable, see [8], we again always have N particles, but a particle at y gives birth to an infinite number of offspring distributed according to a Poisson point process with intensity $e^{-(x-y)}$. Again we keep the right-most N. In this case, it is reasonably straightforward to compute, see [6], that when N is large the genealogy follows the Bolthausen-Sznitman coalescent [3] and the most recent common ancestor is $O(\log N)$ generations in the past. Berestycki, Berestycki, and Schweinsberg [2] have proved a similar result for a system introduced many years ago by Kesten [24]. In that system, Brownian motions with drift $-\sqrt{2}$ on $(0, \infty)$ branch into two at rate 1 and are killed when they hit 0. As Kesten showed $-\sqrt{2}$ is the critical value of the drift separating rapid extinction from exponential growth.

While the Brunet-Derrida model is interesting, it is not relevant to the expanding tumor model we want to study, because in that system space represents a continuum of fitness values. Instead, in this paper, we will consider a model closely related to the one introduced by Hallatsheck and Nelson [20]. In our model, there is one cell at each point of $(L^{-1}\mathbb{Z}) \times \{1, \ldots, M\}$, whose cell-type is either 1 or 0. The cells in deme $w \in L^{-1}\mathbb{Z}$ only interact with those in demes $w - L^{-1}$ and $w + L^{-1}$. Hence each cell $x = (w, i)$ has $2M$ neighbors. Type-0 cells reproduce at rate $2Mr_n$, type-1 cell at rate $2M(r_n + \theta R_n^{-1})$. When reproduction occurs the offspring replaces a neighbor chosen uniformly at random. In the terminology of evolutionary games, this is birth-death updating.

Let $\xi_t(x) = \xi_t^n(x)$ be the type of the cell at x at time t. Our (rescaled) biased voter model $(\xi_t)_{t\geq 0}$ can be constructed using two independent families of i.i.d. Poisson processes: $\{P_t^{x,y} : x \sim y\}$ that have rate r_n, and $\{\tilde{P}_t^{x,y} : x \sim y\}$, that have rate θR_n^{-1}. Here we write $y \sim x$ to indicate that y is a neighbor of x. At a jump time of $P_t^{x,y}$, the cell at x is replaced by an offspring of the one at y. At a jump time of $\tilde{P}_t^{x,y}$, the cell at x is replaced by an offspring of the one at y only if y has cell-type 1. The dynamics of $(\xi_t)_{t\geq 0}$ can be described by the equation

$$\xi_t(x) = \xi_0(x) + \sum_{y\sim x} \int_0^t (\xi_{s-}(y) - \xi_{s-}(x))\, dP_s^{x,y} + \sum_{y\sim x} \int_0^t \xi_{s-}(y)(1 - \xi_{s-}(x))\, d\tilde{P}_s^{x,y}. \qquad (2)$$

In the first integral, if $\xi_{s-}(y) = \xi_{s-}(x)$ then nothing happens.

$$\text{If } \xi_{s-}(y) = 1 \text{ and } \xi_{s-}(x) = 0 \text{ then } \xi_s(x) = 1;$$
$$\text{if } \xi_{s-}(y) = 0 \text{ and } \xi_{s-}(x) = 1 \text{ then } \xi_s(x) = 0.$$

In the second integral, nothing happens unless $\xi_{s-}(y) = 1$ and $\xi_{s-}(x) = 0$. In this case $\xi_s(x) = 1$.

We define the approximate density by

$$u_t^n(w) := \frac{1}{M} \sum_{i=1}^{M} \xi_t(w, i)$$

and linearly interpolate between demes to define $u_t^n(w)$ for all $w \in \mathbb{R}$. It is clear that for all $t \geq 0$, we have $0 \leq u_t^n(w) \leq 1$ for all $w \in \mathbb{R}$ and $u_t^n \in C_b(\mathbb{R})$, the set of bounded continuous functions on $\mathbb{R}$. If we equip $C_b(\mathbb{R})$ with the metric

$$\|f\| = \sum_{k=1}^{\infty} 2^{-k} \sup_{|x|\leq k} |f(x)| \qquad (3)$$

i.e., uniform convergence on compact sets, then $C_b(\mathbb{R})$ is Polish and the paths $t \mapsto u_t^n$ are $C_b(\mathbb{R})$ valued and càdlàg.

Theorem 1. *Suppose that as $n \to \infty$, the initial condition u_0^n converges in $C_b(\mathbb{R})$ to f_0 and that:*

(a) $r_n M_n / L_n^2 \to \alpha \in (0, \infty)$

(b) $r_n / L_n \to \gamma \in [0, \infty)$

(c) $M_n / R_n \to \beta \in [0, \infty)$

(d) $L_n \to \infty$ *and* $L_n R_n \to \infty$

Then the approximate density process $(u_t^n)_{t \geq 0}$ converges in distribution in $D([0,\infty), C_b(\mathbb{R}))$, as $n \to \infty$, to a continuous $C_b(\mathbb{R})$ valued process $(u_t)_{t\geq 0}$ which is the weak solution to the (stochastic) partial differential equation

$$\partial_t u = \alpha\, \Delta u + 2\theta\, \beta\, u(1-u) + |4\gamma\, u(1-u)|^{1/2}\, \dot{W} \tag{4}$$

with initial condition $u_0 = f_0$. Here $\dot{W}$ is the space-time white noise on $[0,\infty) \times \mathbb{R}$.

This result is a straight forward generalization of a result of Mueller and Tribe [32]. They considered a long range voter model on $\mathbb{Z}/n$ in which two voters are neighbors if $|x-y| \leq \sqrt{n}$. Their voters change their opinion at rate $O(n)$ and imitate the opinion of a neighbor chosen at random. More precisely, for each of the $2n^{1/2}$ neighbors, they adopt the opinion of that neighbor at rate $n^{1/2}$ if it is 0 and at rate $n^{1/2} + \theta n^{-1/2}$ if it is 1. Their model corresponds roughly in our situation to $L_n = M_n = R_n = r_n = n^{1/2}$. Their limit is

$$\partial_t u = \frac{1}{6}\, \Delta u + 2\theta\, u(1-u) + |4\, u(1-u)|^{1/2}\, \dot{W}.$$

$\beta = \gamma = 1$ while the 1/6 comes from the fact that the variance of the uniform distribution on $[-1,1]$ is 1/3.

1.2 Duality

By using methods of Durrett and Restrepo [14], we can study the limiting behavior of the the dual branching coalescing random walks. They considered a sequence of voter models on $\mathbb{Z}$ in which voters change their opinion at rate 1 and a voter at x imitates the one at $x+z$ with probability $q^n(z)$ where

1. $q^n(z) = q^n(-z)$,
2. $\sum_{z\in\mathbb{Z}} z^2 q^n(z) = \sigma_n^2 n$ with $\sigma_n \to \sigma \in (0,\infty)$,
3. there is an $h > 0$, independent of n so that $q^n(z) \geq h/n^{1/2}$ for $|z| \leq n^{1/2}$,
4. $q^n(z) \leq C \exp(-c|z|/n^{1/2})$.

Theorem 2. [14, Theorem 3] Consider the voter model on $\mathbb{Z}$ with dispersal kernel q^n satisfying assumptions 1-4 above. Let t_0 be the coalescence time of the lineages starting at 0 and at L_n where $L_n/\sigma n \to x_0 \geq 0$. Then $2t_0/n$ converges in distribution to $\ell_0^{-1}(\sigma\tau/2)$ where ℓ_0 is the local time at 0 of a standard Brownian motion started from x_0 and τ is an independent mean 1 exponential random variable.

The first step in generalizing this result is to describe the dual process in our setting in more detail. To do this, we recall the graphical representation of the biased voter model introduced by Harris (1976) and developed by Griffeath (1978). The ingredients are arrows that spread fluid in the direction of their orientation, and δ's which are dams that stop fluid. At times s of $\tilde{P}_{x,y}$ we draw an arrow from $(s,y) \to (s,x)$. At times s of $P_{x,y}$ we draw an arrow from $(s,y) \to (s,x)$ and put a δ at $(s-,x)$. Intuitively, we inject fluid into the bottom of the graphical representation at the sites of the configuration that are 1 and let it *flow up.* A site x is in state 1 at time t if and only if it can be reached by fluid. If there is fluid at y at time s and an arrow (with no δ) from $(s,y) \to (s,x)$, the fluid will spread to x, i.e., there is a birth at x if it is in state 0. If x is already occupied no change occurs.

If there is an arrow-δ from $(s,y) \to (s,x)$, then a little thought reveals

before		after	
y	x	y	x
1	0	1	1
0	1	0	0
1	1	1	1
0	0	0	0

In the first case the fluid spreads from y to x as before. In the second, there is no fluid to be spread but the dam stops the fluid at x. In the third the dam stops the fluid at x, but it replaced by fluid from y. In the fourth, there is no fluid so nothing happens. Thus the effect in all cases is that x imitates y.

To define the dual, we inject fluid at x at time t and let it *flow down.* It is again stopped by dams but now moves across arrows in the direction OPPOSITE to their orientation. Given $z = (w,i)$ we define the dual process $\zeta_s^{t,z}$, $0 \le s \le t$ to be the set of sites at time $t-s$ that can be reached by fluid starting at z. A little thought shows that $\zeta_0^{t,z} = \{z\}$ and follows the following rules:

- If a particle in $\zeta_s^{t,z}$ is at x and an arrival in $P^{x,y}$ occurs at time $t-s$ then the particle jumps to y.
- If a particle in $\zeta_s^{t,z}$ is at x and an arrival in $\tilde{P}^{x,y}$ occurs at time $t-s$ then the particle gives birth to a new particle at y.
- If a jumping particle or an offspring lands on another particle in $\zeta_s^{t,z}$, then the two particles coalesce to 1.

For a picture see Figure 1. There $\zeta_t^{t,1} = \{-2,-1,1,3\}$. It follows from the definition that

Lemma 1. *$\xi_t(z) = 1$ if and only if $\xi_0(x) = 1$ for some $x \in \zeta_t^{t,z}$.*

To extend the definition to a collection of sites A, we let $\zeta_s^{t,A} = \cup_{z\in A}\zeta_s^{t,z}$. We defined our dual process starting at a fixed time t so that the relationship in Lemma 1 holds with probability 1. If we have two times $t < t'$ then the distributions of $\zeta_s^{t,A}$ and $\zeta_s^{t',A}$ agree up to time t, so the Kolmogorov extension theorem implies that we can define a process ζ_s^A for all time so that the distribution agrees with $\zeta_s^{t,A}$ up to time t.

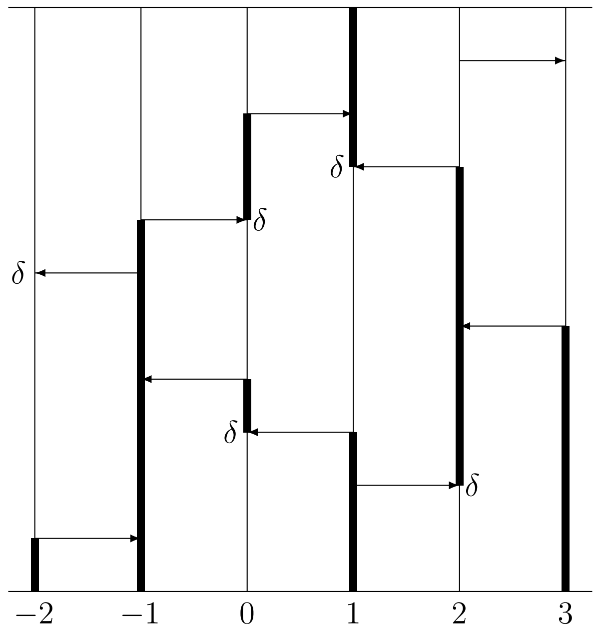

Figure 1: Duality for the biased voter model.

Theorem 3. *As $n \to \infty$, the dual process converges to a limit in which branching occurs at rate $\theta\beta$, particles move according to Brownian motions with variance $2\alpha t$, and two particles coalesce when the local time at 0 of the difference between their locations exceeds $\alpha\zeta/\gamma$, where ζ is a mean one exponential independent of the particle motions. If $\gamma = 0$ there is no coalescence.*

If we let χ_t^A denote the spatial locations of the particles in the limit process in Theorem 3, then Lemma 1 and Theorem 3 imply

$$\mathbb{E} \prod_{x \in A} (1 - u_t(x)) = \mathbb{E} \prod_{y \in \chi_t^A} (1 - u_0(y)). \tag{5}$$

Here A can be a multi-set, e.g., $\{a, a, a, b, b\}$. In this case duality says

$$\mathbb{E}[(1 - u_t(a))^3 (1 - u_t(b)^2] = \mathbb{E} \prod_{y \in \chi_t^A} (1 - u_0(y)).$$

The duality relationship described in the last paragraph is not new. In 1986, Shiga and Uchiyama [37] introduced it to study a collection of Wright-Fisher diffusions coupled by migrations. Shiga [38] used it to show uniqueness in law for (4). Doering, Mueller, and Smerekea [11] gave a simpler description and derivation of the dual. See also Hobson and Tribe [22].

The dual process gives us the set of possible ancestors of the particle at x at time t. To determine the actual ancestor and the ancestral lineage, we assign an ordering to the particles in the dual in such a way that the first occupied site in the list will be the ancestor. The rules are as follows:

- If particle i jumps and there is no coalescence, then its location changes but the order does not.
- If particle i jumps and coalesces with j then the particle with the higher index is removed from the dual. The surviving particle has its location updated. The remaining particles are reindexed.
- If the i particle gives birth then the new particle is labeled i, while all particles with indices $j \geq i$ have their indices increased by 1.

To explain the definition, we will work through the example drawn in Figure 1. The succeisve states are

$$\begin{array}{l} \{1\}, \quad \{0,1\}, \quad \{0,2\}, \quad \{-1,2\}, \quad \{-1,3,2\}, \\ \{0,-1,3,2\}, \quad \{1,-1,3,2\}, \quad \{1,-1,3\}, \quad \{1,-2,-1,3\} \end{array}$$

To see this note that under our rules, at the point where the dual jumps from $\{1\} \to \{0,1\}$, if there is a particle at 0 it will give birth and replace any particle at 1. The next two events involve voting, so the affected particles move but do not change their position in the ordering. The next event is an arrow from 3 to 2. A particle at 3 will replace one at 2, so 3 is inserted in the list before 2. The next novel event is the seventh transition when the particles at 2 and 1 coalesce: at that time we drop the lower ranked particle.

Neuhauser [34] used a similar dual to show that in the multitype contact process, if the particle death rates are equal, then the one with the higher birth rate takes over the system. She was concerned with limits starting from translation invariant initial distributions, so the top ranked particle, which she called the *distinguished particle* had positive probability to land on a site occupied by the type with the higher birth rate. In our situation the distinguished particle will jump left or right with equal probability due to birth and voter events. Thus if we sample an individual at time t near the predicted location of the front, it is unlikely its lineage will land on an occupied site, so the true ancestor will be some particle that is not highly ranked. Because of this, it seems difficult to use the ordered dual to obtain detailed information about the motion of the lineage in our model. However, since coalescences of particles in the branching coalescing random walk occur when local times exceed independent exponentially distributed levels, it is clear that we do not get simultaneous coalescences as in the Bolthausen-Sznitman coalescent, and the times between coalescences are $O(1)$.

1.3 Tracer Dynamics

Hallatschek and Nelson [20] have an interesting approach to studying the ancestral lineage of a particle based on “tracer dynamics”. Think of our expanding population as a fluid and inject a small amount of red liquid at time 0. The locations of the red fluid at time t will identify the locations of their progeny. In particle terms, we will have states 0, 1, and 1^* where the * indicates being labeled by the tracer. To construct the labeled process it is convenient to use the graphical representation. If there is an arrow-δ from y to x then x adopts the state of y. If there is an arrow from y to x and y is in state 1 or 1^* then x will adopt the state of y, but nothing happens if y is in state 0. In fluid terms, the color of fluid at y replaces that at x.

To do computations, we let $\eta_t(x) = 1$ if the individual at x at time t is labeled and 0 otherwise. We only label type 1's, so $\eta_t(x) \le \xi_t(x)$. A noted above $\eta_t(x)$ can be computed using the version of the dual $\zeta_s^{t,x}$ in which the points are ordered. That is, $\eta_t(x) = 1$ if the first occupied site in the list is labeled. More formally, if $\zeta_t^{t,x} = \{y_1, y_2, \dots y_K\}$ (note that K and $y_1, \dots y_K$ are random variables), then

$$P(\xi_t(x) = 0) = E\left(\prod_{i=1}^{K}(1 - \xi_0(y_i))\right) \equiv EF(\zeta_t^x)$$

$$P(\eta_t(x) = 1) = E\left(\sum_{j=1}^{K}\zeta_0(y_j))\prod_{i=1}^{j-1}(1 - \xi_0(y_i))\right) \equiv EG(\zeta_t^x)$$

In the same way one computes

$$\begin{aligned} P(\xi_t(x_1) = 0, \dots \xi_t(x_m) = 0, \eta_t(x_{m+1}) = 1, \dots, \eta_t(x_n) = 1) \\ = E\left[\prod_{i=1}^{m} F(\zeta_t^{x_i}) \prod_{i=m+1}^{n} G(\zeta_t^{x_i})\right] \end{aligned} \tag{6}$$

A standard argument shows that the probabilities just computed determine the distribution of (ξ_t, η_t).

To obtain an idea of the working of the limit process (and to prove tightness), we note that in terms of the previously defined Poisson processes,

$$\begin{aligned} \eta_t(x) - \eta_0(x) = \sum_{y\sim x}\int_0^t (\eta_{s-}(y) - \eta_{s-}(x))\, dP_s^{x,y} \\ + \sum_{y\sim x}\int_0^t \eta_{s-}(y)(1 - \eta_{s-}(x)) - \xi_{s-}(y)(1 - \eta_{s-}(y))\eta_{s-}(x)\, d\tilde{P}_s^{x,y}. \end{aligned} \tag{7}$$

The first term gives the voter interactions. For the second term, note that if y is in state 1^* and x is not, the number of 1^*'s will increase by 1, while if x is in state 1^* and y is in state 1 ($\xi_{s-}(y) = 1$ and $\eta_{s-}(y) = 0$), the number will decrease by 1.

As before, we define the approximate density for the labeled particles by

$$\ell_t^n(w) := \frac{1}{M}\sum_{i=1}^{M}\eta_t(w, i)$$

and linearly interpolate to obtain a function $\ell_t^n(w)$ for all $w \in \mathbb{R}$.

Theorem 4. *Suppose that as $n \to \infty$, the conditions on r_n, R_n, M_n and L_n in Theorem 1 hold, and that the initial condition (u_0^n, ℓ_0^n) converges in $C_b(\mathbb{R}) \times C_b(\mathbb{R})$ to (f_0, g_0). Then the pair of approximate densities $(u_t^n,\, \ell_t^n)_{t\geq 0}$ converges in distribution in $D([0,\infty), C_b(\mathbb{R}) \times C_b(\mathbb{R}))$ to a continuous $C_b(\mathbb{R}) \times C_b(\mathbb{R})$ valued process $(u_t,\, \ell_t)_{t\geq 0}$ which is the weak solution to the coupled (stochastic) partial differential equations*

$$\partial_t u = \alpha\, \Delta u + 2\theta\, \beta\, u(1-u) + |4\gamma\, \ell(1-u)|^{1/2}\, \dot{W}^0 + |4\gamma\, (u-\ell)(1-u)|^{1/2}\, \dot{W}^1$$
$$\partial_t \ell = \alpha\, \Delta \ell + 2\theta\, \beta\, \ell\, (1-u) + |4\gamma\, \ell(1-u)|^{1/2}\, \dot{W}^0 + |4\gamma\, \ell\, (u-\ell)|^{1/2}\, \dot{W}^2$$

with initial condition $(u_0,\, \ell_0) = (f_0,\, g_0)$, where $\dot{W}^i$, $i = 0, 1, 2$ are three independent space-time white noises on $[0,\infty) \times \mathbb{R}$.

As the proof shows the three noises $\dot{W}^0$, $\dot{W}^1$ and $\dot{W}^2$ refer to voting interactions between 1^* and 0, 1 and 0, and 1^* and 1 respectively. The drift in ℓ_t is $2\theta\beta\, \ell(1-u)$ because labeled particles only have a selective advantage in competition with those of type 0.

We will prove our result by showing that the sequence of approximating processes is tight. To conclude that there is weak convergence, we need to show

Lemma 2. *When $\gamma > 0$, the solution of the coupled SPDE in Theorem 4 is unique in law.*

To do this we use our duality function (6). Details are in Section 8.

In Theorems 1, 3 and 4, the assumption $\alpha > 0$ is used crucially in the proof of tightness, but we allow β or γ to be 0. The deterministic regime ($\gamma = 0$) occurs, for instance, if $r_n = n^{1/a}$, $L_n = n^{1/b}$, $M_n = \lceil \alpha n^{2/b - 1/a} \rceil$ and $R_n = M_n/\beta$, where $2a > b > a > 0$. When $\gamma = 0$ the limiting process is a PDE

$$\partial_t u = \alpha\, \Delta u + 2\theta\, \beta\, u(1-u)$$
$$\partial_t \ell = \alpha\, \Delta \ell + 2\theta\, \beta\, \ell\, (1-u).$$

This PDE obviously has a unique weak solution (solve the first equation and then solve the second), but if one wants, this can be proved using duality.

1.4 Lineage dynamics

Using tracer dynamics Hallatschek and Nelson [20], see page 163 and Appendix A, derived the probability density $G(x,t|x',t')$ that an individual at x' at time t' is descended from an ancestor at x at time $t < t'$. They assumed that the population density in a frame moving with velocity v is

$$\partial_t u(x,t) = D\partial_x^2 u(x,t) + v\partial_x u(x,t) + K(x,t)$$

where, for instance, $K(x,t) = su(u_\infty - u) + \epsilon\sqrt{u(u_\infty - u)}Z$ with Z being a space-time White noise, and we have changed their c to u. They concluded, see their (3), that

$$\partial_t G(x,t|x',t') = -\partial_x J(x,t|x',t')$$
$$J(x,t|x',t') = -D\partial_x G + \{v + 2D\partial_x \log[u(x,t)]\}G.$$

Here (x', t') is thought of as the initial condition and (x, t) as the finial condition, so this is the forward equation

$$\partial_t G = D\partial_x^2 G - \partial_x(\{v + 2D\partial_x \log[u(x,t)]\}G) \tag{8}$$

and the drift in the diffusion process is $v + 2D\partial_x \log[c(x,t)]$.

As we will now show, a closely related equation "follows" from Theorem 4. We used quotation marks because we use the nonexistent Itô's formula for the SPDE in Theorem 4 and apply it to the function $f(\ell, u) = \ell/u$ which is not continuous at $(0, 0)$. Suppose (u, ℓ) solves the SPDE and let $\rho = \ell/u$. By calculus and Theorem 4,

$$\begin{aligned}\partial_t \rho = & \frac{1}{u}\Big[\alpha\,\Delta\ell + 2\theta\,\beta\,\ell\,(1-u) + |4\gamma\ell(1-u)|^{1/2}\,\dot W^0 + |4\gamma\ell(u-\ell)|^{1/2}\,\dot W^2\Big] \\ & - \frac{\ell}{u^2}\Big[\alpha\,\Delta u + 2\theta\,\beta\,u(1-u) + |4\gamma\ell(1-u)|^{1/2}\,\dot W^0 + |4\gamma(u-\ell)(1-u)|^{1/2}\,\dot W^1\Big].\end{aligned}$$

The terms involving $\theta\beta$ cancel. To combine the Laplacian terms we use the formula

$$\Delta\left(\frac{\ell}{u}\right) = \frac{\Delta\ell}{u} - \frac{\ell\,\Delta u}{u^2} - 2\,\frac{\partial_x u}{u}\cdot\partial_x\left(\frac{\ell}{u}\right).$$

To add up the noises we note that since the W^i are independent, the variances add up to

$$\begin{aligned} & \frac{(u-\ell)^2}{u^4}\ell(1-u) + \frac{\ell(u-\ell)}{u^2} + \frac{\ell^4(u-\ell)(1-u)}{u^4} \\ = & \frac{\ell(u-\ell)}{u^2}\left[\frac{u-\ell)(1-u)}{u^2} + \frac{u^2}{u^2} + \frac{\ell(1-u)}{u^2}\right] = \frac{\ell(u-\ell)}{u^3}\end{aligned}$$

since $(u-\ell)(1-u) + \ell(1-u) + u^2 = u(1-u) + u^2 = u$. Combining our calculations,

$$\partial_t \rho \;=\; \alpha\,\Delta\rho + 2\alpha\,\partial_x \log u\cdot\partial_x\rho \;+\; |4\gamma\,\rho\,(1-\rho)/u|^{1/2}\,\dot W \tag{9}$$

for some white noise $\dot W$. To compare (9) with (8), note that their $D = \alpha$, they work in a moving reference frame and their equation is for a fixed realization of the total population size; while ours is in a fixed reference frame and does not condition on $u(t, x)$ and hence retains the fluctuation term $|4\gamma\,\rho\,(1-\rho)/u|^{1/2}\,\dot W$.

Equations (8) and (9) both contain drift terms of the form $\partial_x \log[u(x, t)]$. It is a well-known fact that solutions of the SPDE in (4) are Hölder continuous with exponent $1/2 - \epsilon$ in space and $1/4 - \epsilon$ in time, so it is not clear how to make sense of these equations. The fact that $\eta_t(x) \le \xi_t(x)$ means that $\ell \le u$, so in computing ℓ/u we will never divide a positive number by 0. However solutions to u have compact support [31], so it is not clear if the ratio of densities ℓ^n/u^n will be tight.

Forgetting about technical problems and returning to the realm of physical intuition, the drift in equation (8) has two competing parts. The first term v tends to push the lineage into the tip of the wave and is a consequence of the moving frame of reference. In terms of our earlier heuristics, for lineages far from the front there are few opportunities for branching so the genealogies perform Brownian motions while (as we move backwards in time) the front catches up. The second term reflects the fact that going forward in time, particles near the front move from regions of high density to low density due to branching. Hence genealogic lineages drift from the wavefront back toward the center of the wave.

To visualize the movement of the particle in the moving reference frame we introduce the time dependent potential

$$V(x,t) = -vx - 2\alpha \log u(x,t)$$

and note, as [20] do in their formula (A.3), that the drift in (8) is $\partial_x V(x,t)$. To analyze the asymptotic behavior of the motion of genealogies near the boundary, it is natural to assume that (in the moving frame of reference) we are in steady state

$$u(x,t) \approx u_{st}(x).$$

Muller and Sowers [31] have shown that if the constant γ is small enough then the stationary state exists and is the limit for all initial states that are $= 1$ for $x \le a$ and $= 0$ for $x \ge b$, for some $a, b \in \mathbb{R}$. This and the previous calculation suggest that the steady state location of the genealogy (in the moving reference frame) is

$$P_{st}(x) = \lim_{t' \to \infty} G(x,0|x',t') = c\, u_{st}(x)^2 \exp(vx/D),$$

where c is a normalization constant, see (6) in [20]. In appendix C, [20] consider the special case of a deterministic wave u and show that unfortunately, even for the slowest wave speed, $2\sqrt{Ds}$, $u_{st}(x)^2 \exp(vx/D)$ is not integrable. In general, knowing u_{st} is not enough to compute P_{st}, one must also consider the evolutionary history that brought us to the current state. Making sense of equation (8) for the ancestral lineages seems difficult, but since these lineages are embedded in the limit of the branching coalescing random walk their behavior cannot be too pathological.

1.5 Concluding remarks

Theorem 1 is a special case of Theorem 4. Theorem 3 is proved in Section 2. Most of the difficulty in the proof of Theorem 2 in Durrett and Restrepo is due to the generality of the interaction kernel. Since spatial movement of genealogies in our process is a simple random walk, the result here can be proved easily using the argument in Section 2 of [14].

The rest of the paper is devoted to the proof for Theorem 4. In Section 3 we convert the stochastic integral representations in (2) and (7) into approximate martingale problems. This is now a common approach in the study of scaling limits of particle systems, see [13, 10, 12]. The calculations for the u equation are almost identical to those in Section 3 of Mueller and Tribe [32], but some minor changes are needed to study the joint distribution (ℓ, u).

In Sections 5–7 we prove tightness. Again many of the ideas come from [32], but since they only write out the details for their contact process limit theorem, and we have to prove the joint convergence, we have written out the details. See Kleim [25] for a recent example of convergence of rescaled Lotka-Volterra models to a one-dimensional SPDE, this time with a cubic drift term. The main ideas of the tightness proof are given in Section 5. Two lemmas that require a lot of computation are proved in Section 6. Some nonstandard random walk estimates are proved in Section 7. Finally Lemma 2, which establishes distributional uniqueness for the coupled SPDE by using a duality based on (6), is proved in Section 8.

Acknowledgement. We would like to thank Carl Mueller and Edwin Perkins for their help while we were writing this paper.

2 Proof of Theorem 3

Let S_t^n be a random walk that jumps from x to $x \pm 1/L_n$ at rate $2r_n M_n$. The reader should think of the difference of the location of two genealogies, but we allow the two genealogies to move independently even after they hit. There are M_n cells at each deme and $\tilde{P}_{xy}$ has rate θR_n^{-1}. So the branching rate at each deme is $\theta M_n \, R_n^{-1} \to \theta\beta$. On other hand, an elementary computation shows that

$$|S_t^n| - \frac{4r_n M_n}{L_n} \int_0^t 1_{(S_s^n=0)} \, ds$$

is martingale. As $t \to \infty$, $|S_t^n|$ converges to the absolute value of a Brownian motion B_t with variance $4\alpha t$, so for reasons explained in Section 2 of [14], the second term converges to $L_0(t)$, the local time at 0 for the limiting Brownian motion, which is defined by the property that $|B_t| - L_0(t)$ is a martingale.

The sojurn times at 0 are independent and exponential with rate $4r_n M_n$, so the number of visits to 0 up to time t

$$N_t^n \sim 4r_n M_n \int_0^t 1_{(S_s^n=0)} \, ds$$

and it follows that $N_t^n/L_n \to L_0(t)$. On each visit to 0, the two particles have a probability $1/M_n$ to coalesce. Our assumptions imply that

$$\frac{M_n}{L_n} \cdot \frac{r_n}{L_n} \to \alpha \quad \text{so} \quad \frac{L_n}{M_n} \to \frac{\gamma}{\alpha}$$

and the desired result follows.

3 Approximate Martingale Problems

For simplicity we drop the subscript n's on M, L, R and r. We leave the superscript n in u_t^n and ℓ_t^n to distinguish the approximating processes from their limits. We write

$$\langle f, g \rangle := \frac{1}{L} \sum_{w \in L^{-1}\mathbb{Z}} f(w) g(w)$$

whenever it is well-defined and adopt the convention that $\phi(x) := \phi(w)$ when $x = (w, i)$.

3.1 Type 1 particles

Let $\phi : [0, \infty) \times L^{-1}\mathbb{Z} \to \mathbb{R}$ be such that $t \mapsto \phi_t(x)$ is continuously differentiable and $\int_0^T \langle |\phi_s| + \phi_s^2 + |\partial_s \phi_s|, 1 \rangle \, ds < \infty$. Applying integration by parts to $\xi_t(x)\phi_t(x)$, using (2), and summing over x, we obtain for all $t \in [0.T]$,

$$\begin{aligned}\langle u^n_t,\, \phi_t\rangle - \langle u^n_0,\, \phi_0\rangle &- \int_0^t \langle u^n_s, \partial_s \phi_s\rangle ds \\ &= (ML)^{-1} \sum_x \sum_{y\sim x} \int_0^t (\xi_{s-}(y) - \xi_{s-}(x))\phi_s(x)\, dP^{x,y}_s \qquad (10)\\ &+ (ML)^{-1} \sum_x \sum_{y\sim x} \int_0^t \xi_{s-}(y)(1-\xi_{s-}(x))\phi_s(x)\, d\tilde{P}^{x,y}_s. \qquad (11)\end{aligned}$$

Drift term. We break (11) into an average term and a fluctuation term

$$(ML)^{-1} \sum_x \sum_{y\sim x} \int_0^t \xi_{s-}(y)(1-\xi_{s-}(x))\phi_s(x)\, \theta R^{-1}\, ds \qquad (12)$$

$$+ (ML)^{-1} \sum_x \sum_{y\sim x} \int_0^t \xi_{s-}(y)(1-\xi_{s-}(x))\phi_s(x)\, (d\tilde{P}^{x,y}_s - \theta R^{-1}\, ds). \qquad (13)$$

Recalling the definition of the density, (12) becomes

$$\theta \cdot \frac{M}{R} \cdot \frac{1}{L} \sum_{w\in L^{-1}\mathbb{Z}} \int_0^t [u^n_{s-}(w-L^{-1}) + u^n_{s-}(w+L^{-1})](1-u^n_{s-}(w))\phi_s(w)\, ds$$

Since $M/R \to \beta$, this converges to

$$\theta\beta \int_0^t \int_{\mathbb{R}} 2u_s(w)(1-u_s(w))\phi_s(w)\, dw\, ds$$

as $n \to \infty$. Here and in what follows the claimed convergences follow once we have proved C-tightness. We have established convergence of finite dimensional distributions so the sequences of processes converge, and we can use Skorokhod's theorem to show they converge almost surely.

The second term (13) is a martingale $E^{(2)}_t(\phi)$ with

$$\begin{aligned}\langle E^{(2)}(\phi)\rangle_t &\le \frac{\theta}{R(ML)^2} \sum_x \sum_{y\sim x} \int_0^t \phi^2_s(x)\, ds \\ &\le \frac{2\theta}{LR} \int_0^t \langle 1,\, \phi^2_s\rangle\, ds \to 0 \qquad \text{since } LR \to \infty.\end{aligned}$$

White noises. We can rewrite (10) as

$$(ML)^{-1} \sum_x \sum_{y\sim x} \int_0^t \{\xi_{s-}(y)[1-\xi_{s-}(x)] - \xi_{s-}(x)[1-\xi_{s-}(y)\}\phi_s(x)\, dP^{x,y}_s. \qquad (14)$$

We now rewrite the integrand as

$$\xi_{s-}(y)[1-\xi_{s-}(x)]\phi_s(y) - \xi_{s-}(x)[1-\xi_{s-}(y)]\phi_s(x) \qquad (15)$$

$$+ \xi_{s-}(y)[1-\xi_{s-}(x)](\phi_s(x) - \phi_s(y)). \qquad (16)$$

We work first with (15). Interchanging the roles of x and y in the double sum, this part of (14) becomes

$$(ML)^{-1} \sum_x \sum_{y \sim x} \int_0^t [1 - \xi_{s-}(y)] \xi_{s-}(x) \phi_s(x) \, (dP_s^{y,x} - dP_s^{x,y}). \tag{17}$$

To prepare for treating the joint SPDE we split (17) into $Z_t^0(\phi) + Z_t^1(\phi)$ where

$$Z_t^0(\phi) = (ML)^{-1} \sum_x \sum_{y \sim x} \int_0^t [1 - \xi_{s-}(y) \eta_{s-}(x) \phi_s(x) \, (dP_s^{y,x} - dP_s^{x,y}), \tag{18}$$

$$Z_t^1(\phi) = (ML)^{-1} \sum_x \sum_{y \sim x} \int_0^t [1 - \xi_{s-}(y)](\xi_{s-}(x) - \eta_{s-}(x)) \phi_s(x) \, (dP_s^{y,x} - dP_s^{x,y}). \tag{19}$$

These two martingale terms use the same Poisson processes but the product of their integrands is 0 (since $\iota_t(x)\eta_t(x)$ vanishes), so they are uncorrelated.

The variance process $\langle P^{x,y} - P^{y,x} \rangle_t = 2rt$. Hence ignoring the difference between $\phi_s(x)$ and $\phi_s(y)$, we have

$$\langle Z^0(\phi) \rangle_t = 4rL^{-2} \int_0^t \sum_{w \in L^{-1}\mathbb{Z}} \ell_{s-}^n(w)(1 - u_{s-}^n(w)) \phi_s(w)^2 \, ds + o(1)$$

which converges to

$$4\gamma \int_0^t \int_{\mathbb{R}} \ell_s(w)(1 - u_s(w)) \phi_s(w)^2 \, dw \, ds \quad \text{since } r/L \to \gamma.$$

Similarly, $\langle Z^1(\phi) \rangle_t$ converges to

$$4\gamma \int_0^t \int_{\mathbb{R}} (u_s(w) - \ell_s(w)) \, (1 - u_s(w)) \phi_s(w)^2 \, dw \, ds.$$

Laplacian Term. We denote the discrete gradient and the discrete Laplacian respectively by

$$\nabla_L f(w) := L \Big(f(w + L^{-1}) - f(w) \Big) \tag{20}$$

$$\Delta_L f(w) := L^2 \Big(f(w + L^{-1}) + f(w - L^{-1}) - 2f(w) \Big). \tag{21}$$

We break (16) into an average term and a fluctuation term

$$(ML)^{-1} \sum_x \sum_{y \sim x} \int_0^t \xi_{s-}(y) \xi_{s-}^c(x) [\phi_s(x) - \phi_s(y)] \, r \, ds \tag{22}$$

$$+ (ML)^{-1} \sum_x \sum_{y \sim x} \int_0^t \xi_{s-}(y) \xi_{s-}^c(x) [\phi_s(y) - \phi_s(x)] (dP_s^{x,y} - r \, ds). \tag{23}$$

We can replace ξ^c_{s-} by 1 in (22) without changing its value. Doing the double sum over y and then over $x \sim y$ the above is

$$\frac{rM}{L^2} \cdot \frac{1}{L} \sum_w u^n_{s-}(w) \Delta_L \phi_s(w) = \frac{rM}{L^2} \langle u^n_{s-},\, \Delta_L \phi_s \rangle.$$

By assumption $rM/L^2 \to \alpha$, so this term converges to $\alpha \int_{\mathbb{R}} u_s \, \Delta\phi_s$. The other term, (23), is a martingale $E^{(1)}_t(\phi)$ with

$$\langle E^{(1)}(\phi) \rangle_t \le \frac{r}{(ML)^2} \sum_x \sum_{y \sim x} \int_0^t (\phi_s(x) - \phi_s(y))^2 \, ds \tag{24}$$

$$= \frac{2r}{L^3} \int_0^t \langle 1,\, |\nabla_L \phi_s|^2 \rangle \, ds \to 0 \tag{25}$$

since $r/L \to \gamma$ and $L \to \infty$.

Combining our calculations, we see that in the limit $n \to \infty$,

$$\int_{\mathbb{R}} u_t(w)\, \phi_t(w)\, dw - u_0(w)\, \phi_0(w)\, dw - \int_0^t \int_{\mathbb{R}} u_s(w)\, \partial_s \phi_s(w)\, dw\, ds \tag{26}$$

$$- \int_0^t \int_{\mathbb{R}} \alpha u_s(w) \Delta \phi_s(w) - 2\theta\beta\, u_s(w)(1 - u_s(w))\, \phi_s(w)\, dw\, ds \tag{27}$$

is a martingale with quadratic variation

$$4\gamma \int_0^t \int_{\mathbb{R}} u_s(w)(1 - u_s(w)) \phi_s(w)\, dw\, ds$$

which is the martingale problem formulation of (4).

3.2 Labeled Particles

Arguing as for the type 1 particles while using (7) instead of (2), we get

$$\langle \ell^n_t, \phi_t \rangle - \langle \ell^n_0, \phi_0 \rangle - \int_0^t \langle \ell^n_s, \partial_s \phi_s \rangle$$

$$= (ML)^{-1} \sum_x \sum_{y \sim x} \int_0^t (\eta_{s-}(y) - \eta_{s-}(x)) \phi_s(x)\, dP^{x,y}_s \tag{28}$$

$$+ (ML)^{-1} \sum_x \sum_{y \sim x} \int_0^t [\eta_{s-}(y) - \xi_{s-}(y) \eta_{s-}(x)] \phi_s(x)\, d\tilde{P}^{x,y}_s, \tag{29}$$

where we have simplified the second term of (7) using $\xi_{s-}(y)\eta_{s-}(y) = \eta_{s-}(y)$.

Drift term. Breaking the second term (29) into an average term and a fluctuation term as before, we conclude that as $n \to \infty$, the average term is

$$\frac{\theta M}{L R_n} \sum_{w \in L^{-1}\mathbb{Z}} \int_0^t \big[\ell^n_s(w + L^{-1}) - u^n_s(w + L^{-1}) \ell^n_s(w)$$

$$+ \ell^n_s(w - L^{-1}) - u^n_s(w - L^{-1}) \ell^n_s(w) \big] \phi_s(w)\, ds$$

which converges to

$$\theta\beta \int_0^t \int_{\mathbb{R}} 2\ell_s(w)(1-u_s(w))\phi_s(w)\, dw\, ds.$$

White noises. We again change the integrand in (28) to

$$\{\eta_{s-}(y)[1-\eta_{s-}(x)] - \eta_{s-}(x)[1-\eta_{s-}(y)]\}\, \phi_s(x)$$

and then split it into two parts as in (15) and (16). That is, we rewrite the integrand as

$$\eta_{s-}(y)[1-\eta_{s-}(x)]\eta_s(y) - \eta_{s-}(x)[1-\eta_{s-}(y)]\phi_s(x) \tag{30}$$

$$+\, \eta_{s-}(y)[1-\eta_{s-}(x)](\phi_s(x)-\phi_s(y)). \tag{31}$$

Arguing as in the previous, we obtain the following sum coming from (30).

$$(ML)^{-1}\sum_x \sum_{y\sim x} \int_0^t [1-\xi_{s-}(y)]\eta_{s-}(x)\phi_s(x)\, (dP_s^{y,x} - dP_s^{x,y})$$

$$+(ML)^{-1}\sum_x \sum_{y\sim x} \int_0^t [\xi_{s-}(y)-\eta_{s-}(y)]\eta_{s-}(x)\phi_s(x)\, (dP_s^{y,x} - dP_s^{x,y}). \tag{32}$$

The first noise is the same as $Z_t^0(\phi)$ in (18) while the second noise, denoted by $Z_t^2(\phi)$, has variance converging to

$$4\gamma \int_0^t \int_{\mathbb{R}} \ell_s(w)\, (u_s(w)-\ell_s(w))\, \phi_s(w)^2\, dw\, ds.$$

The product of any two of the three integrands in $Z_t^i(\phi)$ $(i=0,1,2)$ is 0, so these three martingales are uncorrelated.

Laplacian term. Breaking (31) into an average term and a fluctuation term as was done for (16), we see that as $n\to\infty$, the average term

$$\frac{rM}{L^2}\langle \ell_{s-}^n, \Delta_L \phi_s\rangle \to \alpha \int_{\mathbb{R}} \ell_s\, \Delta\phi_s.$$

Combining our calculations, we see that in the limit, for any ϕ, $\psi \in C_c^{1,2}([0,\infty)\times\mathbb{R})$,

$$\begin{aligned}
&\int_{\mathbb{R}} (u_t\,\phi_t - u_0\,\phi_0 + \ell_t\,\psi_t - \ell_0\,\psi_0)\, dw \\
&- \alpha \int_0^t \int_{\mathbb{R}} u_s\,(\partial_s\phi_s + \Delta\phi_s) + \ell_s\,(\partial_s\psi_s + \Delta\psi_s)\, dw\, ds \\
&- 2\,\theta\,\beta \int_0^t \int_{\mathbb{R}} u_s\,(1-u_s)\,\phi_s \;+\; \ell_s\,(1-u_s)\,\psi_s\, dw\, ds
\end{aligned} \tag{33}$$

is a continuous martingale with quadratic variation

$$4\,\gamma \int_0^t \int_{\mathbb{R}} u_s\,(1-u_s)\,\phi_s^2 + \ell_s\,(1-\ell_s)\,\psi_s^2 \;+\; 2\,\phi_s\,\psi_s\,\ell_s\,(1-u_s)\, dw\, ds. \tag{34}$$

Any sub-sequential limit $(u,\,\ell)$ solves this martingale problem. It is standard (see p. 536-537 in [32]) to show that $(u,\,\ell)$ then solves the coupled SPDE in Theorem 4 weakly, with respect to some white noises.

4 Green's function representation

As remarked earlier our proof follows the approach in [32]. The first step is to prove the analogue of their (2.11). Observe that $u_t^n(z) = \langle u_t^n,\, L\,\mathbf{1}_z\rangle$ and $\ell_t^n(z) = \langle \ell_t^n,\, L\,\mathbf{1}_z\rangle$, where $\mathbf{1}_z$ is the function on $L^{-1}\mathbb{Z}$ which is 1 at z and zero elsewhere. Let $\alpha_n = r_n M L^{-2}$ which converges to α as $n \to \infty$ and let

$$p_t^n(w) := L\,\mathbb{P}(X_t^n = w \,|\, X_0^n = 0) \tag{35}$$

be the transition probability of the simple random walk $(X_t^{(n)})_{t\geq 0}$ on $L^{-1}\mathbb{Z}$ with jump rate $2L^2$, so that it converges to $p_t(w)$ the transition density of Brownian motion run at rate 2. Let $\{P_t^n\}_{t\geq 0}$ be the associated semigroup which has generator the discrete Laplacian Δ_L defined in (21).

Applying the approximate martingale problems with test function

$$\phi_s^{t,z}(w) = \begin{cases} p^n_{\alpha_n(t-s)}(w-z) & \text{for } s \in [0,t], \\ 0 & \text{otherwise} \end{cases} \tag{36}$$

and using the facts that $\partial_s \phi_s + \alpha_n \Delta_L \phi_s = 0$ and $\langle u_0^n, \phi_0^{t,z}\rangle = P^n_{\alpha_n t} u_0^n(z)$, we have

$$u_t^n(z) = P^n_{\alpha_n t} u_0^n(z) + Y_t(\phi) + Z_t(\phi) + E_t^1(\phi) + E_t^2(\phi) \tag{37}$$

for $t \geq 0$ and $z \in L^{-1}\mathbb{Z}$. Here $Z_t(\phi)$, $E_t^{(1)}(\phi)$ and $E_t^{(2)}(\phi)$ are martingales defined in (17), (23) and (13) respectively. To describe the other term, we let $\beta_n = MR^{-1}$ which converges to β, and let

$$Y_t(\phi) := \frac{\theta\beta_n}{L}\int_0^t \sum_{w\in L^{-1}\mathbb{Z}} [u_s^n(w-L^{-1}) + u_s^n(w+L^{-1})](1-u_s^n(w))\phi_s(w)\,ds. \tag{38}$$

Repeating the last argument for the unlabeled particles,

$$\ell_t^n(z) = P^n_{\alpha_n t}\ell_0^n(z) + Y_t^\ell(\phi) + Z_t^\ell(\phi) + E_t^{\ell,1}(\phi) + E_t^{\ell,2}(\phi). \tag{39}$$

Here $Z_t^{(\ell)}(\phi) := Z_t^0(\phi) + Z_t^2(\phi)$ is given by (32), $E_t^{(\ell,2)}(\phi)$ is obtained by replacing $(\xi,\, \xi^c)$ by $(\eta,\, \eta^c)$ in (23), $E_t^{(\ell,2)}(\phi)$ is the fluctuation term corresponding to (29). The remaining term is

$$Y_t^\ell(\phi) = \frac{\theta\beta_n}{L}\int_0^t \sum_{w\in L^{-1}\mathbb{Z}} \Big[\ell_s^n(w+L^{-1}) - u_s^n(w+L^{-1})\ell_s^n(w) + \ell_s^n(w-L^{-1}) - u_s^n(w-L^{-1})\ell_s^n(w)\Big]\,\phi_s(w)\,ds.$$

5 Tightness

Recall that a sequence of probability measures is said to be C-tight, if it is tight in D and any subsequential limit has a continuous version. The goal of this section is to prove:

Theorem 5. *Suppose the assumptions in Theorem 1 hold. Then the sequence* $\{(u^n,\,\ell^n)\}_{n\geq 1}$ *is* C*-tight in* $D([0,T],\,C_b(\mathbb{R})\times C_b(\mathbb{R}))$ *for every* $T>0$.

Proof. The compact containment condition (condition (a) of Theorem 7.2 in [16, Chapter 3]) holds trivially as $0\leq \ell^n\leq u^n\leq 1$. Since $C_b(\mathbb{R})\times C_b(\mathbb{R})$ is equipped with the product metric, the desired C-tightness follows once we can show that for any $\epsilon>0$, one has

$$\lim_{\delta\to 0}\limsup_{n\to\infty}\mathbb{P}\left(\sup_{\substack{t_1-t_2<\delta\\ 0\leq t_2\leq t_1\leq T}}\left\|u^n_{t_1}-u^n_{t_2}\right\|>\epsilon\right)=0, \tag{40}$$

$$\lim_{\delta\to 0}\limsup_{n\to\infty}\mathbb{P}\left(\sup_{\substack{t_1-t_2<\delta\\ 0\leq t_2\leq t_1\leq T}}\left\|\ell^n_{t_1}-\ell^n_{t_2}\right\|>\epsilon\right)=0. \tag{41}$$

Here and in what follows the norm is the one defined in (3). It is enough to show that (40) holds with u^n replaced by any term in the decomposition given in (37), and that (41) holds with ℓ^n replaced by any term in (39).

First term in (37) **and** (39)**.** By standard coupling arguments for simple random walk, we can check as in Lemma 7(b) of [32] that, upon linearly interpolating $P^n_{\alpha_n t}u^n_0(z)$ in space, we have

$$\sup_{t\in[0,T]}\|P^n_{\alpha_n t}u^n_0-P_{\alpha_n t}f_0\|\to 0\quad\text{as } n\to\infty, \tag{42}$$

where $\{P_t\}_{t\geq 0}$ is the semigroup for the Brownian motion in $\mathbb{R}$ running at rate 2. This implies, by the continuity of the semigroup P_t, that (40) holds with u^n_t replaced by $P^n_{\alpha_n t}u^n_0$. By the same reasoning, we have

$$\sup_{t\in[0,T]}\|P^n_{\alpha_n t}\ell^n_0-P_{\alpha_n t}g_0\|\to 0\quad\text{as } n\to\infty \tag{43}$$

Remaining terms in (37) **and** (39)**.** For simplicity, we write

$$\begin{aligned}\widehat{u}_t(z)&:=Y_t(\phi)+Z_t(\phi)+E^{(1)}_t(\phi)+E^{(2)}_t(\phi),\\ \widehat{\ell}^n_t(z)&:=Y^{(\ell)}_t(\phi)+Z^{(\ell)}_t(\phi)+E^{(\ell,1)}_t(\phi)+E^{(\ell,2)}_t(\phi).\end{aligned}$$

The next moment estimate for space and time increments is similar to Lemma 6 in [32], but ours implies Hölder continuity of the limits with exponent $<1/2$ in space and $<1/4$ in time.

Lemma 3. *For any* $p\geq 2$ *and* $T\geq 0$*, there exists a constant* $C(T,p)>0$ *such that*

$$\mathbb{E}|\widehat{u}^n_{t_1}(z_1)-\widehat{u}^n_{t_2}(z_2)|^p\leq C_{T,p}\left(|t_1-t_2|^{p/4}+|z_1-z_2|^{p/2}+M^{-p}\right) \tag{44}$$

$$\mathbb{E}|\widehat{\ell}^n_{t_1}(z_1)-\widehat{\ell}^n_{t_2}(z_2)|^p\leq C_{T,p}\left(|t_1-t_2|^{p/4}+|z_1-z_2|^{p/2}+M^{-p}\right) \tag{45}$$

for all $0\leq t_2\leq t_1\leq T$*,* z_1*,* $z_2\in L^{-1}\mathbb{Z}$ *and* $n\geq 1$.

The proof of this result is postponed to the next section since it requires a number of computations. We now argue that (44) implies (40) holds for $\widehat{u}^n$. This idea is described in the paragraph before Lemma 7 in [32] and page 648 of [25]: we approximate the càdlàg process $\widehat{u}^n$ by a *continuous* process $\tilde{u}$ and invoke a tightness criterion inspired by Kolmogorov's continuity theorem.

Lemma 4. *Define $\tilde{u}^n \in C([0,\infty), C_b(\mathbb{R}))$ by $\tilde{u}_t = \widehat{u}_t$ on the grid $t \in \theta_n \mathbb{Z}_+$ and then linearly interpolate in t for each $w \in \mathbb{R}$. Suppose $\theta_n > M^{-4}$ and $\lim_{n\to\infty} \theta_n = 0$. Then there exists $n_0 \in \mathbb{N}$ such that for any $p \geq 2$, $T \geq 0$ and $K \geq 0$,*

$$\mathbb{E}|\tilde{u}^n_{t_1}(z_1) - \tilde{u}^n_{t_2}(z_2)|^p \leq C_{T,p,K}\Big(|t_1 - t_2|^{p/4} + |z_1 - z_2|^{p/2}\Big), \tag{46}$$

for all $0 \leq t_2 \leq t_1 \leq T$, $z_i \in \mathbb{R}$ with $|z_i| \leq K$ ($i = 1, 2$) and $n \geq n_0$.

By a standard argument (see, for example, Problems 2.2.9 and 2.4.11 of [23]), one can show that (46) implies that (40) holds when u is replaced by $\tilde{u}^n$. Finally, by the reasoning in the proof of Lemma 7(a) in [32], there is a $\sigma > 0$ so that

$$\limsup_{n\to\infty} \mathbb{P}\left(\sup_{t\in[0,T]} \|\tilde{u}^n_t - \widehat{u}^n_t\| \geq n^{-\sigma}\right) = 0.$$

Therefore (40) holds for $\widehat{u}^n$. By the same argument, (45) implies (41) holds for $\widehat{\ell}^n$.

The proof for Theorem 5 will be complete once Lemmas 3 and 4 are proved. □

6 Proofs of Lemmas 3 and 4

Proof of Lemma 3. We prove only (44) for unlabeled particles. The proof of (45) for labeled particles is similar. The basic ingredients are the following estimates of time and space increments of the transition probability of simple random walk. Namely, there exists a constant $C > 0$ *independent of* n such that

$$0 \leq \int_0^T p^n_s(0)\, ds \leq C\sqrt{T}, \tag{47}$$

$$0 \leq \int_0^\infty p^n_s(0) - p^n_{s+\theta}(0)\, ds \leq C\sqrt{\theta}, \tag{48}$$

$$0 \leq \int_0^\infty p^n_s(0) - p^n_s(z)\, ds \leq C\,|z| \tag{49}$$

$$\int_0^T \frac{1}{L}\sum_w |p^n_{s+\theta}(w) - p^n_s(w)|\, ds \ \leq\ C\sqrt{T\theta}, \tag{50}$$

$$\int_0^T \frac{1}{L}\sum_w |p^n_s(w) - p^n_s(z+w)|\, ds \ \leq\ C\sqrt{T}\,|z| \tag{51}$$

for $\theta \geq 0$, $z \in L^{-1}\mathbb{Z}$ and $T > 0$. These estimates can be either found or deduced from the standard methods described in Chapter 2 in [29]. For completeness, we give precise references and missing details in the next section.

We will show that each of the four terms of $\widehat{u}^n$ satisfies (44). To simplify notation, we assume, without loss of generality for the proof, that $\alpha_n \equiv 1$. First, we deal with the process Y that has no jumps. To reduce the size of the formulas we let

$$v^n_s(w) = [u^n_s(w - L^{-1}) + u^n_s(w + L^{-1})](1 - u^n_s(w)).$$

Using the definitions of Y (38) and of our test function (36) we have for $t_1 > t_2$

$$
\begin{aligned}
Y_{t_1}(\phi^{t_1,z_1}) - Y_{t_2}(\phi^{t_2,z_2}) &= \theta\beta_n \int_{t_2}^{t_1} \frac{1}{L}\sum_w v_s^n(w) p_{t_1-s}^n(z_1-w)\,ds \\
&\quad + \theta\beta_n \int_0^{t_2} \frac{1}{L}\sum_w v_s^n(w)[p_{t_1-s}^n(z_1-w) - p_{t_2-s}^n(z_2-w)]\,ds \\
&\equiv \theta\beta_n(\Theta_1(Y) + \Theta_2(Y)).
\end{aligned}
$$

The sums are over $w \in L^{-1}\mathbb{Z}$, which we have omitted to simplify notation. Since $0 \le u^n \le 1$ and $L^{-1}\sum_w p_t^n(z-w) = 1$ (recall the definition in (35)), we have

$$
\mathbb{E}|\Theta_1(Y)|^p \le 2^p\,(t_1-t_2)^p. \tag{52}
$$

By the triangle inequality and the translation invariance and symmetry of the transition density, we obtain

$$
\begin{aligned}
&\int_0^{t_2} \frac{1}{L}\sum_w |p_{t_1-s}^n(z_1-w) - p_{t_2-s}^n(z_2-w)|\,ds \\
&\le \int_0^{t_2} \frac{1}{L}\sum_w \left(|p_{t_1-s}^n(z_1-z_2+w) - p_{t_1-s}^n(w)| + |p_{t_1-s}^n(w) - p_{t_2-s}^n(w)|\right) ds \\
&\le C_T|z_1-z_2| + C\sqrt{t_1-t_2}
\end{aligned}
$$

by (51) and (50). Since $0 \le v^n \le 2$,

$$
\mathbb{E}\big[\,|\Theta_2(Y)|^p\,\big] \le C_{T,p}\left(\sqrt{t_1-t_2} + |z_1-z_2|\right)^p \tag{53}
$$

for all $0 \le t_2 \le t_1 \le T$, z_1, $z_2 \in L^{-1}\mathbb{Z}$ and $n \ge 1$.

It remains to consider $E_t^{(1)}(\phi)$, $E_t^{(2)}(\phi)$ and $Z_t(\phi)$. Note that for each of them, the largest possible jump is bounded almost surely by

$$
2(ML)^{-1}\sup_{s\ge 0}\|\phi_s\|_\infty \le 2\,M^{-1}.
$$

We shall employ a version of the Burkholder-Davis-Gundy inequality stated at the bottom of page 527 in [32]. Namely, for any càdlàg martingale X with $X_0 = 0$ and for $p \ge 2$,

$$
\mathbb{E}[\sup_{s\in[0,t]}|X_s|^p\,] \le C(p)\,\mathbb{E}\Big[\langle X\rangle_t^{p/2} + \sup_{s\in[0,t]}|X_s - X_{s-}|^p\Big], \quad t \ge 0. \tag{54}
$$

Writing $N_s^{x,y}$ for the compensated Poisson process $P_s^{x,y} - r_n s$, and $\sum_{x,y\sim x}$ for the double sum $\sum_x\sum_{y\sim x}$, we decompose

$$
\begin{aligned}
&E_{t_1}^{(1)}(\phi^{t_1,z_1}) - E_{t_2}^{(1)}(\phi^{t_2,z_2}) \\
&\qquad = \frac{1}{ML}\sum_{x,y\sim x}\int_{t_2}^{t_1} \xi_{s-}(y)\xi_{s-}(x)\left(p_{t_1-s}^n(z_1-y) - p_{t_1-s}^n(z_1-x)\right) dN_s^{x,y} \\
&\qquad + \frac{1}{ML}\sum_{x\sim y}\int_0^{t_2} \xi_{s-}(y)\xi_{s-}(x)\big(p_{t_1-s}^n(z_1-y) - p_{t_1-s}^n(z_1-x) \\
&\qquad\qquad\qquad - p_{t_2-s}^n(z_2-y) + p_{t_2-s}^n(z_2-x)\big)\,dN_s^{x,y} \\
&\qquad \equiv \Theta_1(E^{(1)}) \,+\, \Theta_2(E^{(1)}).
\end{aligned}
$$

Writing $\tilde{N}_s^{x,y}$ for the compensated Poisson process $\tilde{P}_s^{x,y} - \theta R_n^{-1} s$ we have

$$
\begin{aligned}
&E_{t_1}^{(2)}(\phi^{t_1,z_1}) - E_{t_2}^{(2)}(\phi^{t_2,z_2}) \\
&\qquad = \frac{1}{ML} \sum_{x,y\sim x} \int_{t_2}^{t_1} \xi_{s-}(y)(1-\xi_{s-}(x))\, p_{t_1-s}^n(z_1-x)\big)\, d\tilde{N}_s^{x,y} \\
&\qquad + \frac{1}{ML} \sum_{x\sim y} \int_0^{t_2} \xi_{s-}(y)(1-\xi_{s-}(x))[p_{t_1-s}^n(z_1-x) - p_{t_2-s}^n(z_2-x)]\, d\tilde{N}_s^{x,y} \\
&\qquad \equiv \Theta_1(E^{(1)}) \,+\, \Theta_2(E^{(1)}).
\end{aligned}
$$

Finally, we have

$$
\begin{aligned}
&Z_{t_1}(\phi^{t_1,z_1}) - Z_{t_2}(\phi^{t_2,z_2}) \\
&= \frac{1}{ML} \sum_{x,y\sim x} \int_{t_2}^{t_1} (1-\xi_{s-}(y))\xi_{s-}(x) p_{t_1-s}^n(z_1-x)\, d(N_s^{y,x} - N_s^{x,y}) \\
&+ \frac{1}{ML} \sum_{x,y\sim x} \int_0^{t_2} (1-\xi_{s-}(y))\xi_{s-}(x)[p_{t_1-s}^n(z_1-x) - p_{t_2-s}^n(z_2-x)]\, d(N_s^{y,x} - N_s^{x,y}) \\
&\equiv \Theta_1(Z) \,+\, \Theta_2(Z).
\end{aligned}
$$

Once we use $0 \le \xi \le 1$ to simplify the integrands the three expressions have a similar structure. E^1 will be the smallest since it has a difference of transition densities at adjacent sites, so we begin by estimating E^2. To estimate $\Theta_1(E^2)$, we let

$$
X_t^1 = \frac{1}{ML} \sum_{x,y\sim x} \int_0^t \xi_{s-}(y)\xi_{s-}(x) p_{t_1-t_2-s}^n(z_1-x)\big)\, dN_s^{x,y}.
$$

By Markov property of $(\xi_t)_{t\ge 0}$ and the stationarity of the compensated Poisson process,

$$
\mathbb{E}(|\Theta_1(E^2)|^p) = \mathbb{E}\,\mathbb{E}^{\xi_{t_2}}(|X_{t_1-t_2}|^p),
$$

where $\mathbb{E}^{\xi_{t_2}}$ is the expectation w.r.t. the law of ξ starting at ξ_{t_2}.

To prepare for the next calculation we note that using the symmetry of the transition density and the Chapman-Kolmogorov equation

$$
L^{-1} \sum_w [p_u^n(w)]^2 = L^{-1} \sum_w p_u^n(w) p_u^n(-w) = 2p_{2u}(0). \tag{55}
$$

The predictable bracket process of X^1 is

$$
\langle X^1 \rangle_t = \frac{\theta}{R_n(ML)^2} \sum_{x,y\sim x} \int_0^t \xi_{s-}(y)\xi_{s-}(x)[p_{t_1-t_2-s}^n(z_1-x)]^2\, ds.
$$

Since there are $2M$ values of y for each x and M values of x for each $w \in L^{-1}\mathbb{Z}$, if we let $c_n = 2\theta/R_n L$ then the above is

$$
\begin{aligned}
&\le c_n \int_0^t L^{-1} \sum_w [p_{t_1-t_2-s}^n(w)]^2\, ds \\
&= 2c_n \int_0^t p_{2(t_1-t_2-s)}^n(0)\, ds \le C(t_1-t_2)^{1/2},
\end{aligned}
$$

where in the second step we have used the (55) and in the last step we used (49), and the fact that $R_N L \to \infty$.

If we do the calculation for Z then $c_n = 2r_n/L \to \gamma$ so we get the same upper bound. In E^1, $c_n = r_n/L \to \gamma$ but we have $p^n_{t_1-s}(z_1-y) - p^n_{t_1-s}(z_1-x)$ instead of a singe p, so

$$L^{-1}\sum_w [p^n_{t_1-t_2-s}(w) - p^n_{t_1-t_2-s}(w-L^{-1})]^2 = 2p^n_{2(t_1-t_2-s)}(0) - 2p^n_{2(t_1-t_2-s)}(L^{-1}).$$

If we used (51) now we would get an upper bound of $C_T L^{-1}$ Ignoring the cancellation we get the same upper bound. Using (54) now, we have

$$\mathbb{E}(|\Theta_1|^p) \le C_p(|t_1-t_2|^{p/4} + M^{-p}). \tag{56}$$

for E^1, E^2 and Z.

Similarly, $\mathbb{E}(|\Theta_2(E^2)|^p) = \mathbb{E}(|X^2_{t_2}|^p)$, where

$$X^2_t = \frac{1}{ML}\sum_{x,y\sim x}\int_0^t \xi_{s-}(y)(1-\xi_{s-}(x))[p^n_{t_1-s}(z_1-x) - p^n_{t_2-s}(z_2-x)], d\tilde{N}^{x,y}_s$$

is a càdlàg martingale for $t \le t_2$ with predictable bracket process

$$\langle X^2\rangle_t \le c_n \int_0^t L^{-1}\sum_w [p^n_{t_1-s}(z_1-w) - p^n_{t_2-s}(z_2-w)]^2.$$

Arguing as before using (55) we get

$$\le c_n \int_0^t p^n_{2(t_1-s)}(0) + p^n_{2(t_2-s)}(0) - 2p^n_{t_2+t_1-2s}(z_1-z_2)$$

a result that also holds for E^1 and Z. Adding and subtracting $2p^n_{t_2+t_1-2s}(0)$ and using (50) and (51) the above is

$$\le C\sqrt{t_1-t_2} + C_T|z_1-z_2|).$$

Using (54) now, we have

$$\mathbb{E}(|\Theta_2|^p) \le C_{p,T}(|t_1-t_2|^{p/4} + |z_1-z_2|^{p/2} + M^{-p}). \tag{57}$$

which holds for E^1, E^2 and Z and the proof is complete. □

Proof of Lemma 4. It suffices to consider the case $|t_1-t_2|^{1/4} \le M^{-1}$ and $|z_1-z_2|^{1/2} \le M^{-1}$, since otherwise Lemma 3 easily implies (46). The triangle inequality gives

$$|\tilde{u}^n_{t_1}(z_1) - \tilde{u}^n_{t_2}(z_2)| \le |\tilde{u}^n_{t_1}(z_1) - \tilde{u}^n_{t_2}(z_1)| + |\tilde{u}^n_{t_2}(z_1) - \tilde{u}^n_{t_2}(z_2)|. \tag{58}$$

We first estimate the time difference on the right. Write $s_k := k\theta_n$ for $k \in \mathbb{Z}_+$. Since $|t_1-t_2| \le M^{-4} < \theta_n$, we have either $s_k \le t_2 < t_1 \le s_{k+1}$ for some k or $t_2 < s_k < t_1$ for some k. Since $\tilde{u}^n$ is linear between grid points, in either case

$$|\tilde{u}^n_{t_1}(z) - \tilde{u}^n_{t_2}(z)| \le 2\Big(|\widehat{u}^n_{s_{k+1}}(z) - \widehat{u}^n_{s_k}(z)| \vee |\widehat{u}^n_{s_k}(z) - \widehat{u}^n_{s_{k-1}}(z)|\Big)\,\frac{|t_1-t_2|}{\theta_n}.$$

Hence Lemma 3, the assumption $M^{-4} < \theta_n$ and $|t_1 - t_2| \le M^{-4}$ imply that

$$\begin{aligned}\mathbb{E}|\tilde{u}^n_{t_1}(z) - \tilde{u}^n_{t_2}(z)|^p &\le C_{T,p}\,\frac{|t_1 - t_2|^p}{\theta_n^p}\big(\theta_n^{p/4} + M^{-p}\big)\\ &\le C_{T,p}\left(\frac{|t_1 - t_2|}{\theta_n^{3/4}}\right)^p \le C_{T,p}\,|t_1 - t_2|^{p/4}.\end{aligned}$$

Next, we estimate the space difference on the right of (58). Take n large so that $M^{-1} < (1+\gamma)L^{-1}$ and $(1+\gamma)^2 L^{-2} < L^{-1}$ (this is possible by our assumptions on scalings, even if $\gamma = 0$). Then $|z_1 - z_2| < (1+\gamma)^2 L^{-2} < L^{-1}$. By almost the same argument used above, we easily obtain

$$\begin{aligned}\mathbb{E}|\tilde{u}^n_t(z_1) - \tilde{u}^n_t(z_2)|^p &\le C(T,p)\,|z_1 - z_2|^p\,L^p\big(L^{-p/2} + M^{-p}\big)\\ &\le C(T,p)\,|z_1 - z_2|^p\,(L^{p/2} + (1+\gamma)^p)\\ &\le C(T,p)\Big[\,(1+\gamma)^{p/2}\,|z_1 - z_2|^{3p/4} + (1+\gamma)^p\,|z_1 - z_2|^p\,\Big].\end{aligned}$$

The proof of Lemma 4 is complete. □

7 Random walk estimates

The first two, (47) and (48), follow directly from the local central limit theorem (LCLT) (see, for example, Proposition 2.5.6 in [29]). (50) follows from (51) and the Chapman-Kolmogorov equation: the integrand can be written as

$$p^n_{s+\theta}(w) - p^n_s(w) = \frac{1}{L}\sum_{z\in L^{-1}\mathbb{Z}} p^n_\theta(z)\,\big(p^n_s(w-z) - p^n_s(w)\big).$$

It remains to prove (49) and (51).

By scaling, $p^n_t(w) = L\,p_{L^2 t}(Lw)$ where $p_t(k)$ is the transition density of the simple random walk on $\mathbb{Z}$. The integral of (49) is therefore

$$\frac{1}{L}\int_0^\infty p_s(0) - p_s(Lz)\,ds.$$

Splitting this integral into two parts according to whether $s \le L|z|^2$ or $s > L|z|^2$, the first part is bounded by $L^{-1}\int_0^{L|z|^2} p_s(0)\,ds \le C|z|/\sqrt{L}$ according to (47). The second part is bounded by $C|z|$ by the LCLT.

Formula (51) is similar to Proposition 2.4.1 in [29] which says that

$$\sum_{k\in\mathbb{Z}}\big|q_m(k) - q_m(k+j)\big| \le \frac{C\,j}{\sqrt{m}}, \tag{59}$$

where $q_m(k)$ is the transition density for the discrete time simple random walk on $\mathbb{Z}$. Hence, by using scaling and an independent Poisson process N_t, we rewrite the left hand side of

(51) as

$$\frac{1}{L^2}\int_0^{L^2T}\sum_{k\in\mathbb{Z}}|p_s(k)-p_s(k+Lz)|\,ds$$
$$=\frac{1}{L^2}\int_0^{L^2T}\sum_{k\in\mathbb{Z}}\Big|\sum_{m=0}^{\infty}\mathbb{P}(N_s=m)\big(q_m(k)-q_m(k+Lz)\big)\Big|\,ds$$

which is at most

$$\frac{C\,|z|}{L}\int_0^{L^2T}\sum_{m=0}^{\infty}\frac{\mathbb{P}(N_s=m)}{\sqrt{m}}\,ds$$

by (59). Arguing as in the proof of Proposition 2.5.6 in [29] by using Proposition 2.5.5 (the LCLT for Poisson processes), we obtain that the integral is of order $\sqrt{L^2T}$ and hence (51) holds.

8 Proof of Lemma 2

To prepare for the proof for the SPDE, we begin by considering the diffusion process

$$dU=\beta U(1-U)\,dt+\sigma\sqrt{U(1-U)}\,dB. \tag{60}$$

Following the approach of Doering, Mueller and Smereka [11], we change variables $Z=1-U$ to get (recall $dZ=-dU$)

$$dZ=-\beta Z(1-Z)\,dt-\sigma\sqrt{Z(1-Z)}\,dB. \tag{61}$$

The minus in front of the diffusion term is not important here but it will be in the next calculation in (64). Using Itô's formula and ignoring the martingale terms

$$\begin{aligned}\text{drift}(Z^m)&=mZ^{m-1}(-\beta Z(1-Z))+m(m-1)Z^{m-2}\frac{\sigma^2}{2}Z(1-Z)\\&=\beta m[Z^{m+1}-Z^m]+\frac{\sigma^2m(m-1)}{2}[Z^{m-1}-Z^m]).\end{aligned}$$

Let $N(t)$ be a Markov process with Q matrix

$$\begin{aligned}&Q_{m,m+1}=\beta m \qquad Q_{m,m-1}=\sigma^2\frac{m(m-1)}{2}\\&Q_{m,m}=-\beta m-\sigma^2\frac{m(m-1)}{2}.\end{aligned} \tag{62}$$

Combining our calculations,

$$\frac{d}{dt}EZ^m=\sum_n Q_{m,n}EZ^n$$

Letting $P_{\ell,m}(t)=\mathbb{P}(N(t)=m\,|\,N(0)=\ell)$ be the transition probabilities, we have

Lemma 5. *For fixed $T > 0$ and $\ell \geq 1$, $M_t = \sum_{m=1}^{\infty} P_{\ell,m}(T-t)Z^m(t)$ is a martingale.*

Proof. Differentiating we have

$$\begin{aligned}
\frac{d}{dt}EM_t &= \sum_m EZ^m(t)\frac{d}{dt}P_{\ell,m}(T-t) + P_{\ell,m}(T-t)\frac{d}{dt}EZ^m(t) \\
&= \sum_m -EZ^m(t)\sum_n P_{\ell,n}(T-t)Q_{n,m} + P_{\ell,m}(T-t)\sum_n Q_{m,n}EZ^n(t) = 0
\end{aligned}$$

if we interchange the roles of m and n in the second sum. □

From Lemma 5 we get

$$EZ^{N(0)}(T) = EZ^{N(T)}(0). \tag{63}$$

Now consider the system

$$\begin{aligned}
dZ &= -\beta Z(1-Z)\,dt - \sigma\sqrt{VZ}\,dB^0 - \sigma\sqrt{Z(1-Z-V)}\,dB^1, \\
dV &= \beta VZ\,dt + \sigma\sqrt{VZ}\,dB^0 + \sigma\sqrt{V(1-V-Z)}\,dB^2.
\end{aligned} \tag{64}$$

where the B^i are independent Brownian motions. To get our second dual function, we consider $Y_n = \sum_{m=1}^{n} Z^{m-1}V$. Using Itô's formula and for the second term recall (61),

$$\begin{aligned}
\text{drift}\left(\sum_{m=1}^{n} Z^{m-1}V\right) &= V\sum_{m=2}^{n}(m-1)Z^{m-2}(-\beta Z(1-Z)) \\
&\quad + \frac{\sigma^2}{2}V\sum_{m=3}^{n}(m-1)(m-2)Z^{m-3}Z(1-Z) \\
&\quad + \sum_{m=1}^{n} Z^{m-1}\beta ZV - \sigma^2\sum_{m=2}^{n}(m-1)Z^{m-2}VZ,
\end{aligned}$$

where the last term comes from the fact that the covariance of Z and V is $-VZ$. Collecting the terms with β, and changing variables $k = m-1$ in the second sum we get

$$= \beta V\left[\sum_{m=2}^{n}(m-1)Z^m - \sum_{k=1}^{n-1} kZ^k + \sum_{m=1}^{n} Z^m\right].$$

Adding the third sum to the first

$$\begin{aligned}
&= \beta V\left[\sum_{m=1}^{n} mZ^m - \sum_{k=1}^{n-1} kZ^k\right] \\
&= n\beta VZ^n = n\beta\left(\sum_{m=1}^{n+1} Z^{m-1}V - \sum_{m=1}^{n} Z^{m-1}V\right)
\end{aligned}$$

which corresponds to jumps from n to $n+1$ at rate βn. Collecting the terms with σ^2 and changing variables to have Z^{k-1}, we get

$$= \sigma^2 V\left[\sum_{k=2}^{n-1}\frac{k(k-1)}{2}Z^{k-1} - \sum_{k=3}^{n}\frac{(k-1)(k-2)}{2}Z^{k-1} - \sum_{k=2}^{n}(k-1)Z^{k-1}\right].$$

Moving terms $k = 2$ to $n-1$ from the last sum into the first one

$$= \sigma^2 V \left[\sum_{k=2}^{n-1} \frac{(k-2)(k-1)}{2} Z^{k-1} - \sum_{k=3}^{n} \frac{(k-1)(k-2)}{2} Z^{k-1} - (n-1)Z^{n-1} \right].$$

The $k = 2$ term in the first sum vanishes. Terms 3 to $n-1$ in the first sum cancel with the second sum leaving

$$= -\sigma^2 V \frac{n(n-1)}{2} Z^{n-1} = \sigma^2 \frac{n(n-1)}{2} \left(\sum_{m=1}^{n-1} Z^{m-1} V - \sum_{m=1}^{n} Z^{m-1} V \right)$$

which corresponds to jumps n to $n-1$ at rate $\sigma^2 n(n-1)/2$.

Combining our calculations

$$\frac{d}{dt} EY_n = \sum_m Q_{m,n} EY_n.$$

where $Q_{m,n}$ is given in (62). Using Lemma 5 again,

$$E\left(\sum_{m=1}^{N(0)} Z^{m-1}(T)V(T) \right) = E\left(\sum_{m=1}^{N(T)} Z^{m-1}(0)V(0) \right).$$

8.1 Duality for the Wright-Fisher SPDE

We begin by proving the duality result for the equation for single SPDE (4). This is a known result due to Shiga [38]. However, he did not give many details and we need to generalize his result to our coupled SPDE, so we will follow the approach of Athreya and Tribe [1]. Let $z = 1 - u$. Define $\bar{z}_t(x) = \int z_t(y)p^\epsilon(y-x)\,dy$ where p^ϵ is the normal density with mean 0 and variance ϵ. Noting that $x \to \bar{z}_t(x)$ is smooth and using the weak formulation of (4) with test function $\phi^{\epsilon,x}(y) = p^\epsilon(y-x)$, we have

$$\begin{aligned} \bar{z}_t(x) - \bar{z}_0(x) = & \int_0^t \alpha \Delta \bar{z}_s(x)\,ds \\ & - \theta\beta \int_0^t \int 2z_s(y)(1-z_s(y))p^\epsilon(y-x)\,dy\,ds \\ & + \int_0^t \int \sqrt{4\gamma z_s(y)(1-z_s(y))}p^\epsilon(y-x)\,dW. \end{aligned} \tag{65}$$

Using Itô's formula (each $\bar{z}_t(x)$ is a semi-martingale so this is legitimate) and writing $\mathcal{L}_z$ for the generator of $(\bar{z}_t(x_1), \dots, \bar{z}(x_n))$ with $x_1, \dots x_n$ fixed, we see that (ignoring the martingale terms)

$$\text{drift}\left(\mathcal{L}_{\bar z}\prod_i \bar z_t(x_i)\right) = \sum_{i=1}^n \prod_{j\neq i} \bar z_t(x_j)\alpha\Delta \bar z_t(x_i)$$
$$+ 2\theta\beta\sum_{i=1}^n\prod_{j\neq i}\bar z_t(x_j)\int [z_t^2(y) - z_t(y)]p^\epsilon(y-x_i)\,dy \tag{66}$$
$$+ 4\gamma\sum_{i=1}^{n-1}\sum_{j=i+1}^n\prod_{k\neq i,j}\bar z_t(x_k)\int [z_t(y)(1-z_t(y))]p^\epsilon(y-x_i)p^\epsilon(y-x_j)\,dy.$$

The dual process is a system of branching coalescing Brownian particles. During their lifetime the particles are Brownian motions run at rate 2α with each giving birth at rate $2\theta\beta$. In addition, for $i<j$, particle j is killed by particle i at rate $4\gamma L_t^{i,j}$ where $L_t^{i,j}$ denotes the local time of the process $x_j - x_i$ at 0. Writing $\mathcal{L}_x$ for the generator, we have

$$\text{drift}\left(\mathcal{L}_x\prod_i \bar z_t(x_i)\right) = \sum_{i=1}^n \prod_{j\neq i} \bar z_t(x_j)\alpha\Delta \bar z_t(x_i)$$
$$+ 2\theta\beta\sum_{i=1}^n\prod_{j\neq i}\bar z_t(x_j)\cdot[\bar z_t^2(x_i) - \bar z_t(x_i)] \tag{67}$$
$$+ 4\gamma\sum_{i=1}^{n-1}\sum_{j=i+1}^n\prod_{k\neq i,j}\bar z_t(x_k)\cdot[\bar z_t(x_i)(1-\bar z_t(x_j))]\,\delta_{\{x_j=x_i\}}$$

in which we used the formal notation $dL_t^{i,j} = \delta_{\{x_j(t)=x_i(t)\}}\,dt$. The precise meaning of the last term involves integration w.r.t. local times and is explained in (71).

We now follow Proposition 1 in [1] to use the duality method encapsulated in Theorem 4.4.11 of Ethier and Kurtz [16]. In their notation $\alpha=\beta=0$.

$$F(\bar z, x) = \prod_{i=1}^n \bar z(x_i) \quad \text{if } x = (x_1,\cdots,x_n).$$

They suppose $F(\bar z_t, x) - \int_0^t G(\bar z_s, x)\,ds$ and $F(\bar z, x(t)) - \int_0^t H(\bar z, x(s))\,ds$ are martingales and conclude that for $t \geq 0$,

$$\mathbb{E}F(\bar z_t, x(0)) - \mathbb{E}F(\bar z_0, x(t)) = \mathbb{E}\int_0^t G(\bar z_{t-s}, x(s)) - H(\bar z_{t-s}, x(s))\,ds. \tag{68}$$

In our situation, (68) holds with $G(\bar z, x) = \mathcal{L}_{\bar z}F(\bar z, x)$ and $H(\bar z, x) = \mathcal{L}_x F(\bar z, x)$.

By the continuity of $x \mapsto z_t(x)$, we have $F(\bar z_t, x(0)) \to F(z_t, x(0))$ and $F(\bar z_0, x(t)) \to F(z_0, x(t))$ a.s. as $\epsilon \to 0$, so using the bounded convergence theorem,

$$\mathbb{E}F(\bar z_t, x(0)) - \mathbb{E}F(\bar z_0, x(t)) \to \mathbb{E}F(z_t, x(0)) - \mathbb{E}F(z_0, x(t)).$$

To prove the desired duality formula

$$\mathbb{E}\prod_{i=1}^{n(0)} z_t(x_i(0)) = \mathbb{E}\prod_{i=1}^{n(t)} z_0(x_i(t)), \quad t\geq 0, \tag{69}$$

it remains to argue that the RHS of (68) tends to zero as $\epsilon \to 0$.

The first term in (66) agrees with that of (67). For the second terms, omitting $2\theta\beta$, the integrand of the difference is

$$\mathbb{E}\int_0^t \sum_{i=1}^{n(s)} \prod_{j\neq i} \bar{z}_{t-s}(x_j(s))\Big[\int z_{t-s}^2(y)p^\epsilon(y-x_i(s))\,dy - \bar{z}_{t-s}^2(x_i(s))\Big]\,ds \to 0 \tag{70}$$

a.s. for $s \in (0,t)$, by continuity of $y \mapsto z_s(y)$ and dominated convergence. The contribution to (68) from the third term of (67) is (omitting 4γ)

$$\mathbb{E}\int_0^t \sum_{i=1}^{n(s)-1} \sum_{j=i+1}^{n(s)} \prod_{k\neq i,j} \bar{z}_{t-s}(x_k(s)) \cdot [\bar{z}_{t-s}(x_i(s))\big(1-\bar{z}_{t-s}(x_j(s))\big)]\,dL_s^{i,j} \tag{71}$$

which converges, by dominated convergence to

$$\mathbb{E}\int_0^t \sum_{i=1}^{n(s)-1} \sum_{j=i+1}^{n(s)} \prod_{k\neq i,j} z_{t-s}(x_k(s)) \cdot [z_{t-s}(x_i(s))\big(1-z_{t-s}(x_j(s))\big)]\,dL_s^{i,j}. \tag{72}$$

Finally, we consider the contribution to (68) from the third term of (66). After the substitution $y \mapsto y + x_i(s)$, we have (omitting 4γ)

$$\mathbb{E}\int_0^t \sum_{i=1}^{n(s)-1} \sum_{j=i+1}^{n(s)} \int p^\epsilon(y)p^\epsilon(y+x_i(s)-x_j(s))\,Y_{s,t}^{i,j}(y)\,dy\,ds, \tag{73}$$

where for any i, j,

$$Y_{s,t}^{i,j}(y) = \prod_{k\neq i,j} \bar{z}_{t-s}(x_k(s)) \cdot [z_{t-s}(y+x_i(s))(1-z_{t-s}(y+x_i(s)))].$$

At this point we would like to apply Lemma 2 of [1], to obtain

$$\int_0^t p^\epsilon(y+x_i(s)-x_j(s))\,Y_{s,t}^{i,j}(y)\,ds = \int\int_0^t p^\epsilon(y+z)\,Y_{s,t}^{i,j}(y)\,dL_s^{i,j,z}\,dz, \tag{74}$$

where $L_t^{i,j,z}$ denotes the local time of the process $x_j - x_i$ at z. Their formula asssumes Y is predictable, so we substitute $s-$ for s and note that this does not change the integral in (73).

Putting (74) into (73), then using the continuity of the local time and that of $Y_{s,t}^{i,j}(y)$ (see details in pages 1724–1725 of [1]), the integrand of (73) converges a.s. to the integrand of (72). Convergence of expectations then follows from dominated convergence and the proof of (69) is complete.

8.2 Duality for the coupled SPDE

To prove the duality formula for the coupled SPDE, we order the particles in $x_i(t)$, $i \le n(t)$ in such a way that (i) when two particles coalesce we keep the smaller index and (ii) immediately after particle i gives birth, its offspring has index i and all particles with index $\ge i$ (including the one that just game birth) increase their index by 1. Adding the number of particles in the dual as another variable to help clarify things, we let

$$F_2((z,\ell),(x,n)) = \sum_{j=1}^{n} \prod_{i=1}^{j-1} z(x_j) \cdot \ell_t(x_j).$$

The following duality formula is motivated by (6):

$$\mathbb{E}F_2((z_t,\ell_t),(x_0,n_0)) = \mathbb{E}F_2((z_0,\ell_0),(x_t,n_t)) \tag{75}$$

Once we have shown (75) the uniqueness claimed in Lemma 2 follows, since it allows us to conclude that the distribution at a fixed time is unique and uniqueness of the law of the process follows from Theorem 4.2 in Chapter 4 of Ethier and Kurtz [16]. To prove (75), we let $\bar{\ell}_t(x) = \int \ell_t(y) p^\epsilon(y-x)\,dy$ and note

$$\begin{aligned}
\bar{z}_t(x) - \bar{z}_0(x) &= \int_0^t \alpha\Delta\bar{z}_s(x)\,ds - 2\theta\beta\int_0^t\int z_s(y)(1-z_s(y))p^\epsilon(y-x)\,dy\,ds \\
&\quad - \int_0^t\int \sqrt{4\gamma z_s(y)\ell_s(y)}\,p^\epsilon(y-x)dW^0 \\
&\quad - \int_0^t\int \sqrt{4\gamma z_s(y)(1-z_s(y)-\ell_s(y))}\,p^\epsilon(y-x)dW \\
\bar{\ell}_t(x) - \bar{\ell}_0(x) &= \int_0^t \alpha\Delta\bar{\ell}_s(x)\,ds + 2\theta\beta\int_0^t\int z_s(y)\ell_s(y)p^\epsilon(y-x)\,dy\,ds \\
&\quad + \int_0^t\int \sqrt{4\gamma z_s(y)\ell_s(y)}\,p^\epsilon(y-x)dW^0 \\
&\quad + \int_0^t\int \sqrt{4\gamma \ell_s(y)(1-z_s(y)-\ell_s(y))}\,p^\epsilon(y-x)dW^2.
\end{aligned}$$

Writing $\mathcal{L}_{\bar{z},\bar{\ell}}$ for the generator of $(\bar{z}_t(x_1), \dots, \bar{z}(x_{k-1}), \bar{\ell}(x_k))$ with $x_1, \dots x_n$ and n fixed,

$$\text{drift}\left(\mathcal{L}_{\bar{z},\bar{\ell}} F_2((\bar{z}_t, \bar{\ell}_t), (x, n))\right)$$

$$= \sum_{k=2}^{n} \sum_{i=1}^{k-1} \prod_{\substack{1 \le j \le k \\ j \ne i}} \bar{z}_t(x_j) \cdot \alpha \Delta \bar{z}_t(x_i) \cdot \bar{\ell}_t(x_k) + \prod_{j=1}^{k-1} \bar{z}_t(x_j) \alpha \Delta \bar{\ell}_t(x_k) \tag{76}$$

$$+ \sum_{k=2}^{n} 2\theta\beta \sum_{i=1}^{k-1} \prod_{j \ne i} \bar{z}_t(x_j) \cdot \bar{\ell}_t(x_k) \int [z_t^2(y) - z_t(y)] p^\epsilon(y - x_i)\, dy \tag{77}$$

$$+ \sum_{k=1}^{n} 2\theta\beta \prod_{j=1}^{k-1} \bar{z}_t(x_j) \cdot \int z_t(y) \ell_t(y) p^\epsilon(y - x_k)\, dy \tag{78}$$

$$+ \sum_{k=3}^{n} 4\gamma \sum_{j=2}^{k-1} \sum_{i=1}^{j-1} \prod_{\substack{1 \le h \le k-1 \\ h \ne i,j}} \bar{z}_t(x_h) \cdot \bar{\ell}_t(x_k) \int [z_t(y)(1 - z_t(y))] p^\epsilon(y - x_i) p^\epsilon(y - x_j)\, dy \tag{79}$$

$$- \sum_{k=2}^{n} 4\gamma \sum_{i=1}^{k-1} \prod_{\substack{1 \le h \le k-1 \\ h \ne i}} \bar{z}_t(x_h) \int [z_t(y) \ell_t(y)] p^\epsilon(y - x_i) p^\epsilon(y - x_k)\, dy. \tag{80}$$

Writing $\mathcal{L}_{x,n}$ for the generator of the dual ordered particle system, we want to compute

$$\text{drift}\left(\mathcal{L}_{x,n} F_2((\bar{z}, \bar{\ell}), (x, n))\right). \tag{81}$$

and to show, as in (68), that

$$\mathbb{E} \int_0^t \mathcal{L}_{\bar{z},\bar{\ell}} F_2((\bar{z}_{t-s}, \bar{\ell}_{t-s}), (x(s), n(s))) - \mathcal{L}_{x,n} F_2((\bar{z}_{t-s}, \bar{\ell}_{t-s}), (x(s), n(s)))\, ds \to 0. \tag{82}$$

The terms coming from particle Brownian motions are

$$\sum_{k=1}^{n} \left(\sum_{i=1}^{k-1} \prod_{\substack{1 \le j \le k-1 \\ j \ne i}} \bar{z}_t(x_j) \alpha \Delta \bar{z}_t(x_i) \cdot \bar{\ell}_t(x_k) + \prod_{i=1}^{k-1} \bar{z}_t(x_i) \cdot \Delta \bar{\ell}_t(x_k) \right) \tag{83}$$

which agree with (76), and hence cancel in (82).

Birth terms. Given a vector $x = (x_1, \dots x_n)$, let $x^i = (x_1, \dots x_i, x_i, \dots x_n)$ be the vector of length $n+1$ with the i coordinate duplicated. The total change in the drift (81) due to births is (omitting the $2\theta\beta$)

$$\sum_{i=1}^{n} \left[\sum_{k=i+1}^{n+1} \prod_{j=1}^{k-1} \bar{z}_t(x_j^i) \ell_t(x_k^i) - \sum_{k=i+1}^{n} \prod_{j=1}^{k-1} \bar{z}_t(x_j) \ell_t(x_k) \right]. \tag{84}$$

Here i is the location of the duplication and there is no change in the drift for terms with $k \le i$. The difference

$$\sum_{k=i+2}^{n+1} - \sum_{k=i+1}^{n} = [\bar{z}_t(x_i) - 1] \sum_{k=i+1}^{n} \prod_{j=1}^{k-1} \bar{z}_t(x_j) \ell_t(x_k) \tag{85}$$

since when $k \geq i+2$, we have $x^i_{k+1} = x_k$ and $\bar{z}_t(x_i)$ appears twice in the product. After we interchange the order of the summation this agrees with (77) if we replace $p^\epsilon(y - x_i)$ in that formula by a pointmass at x_i. The remaining term ($k = i+1$) in the first sum in (84) is

$$\sum_{i=1}^n \prod_{j=1}^i \bar{z}_t(x_j)\ell_t(x_i) \tag{86}$$

since $x^i_i = x^i_{i+1} = x_i$. This agrees with (78) if we again replace $p^\epsilon(y - x_i)$ by a pointmass at x_i. As we argued in (70) it follows that

$$\mathbb{E} \int_0^t (77) - (85) + (78) - (86)\, ds \to 0$$

Killing terms. Given a vector $x = (x_1, \dots x_n)$ let $\hat{x}^j = (x_1, \dots x_{j-1}, x_{j+1}, \dots x_n)$ be the vector of length $n-1$ with the j coordinate removed. The total change in the drift (81) due to deaths is (omitting the 4γ)

$$\sum_{j=2}^n \sum_{i=1}^{j-1} \left[\sum_{k=j}^{n-1} \prod_{h=1}^{k-1} \bar{z}_t(\hat{x}^j_h)\ell_t(\hat{x}^j_k) - \sum_{k=j}^n \prod_{h=1}^{k-1} \bar{z}_t(x_h)\ell_t(x_k) \right] \delta_{\{x_j = x_i\}}. \tag{87}$$

Here j is the location of the deletion and there is no change in the drift for terms with $k < j$. When $j \leq k$, $\hat{x}^j_k = x_{k+1}$ so we have

$$\sum_{k=j}^{n-1} - \sum_{k=j+1}^n = \sum_{k=j+1}^n \prod_{\substack{1 \leq h \leq k-1 \\ h \neq j}} \bar{z}_t(x_h)[1 - \bar{z}_t(x_j)]\bar{\ell}_t(x_k).$$

The remaining term coming from $k = j$ in the second sum is

$$-\prod_{h=1}^{j-1} \bar{z}_t(x_h)\bar{\ell}_t(x_j).$$

Using these results in (87) and noting that in the first case $j = n$ is impossible, we have

$$\begin{aligned} &\sum_{j=2}^{n-1} \sum_{i=1}^{j-1} \sum_{k=j+1}^n \prod_{\substack{1 \leq h \leq k-1 \\ h \neq i,j}} \bar{z}_t(x_h) \cdot \bar{\ell}_t(x_k) \cdot \bar{z}_t(x_i)[1 - \bar{z}_t(x_j)]\, \delta_{\{x_j = x_i\}} \qquad (88) \\ &- \sum_{j=2}^n \sum_{i=1}^{j-1} \prod_{1 \leq h \leq j-1, h \neq i} \bar{z}_t(x_h) \cdot \bar{z}_t(x_i)\bar{\ell}_t(x_j)\, \delta_{\{x_j = x_i\}}. \end{aligned}$$

where in the second sum we have changed j to k, and in both terms we have removed an additional term from the product over h. On other hand, interchanging the order of summation in (79)

$$\sum_{k=3}^n \sum_{j=2}^{k-1} \sum_{i=1}^{j-1} = \sum_{j=2}^{n-1} \sum_{k=j+1}^n \sum_{i=1}^{j-1} = \sum_{j=2}^{n-1} \sum_{i=1}^{j-1} \sum_{k=j+1}^n$$

Hence formulas (79) and (80) become

$$\sum_{j=2}^{n-1}\sum_{i=1}^{j-1}\sum_{k=j+1}^{n}\prod_{\substack{1\le h\le k-1\\ h\ne i,j}}\bar{z}_t(x_h)\cdot\bar{\ell}_t(x_k)\int[z_t(y)(1-z_t(y))]p^\epsilon(y-x_i)p^\epsilon(y-x_j)\,dy \tag{89}$$
$$-\sum_{k=2}^{n}4\gamma\sum_{i=1}^{k-1}\prod_{\substack{1\le h\le k-1\\ h\ne i}}\bar{z}_t(x_h)\int[z_t(y)\ell_t(y)]p^\epsilon(y-x_i)p^\epsilon(y-x_k)\,dy.$$

Arguing as in (71)–(74) now completes the proof of (82). The proof of (75) is complete, and hence that of Lemma 2.

References

[1] Athreya, S., and Tribe, R. (2000) Uniqueness for a class of one dimensional stochastic PDEs using moment duality. *Ann. Probab.* 28, 1711–1734

[2] Berestycki, J., Berestycki, N., and Schweinsberg, J. (2013) The genealogy of branching Brownian motion with absorption. *Ann. Probab.* 41, 527–618

[3] Bolthausen, E., and Sznitman, A.S. (1998) On Ruelle's probability cascades and an abstract cavity method. *Comm. Math. Phys.* 197, 247–276

[4] Bramson, M., and Griffeath, D. (1980) On the Williams-Bjerknes tumour growth model. II. *Math. Proc. Cambridge Philos. Soc.* 88, 339–357

[5] Bramson, M., and Griffeath, D. (1981) On the Williams-Bjerknes tumour growth model. I. *Ann. Probab. 9*, 173–185

[6] Brunet, E., and Derrida, B. (2012) Genealogies in simple models of evolution. arXiv:1202.5997

[7] Brunet, E., Derrida, B., Mueller, A.H., and Munier, S. (2006) Noisy traveling waves: Effect of selection on genealogies. *Europhys. Letters* 76, 1–7

[8] Brunet, E., Derrida, B., Mueller, A.H., and Munier, S. (2007) Effect of selection on ancestry: An exactly soluble case and its phenomenological generalziation. *Physical Review E.* 76, paper 041104

[9] Cavalli-Sforza, L.L., Menozzi, P., and Piazza, A. (1993) Demic expansion and human evolution. *Science.* 259, 639–646

[10] Cox, T., Durrett, R., and Perkins, E.A. (2000) Rescaled voter models converge to super-Brownian motion. *Ann. Probab.* **28** (2000), 185–234

[11] Doering, C.R., Mueller, C., and Smerka, P. (2003) Interacting particles,the stochastic Fisher - Kolmogorov - Petrovsky - Piscounov equation, and duality. *Physica A.* 325, 243-259

[12] Durrett, R., Mytnik, L., and Perkins, E.A. (2005) Competing super-Brownian motions as limits of interacting particle systems. *Electron. J. of Probab.* 10, 1147–1220

[13] Durrett, R., and Perkins, E.A. (1999) Rescaled contact processes converge to super-Brownian motion in two or more dimensions. *Probab. Theory Related Fields.* 114 (1999), 309–399

[14] Durrett, R., and Restrepo, M. (2008) One-dimensional stepping stone models, sardine genetics and Brownian local time. *Ann. Appl. Probab.* 18, 334–358

[15] Edmonds, C.A., Lille, A.S., and Cavalli-Sforza, L.L. (2004) Mutations in the wave front of an expanding population. *Proc. Natl. Acad. Sci.* 101, 975–979

[16] S.N. Ethier, S.N. and Kurtz, T.G. *Markov processes. Characterization and Convergence.* Wiley, New York, 1986. MR0838085.

[17] Gouéré, J.P. (2007) Shape of territories in some competing growth models. *Ann. Appl. Probab.* 17, 1273-1305

[18] Häggström, O., and Pemantle, R. (1998) First passage percolation and a model for competing spatial growth. *J. Appl. Probab.* 35, 683–692

[19] Hallatschek, O., Hersen, P., Ramanathan, S., and Nelson, D.R. (2007) Genetic drift at expanding frontiers promotes gene segregation. *Proc. Natl. Acad. Sci.* 104, 19926–19930

[20] Hallatshek, O., and Nelson, D.R. (2008) Gene surfing in expanding populations. *Theor. Pop. Biol.* 73, 158–170

[21] Hallatschek, O., and Nelson, D.R. (2010) Life at the front of an expanding population. *Evolution.* 64, 193–206

[22] Hobson, T., and Tribe, R. (2005) On the duality between coalescing Brownian particles and the heat equation driven by Fisher-Wright noise. *Electron. Commun. Probab.* 10 (2005), 136–145

[23] Karatzas, I. and Shreve, S.E. (1991) *Brownian Motion and Stochastic Calculus.* Graduate Texts in Mathematics, Springer

[24] Kesten, H. (1978). Branching Brownian motion with absorption. *Stoch. Proc. Appl.* 7, 9-47

[25] Kliem, S. (2011) Convergence of rescaled competing species processes to a class of SPDEs. *Electron. J. Probab.* 16, 618–657

[26] Klopfstein, S., Currat, M., and Excofier, L. (2006) The fate of mutations surfing on the wave of a range expansion. *Mol. Biol. Evol.* 223, 482–490

[27] Korolev, K.S., Müller, M.J.I., Karahan, N., Murray, A.W., Hallatschek, O., and Nelson, D.R. (2012) Selective sweeps in growing microbial colonies. *Physical Biology.* 9, paper 026008

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

[29] Lawler, G. F., and Limic, V. (2010) *Random walk: a modern introduction.* Cambridge University Press.

[30] Lehe, R., Hallatschek, O., and Nelson, D.R. (2012) The rate of beneficial mutations surfing on the wave of a range expansion. *PLoS Comp. Biol.* 8, paper e1002447

[31] Mueller, C., and Sowers, A.B. (1995) Random traveling waves for the KPP equation with noise. *Journal Functional Analysis.* 128, 439–498

[32] Mueller, C., and Tribe, R. (1995) Stochastic p.d.e.'s arising from the long range contact process and long range voter model. *Probab. Theory Related Fields.* 102, 519–545

[33] Neher, R.A., and Hallatschek, O. (2013) Genealogies of rapidly adapting populations. *Proc. Natl. Acad. Sci.* 110, 437–442

[34] Neuhauser, C. (1992) Ergodic theorems for the multitype contact process. *Probab. Theory Related Fields.* 91, 467–506

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

[36] Nullmeier, J., and Hallatschek, O. (2012) The coalescent in boundary limited range expansions. *Evolution.* 67, 1307–1320

[37] Shiga, T., and Uchiyama, K. (1986) Stationary states and their stability of the steppping stone model involving mutation and selection. *Probab. Theory Related Fields.* 73, 87–117

[38] Shiga, T. (1988) Stepping stone models in population genetics and population dynamics. Pages 345355 in *Stochastic Processes in Physics and Engineering.* Reidel, Dordrecht

[39] Sottoriva, A. et al. (2015) A Big Bang model of human colorectal tumor growth. *Nature Genetics.* 47, 209–216

[40] Walsh, J. B. (1986) An introduction to stochastic partial differential equations. Pages 265–439 in *École d'été de probabilités de Saint-Flour.* Lecture Notes in Math., 1180, Springer, Berlin.