On the width of hybrid zones

by Rick Durrett[*] and Iljana Zähle [†]

May 19, 2006

Abstract

Hybrid zones occur when two species are found in close proximity and interbreeding occurs, but the species characteristics remain distinct. These systems have been treated in the biology literature using partial differential equations models. Here we investigate a stochastic spatial model and prove the existence of a stationary distribution that represents the hybrid zone in equilibrium. We calculate the width of the hybrid zone, which agrees with the PDE formula only in dimensions $d \geq 3$. Our results also give insight into properties of hybrid zones in patchy environments.

1   Introduction

In a number of situations one finds "hybrid zones": two adjacent regions in space that contain relatively homogeneous populations, which differ from each other and are separated by a narrow zone in which hybrids are found. There are literally hundreds of such examples. Barton and Hewitt's (1985) survey lists 150 that have been studied. A text book example is the common house mouse in Denmark (see Hunt and Selander 1973) which exists in the form *Mus musculus* in the North and in the form *M. domesticus* in the South and along parts of the Western coast. Syzmura and Barton (1986) have studied the fire bellied toads *Bombina bombina* and *B. variegata* in southern Poland. The motivation for this study comes from work Harrison has done (Harrison 1986, Rand and Harrison 1989, Harrison and Rand 1989, Harrison and Bogdanowicz 1997) investigating two eastern North American field crickets, *Gryllus Pennsylvanicus* and *G. firmus*, which hybridize along a zone of contact that extends from New England to Virginia.

Partial differential equations.   There are two commonly quoted explanations for hybrid zones, both of which assume that the underlying genetics is controlled by a single locus with two alleles. Haldane (1948) advocated the explanation that one type is more fit in one region and the second is more fit in the complement. Slatkin (1973) analyzed this situation using a partial differential equation. Letting $\mathrm{sgn}(x) = 1$ for

$x > 0$, $\text{sgn}(x) = -1$ for $x < 0$ and $\text{sgn}(x) = 0$ for $x = 0$ and letting s denote the selective advantage of a type in its favored environment, we can write his PDE as

$$\frac{\partial u}{\partial t} = \sigma^2 \frac{\partial^2 u}{\partial x^2} + \text{sgn}(x)su(x)(1 - u(x)) \tag{1.1}$$

where x is a one-dimensional spatial variable, $u(x)$ is the density of the type that is favored on the right half-line, and σ^2 is the variance of the displacement of offspring from their parents. To connect our equation with Slatkin's we write $v = 2u - 1$ or $u = (v + 1)/2$ to get

$$\frac{\partial v}{\partial t} = 2 \cdot \left(\frac{\sigma^2}{2} \frac{\partial^2 v}{\partial x^2} + \frac{1}{4} \text{sgn}(x)s(1 - v^2(x)) \right)$$

so if $A(x)$ is the equilibrium distribution

$$\sigma^2 \frac{\partial^2 A}{\partial x^2} + \frac{1}{2} \text{sgn}(x)s(1 - A^2(x)) = 0$$

Changing variables $y = x\sqrt{s}/\sigma$, Slatkin nondimensionalizes the equation, see his (8), as

$$0 = \frac{\partial^2 A}{\partial y^2} + \frac{1}{2} \text{sgn}(y)(1 - A^2(y))$$

The equilibrium distribution $A(y)$ has $A(0) = 0$ by symmetry and, see his (9), for $y > 0$

$$A(y) = -2 + 3\tanh^2(y/2 + c) \qquad \text{where } c = \tanh^{-1}\sqrt{2/3}$$

The solution looks a little strange, but is not hard to verify once one realizes $f(y) = \tanh(y)$ has $f' = 1 - f^2$. The choice of c makes $A(0) = 0$. Since after the change of variables $y = x\sqrt{s}/\sigma$, the equilibrium does not depend on σ and s, the hybrid zone has width of order $\sqrt{s}/\sigma$.

Barton (1979) has argued (see also Barton and Gale 1993) that hybrid zones are maintained by selection against hybrids, i.e., the three genotypes AA, Aa and aa have relative fitnesses 1, $1 - s$, and 1. In this case the PDE is

$$\frac{\partial u}{\partial t} = \frac{\sigma^2}{2} \frac{\partial^2 u}{\partial x^2} + \gamma(x)su(x)(1 - u(x))(2u(x) - 1) \tag{1.2}$$

and the equilibrium solution is (see also Bazykin 1969)

$$v(x) = \frac{1}{2}\left(1 + \tanh(x\sqrt{s/2\sigma^2})\right)$$

so again the hybrid zone has width of order $\sqrt{s}/\sigma$.

Spatial model. Here our goal is not to try to differentiate between these two hypotheses (for work in that direction see e.g., Kruuk, Baird, Gale, and Barton 1999), but instead to see how the predictions change when the PDE is replaced by a stochastic spatial model. Following Cox and Durrett (1995) and Durrett, Buttel, and Harrison (2000) we introduce a general model. Let $q(z)$ be an irreducible probability distribution on $\mathbb{Z}^d$ with

- $q(0) = 0$, $q(x) = 0$ if $\sup_i |x_i| \geq R$,

- symmetry with respect to reflection in each axis

- covariance $\sigma^2 I$

It is natural to assume that q has the same symmetries as $\mathbb{Z}^d$, which implies the second and third assumptions. With a little more work, the finite range assumption could be weakened, but it helps make the arguments simple.

At the locus under selection, the genetics may be haploid or diploid, so the state at time t is either $\eta_t : \mathbb{Z}^d \to \{A, a\}$ or $\eta_t : \mathbb{Z}^d \times \{1, 2\} \to \{A, a\}$. Using the population dynamics of the Moran model, each site attempts to change its state at rate 1. If x is the next site to change state then in the haploid case a parent is chosen according to $q(y - x)$, while in the diploid case we twice choose (with replacement) a parent according to $q(y - x)$ and one of their chromosomes at random. To determine whether the individual at x will be replaced, we compute the relative fitness $\phi \leq 1$ of the proposed new individual in the environment at x and do the replacement if a uniformly distributed random variable $U < \phi$.

Cox and Durrett (1995) considered the neutral case $s = 0$, starting from $\eta_0(x) = (a, a)$ for $x_1 \geq 0$ and $\eta_0(x) = (A, A)$ for $x_1 < 0$. Let ν_ρ be the limiting state starting from product measure with a density ρ of A's, let Φ be the standard normal distribution function, and let θ_x be the operator that shifts the configuration to the right by x, i.e., $(\theta_x \eta)(y) = \eta(y - x)$. They showed that if one considers a sequence $x_t \in \mathbb{Z}^d$ and $x_t/\sqrt{t} \to x$ then

$$\theta_x \eta_t \Rightarrow \nu_{\Phi(-x_1/\sigma)}$$

where x_1 is the first coordinate of x. In words, the density of A's evolves as predicted by the central limit theorem. In $d \geq 3$ the ν_ρ are extremal stationary distributions, while in $d \leq 2$, $\nu_\rho = \rho \delta_{AA} + (1 - \rho) \delta_{aa}$. In $d = 2$ the density of heterozygotes

$$P(\eta_t(x_t, 1) \neq \eta_t(x_t, 2)) \sim C\Phi(-x_1/\sigma)(1 - \Phi(-x_1/\sigma))/\log t$$

In the $d = 1$ haploid case they showed that if $\ell_t = \inf\{t : \eta_t(x) = a\}$ and $r_t = \sup\{t : \eta_t(x) = A\}$, then $r_t - \ell_t$ converges in distribution to a limit.

Durrett, Buttel, and Harrison (2000) studied the diploid model on $\mathbb{Z}^2$ and added a second linked neutral locus with recombination probability r between the two loci. Using simulation they investigated the allele frequency in equilibrium and the decay with time of linkage disequilibrium (i.e., the correlation between the selected and neutral loci). The second quantity is difficult to study even with nonrigorous methods. Barton (1986) proposed an approximation for weak selection based on a diffusion approximation, but as Figure 8 of Durrett, Buttel, and Harrison (2000) shows, it is not very accurate.

Here, we will focus on the allele frequency at the selected locus. For simplicity, we will restrict our attention to the haploid case. To return to the usual notation of interacting particle systems we will write 1 for A, 0 for a, and replace s by β (since t and s are commonly used for time). Let η_t be the nonhomogeneous biased voter model on $\mathbb{Z}^d$ where 1's are favored on $(0, \infty) \times \mathbb{Z}^{d-1}$ and 0's are favored on $(-\infty, 0] \times \mathbb{Z}^{d-1}$. If

the system is in state η then at x with $x_1 \leq 0$

$$0 \longrightarrow 1 \qquad \text{at rate } \sum_{y \in \mathbb{Z}^d} q(y - x)\eta(y)$$

$$1 \longrightarrow 0 \qquad \text{at rate } (1 + \beta) \sum_{y \in \mathbb{Z}^d} q(y - x)(1 - \eta(y))$$

and at x with $x_1 > 0$

$$0 \longrightarrow 1 \qquad \text{at rate } (1 + \beta) \sum_{y \in \mathbb{Z}^d} q(y - x)\eta(y)$$

$$1 \longrightarrow 0 \qquad \text{at rate } \sum_{y \in \mathbb{Z}^d} q(y - x)(1 - \eta(y))$$

We start with $\eta_0(x) = 0$ if $x_1 \leq 0$ and $\eta_0(x) = 1$ if $x_1 > 0$. Let

$$w_\beta = \begin{cases} 1/\sqrt{\beta} & d \geq 3 \\ \sqrt{\frac{1}{\beta}\log(1/\beta)} & d = 2 \\ 1/\beta & d = 1 \end{cases} \tag{1.3}$$

Theorem 1 *There is a nontrivial stationary distribution, ξ_∞. The width of the hybrid zone in ξ_∞ is $O(w_\beta)$. That is,*

$$\lim_{k \to \infty} \liminf_{\beta \to 0} \inf_{x_1 \geq k w_\beta} P(\xi_\infty(x) = 1) = 1$$

$$\lim_{k \to \infty} \limsup_{\beta \to 0} \sup_{x_1 \leq -k w_\beta} P(\xi_\infty(x) = 1) = 0$$

In $d \geq 2$ or in the one dimensional nearest neighbor case (i.e., $q(1) = q(-1) = 1/2$) we also have

$$\limsup_{\beta \to 0} \sup_{0 \leq x_1 \leq w_\beta} P(\xi_\infty(x) = 1) < 1$$

To prove Theorems 1, let $\tilde{\eta}_t$ be the system that has permanent 0's at all x with $x_1 \leq 0$, which is also started from $\tilde{\eta}_0(x) = 0$ if $x_1 \leq 0$ and $\tilde{\eta}_0(x) = 1$ if $x_1 > 0$. Obviously $P(\eta_t(x) = 1) \geq P(\tilde{\eta}_t(x) = 1)$ for $x_1 > 0$. $\tilde{\eta}_t$ is a biased voter model so, see Griffeath (1978) or Liggett (1985), it has a dual process $\tilde{\xi}_t$, which is a system of coalescing branching random walks with migration rate 1 and branching rate β where particles are killed on $(-\infty, 0] \times \mathbb{Z}^{d-1}$. At a migration event the particle jumps according to q. At a branching event a new particle is born at a site randomly chosen according to q. When two particles come to the same site they coalesce.

The duality gives us $P(\tilde{\eta}_t(x) = 1) = P(\tilde{\xi}_t^x(\mathbb{Z}^d) > 0)$ for $x_1 > 0$. Here $\tilde{\xi}_t^x$ is the system started with one particle at x. To prove our results we will use a block construction in which the blocks have size $O(w_\beta)$. Once this is done the existence of a nontrivial stationary distribution and the lower bound follow easily. The approach we use to the block construction depends on the dimension. In $d \geq 2$ let

$$\tau(\beta) = \begin{cases} \frac{1}{\sqrt{\beta}} & d \geq 3 \\ \\ \frac{1}{\beta\sqrt{\log(\frac{1}{\beta})}} & d = 2 \end{cases} \tag{1.4}$$

4

Let $\hat{\xi}_t$ be a pruned modification of $\tilde{\xi}_t$ in which new born particles have mass 0 until time $\tau(\beta)$ and then are given mass 1 if they have not coalesced with another particle. Let $\hat{\xi}_t^\beta$ be the rescaled modification of $\hat{\xi}_t$, i.e., $\hat{\xi}_t^\beta(\cdot) = \hat{\xi}_{h(\beta)t}(\sqrt{h(\beta)}\,\cdot)$, where

$$ h(\beta) = \begin{cases} \frac{1}{\beta} & d \geq 3 \\[2mm] \frac{1}{\beta}\log(\frac{1}{\beta}) & d = 2 \end{cases} $$

$h(\beta)$ is the height of blocks in the block construction.

The key step in the proof is to show that $\hat{\xi}_t^\beta$ converges weakly to ζ_t be the branching Brownian motion with branching rate $2\pi\sigma^2$ for $d = 2$ and branching rate γ for $d \geq 3$, where γ is the escape probability of the random walk. This is easy to see in $d \geq 3$: a newly born particle coalesces with its parent with probability $1 - \gamma$ and if it does so the coalescence takes $O(1)$. In $d = 2$ the coalescence time distribution has a tail that is $\sim C/(\log t)$, so the fraction that take time $> 1/\beta$ is $O(1/\log(\frac{1}{\beta}))$, and we have to run the process for time of order $h(\beta)$ to a branching that does not coalesce by time $1/\beta$.

The result in $d = 1$ is easy to see if we consider the nearest neighbor case. In this situation if $\ell_t = \inf\{x : \xi_t(x) = 1\}$ then all sites $\geq \ell_t$ are 1 and those $< \ell_t$ are 0. ℓ_t is a nearest neighbor random walk with drift $-\beta$ when it is > 0 and β if it is < 0. Speeding time up we have $\ell_{t/\beta^2}/\beta \Rightarrow L_t$ the solution of the stochastic differential equation

$$ dL_t = dB_t - \operatorname{sgn}(L_t)\, dt $$

To prove the result in $d = 1$ for a general symmetric finite range distribution, we use duality. Too many particles are lost to coalescence so we take a different approach. The key idea is that if we follow branching arrows when they go to the right we get a random walk with positive drift. By following branching when it takes us closer to a target, we get a process that moves linearly toward a target and then stays close to it.

Mosaic hybrid zones. In the cricket hybrid zone, the frequencies of the two species show a patchy pattern and are strongly correlated with soil type (Harrison 1986, Rand and Harrison 1989, Harrison and Rand 1989). Durrett, Buttel, and Harrison (2003) have investigated by simulation a model in which the environment is random and patchy. Their results show that for a patch to be visible in the stationary distribution it must exceed a critical size. This observation is not new. Slatkin (1973), see page 742 investigated the situation where A's are favored on an interval $(-L, L)$ and a's are favored on the rest of the space. A's can only persist in equilibrium if $L > 0.5\ell$ where $\ell = \sqrt{s}/\sigma$ is the hybrid zone width. As a consequence of our proof of Theorem 1 by the block construction, we can conclude that if patches contain a d-dimensional cube with side that is a large multiple of w_β, then A's will persist for a long time even if all of the environment outside favors a's.

2   Proof of Theorem 1 in $d \geq 2$

Since the pruned dual process $\hat{\xi}_t$ is started with finitely many particles and since we are interested only in a finite time horizon T we can think of the process as follows

$$ \hat{X}_t = (\hat{X}_t^1, \ldots, \hat{X}_t^k, \hat{X}_t^{k+1}, \ldots, \hat{X}_t^N, \infty, \ldots) \tag{2.1} $$

Here the first k entries represent the initial particles. For them $\hat{X}_t^m \in \mathcal{D}([0,T], \mathbb{R}^d)$ is the path followed by the m-th particle until time t and then continued to time T as a constant. Let τ_n be the time the n-th particle that survives for time $\tau(\beta)$ is born. Let $\hat{\tau}_n = \tau_n + \tau(\beta)$. Then the $k+n$-th entry is ∞ until time $\hat{\tau}_n$ and afterwards it is the path followed by this particle and its ancestors. Whenever two particles coalesce the path of the particle with the higher number is replaced by ∞. Let $\hat{X}_t^\beta = (\hat{X}_t^{1,\beta}, \dots, \hat{X}_t^{k,\beta}, \hat{X}_t^{k+1,\beta}, \dots, \hat{X}_t^{N,\beta}, \infty, \dots)$ be the rescaled modification, $\hat{\xi}_t^\beta(\cdot) = \hat{\xi}_{h(\beta)t}(\sqrt{h(\beta)}\cdot)$.

Analogously we can interpret the branching Brownian motion ζ_t as

$$Y_t = (Y_t^1, \dots, Y_t^k, Y_t^{k+1}, \dots, Y_t^N, \infty, \dots)$$

$\hat{X}$, $\hat{X}^\beta$ and Y are processes taking values in

$$\mathcal{D} = \{(y_1, y_2, \dots) : \exists k_0 \forall k \le k_0, y_k \in D([0,T], \mathbb{R}^d) \cup \{\infty\}; \forall k > k_0, y_k = \infty\} \qquad (2.2)$$

with the Skorokhod topology modified in the obvious way to account for the fact that the sample paths have an initial segment $= \infty$.

Lemma 2.1 *Let $d \ge 2$ and $x_1, \dots, x_k \in \mathbb{R}^d$ and define*

$$x_i^\beta = h(\beta)^{-1/2}[h(\beta)^{1/2}x_i]$$

where $[x]$ is the nearest point to x in $\mathbb{Z}^d$. Let $\hat{X}_0^\beta = (x_1^\beta, \dots, x_k^\beta, \infty, \dots)$ and $Y_0 = (x_1, \dots, x_k, \infty, \dots)$. Then as $\beta \to 0$

$$\hat{X}_T^\beta \xrightarrow{d} Y_T$$

Proof (a) Consider the case $d \ge 3$. Let ξ_t be a pruned branching random walk with migration rate 1 and branching rate β. Pruned means that a new born particle has mass 0 until time $\tau(\beta) = 1/\sqrt{\beta}$ and its gets mass 1 if it has not coalesced with its parent. At a migration event the particle jumps according to q. At a branching event a new particle is born at a site randomly chosen according to q. The difference between ξ and $\hat{\xi}$ is that in ξ particles do not coalesce if they meet. Again we interpret the process ξ_t as $X_t = (X_t^1, \dots, X_t^k, X_t^{k+1}, \dots, X_t^N, \infty, \dots)$ analogously to (2.1).

Let ξ_t^β be the rescaled modification of ξ_t, i.e., $\xi_t^\beta(\cdot) = \xi_{t/\beta}(\cdot/\sqrt{\beta})$. Each single random walk converges to Brownian motion. Let Z^1 and Z^2 be two independent random walks with jump rate 1 and jump kernel q both starting at 0. Let P_q indicate that one random walk is started in the origin and the starting position of the other one is chosen according to q. Let $T_0 = \inf\{t : Z_t^1 = Z_t^2\}$ and $\gamma = P_q(T_0 = \infty)$. Since $\gamma > 0$ in $d \ge 3$, Chebyshev's inequality implies

$$P_q\left(|Z_{1/\sqrt{\beta}}^1 - Z_{1/\sqrt{\beta}}^2| \ge \beta^{-1/3} \middle| T_0 > 1/\sqrt{\beta}\right)$$
$$\le \frac{1}{\gamma}P_q\left(|Z_{1/\sqrt{\beta}}^1 - Z_{1/\sqrt{\beta}}^2| \ge \beta^{-1/3}\right) \le C\beta^{2/3}\beta^{-1/2} \to 0 \qquad (2.3)$$

This means if the new born particle survives until time $1/\sqrt{\beta}$ then it is likely to be $\le \beta^{-1/3} = o(\beta^{-1/2})$ from the parent particle.

In the unscaled system the branching rate is β, but not every branching event is successful, namely only those where the new born particle does not coalesce with its parent particle by time $1/\sqrt{\beta}$. Let $T_0 = \inf\{t : Z_t^1 = Z_t^2\}$. We observe that

$$P_q\left(T_0 \geq 1/\sqrt{\beta}\right) \longrightarrow \gamma$$

Hence the rate of successful branching events in the scaled system converges to γ. From this we can conclude that the pruned branching random walk converges to the branching Brownian motion

$$X_T^\beta \xrightarrow{d} Y_T$$

It remains to show that the error we make by replacing the pruned dual process $\hat{\xi}^\beta$ with the pruned branching random walk ξ^β is small. There are two possible errors: a new born particle could coalesce with another particle (not with the parent particle) by time $1/\sqrt{\beta}$, or there could be coalescing events after time $1/\sqrt{\beta}$.

Let $\varepsilon > 0$. The number N_T of all particles in the system $\hat{\xi}^\beta$ till time T (all initial and all new-born) is dominated by a branching process with branching rate 1 and death rate 0. Hence $EN_T \leq ke^T$, so if M is large enough then $P(N_T > M) \leq \varepsilon$. We just have to be concerned with the case where $N_T \leq M$. Let $\bar{Z}_t = Z_t^1 - Z_t^2$ which is then a random walk with jump rate 2 and kernel q. Let $\delta > 0$ such that $(\frac{3}{2}M^3 + 3M^2)\delta < \varepsilon$. The reason for this choice will become clear as the proof goes on. Choose J large enough such that

(i) $P^x\left(\bar{Z}_t = 0 \text{ for some } t\right) < \delta$ for $|x| > J$

(ii) $P^x\left(|\bar{Z}_t| \leq J \text{ for some } t \geq J^8\right) < \delta$ for all x

The first statement follows easily from transience of the random walk. To prove the second statement, we decompose according to T_y the first visit to y after time J^8:

$$\int_{J^8}^\infty ds\, P^x(\bar{Z}_s = y) = \int_{J^8}^\infty ds \int_{J^8}^s P^x(T_y \in dt) P^0(\bar{Z}_{s-t} = 0)$$

$$= P^x\left(\bar{Z}_t = y \text{ for some } t \geq J^8\right) \int_0^\infty ds\, P^0(\bar{Z}_s = 0)$$

Summing over $|y| \leq J$

$$P^x\left(|\bar{Z}_t| \leq J \text{ for some } t \geq J^8\right) \leq C \sum_{|y| \leq J} \int_{J^8}^\infty ds\, P^x(\bar{Z}_s = y)$$

By the local central limit theorem we get

$$P^x\left(|\bar{Z}_t| \leq J \text{ for some } t \geq J^8\right) \leq C \sum_{|y| \leq J} \int_{J^8}^\infty s^{-d/2}\, ds = C J^d J^{-4(d-2)} = C J^{-3d+8}$$

which $\to 0$ since $d \geq 3$.

In the following argument we sum up an estimate for the error probability, e, which we update in each step. If there are k initial particles separated by J then by (i) the

coalescing probability for each pair is smaller then δ. Since there are at most $\frac{1}{2}M^2$ pairs, now $e = \frac{1}{2}M^2\delta$.

Let τ be the time of the first branching in the unscaled system. With high probability this time is larger than J^8, since for small β

$$P(\tau < J^8) \leq M(1 - e^{-\beta J^8}) < M\delta$$

Now $e = \frac{1}{2}M^2\delta + M\delta$.

Using (ii) at time τ the initial particles are separated by J and the new born particle (which has mass 0 at that time) is separated from all other particles except for the parent particle by J with high probability ($e = \frac{1}{2}M^2\delta + M\delta + \frac{1}{2}M^2\delta$). That means for all particles except for the parent particle the probability that this particle coalesces with the new particle is smaller than δ ($e = M^2\delta + M\delta + M\delta$). Again by (ii) after time $\tau + J^8$ all particles are separated by J with high probability ($e = M^2\delta + 2M\delta + \frac{1}{2}M^2\delta$). If the new particle does not coalesce with the parent particle by time $1/\sqrt{\beta}$ it gets mass 1 at this time, but at this time it is separated from all other particles by J, since $1/\sqrt{\beta} > J^8$ for small β. Thus the coalescing probability for each pair of particles containing the new one is smaller than δ by (i) ($e = \frac{3}{2}M^2\delta + 2M\delta + M\delta$). Furthermore, we are in the same setting as in the beginning now with $k+1$ particles. Now we repeat the argument at most M times. Hence $e = M(\frac{3}{2}M^2 + 3M)\delta$ and thus $e < \varepsilon$. Thus the error probability caused by replacing $\hat{X}^\beta$ with X^β is small. This completes the proof of lemma 2.1 for $d \geq 3$.

(b) Consider the case $d = 2$. Let ξ_t be a pruned branching random walk with migration rate 1 and branching rate $2\pi\sigma^2$. Pruned means that a new born particle has mass 0 until time $\tau(\beta) = 1/(\beta\sqrt{\log(1/\beta)})$ and its gets mass 1 if it has not coalesced with its parent. At a migration event the particle jumps according to q. At a branching event a new particle is born at a site randomly chosen according to q. The difference between ξ and $\hat{\xi}$ is that in ξ particles do not coalesce if they meet. Again we interpret the process ξ_t as $X_t = (X_t^1, \ldots, X_t^k, X_t^{k+1}, \ldots, X_t^N, \infty, \ldots)$ analogously to (2.1).

Let ξ_t^β be the rescaled modification of ξ_t, i.e., $\xi_t^\beta(\cdot) = \xi_{h(\beta)t}(\sqrt{h(\beta)}\cdot)$. Let Z^1 and Z^2 be two independent random walks with jump rate 1 and jump kernel q both starting at 0. Let P_q indicate that one random walk is started in the origin and the starting position of the other one is chosen according to q. Let $T_0 = \inf\{t : Z_t^1 = Z_t^2\}$. By Zähle, Cox and Durrett (2003) Lemma 3.1 we know $P_q(T_0 > t) \sim 2\pi\sigma^2(\log t)^{-1}$. Hence

$$P_q\Big(T_0 > \tau(\beta)\Big) \sim \frac{2\pi\sigma^2}{\log\frac{1}{\beta} - \frac{1}{2}\log\log\frac{1}{\beta}} \tag{2.4}$$

By Chebyshev's inequality we have

$$P_q\Big(|Z_{\tau(\beta)}^1 - Z_{\tau(\beta)}^2| \geq \beta^{-1/2}(\log(1/\beta))^{1/3}\Big|T_0 > \tau(\beta)\Big)$$

$$\leq \frac{1}{P_q\Big(T_0 > \tau(\beta)\Big)}P_q\Big(|Z_{\tau(\beta)}^1 - Z_{\tau(\beta)}^2| \geq \beta^{-1/2}(\log(1/\beta))^{1/3}\Big) \tag{2.5}$$

$$\leq C(\log(1/\beta))\frac{\beta}{(\log(1/\beta))^{2/3}} \cdot \frac{1}{\beta(\log(1/\beta))^{1/2}} = C(\log(1/\beta))^{-1/6} \tag{2.6}$$

8

This means if the new born particle survives until time $\tau(\beta)$ then it is likely to be $\leq \beta^{-1/2}(\log(1/\beta))^{1/3} = o(\tau(\beta)^{1/2})$ from the parent particle.

On the other hand they are not too close by the following argument using the local central limit theorem

$$
\begin{aligned}
P_q\Big(|Z^1_{\tau(\beta)} - Z^2_{\tau(\beta)}| &\leq \frac{1}{\sqrt{\beta}(\log(1/\beta))}\Big| T_0 > \tau(\beta)\Big) \\
&\leq \frac{1}{P_q\Big(T_0 > \tau(\beta)\Big)} P_q\Big(|Z^1_{\tau(\beta)} - Z^2_{\tau(\beta)}| \leq \frac{1}{\sqrt{\beta}(\log(1/\beta))}\Big) \\
&\leq C(\log(1/\beta))\left(\frac{1}{\sqrt{\beta}(\log(1/\beta))}\right)^2 \frac{1}{\tau(\beta)} \\
&= C(\log(1/\beta))\frac{1}{\beta(\log(1/\beta))^2}\beta(\log(1/\beta))^{1/2} \leq C(\log(1/\beta))^{-1/2} \quad (2.7)
\end{aligned}
$$

In the unscaled system the branching rate is β, but not every branching event is successful, namely only those where the new born particle does not coalesce with its parent particle by time $\tau(\beta)$. By (2.4) we see that the rate of successful branching events in the scaled system converges to $2\pi\sigma^2$. From this we can conclude that the pruned branching random walk converges to the branching Brownian motion

$$
X^\beta_T \xrightarrow{d} Y_T
$$

It remains to show that the error we make by replacing pruned dual process $\hat{\xi}^\beta$ with the pruned branching random walk ξ^β is small. Again there are two possible errors: a new born particle could coalesce with another particle (not with the parent particle) by time $\tau(\beta)$, or there could be coalescing events after time $\tau(\beta)$. First of all we observe that the number N_T of all particles in the system $\hat{\xi}^\beta$ till time T is dominated by a branching process with branching rate $c = 3\pi\sigma^2$ and death rate 0 if β is small enough. Hence $EN_T \leq ke^{cT}$, so if M is large enough then $P(N_T > M) \leq \varepsilon$. We just have to be concerned with the case where $N_T \leq M$.

Let $\bar{Z}_t = Z^1_t - Z^2_t$ which is then a random walk with jump rate 2 and kernel q. If $x = x(\beta)$ with $|x(\beta)| > 1/(\sqrt{\beta}\log\frac{1}{\beta})$ then as $\beta \to 0$

$$
P^x\Big(\bar{Z}_t = 0 \text{ for some } t \leq h(\beta)T\Big) \longrightarrow 0 \quad (2.8)
$$

This can be seen by the following argument. Consider the potential kernel $A(x) = \sum_{k=0}^\infty [q^k(0) - q^k(x)]$. In Fukai and Uchiyama (1996) one can find the following estimate

$$
A(x) = c\log|x| + O(1) \quad (2.9)
$$

On the other side $\sum_y q(y - x)A(y) - A(x) = \delta(0, x)$, thus if S_n is a random walk with jump distribution q, $E[A(S_{n+1})|F_n] = \sum_x A(S_n + x)q(x) = \sum_y A(y)q(y - S_n) = A(S_n) + \delta(0, S_n)$. That means $A(S_n)$ stopped at the first time when $S_n = 0$ is a martingale. Let α be the probability that the random walk started in $x(\beta)$ hits the ball of radius $2K$

9

before it exits the ball of radius $\frac{1}{\sqrt{\beta}} \log \frac{1}{\beta}$. By the optional sampling theorem together with the last estimate for $A(x)$ we get

$$\frac{1}{2} \log \frac{1}{\beta} - \log \log \frac{1}{\beta} + O(1) = \alpha \Big[\log(2K) + O(1) \Big] + (1 - \alpha) \Big[\frac{1}{2} \log \frac{1}{\beta} + \log \log \frac{1}{\beta} + O(1) \Big]$$

Then

$$\alpha \frac{1}{2} \log \frac{1}{\beta} \leq \log \log \frac{1}{\beta} + O(1)$$

which means that α converges to 0 as $\beta \to 0$. This proves (2.8).

If the initial particles are separated by $1/(\sqrt{\beta} \log \frac{1}{\beta})$ then they do not coalesce by time $h(\beta)T$ with high probability. This follows from (2.8). At the time a new born particle gets non-zero mass (that means it has already survived until time $\tau(\beta)$) it is separated from the parent particle by $1/(\sqrt{\beta} \log \frac{1}{\beta})$. This follows from (2.7). Again by an induction argument similar to the case $d \geq 3$ we see that the error probability is small. This completes the proof. $\qquad \square$

Result for Branching Brownian Motion. Let $e_1 = (1, 0, \ldots, 0)$, $I = [L, 2L]^d$ and $I_m = mLe_1 + I$. Let ζ_t be the branching Brownian motion with branching rate

$$\mu = \begin{cases} \gamma & d \geq 3 \\ 2\pi\sigma^2 & d = 2 \end{cases}$$

where $\gamma = P_q(T_0 = \infty)$ is the escape probability of the random walk. Let $\bar{\zeta}_t$ be a modification of ζ_t in which particles are killed when they land outside $[0, 4L]^d$. It is easy to see (e.g., by checking that two sides of the equation satisfy the same differential equation and initial condition) that

$$E[\bar{\zeta}_t^x(A)] = e^{\mu t} P(\bar{B}_t^x \in A)$$

where $\bar{B}_t^x$ is a Brownian motion that starts at x and is killed when it lands outside $[0, 4L]^d$.

By the scaling invariance of Brownian motion we get

$$P\big(\bar{B}_{L^2}^x \in I_0\big) = P^{x/L}\big(B_s \in [0, 4]^d \text{ for } s \leq 1, B_1 \in [1, 2]^d\big)$$

where B_t is Brownian motion. Hence

$$\liminf_{L \to \infty} \inf_{x \in [L, 2L]^d} P\big(\bar{B}_{L^2}^x \in I_0\big) \geq \inf_{\vartheta \in [1,2]^d} P^\vartheta\big(B_s \in [0, 4]^d \text{ for } s \leq 1, B_1 \in [1, 2]^d\big) > 0$$

We can pick L large enough so that

$$\inf_{x \in [L, 2L]^d} E[\bar{\zeta}_{L^2}^x(I_0)] \geq 2$$

Then obviously for $A \subset [L, 2L]^d$

$$E[\bar{\zeta}_{L^2}^A(I_0)] \geq 2|A|$$

10

where $\bar{\zeta}_t^A$ denotes the modified branching Brownian motion started with one particle at each point of A. Since $E[\bar{\zeta}_{L^2}^x(I_0)^2] \leq E[\zeta_{L^2}^x(\mathbb{R}^d)^2] = c_L < \infty$ and since different particles act independently

$$\text{var}\,[\bar{\zeta}_{L^2}^A(I_0)] \leq |A| c_L$$

it follows by Chebyshev's inequality that if $A \subset [L, 2L]^d$ has $|A| \geq K$ then

$$P\big(\bar{\zeta}_{L^2}^A(I_0) < K\big) \leq \frac{c_L}{K}$$

The same arguments work also for I_1 instead of I_0.

For the given L and a given ε we can choose K large enough such that

$$P\big(\bar{\zeta}_{L^2}^A(I_i) < K\big) \leq \varepsilon \qquad (i = 0, 1) \tag{2.10}$$

for any $A \subset [L, 2L]^d$ with $|A| \geq K$.

Continuity Argument. Recall $\hat{X}_t^\beta$, Y_t and $\mathcal{D}$. By Theorem 2.1 we know $\hat{X}_T^\beta \xrightarrow{d} Y_T$. Let

$$F_k : \mathcal{D} \longrightarrow \mathbb{N}, \quad (y_1, y_2, \dots) \longrightarrow \sum_i 1_{G_i \cap \{y_i(L^2) \in I_k\}}$$

where $G_i = \{y_i(t) \in [0, 4L]^d \cup \{\infty\}$ for $0 \leq t \leq L^2\}$. In words, F_k counts the paths which stay in $[0, 4L]^d$ and end up in I_k. Since F_k is continuous almost surely with respect to the limit distribution, the continuous mapping theorem yields

$$F_k(\hat{X}^\beta) \xrightarrow{d} F_k(Y) \tag{2.11}$$

Let $\bar{\hat{\xi}}_t^\beta$ resp. $\bar{\zeta}_t$ be a modification of $\hat{\xi}_t^\beta$ resp. ζ_t in which particles are killed when they land outside $[0, 4L]^d$. Let $\bar{\hat{\xi}}^{\beta,A}$ resp. $\bar{\zeta}_t^A$ be this system started with one particle at each point of A. If β is small enough then by (2.11) and (2.10) we get

$$P\big(\bar{\hat{\xi}}_{L^2}^{\beta,A}(I_k) < K\big) \leq P\big(\bar{\zeta}_{L^2}^A(I_k) < K\big) + \varepsilon \leq 2\varepsilon \qquad (k = 0, 1) \tag{2.12}$$

Block Construction. At this point we have shown that if the rescaled pruned dual process has $\geq K$ particles in I_0 then in the system with particles killed outside $[0, 4L]^d$ then with probability $\geq 1 - 4\varepsilon$ there will be $\geq K$ particles in I_0 and I_1 at time $h(\beta)$. The existence of a stationary distribution now follows from a comparison with oriented percolation, as explained for example in Section 4 of Durrett (1995).

Lower bound. As proved in 2.10, for the given L and a given ε we can choose K large enough such that

$$P\big(\bar{\zeta}_{L^2}^A(I_i) < K\big) \leq \varepsilon \qquad (i = 0, 1)$$

for any $A \subset [L, 2L]^d$ with $|A| \geq K$. Results of Asmussen and Kaplan (1976) imply that for a branching Brownian motion started at the center of I_0 the number of particles in I_0 at time t is almost surely

$$\sim \frac{c_d L^d}{t^{d/2}} e^{\mu t} W$$

11

where W is a random variable that gives the overall growth rate of the branching process. From this it follows that if t is large enough the probability of $\leq K$ particles in I_0 in the branching Brownian motion is $\leq \varepsilon$. By the continuity results we see that if β is small this probability is $\leq 2\varepsilon$ for a rescaled branching random walk with no killing in the left half-space. It follows that if x_1 is large then the probability of $< K$ particles in $x + I_0$ is $\leq 3\varepsilon$ for the branching process started at x and killing in the left half space. Having shown that the block event probability is $\geq 1 - 6\varepsilon$ in our 1-dependent block construction it follows from Theorem 4.2 in Durrett (1995) that the probability of percolation in the block construction is $\geq 1 - 55(6\varepsilon)^{1/9}$, which gives the lower bound.

Upper bound. For starting points of the rescaled pruned dual process that have $0 \leq x_1 \leq 1$, there is positive probability that there is no branching before time 1 and that hits the left half line by that time. Symmetry implies that for points in the left half plane the probability in the stationary distribution is $\leq 1/2$ and the desired result follows.

3   Proof Of Theorem 1 in $d = 1$

Let $\tilde{\xi}_t^\beta = \tilde{\xi}_{t/\beta^2}(\cdot/\beta)$ be the rescaled dual process. Let $I = [L, 2L]$ and $I_m = mL + I$. We again consider the rescaled dual with particles killed outside $[0, 4L]$. However, this time we assume that there is at least one particle in I_0 and will show that with high probability at time L^2 there will be at least one in I_0 and at least one in I_1.

Consider first the modification of the dual process in which we always follow a particle born to the right. In the rescaled process jumps occur at rate $\beta(1/\beta^2)$ and have expected size $\beta\nu$ where $\nu = \sum_{x>0} xq(x)$. Since $\beta \to 0$ the infinitesimal variance is asymptotically $\sigma^2 = \sum_x x^2 q(x)$. Since the modified dual process has independent increments it is easy to see that it converges to $\sigma B_t + \nu t$.

In time L^2 the modified dual process will move an average distance of νL^2. To get the particles to stay where we want them we declare $x_m = (m + 3/2)L$ to be the target in I_m and at branching points of the dual we follow the particle that is closer to the target if the current location in the rescaled dual is $< x_m - 1$ or $> x_m + 1$. At $x_m + y$ with $y \in [-1, y + 1]$ we follow a birth to the left with probability $(y + 1)/2$ and a birth to the right with probability $(1 - y)/2$. The infinitesimal mean and variance of the approximating chains converge so using results in Section 8.7 of Durrett (1996) the controlled dual process converges to the solution of

$$dZ_t = b(X_t)\, dt + \sigma B_t$$

where $b(x) = -\nu$ for $x \leq -1$, $b(x) = \nu$ for $x \geq 1$ and is $b(x)$ is linear on $[-1, 1]$.

If we let $\hat{Z}_t$ be the diffusion restricted to $[1, \infty)$ with a reflecting barrier at 1, and $\alpha = 2\nu/\sigma^2$ then $e^{\alpha x}$ is a harmonic function on $(1, \infty)$. Suppose now that our diffusion starts from a point $y \in [x_0 + 1, 2L]$. The probability it hits $x_0 + 1$ before $5L/2$ is

$$\geq \frac{e^{\alpha 5L/2} - e^{\alpha 2L}}{e^{\alpha 5L/2} - e^{\alpha(x_0+1)}} \geq 1 - e^{-\alpha L/2}$$

Since the drift is -1 before we hit $x_0 + 1$, this can take time longer than L only if $\sigma B_L > L/2$, an event with probability $\leq \exp(-cL)$. Once $Z_t \leq x_0 + 1$, its movements to the right can be bounded by comparison with $\hat{Z}_t$. Starting from $x_0 + 2$ the probability of hitting $2L$ before $x_0 + 1$ is

$$= \frac{e^{2\alpha} - e^{\alpha}}{e^{\alpha(L/2)} - e^{\alpha}}$$

Each return trip from $x_0 + 2$ to $x_0 + 1$ takes time ≥ 1 with positive probability, so if Z_t starts $\leq x_0 + 1$ the probability of hitting $2L$ before time L^2 tends to 0.

In the same way we can estimate the probability that a particle starting from I_0 hits $L/2$ before $[x_0 - 1, \infty)$, the probability this takes longer than L units of time, and the probability of hitting L within L^2 units of time after reaching $[x_0 - 1, \infty)$. Similar arguments can be used to estimate the movement from I_0 the $[x_1 - 1, x_1 + 1]$ and to show that after the interval is hit, the particle stays in I_1 for L^2 units of time with high probability. All of these calculations have been done for the limiting system but applying the continuity argument gives the desired block event for the controlled dual process.

Lower bound. The estimates on the success of the block construction $\to 1$ as $L \to \infty$. This implies that the survival probability of the rescaled dual started from locations in $[L, 2L]^d$ tends to 1 as $L \to \infty$ which gives the lower bound on the density.

Upper bound in the nearest neighbor case. As remarked in the introduction, in this situation if $\ell_t = \inf\{x : \xi_t(x) = 1\}$ then all sites $\geq \ell_t$ are 1 and those $< \ell_t$ are 0. ℓ_t is a nearest neighbor random walk with drift $-\beta$ when it is > 0 and β if it is < 0. Speeding time up we have $\ell_{t/\beta^2}/\beta \Rightarrow L_t$ the solution of the stochastic differential equation

$$dL_t = dB_t - \text{sgn}\,(L_t)\,dt$$

The upper bound follows immediately from this.

References

Asmusssen, S., and Kaplan, N. (1976) Branching random walks, I+II. *Stoch. Process Appl.* 1, 1–31

Barton, N.H. (1979) The dynamics of hybrid zones. *Heredity.* 43, 341–359

Barton, N.H. (1986) The effect of linkage and density dependent regulation on gene flow. *Heredity.* 43, 333–339

Barton, N.H., and Hewitt, G.M. (1985) Analysis of hybrid zones. *Ann. REv. Ecol. Syst.* 16, 113–148

Barton, N.H and Gale, K.S. (1997) Genetic analysis of hybrid zones. Patges 13–45 in *Hybrid Zones and Evolutionary Process.* Edited by R.G. Harrison, Oxford University Press

Bazykin, A.D. (1969) A hypothetical mechanism of speciation. *Evolution.* 23, 685–687

Cox, J.T., and Durrett, R. (1995) Hybrid zones and voter model interfaces. *Bernoulli.* 1, 343–370

Durrett, R. (1995) Ten lectures on particle systems. Pages 97–201 in *Ecole d'Eté de Probabilitiés de Saint-Flour 1993.* Springer Lecture Notes in Math, 1608

Durrett, R. (1996) *Stochastic Calculus.* CRC Press, Boca Raton, Florida

Durrett, R., Buttel, L., and Harrison, R. (2000) Spatial hybrid zones. *Heredity.* 84, 9–19

Durrett, R., Buttel, L., and Harrison, R. (2003) Genetic structure of mosaic hybrid zones. Preprint.

Ewens, W.J. (2004) *Mathematical Population Genetics.* Springer-Verlag, New York

Fukai, Y. and Uchiyama, K. (1996) Potential kernel for two-dimensional random walk. *Ann. Probab.* 24, 1979–1992

Griffeath, D. (1978) *Additive and cancellative interacting particle systems.* Springer Lecture Notes in Math, 714

Haldane, J.B.S. (1948) the theory of a cline. *Genetics.* 48, 277–284

Harrison, R.G. (1986) Pattern and process in a narrow hybrid zone. *Heredity* 56, 337–349

Harrison, R.G., and Rand, D.M. (1989) Mosaic hybrid zones and the nature of species boundaries. Pages 111–133 in *Speciation and its consequences.* Edited by D. Ottie and J.A. Endler. Sinauer, Sunderland, MA

Harrison, R.G., and Bogdanowicz, S.M. (1997) Patterns of variation and linkage disequilibrium in a field cricket hybrid zone. *Evolution.* 51, 493–505

Hunt, W.G., and Selander, R.K. (1973) Biochemical genetics of hybridization in European house mice. *Heredity.* 31, 11–33

Kruuk, L.E.B., Baird, S.J.E., Gale, K.S., and Barton, N.H. (1999) A comparison of multilocus clines maintained by environmental adaptation or by selection against hybrids. *Genetics.* 153, 1959–1971

Liggett, T.M. (1985) *Interacting Particle Systems.* Springer-Verlag, New York

Rand, D.M., and Harrison, R.G. (1989) Ecological genetics of a mosaic hybrid zone: mitochondrial, nuclear, and repproductive differentiation of crickets by soil type. *Evolution.* 43, 432–449

Slatkin, M. (1973) Gene flow and selection in a cline. *Genetics.* 75, 733–756

Syzmura, J.M., and Barton, N.H. (1986) Genetic analysis of a hybrid zone between the fire-bellied toads *Bombina bombina* and *B. variegata* near Cracow in southern Poland. *Evolution.* 40, 1141–1159