

VOTER MODEL AND BIASED VOTER MODEL IN HETEROGENEOUS ENVIRONMENTS

N. LANCHIER,* *Université de Rouen*

C. NEUHAUSER,** *University of Minnesota*

Abstract

With the rapid adoption of transgenic crops, gene flow from transgenic crops to wild relatives through pollen dispersal is of significant concern and warrants both empirical and theoretical studies to assess the risk of introduction of transgenes into wild populations. We propose to use the (biased) voter model in a heterogeneous environment to investigate the effects of recurrent gene flow from transgenic crop to wild relatives. The model is defined on the d -dimensional integer lattice that is divided into two parts, Δ and $\mathbb{Z}^d \setminus \Delta$. Individuals carrying the transgene and individuals carrying the wild type gene compete according to the evolution rules of a (biased) voter model on $\mathbb{Z}^d \setminus \Delta$, while the process is conditioned to have only individuals carrying the transgene on Δ . Our main findings suggest that unless transgenes confer increased fitness in wild relatives, introgression of transgenes into populations of wild plants is slow and may even be reversible without intervention. Our study also addresses the effects of different spatial planting patterns of transgenic crops on the rate of introgression.

Keywords: Interacting particle systems; voter model interfaces; random walk; transgenic crops

2000 Mathematics Subject Classification: Primary 60K35; 82C22

1. Introduction

The technology of inserting genes into the genomes of organisms for commercial applications was developed in the early 1980s. Since then, genetic engineering has become a standard tool of gene technology for drug and agricultural product design. Genetically modified crop plants (transgenic crops) are now being developed to resist insect herbivory, herbicides, or viral pathogens. A well-known example is Bt maize that is engineered to produce a toxin that is lethal to the larval stage of maize's main insect pest, the European corn borer *Ostrinia nubilalis* (Hübner) (Lepidoptera: Crambidae). Bt maize has been commercially available since the mid-1990s and its use is widespread. Other widely used genetically engineered crops include cotton and soybeans, which have both been engineered to be herbicide tolerant.

Environmental safety of transgenic crops is a major concern, particularly the potential of gene transfer from transgenic crops into natural populations of close relatives through pollination. That this concern is not just a theoretical possibility was demonstrated by Watrud *et al.* (2004) in a recent study of gene flow among the wind-pollinated perennial grass *Agrostis stolonifera* and

Received 13 April 2005; revision received 31 May 2007.

* Current address: Department of Mathematics and Statistics, Arizona State University, Tempe, AZ 85287, USA.
Email address: lanchier@math.asu.edu

** Postal address: Ecology, Evolution and Behavior, University of Minnesota, 1987 Upper Buford Circle, St Paul, MN 55108, USA.

Partially supported by NSF Grants DMS-00-72262 and DMS-00-83468.

some of its close relatives. It has been genetically modified to express resistance to glyphosate, the active ingredient of Roundup herbicide (Monsanto, St. Louis, MO). Experiments by Watrud *et al.* (2004) showed that gene flow typically occurs within 2 km, but long-distance dispersal events of up to 21 km were also observed.

Crop plants are descended from wild plants and have close relatives among them. It is therefore no surprise that gene flow between crop plants and wild plants has occurred in the past (Ellstrand *et al.* (1999)). However, with novel genes being inserted into the genomes of crop plants to express proteins for specific biological functions, there is increased concern that these novel genes would escape into the wild and confer increased fitness to some species, which could then become aggressive invaders.

Few mathematical models have explored the potential of gene flow from genetically engineered crops to wild relatives; see Huxel (1999), Wolf *et al.* (2001), and Haygood *et al.* (2003). They found that even if selection disfavors the transgene, invasion into the wild population is possible. Their models point to the need for closer monitoring of gene flow from genetically engineered crops into wild populations.

None of the models mentioned above is spatially explicit. Since gene flow is an inherently spatial process, the lack of a spatial component is potentially a serious shortcoming. To begin to understand how the spatial configuration of permanent plots of genetically modified crop plants affect gene flow and invasion of transgenes into adjacent natural populations of close relatives, we propose a spatially explicit, stochastic model in a heterogeneous environment. The environment will reflect the permanent plots of genetically modified organisms embedded in a matrix of wild plants. Since we think of the genetically modified crop plants as planted, gene flow will occur only within the wild plants and from the genetically modified plants to the wild plants. The model is based on the simplest population genetics model, the haploid Wright–Fisher model with selection. The spatial analog of the Wright–Fisher model is the (biased) voter model. We will define the voter model in a heterogeneous environment with gene flow in such a way to address the problem of recurrent gene flow from genetically modified crop plants to their wild relatives.

The heterogeneous voter model is a continuous-time Markov process in which the state at time $t \geq 0$ is $\{0, 1\}^{\mathbb{Z}^d}$. In our context, we think of 1s as representing individuals carrying the transgene and 0s as representing individuals carrying the wild type gene. A site $x \in \mathbb{Z}^d$ is said to be occupied by a 1 or 0 if $\xi_t(x) = 1$ or 0, respectively. We will also use the notation $x \in \xi_t$ to indicate that $\xi_t(x) = 1$. We let $\Delta \subset \mathbb{Z}^d$, and think of Δ as representing the part of the habitat that is occupied by transgenic crop plants. The dynamics are defined as follows. The evolution on Δ is frozen by pretending that $\xi_t(x) = 1$ for any $x \in \Delta$ and $t \geq 0$. To define the evolution on $\mathbb{Z}^d \setminus \Delta$, we let $p(x, y)$ be a transition probability on \mathbb{Z}^d that is translation invariant, i.e. $p(x, y) = q(y - x)$ for some probability density function q ; symmetric, i.e. $q(z) = q(-z)$, $z \in \mathbb{Z}^d$; and such that the probability density $q(z)$ has variance $\sigma^2 < \infty$. If a site $x \in \mathbb{Z}^d \setminus \Delta$ is occupied by a 1 or 0, then, at rate 1 or β , respectively, it picks a site $y \in \mathbb{Z}^d$ with probability $p(x, y)$ and adopts the state of the individual at y . More formally, the state of $x \in \mathbb{Z}^d \setminus \Delta$ flips according to the transition rates

$$\begin{aligned}
 0 &\rightarrow 1 && \text{at rate } \beta \sum_{y \in \mathbb{Z}^d} p(x, y) \xi(y), \\
 1 &\rightarrow 0 && \text{at rate } \sum_{y \in \mathbb{Z}^d} p(x, y) (1 - \xi(y)).
 \end{aligned}$$

Except in the homogeneous case in which $\Delta = \emptyset$, we will assume from now on that $\xi_0(x) = 1$ if and only if $x \in \Delta$. It follows from the verbal definition of the model that $\xi_t(x) = 1$ for all $x \in \Delta$ and $t \geq 0$.

1.1. The homogeneous voter model: $\Delta = \emptyset$

In this case and when $\beta = 1$, the results of Holley and Liggett (1975) reveal a sharp change in behavior depending on the spatial dimension.

Theorem 1. (Holley and Liggett (1975).) *Assume that the process ξ_t has as initial distribution the Bernoulli product measure with density θ . Then, ξ_t converges in distribution to the measure μ as $t \rightarrow \infty$, where*

$$\mu = \begin{cases} (1 - \theta)\delta_0 + \theta\delta_1 & \text{in } d \leq 2, \\ \nu_\theta & \text{in } d \geq 3, \end{cases}$$

where δ_0 and δ_1 concentrate on the ‘all 0’ and the ‘all 1’ configurations, respectively, whereas ν_θ is a nontrivial measure in which $\nu_\theta(\xi_\infty(x) = 1) = \theta$ for any $x \in \mathbb{Z}^d$.

Note that choosing $\Delta \neq \emptyset$ can drastically change the limiting behavior of the process. For instance, in contrast to Theorem 1, if we set $\Delta = \{0\}$ and $d \leq 2$, then the 1, which is initially located at the origin, will produce a cluster that will invade the whole space. This is a straightforward consequence of the duality techniques used to prove Theorem 1 in $d \leq 2$; see Holley and Liggett (1975) for further details. When $\Delta = \emptyset$ and $\beta > 1$, the results of Bramson and Griffeath (1980), (1981) show that, in any dimension, if we let $\Omega_\infty = \{\xi_t \neq \emptyset \text{ for all } t \geq 0\}$, then starting from a configuration where there is a single 1 at the origin, there is a convex set A so that conditioned on the event Ω_∞ we have, for any $\varepsilon > 0$, with probability 1,

$$(1 - \varepsilon)tA \cap \mathbb{Z}^d \subset \xi_t \subset (1 + \varepsilon)tA \cap \mathbb{Z}^d \quad \text{for all sufficiently large } t.$$

Moreover, on the event $\Omega_\infty^c = \{\xi_t = \emptyset \text{ for some } t \geq 0\}$, the process converges to the “all 0” configuration exponentially fast.

We now fix $d = 1$. Due to assumptions on the transition probability $p(x, y)$, the discussion above implies that if $\Delta \neq \emptyset$ then $\xi_t \Rightarrow \delta_1$. The first question we would like to answer is: what is the effect of the geometry of Δ on the speed of convergence to the ‘all 1’ configuration? The first step is to investigate the process starting from $\xi_0(x) = \mathbf{1}_{\{y \leq 0\}}(x)$. When $\Delta = \emptyset$ and $\beta = 1$, the behavior of the process in $d = 1$ at the interface has been studied by Cox and Durrett (1995). To introduce their results, we consider the leftmost-zero and the rightmost-one processes

$$\ell_t = \inf\{x \in \mathbb{Z}: \xi_t(x) = 0\} \quad \text{and} \quad r_t = \sup\{x \in \mathbb{Z}: \xi_t(x) = 1\},$$

respectively. Moreover, we assume, for technical reasons, that

- (i) p is irreducible, i.e. it is possible to get from 0 to any $x \in \mathbb{Z}$ in a finite number of steps with positive probability, and
- (ii) q has finite absolute third moments, i.e. $\sum_{z \in \mathbb{Z}} |z|^3 q(z) < \infty$.

Then $\{r_t - \ell_t : t \geq 0\}$ is stochastically compact, that is

$$\text{for all } \varepsilon_0 > 0 \text{ there exists an } M < \infty \text{ such that } P(r_t - \ell_t \geq M) \leq \varepsilon_0.$$

See Cox and Durrett (1995, Section 4). Moreover, if we denote by $\Phi(x)$ the standard normal distribution function, then

$$\lim_{t \rightarrow \infty} P\left(\frac{r_t}{\sigma\sqrt{t}} \leq x\right) = \lim_{t \rightarrow \infty} P\left(\frac{\ell_t}{\sigma\sqrt{t}} \leq x\right) = \Phi(x), \quad x \in \mathbb{R}, \tag{1}$$

under the assumption that $\beta = 1$.

1.2. The heterogeneous voter model: $\Delta \neq \emptyset$

In the case $\Delta = \mathbb{Z}^-$, where \mathbb{Z}^- denotes the set of nonpositive integers, we will first prove that the family $\{r_t - \ell_t : t \geq 0\}$ is still stochastically compact. Equation (1), however, becomes false since r_t and ℓ_t are now forced to live on the right side of 0. More precisely, we will prove the following theorem.

Theorem 2. *Let $\Delta = \mathbb{Z}^-$ and $\beta = 1$. If $x \geq 0$, then*

$$\lim_{t \rightarrow \infty} P(r_t \geq x\sigma\sqrt{t}) = \lim_{t \rightarrow \infty} P(\ell_t \geq x\sigma\sqrt{t}) = 2(1 - \Phi(x)) = P(X \geq x \mid X \geq 0),$$

where X is Gaussian with mean 0 and variance 1.

In words, particles diffuse to the right revealing a very slow invasion. In view of Theorem 2, our guess is that the more Δ is scattered, the faster the transgene invades the wild population. To make this argument precise, we consider the process on the torus $\varepsilon\mathbb{Z}/\mathbb{Z}$, where $\varepsilon > 0$ is close to 0. Moreover, we speed up time by ε^{-2} and introduce the transition probability

$$p_\varepsilon(x, y) = q(\varepsilon^{-1}(y - x)) \quad \text{for all } x, y \in \varepsilon\mathbb{Z}/\mathbb{Z},$$

to formulate the dynamics. We let $L > 0$ and $K = L^{-1}$ such that both K and $L\varepsilon^{-1}$ are integers, with K even. For any $z \in \{0, 1, \dots, K - 1\}$, we define the subsets

$$D_z = [zL, (z + 1)L) \cap \varepsilon\mathbb{Z} \quad \text{and} \quad \Delta = \bigcup_{z \text{ even}} D_z. \tag{2}$$

The reader will note that since we start the process with $\xi_0(x) = 1$ if and only if $x \in \Delta$, for any choice of L , half of the sites are initially occupied, and half of the sites are initially vacant. Finally, we denote by T_{inv} the stopping time, given by $T_{\text{inv}} = \inf\{t \geq 0 : \xi_t \equiv 1\}$. We refer to T_{inv} as the time to complete invasion.

Theorem 3. *Consider the process ξ_t on the torus. Let Δ be given by (2) and set $\beta = 1$. If q has compact support, i.e. there exists $R > 0$ such that $p(x, y) = 0$ whenever $|x - y| > R$, then*

$$\lim_{\varepsilon \rightarrow 0} P(T_{\text{inv}} \leq t) = (1 - (P(-L \leq Y \leq L))^2)^{1/2L}, \tag{3}$$

where Y is Gaussian with mean 0 and variance $2\sigma^2t$.

Given $z \in \{0, 1, \dots, K - 1\}$ even, $(P(-L \leq Y \leq L))^2$ is the probability that the transgene has not totally invaded the interval D_z at time t . This can be proved by using the duality between the voter model and coalescing random walks. The right-hand side of (3) then follows from the additional fact that the times to invasion of D_0, D_2, \dots, D_{K-2} are independent if $\varepsilon > 0$ is small enough. Theorem 3 gives insights into the effects of the geometry of Δ on the time to complete invasion. Numerical investigations suggest that, for fixed $t \geq 0$, the limit $\lim_{\varepsilon \rightarrow 0} P(T_{\text{inv}} \geq t)$

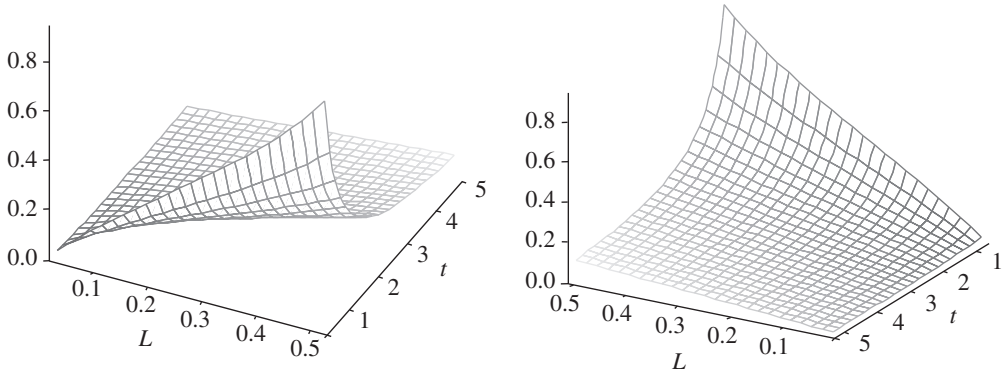


FIGURE 1: $\lim_{\epsilon \rightarrow 0} P(T_{\text{inv}} > t)$ as a function of L and t ($\sigma^2 = 1$).

is an increasing function of the parameter L . See Figure 1 for an example. This implies that the time to complete invasion will be slowest in the extreme case in which $L = \frac{1}{2}$, and faster as the number of components D_z of Δ increases.

Unfortunately, we do not know how to extend Theorem 3 to higher dimensions. To be convinced, however, that the analogous result holds in $d = 2$, we simulated the process on a 200×200 lattice with periodic boundary conditions and simplest nearest neighbor interactions, tiled into 20×20 squares $D_{i,j}$, $i, j = 1, 2, \dots, 10$. Figure 2 shows the time to complete invasion in three different environments. In the chessboard model, Δ is the set of $D_{i,j}$ with $i + j$ even. In the striped model, Δ is the set of $D_{i,j}$ with i even. Finally, in the four-patch model, Δ is the set of $D_{i,j}$ with $i, j = 1, 2, \dots, 5$ or $i, j = 6, 7, \dots, 10$. As suggested by Figure 2, the simulations reveal that T_{inv} is a decreasing function of the number of components D_z of Δ .

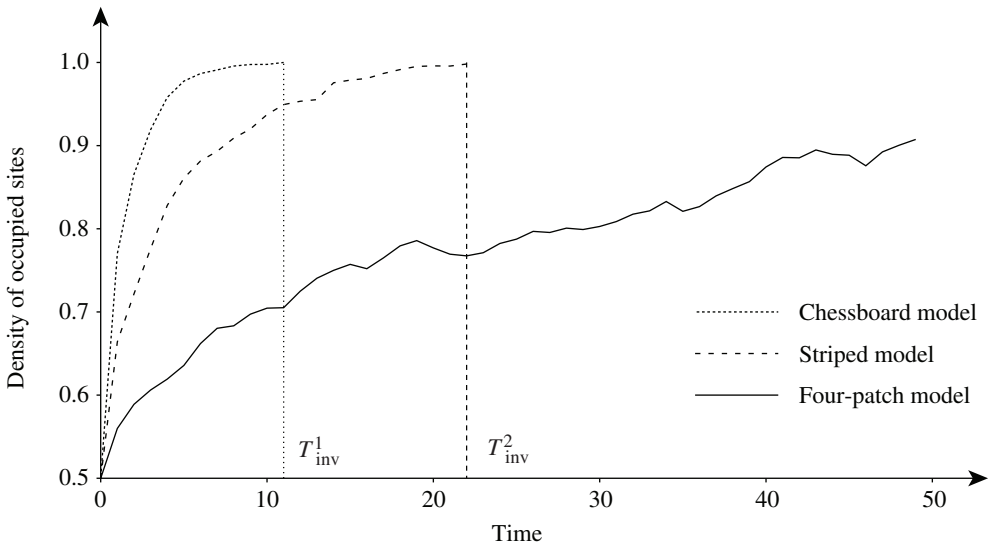


FIGURE 2: Evolution to resistance.

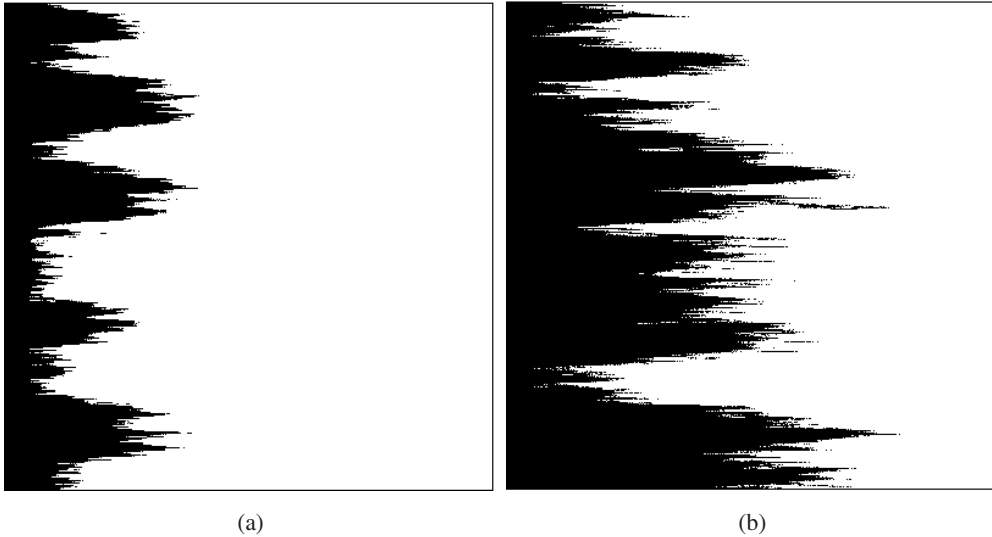


FIGURE 3: Voter model with $\Delta = \mathbb{Z}^-$ in the interval $[-20, 380]$ from time 0 at the top to time 20 000 at the bottom. (a) The range of the interactions is 2 and the number of returns to 0 is 91. (b) The range of the interactions is 4 and the number of returns to 0 is 6.

We now consider the process on \mathbb{Z} , where the particles give birth according to the transition probability $p(x, y)$. Let $\Delta = \mathbb{Z}^-$ and start the process with $\xi_0(x) = 1$ if and only if $x \in \Delta$. Moreover, we assume that q has compact support. The next step is to investigate the properties of the rightmost-one process, r_t , depending on the birth rate β . The limiting behavior of r_t when $\beta = 1$ follows from Theorem 2. By working a little bit more, we can prove that the process comes back to its initial configuration infinitely often (i.o.). See Figure 3 for a picture.

Theorem 4. Assume that q has compact support, i.e. there exists $R > 0$ such that $p(x, y) = 0$ whenever $|x - y| > R$. If $\Delta = \mathbb{Z}^-$ and $\beta = 1$ then $\mathbb{P}(r_t = 0 \text{ i.o.}) = 1$.

To investigate the case in which $\beta \neq 1$, we consider the process with the simplest nearest neighbor interactions, i.e. we let $p(x, y) = \frac{1}{2}$ if $|x - y| = 1$ and 0 otherwise. If $\beta < 1$, Theorem 4 accompanied by a coupling argument implies that $r_t = 0$ i.o.; the process, however, does not exhibit the behavior described in Theorem 2.

Theorem 5. Assume nearest neighbor dispersal, that is, $p(x, y) = \frac{1}{2}$ when $|x - y| = 1$ and 0 otherwise. If $\Delta = \mathbb{Z}^-$ and $\beta < 1$ then $\mathbb{P}(r_t \geq x) \leq \beta^x$ at any time $t \geq 0$.

Finally, if $\beta > 1$, the process behaves like the corresponding biased voter model in a homogeneous environment, namely the process grows linearly in time with a wave speed of order $\beta - 1$. See Figure 4 for a picture.

Theorem 6. Assume nearest neighbor dispersal. If $\Delta = \mathbb{Z}^-$ and $\beta > 1$, there exist $C < \infty$ and $\gamma > 0$ such that, for any $t > 0$ and any $\varepsilon > 0$,

$$\mathbb{P}(|r_t - (\beta - 1)t| \geq \varepsilon t) \leq C e^{-\gamma \varepsilon^2 t}.$$

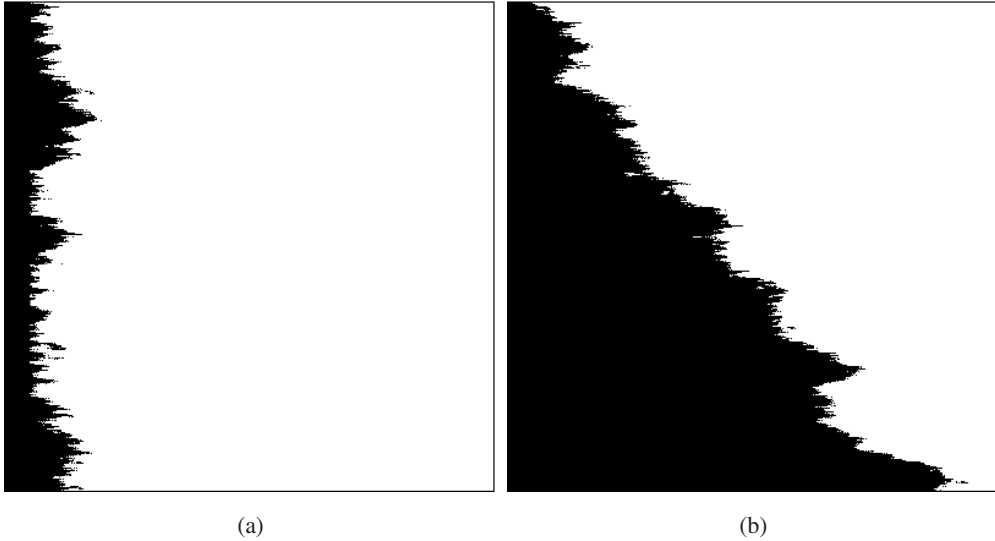


FIGURE 4: Range 2 voter model with $\Delta = \mathbb{Z}^-$ in the interval $[-20, 380]$ from time 0 at the top to time 4000 at the bottom and (a) $\beta = 0.95$, (b) $\beta = 1.05$.

In conclusion, Theorems 2 and 5–6 imply that

1. if $\beta \geq 1$ then, for any $x \geq 0$ and $\varepsilon > 0$, there exists a time $t_0 \geq 0$ such that

$$P(r_t \geq x) \geq 1 - \varepsilon \quad \text{for all } t \geq t_0,$$

2. if $\beta < 1$, then $P(r_t \geq x) \leq \beta^x$ for any $t \geq 0$.

Moreover, Theorems 4–6 exhibit the 0-1 law:

$$P(r_t = 0 \text{ i.o.}) = \begin{cases} 0 & \text{if } \beta > 1, \\ 1 & \text{if } \beta \leq 1. \end{cases}$$

In particular, if $\beta = 1$ and q has compact support, the rightmost-one process r_t converges in probability to ∞ as $t \rightarrow \infty$, but not almost surely.

Our results confirm the need to monitor recurrent gene flow from genetically modified crops to their wild relatives, but our results differ in an important way from the results of nonspatial models analyzed by other researchers (Huxel (1999), Wolf *et al.* (2001), and Haygood *et al.* (2003)) who consistently found that the transgene would eventually go to fixation over a wide range of parameters even if not favored by selection. Both selection and the spatially explicit structure of our model greatly affect the rate and extent of invasion of transgenes into wild populations. If the transgene is favored by selection, it will completely take over a wild population. In the neutral case, the transgene can penetrate a wild population arbitrarily far, which can result in fixation when the wild population is only of limited spatial extent, but the invasion is slow. If the transgene is not favored, it will continue to be present in the wild population and the extent to which it can penetrate into the the wild population depends strongly on the strength of selection. Simulations strongly indicate that the dispersal distance is a key factor in how quickly and to what extent transgenes can invade natural populations. Considering

that pollen can disperse over very large spatial distances, genetic pollution of wild population remains a serious concern, even if the transgene is not favored by selection.

Modeling at this conceptual level can only point out the various factors that affect transgene introgression. It does by no means replace necessary field work and system specific simulation models that are parameterized by field data. However, our work clearly shows that spatially explicit and stochastic models can contribute to a deeper understanding of the problem of recurrent gene flow by identifying key parameters that need to be measured to assess risk.

The paper is organized as follows. Section 2 provides some preliminary results. Section 3 is devoted to the proofs of Theorems 2 and 3, Section 4 to the proof of Theorem 4, and Section 5 to the proofs of Theorems 5 and 6. Figures 3 and 4 provide some visualizations of the effects of dispersal range and selection strength.

2. Duality: preliminary results

We start by constructing the process ξ_t , as defined in Section 1, from a collection of Poisson processes in the case in which $\beta \geq 1$. For each $x \in \mathbb{Z}^d \setminus \Delta$, we let $\{S_n^x : n \geq 1\}$ and $\{T_n^x : n \geq 1\}$ be the arrival times of independent Poisson processes with rate 1 and $\beta - 1$, respectively, and we let $\{U_n^x : n \geq 1\}$ and $\{V_n^x : n \geq 1\}$ be independent and identically distributed sequences with $P(U_n^x = y) = P(V_n^x = y) = p(x, y)$. At times S_n^x we draw an arrow from U_n^x to x and put a δ at x , while at times T_n^x we just draw an arrow from V_n^x to x . Then the process is obtained from the graphical representation as follows. At time S_n^x , the site x imitates the site U_n^x , i.e. becomes occupied by a 1 or 0 if U_n^x is occupied by a 1 or 0, respectively. At time T_n^x , the site x becomes occupied by a 1 if V_n^x is occupied by a 1. An idea of Harris (1972) implies that such a graphical representation can be used to construct the process starting from any initial configuration $\xi_0 : \mathbb{Z}^d \rightarrow \{0, 1\}$. See Figure 5(a) for an illustration. After constructing the graphical representation, we can now define the dual process. We say that there is a path from $(z, 0)$ to (x, t) if there is a sequence of times $s_0 = 0 < s_1 < \dots < s_{n+1} = t$ and spatial locations $x_0 = z, x_1, \dots, x_n = x$ such that the following two conditions hold.

1. For $i = 1, 2, \dots, n$, there is an arrow from x_{i-1} to x_i at time s_i .
2. For $i = 0, 1, \dots, n$, the vertical segment $\{x_i\} \times (s_i, s_{i+1})$ does not contain any δ s.

We say that there exists a dual path from (x, t) to $(z, t - s)$, $0 \leq s \leq t$, if there is a path from $(z, t - s)$ to (x, t) , and we define the dual process starting at (x, t) by setting

$$\hat{\xi}_s^{(x,t)} = \{z \in \mathbb{Z}^d : \text{there is a dual path from } (x, t) \text{ to } (z, t - s)\}$$

for any $0 \leq s \leq t$. See Figure 5(b) for an illustration. The dual process is naturally defined only for $0 \leq s \leq t$, but it is convenient to assume that the Poisson processes involved in the graphical representation are defined for negative times so that the dual process can be defined for all $s \geq 0$. The dual process allows us to deduce the state of site x at time t from the configuration at earlier times. Recall that $x \in \xi_t$ if and only if $\xi_t(x) = 1$. Then,

$$x \in \xi_t \iff \hat{\xi}_s^{(x,t)} \cap \xi_{t-s} \neq \emptyset.$$

See, e.g. Durrett (1988, Chapter 3). In the homogeneous case, $\Delta = \emptyset$, the dual process is a well-known object. If $\beta = 1$, the process $\hat{\xi}_s^{(x,t)}$ performs a continuous-time random walk on \mathbb{Z}^d run at rate 1. Starting the dual process from two sites x and y amounts to running two individual dual processes $\hat{\xi}_s^{(x,t)}$ and $\hat{\xi}_s^{(y,t)}$ independently until their \mathbb{Z}^d coordinates are identical,

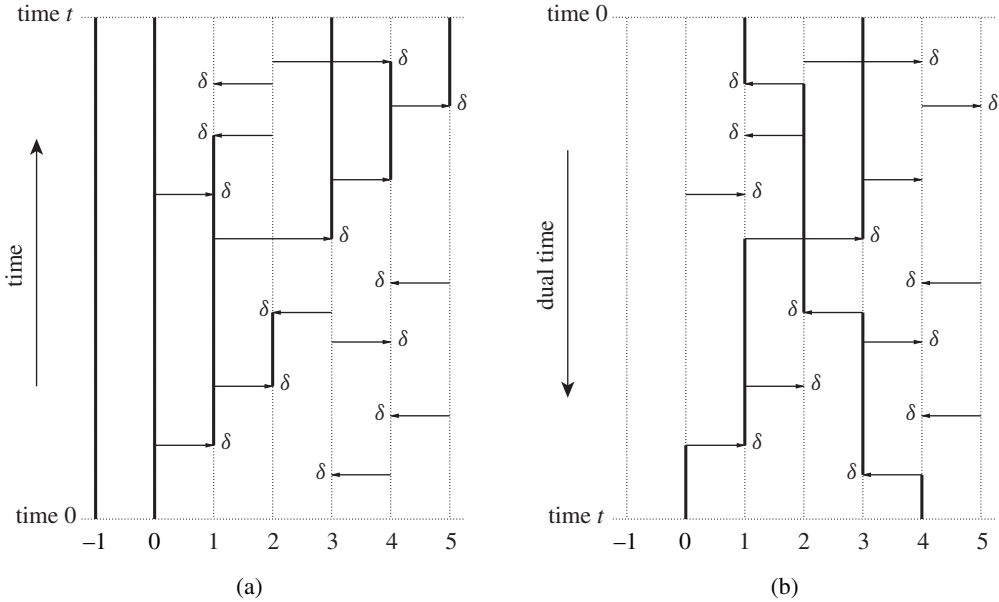


FIGURE 5: (a) Construction of the process from Harris' graphical representation. Black sites refer to sites occupied by a type 1 particle. (b) Construction of the dual process starting from {1, 3} in thick lines. Dual paths are constructed from Harris' graphical representation by going backwards in time and reversing the direction of the arrows. Note that, in our example, an inversion occurs.

at which point they coalesce, i.e. $\hat{\xi}_s^{(x,t)} - \hat{\xi}_s^{(y,t)}$ has 0 as an absorbing state. After coalescing, the two dual processes move together according to the rules of the dual process of a single random walk. This implies a duality between the voter model and coalescing random walks when $\beta = 1$. If $\beta > 1$, the dual process branches when it encounters the tip of a single arrow. More precisely, we start with a single particle at x at time $s = 0$. If a particle in $\hat{\xi}_s^{(x,t)}$ meets a δ at some time $s = t - S_n^z$ then this particle moves to U_n^z . If a particle in $\hat{\xi}_s^{(x,t)}$ meets the tip of a single arrow at time $s = t - T_n^z$ then this particle gives birth to a new particle which is then sent to V_n^z . In conclusion, for $\beta > 1$, there is a duality between the biased voter model and coalescing branching random walks.

In the heterogeneous case, $\Delta \neq \emptyset$, we observe that if $z \in \Delta$ then no arrows point at z . If $\beta = 1$, the process $\hat{\xi}_s^{(x,t)}$ then evolves like a continuous-time random walk until the first time it visits Δ , where it becomes frozen. It is convenient to construct the dual process in the heterogeneous environment from that in a homogeneous environment. If we denote by ζ_t the process constructed from the graphical representation in a homogeneous environment, then ξ_t can be constructed from the graphical representation obtained by removing all the arrows that point at a site $z \in \Delta$. Since both processes have the same graphical representation on $\mathbb{Z}^d \setminus \Delta$, $\hat{\xi}_s^{(x,t)}$ and $\hat{\zeta}_s^{(x,t)}$ move together until they hit Δ when $\hat{\xi}_s^{(x,t)}$ is frozen. In the same way, if $\beta > 1$, the particles in each dual move and give birth together as long as they belong to $\mathbb{Z}^d \setminus \Delta$. When a particle in $\hat{\xi}_s^{(x,t)}$ hits the set Δ it is frozen, whereas the corresponding particle in $\hat{\zeta}_s^{(x,t)}$ continues to move and give birth.

The rest of this section is devoted to the proof of a preliminary result, which is the main ingredient in establishing Theorems 2–4. From now on, we set $d = 1$, let $\Delta = \mathbb{Z}^-$ be the

set of nonpositive integers and assume that $\xi_0(x) = 1$ if and only if $x \in \Delta$. We define the leftmost-zero and the rightmost-one processes by

$$\ell_t = \inf\{x \in \mathbb{Z} : \xi_t(x) = 0\} \quad \text{and} \quad r_t = \sup\{x \in \mathbb{Z} : \xi_t(x) = 1\},$$

respectively.

Proposition 1. *Let $\Delta = \mathbb{Z}^-$ and $\beta = 1$. Then the family $\{r_t - \ell_t : t \geq 0\}$ is stochastically compact. That is, for any $\varepsilon_0 > 0$ there is $M > 0$ so that $P(r_t - \ell_t \geq M) \leq \varepsilon_0$ at any time $t \geq 0$.*

In the case in which $\Delta = \emptyset$, Proposition 1 was proved by Cox and Durrett (1995, Section 4). Their proof relied on a strong analysis of the dual process and random walks estimates. They introduced the following concept, which is key to proving Proposition 1. The event that there are $x, y \in \mathbb{Z}$ such that $x < y$ and $\hat{\xi}_t^{(y,t)} \leq 0 < \hat{\xi}_t^{(x,t)}$ will be called an inversion by time t . See Figure 5(b) for an example of an inversion event. If we let B_t be the number of inversions that occur by time t , then Cox and Durrett (1995) showed that

$$-1 \leq r_t - \ell_t \leq B_t.$$

The proof of Proposition 1 can then be reduced to an analysis of the dual process. See Cox and Durrett (1995, Section 4). To deduce Proposition 1 from the result in the homogeneous case, we will prove that there exists a constant $C < \infty$ such that $E(B_t) \leq C E(B'_t)$, where B'_t denotes the number of inversions by time t for the process ζ_t defined above. To compare the dual processes, we introduce the following notation. For any $x, y \in \mathbb{Z}$, we consider the stopping times

$$\begin{aligned} \sigma_{x,y} &= \inf\{s \geq 0 : \hat{\xi}_s^{(x,t)} = \hat{\xi}_s^{(y,t)} \text{ or } \hat{\xi}_s^{(x,t)}, \hat{\xi}_s^{(y,t)} \leq 0\}, \\ \tau_{x,y} &= \inf\{s \geq 0 : \hat{\xi}_s^{(x,t)} = \hat{\xi}_s^{(y,t)}\}, \\ \text{and } s_x &= \inf\{s \geq 0 : \hat{\xi}_s^{(x,t)} \leq 0\}. \end{aligned}$$

Moreover, we let $\varrho_{x,y} = \min(\sigma_{x,y}, \tau_{x,y})$ and $s_{x,y} = \min(s_x, s_y)$. Note that the law of the stopping times we have just defined does not depend on the starting time t . We will prove Proposition 1 with the help of the forthcoming lemmas. For a better understanding of Lemmas 1–4, we refer the reader to Figure 6, which gives a picture of the stopping times we have just introduced. Finally, we would like to remind the reader that our construction implies that $\hat{\xi}_s^{(x,t)}$ and $\hat{\xi}_s^{(x,t)}$ move together as long as their position is to the right of 0.

Lemma 1. *For any $z \in \mathbb{Z}^-$ and $w \in \mathbb{Z}$, $P(\sigma_{z,w} > 2t) \leq P(\tau_{z,w} > t)$.*

Proof. If $w \leq 0$, the inequality is trivial since in this case $\hat{\xi}_0^{(w,t)} = w \leq 0$ and so $\sigma_{z,w} = 0$. To deal with the case in which $w > 0$, we let

$$X_s = \hat{\xi}_s^{(w,t)} - \hat{\xi}_s^{(z,t)} \quad \text{and} \quad Y_s = \hat{\xi}_s^{(w,t)} - \hat{\xi}_s^{(z,t)}.$$

Since $z \leq 0$ we have $\hat{\xi}_s^{(z,t)} = z$ at any time $s \geq 0$ by construction. This implies that X_s performs a rate-one random walk starting at $w - z$ until time $\sigma_{z,w}$ when X_s is stopped. Conversely, the process Y_s performs a rate-two random walk starting at $w - z$ until time $\tau_{z,w}$ when Y_s hits its absorbing state 0. In particular, since $z \leq 0$,

$$P(X_s > -z \text{ for all } s \leq 2t) = P(Y_s > -z \text{ for all } s \leq t) \leq P(Y_s \neq 0 \text{ for all } s \leq t).$$

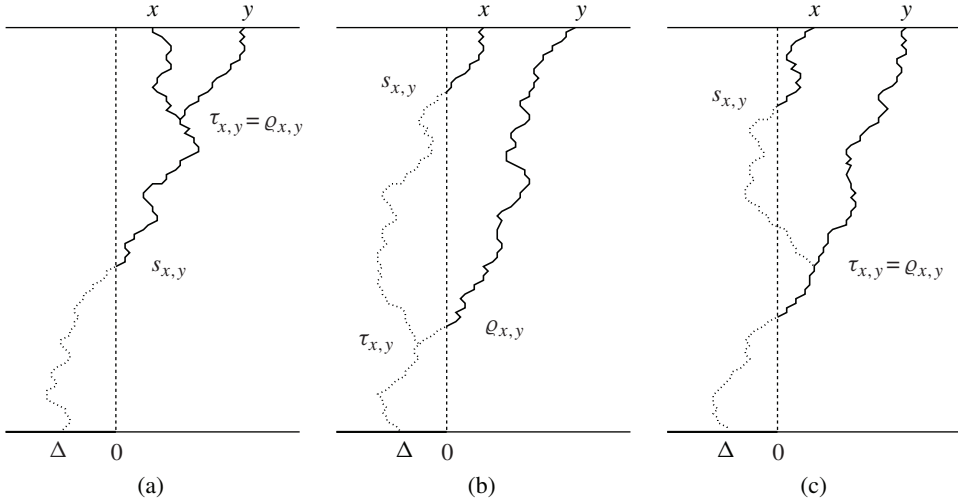


FIGURE 6: Pictures of the dual processes. The solid lines refer to the path of the dual processes moving together. The dotted lines refer to the path of $\hat{\zeta}_s^{(\cdot,t)}$ after $\hat{\zeta}_s^{(\cdot,t)}$ has been frozen.

By observing that

$$\{\sigma_{z,w} > 2t\} = \{X_s > -z \text{ for all } s \leq 2t\} \quad \text{and} \quad \{\tau_{z,w} > t\} = \{Y_s \neq 0 \text{ for all } s \leq t\}$$

the lemma follows.

Lemma 2. For any $x, y \in \mathbb{Z}$, $P(\sigma_{x,y} > 2t; s_{x,y} < \rho_{x,y}) \leq P(\tau_{x,y} > t; s_{x,y} < \rho_{x,y})$.

Proof. We first observe that

$$\begin{aligned} P(\sigma_{x,y} > t; s_{x,y} < \rho_{x,y}) &= P(\sigma_{x,y} > t; s_{x,y} < \rho_{x,y}; s_x < s_y) \\ &\quad + P(\sigma_{x,y} > t; s_{x,y} < \rho_{x,y}; s_x > s_y). \end{aligned}$$

We will prove the assertion for $s_x = s_y$. (See Figures 6(b) and (c) for an illustration.) By the Markov property

$$\begin{aligned} &P(\sigma_{x,y} > 2t; s_x < \rho_{x,y}; s_x < s_y) \\ &= \int_0^\infty P(\sigma_{x,y} > 2t; \rho_{x,y} > s; s_y > s) P^x(s_x \in ds) \\ &= \sum_{z \leq 0} \sum_{w > 0} \int_0^\infty P(\sigma_{x,y} > 2t; \rho_{x,y} > s; s_y > s; \hat{\xi}_s^{(x,t)} = z; \hat{\xi}_s^{(y,t)} = w) P^x(s_x \in ds) \\ &= \sum_{z \leq 0} \sum_{w > 0} \int_0^\infty P(\sigma_{z,w} > 2t - s) \\ &\quad \times P(\rho_{x,y} > s; s_y > s; \hat{\xi}_s^{(x,t)} = z; \hat{\xi}_s^{(y,t)} = w) P^x(s_x \in ds). \end{aligned}$$

Similarly,

$$\begin{aligned} &P(\tau_{x,y} > t; s_x < \varrho_{x,y}; s_x < s_y) \\ &= \sum_{z \leq 0} \sum_{w > 0} \int_0^\infty P(\tau_{z,w} > t - s) \\ &\quad \times P(\varrho_{x,y} > s; s_y > s; \hat{\xi}_s^{(x,t)} = z; \hat{\xi}_s^{(y,t)} = w) P^x(s_x \in ds). \end{aligned}$$

Now, since $\hat{\xi}_s^{(x,t)} = \hat{\zeta}_s^{(x,t)}$ and $\hat{\xi}_s^{(y,t)} = \hat{\zeta}_s^{(y,t)}$ until time $s = s_{x,y}$ it follows that

$$\begin{aligned} &P(\varrho_{x,y} > s; s_y > s; \hat{\xi}_s^{(x,t)} = z; \hat{\xi}_s^{(y,t)} = w; s_x = s) \\ &= P(\varrho_{x,y} > s; s_y > s; \hat{\zeta}_s^{(x,t)} = z; \hat{\zeta}_s^{(y,t)} = w; s_x = s). \end{aligned}$$

Furthermore, Lemma 1 implies that, for $z \leq 0$,

$$P(\sigma_{z,w} > 2t - s) \leq P(\sigma_{z,w} > 2(t - s)) \leq P(\tau_{z,w} > t - s).$$

This completes the proof.

Lemma 3. For any $x, y \in \mathbb{Z}$, $P(\sigma_{x,y} > 2t) \leq P(\tau_{x,y} > t)$.

Proof. We observe that if $s_{x,y} \geq \varrho_{x,y}$ then $\sigma_{x,y} = \tau_{x,y}$ since $\hat{\xi}_s^{(x,t)} = \hat{\zeta}_s^{(x,t)}$ and $\hat{\xi}_s^{(y,t)} = \hat{\zeta}_s^{(y,t)}$ until time $s_{x,y}$. (See Figure 6(a) for an illustration.) In particular,

$$P(\sigma_{x,y} > 2t; s_{x,y} \geq \varrho_{x,y}) \leq P(\sigma_{x,y} > t; s_{x,y} \geq \varrho_{x,y}) = P(\tau_{x,y} > t; s_{x,y} \geq \varrho_{x,y}).$$

The conclusion then follows from Lemma 2.

Lemma 4. For any integer $y > 0$,

$$\sum_x P(\hat{\xi}_t^{(x+y,t)} \leq 0 < \hat{\xi}_t^{(x,t)} \mid \sigma_{x,x+y} > t) \leq \sum_x P(\hat{\zeta}_t^{(x+y,t)} \leq 0 < \hat{\zeta}_t^{(x,t)} \mid \tau_{x,x+y} > t).$$

Proof. We first observe that $P(\hat{\zeta}_t^{(x+y,t)} \leq 0 < \hat{\zeta}_t^{(x,t)}) = P(\hat{\zeta}_t^{(-x,t)} \leq 0 < \hat{\zeta}_t^{(-x-y,t)})$, so that

$$\sum_{x > 0} P(\hat{\zeta}_t^{(x+y,t)} \leq 0 < \hat{\zeta}_t^{(x,t)}) \leq \frac{1}{2} \sum_x P(\hat{\zeta}_t^{(x+y,t)} \leq 0 < \hat{\zeta}_t^{(x,t)}).$$

Since $P(\hat{\xi}_t^{(x,t)} > 0) = 0$ for any $x \leq 0$, it suffices to prove that

$$P(\hat{\xi}_t^{(x+y,t)} \leq 0 < \hat{\xi}_t^{(x,t)} \mid \sigma_{x,x+y} > t) \leq 2P(\hat{\zeta}_t^{(x+y,t)} \leq 0 < \hat{\zeta}_t^{(x,t)} \mid \tau_{x,x+y} > t)$$

for any $x > 0$ and $y > 0$. Now, if an inversion occurs before time t then $s_{x+y} \leq t$, and hence

$$\begin{aligned} &P(\hat{\xi}_t^{(x+y,t)} \leq 0 < \hat{\xi}_t^{(x,t)} \mid \sigma_{x,x+y} > t) \\ &\leq P(\hat{\xi}_t^{(x,t)} > 0; s_{x+y} \leq t \mid \sigma_{x,x+y} > t) \\ &= \int_0^t P(\hat{\xi}_t^{(x,t)} > 0; s_{x+y} = s \mid \sigma_{x,x+y} > t) P^{x+y}(s_{x+y} \in ds) \end{aligned}$$

$$\begin{aligned} &\leq \int_0^t \mathbb{P}(\hat{\xi}_t^{(x,t)} > 0; s_{x+y} = s; \hat{\xi}_t^{(x+y,t)} \geq 0 \mid \tau_{x,x+y} > t) \mathbb{P}^{x+y}(s_{x+y} \in ds) \\ &\quad + \int_0^t \mathbb{P}(\hat{\xi}_t^{(x,t)} > 0; s_{x+y} = s; \hat{\xi}_t^{(x+y,t)} \leq 0 \mid \tau_{x,x+y} > t) \mathbb{P}^{x+y}(s_{x+y} \in ds) \\ &= 2 \mathbb{P}(\hat{\xi}_t^{(x+y,t)} \leq 0 < \hat{\xi}_t^{(x,t)} \mid \tau_{x,x+y} > t). \end{aligned}$$

The last step follows from the fact that if $\hat{\xi}_s^{(x+y,t)}$ becomes frozen, $\hat{\xi}_s^{(x+y,t)}$ can be thought of as starting from a site to the left of 0 and running until time t . This together with the fact that the dispersal probability kernel is symmetric then implies the result. This completes the proof.

By Cox and Durrett (1995, Section 3), there exists a constant α that depends only on $y - x$ such that $\mathbb{P}(\tau_{x,y} > t) \sim \alpha/\sqrt{t}$ as $t \rightarrow \infty$. Lemma 3 then implies that there exists a $C < \infty$ such that, for sufficiently large t , $\mathbb{P}(\sigma_{x,y} > t) \leq C \mathbb{P}(\tau_{x,y} > t)$. This together with Lemma 4 leads to

$$\begin{aligned} \mathbb{E}(B_t) &= \sum_{y>0} \sum_x \mathbb{P}(\hat{\xi}_t^{(x+y,t)} \leq 0 < \hat{\xi}_t^{(x,t)}) \\ &= \sum_{y>0} \sum_x \mathbb{P}(\hat{\xi}_t^{(x+y,t)} \leq 0 < \hat{\xi}_t^{(x,t)} \mid \sigma_{x,x+y} > t) \mathbb{P}(\sigma_{x,x+y} > t) \\ &\leq C \sum_{y>0} \sum_x \mathbb{P}(\hat{\xi}_t^{(x+y,t)} \leq 0 < \hat{\xi}_t^{(x,t)} \mid \tau_{x,x+y} > t) \mathbb{P}(\tau_{x,x+y} > t) \\ &= C \mathbb{E}(B'_t). \end{aligned}$$

Proposition 1 then follows from Proposition 1 of Cox and Durrett (1995).

3. Proofs of Theorems 2 and 3

Proof of Theorem 2. By symmetry, we only need to prove the result for r_t . Let $\lfloor \cdot \rfloor$ denote the integer part. Since the condition $\xi_t(\lfloor x\sigma\sqrt{t} \rfloor) = 1$ implies that $r_t \geq x\sigma\sqrt{t}$, it follows, from the duality properties introduced in Section 2, that

$$\mathbb{P}(r_t \geq x\sigma\sqrt{t}) \geq \mathbb{P}(\xi_t(\lfloor x\sigma\sqrt{t} \rfloor) = 1) = \mathbb{P}(s_{\lfloor x\sigma\sqrt{t} \rfloor} \leq t).$$

Donsker’s Theorem implies that $\zeta_s^{(\lfloor x\sigma\sqrt{t} \rfloor, t)}/\sigma\sqrt{t}$ converges in distribution to W_s^x , a standard Brownian motion starting at x . We denote by τ_0 the first time W_s^x hits 0. Then

$$\lim_{t \rightarrow \infty} \mathbb{P}(r_t \geq x\sigma\sqrt{t}) \geq \lim_{t \rightarrow \infty} \mathbb{P}(s_{x\sigma\sqrt{t}} \leq t) = \mathbb{P}(\tau_0 \leq 1) = 2(1 - \Phi(x)).$$

For a proof of the last equality, see Karlin and Taylor (1975, Section 7.3). To get the opposite inequality, let $\varepsilon_0 > 0$ and refer to Proposition 1 to choose an $M > 0$ such that

$$\mathbb{P}(r_t - \ell_t \geq M) \leq \varepsilon_0.$$

Since $\xi_t(\lfloor x\sigma\sqrt{t} - M \rfloor) = 0$ and $r_t \geq x\sigma\sqrt{t}$ imply that $r_t - \ell_t \geq M$, we obtain

$$\mathbb{P}(r_t \geq x\sigma\sqrt{t}) \leq \mathbb{P}(\xi_t(\lfloor x\sigma\sqrt{t} - M \rfloor) = 1) + \varepsilon_0.$$

Since $\lim_{t \rightarrow \infty} \mathbb{P}(\xi_t(\lfloor x\sigma\sqrt{t} \rfloor) = 1) = \lim_{t \rightarrow \infty} \mathbb{P}(\xi_t(\lfloor x\sigma\sqrt{t} - M \rfloor) = 1)$, the result follows.

Proof of Theorem 3. For any $z \in \{0, 1, \dots, K - 1\}$, define the stopping time, T_z , as

$$T_z = \inf\{t \geq 0: \xi_t \equiv 1 \text{ on } D_z\},$$

where D_z is given by (2). First of all, we observe that if z is even then $D_z \subset \Delta$, which implies that $T_z = 0$. Furthermore, if $x \in D_z$ and $y \in D_{z+2}$ then, since q has compact support, $p_\varepsilon(x, y) = 0$ for $\varepsilon > 0$ sufficiently small. It follows that for small enough ε the stopping times T_z and T_{z+2} are independent. For the rest of the proof we assume that ε is sufficiently small so that the above assertion holds. To estimate T_z for odd z , we now consider the following edge processes:

$$\begin{aligned} \ell_t(z) &= \inf\{x \in D_z: \hat{\xi}_t^{(x,t)} \in D_z\}, \\ r_t(z+1) &= \sup\{x \in D_z: \hat{\xi}_t^{(x,t)} \in D_z\}. \end{aligned}$$

Since $\xi_t \not\equiv 1$ on D_z if and only if $\ell_t(z) < r_t(z+1) + \varepsilon$, we obtain

$$P(T_z > t) = P(\ell_t(z) - r_t(z+1) < \varepsilon).$$

Unfortunately, in view of the properties of the dual process, $\ell_t(z)$ and $r_t(z+1)$ are not a priori independent. To deal with this problem, we set

$$\begin{aligned} r_t(z) &= \sup\{x \in D_z: \hat{\xi}_t^{(x,t)} \in D_{z-1}\}, \\ \ell_t(z+1) &= \inf\{x \in D_z: \hat{\xi}_t^{(x,t)} \in D_{z+1}\}, \end{aligned}$$

and observe that $r_t(z)$ and $\ell_t(z+1)$ move independently as long as $r_t(z) < \ell_t(z+1)$. Moreover, Proposition 1 implies that, for any $z \in \{0, 1, \dots, K - 1\}$,

$$\lim_{\varepsilon \rightarrow 0} P(\ell_t(z) \geq x) = \lim_{\varepsilon \rightarrow 0} P(r_t(z) \geq x).$$

Hence, for odd z ,

$$\lim_{\varepsilon \rightarrow 0} P(T_z > t) = \lim_{\varepsilon \rightarrow 0} P(r_t(z) - \ell_t(z+1) \leq 0).$$

To investigate the process $r_t(z) - \ell_t(z+1)$, we first observe that $r_t(z) - zL$ and $(z+1)L - \ell_t(z+1)$ are identically distributed. Furthermore, by Theorem 2,

$$\lim_{\varepsilon \rightarrow 0} P(r_t(z) - zL \geq x) = \sqrt{\frac{2}{\pi}} \int_{x/\sigma\sqrt{t}}^\infty \exp\left(-\frac{y^2}{2}\right) dy.$$

In particular, $r_t(z) - zL$ and $(z+1)L - \ell_t(z+1)$ have the density

$$\Psi_t(x) = -\frac{\partial}{\partial x} \sqrt{\frac{2}{\pi}} \int_{x/\sigma\sqrt{t}}^\infty \exp\left(-\frac{y^2}{2}\right) dy = \sqrt{\frac{2}{\pi}} \frac{1}{\sigma\sqrt{t}} \exp\left(-\frac{1}{2t} \frac{x^2}{\sigma^2}\right).$$

It follows that, for odd z ,

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} P(T_z > t) &= \lim_{\varepsilon \rightarrow 0} P(r_t(z) - \ell_t(z+1) \leq 0) \\ &= \lim_{\varepsilon \rightarrow 0} P(r_t(z) - zL + (z+1)L - \ell_t(z+1) \leq L) \\ &= \int_0^L (\Psi_t * \Psi_t)(x) dx. \end{aligned}$$

By setting $2z = 2y - x$ and then $u = 2z$, we obtain

$$\begin{aligned}
 (\Psi_t * \Psi_t)(x) &= \int_0^x \Psi_t(y)\Psi_t(x - y) \, dy \\
 &= \frac{2}{\pi} \frac{1}{\sigma^2 t} \int_{-x/2}^{+x/2} \exp\left(-\frac{(2z+x)^2}{8\sigma^2 t}\right) \exp\left(-\frac{(2z-x)^2}{8\sigma^2 t}\right) \, dz \\
 &= \frac{2}{\pi} \frac{1}{\sigma^2 t} \exp\left(-\frac{x^2}{4\sigma^2 t}\right) \int_0^x \exp\left(-\frac{u^2}{4\sigma^2 t}\right) \, du \\
 &= 8f_t(x)(F_t(x) - F_t(0)),
 \end{aligned}$$

where f_t and F_t are the density and the distribution functions of the Gaussian law Y , respectively, with mean 0 and variance $2\sigma^2 t$. Since $f_t(x) = F_t'(x)$, we have

$$\begin{aligned}
 \int_0^L (\Psi_t * \Psi_t)(x) \, dx &= 8 \int_0^L f_t(x)(F_t(x) - F_t(0)) \, dx \\
 &= 8 \int_0^L (F_t(x) - F_t(0))'(F_t(x) - F_t(0)) \, dx \\
 &= 4(F_t(L) - F_t(0))^2 \\
 &= (F_t(L) - F_t(-L))^2.
 \end{aligned}$$

Finally, by using the fact that the stopping times T_z are independent for sufficiently small $\varepsilon > 0$, we can conclude that

$$\begin{aligned}
 \lim_{\varepsilon \rightarrow 0} P(T_{\text{inv}} \leq t) &= \lim_{\varepsilon \rightarrow 0} P(\max_z T_z \leq t) \\
 &= \lim_{\varepsilon \rightarrow 0} \prod_{z \text{ odd}} P(T_z \leq t) \\
 &= \left(1 - \int_0^L (\Psi_t * \Psi_t)(x) \, dx\right)^{1/2L} \\
 &= (1 - (F_t(L) - F_t(-L))^2)^{1/2L},
 \end{aligned}$$

which completes the proof of Theorem 3.

4. Proof of Theorem 4

Recall that q has compact support, and let $R > 0$ such that if $|x - y| > R$ then $p(x, y) = 0$. We start by introducing the stopping times

$$\begin{aligned}
 \tau_s^1 &= \inf\{t \geq s : \ell_t = r_t + 1\}, \\
 \tau_s^2 &= \inf\{t \geq s : \ell_t \leq R\}, \\
 \text{and } \tau_s^3 &= \inf\{t \geq s : r_t = 0\},
 \end{aligned}$$

and, for $x \geq 0$, the interface process

$$\iota_t(x) = \xi_t(x + \ell_t - 1).$$

We observe that, since $r_t = 0$ if and only if $\xi_t(x) = \mathbf{1}_{\{y \leq 0\}}(x)$, it suffices to prove that $P(\tau_s^3 < \infty) = 1$. To do this, we will prove that each of the stopping times introduced above is finite with probability 1. Let $\varepsilon_0 > 0$. Then Proposition 1 provides a constant $M > 0$ such that

$$P(r_s - \ell_s \geq M) \leq \varepsilon_0 \quad \text{for all } s \geq 0.$$

Furthermore, if $r_s - \ell_s < M$, then the number of possible configurations for the process ι_s is at most $2^M < \infty$. In particular, there exists an $\varepsilon_1 > 0$ such that the probability that from time s the process $r_t - \ell_t$ decreases to -1 is greater than or equal to ε_1 . Since $r_s - \ell_s < M$ infinitely often, $P(\tau_s^1 < \infty) = 1$. Let $t_1 = \tau_s^1$ and denote by ζ_t the voter model starting at time t_1 from $\zeta_{t_1}(x) = \xi_{t_1}(x) = \mathbf{1}_{\{y \leq r_{t_1}\}}(x)$. Since $p(x, y) = 0$ if $|x - y| > R$, it is easy to check that ξ_t and ζ_t can be defined on the same space in such a way that $\xi_t = \zeta_t$ until the first time $t_2 = \tau_{t_1}^2$ at which the process ℓ_t visits $[1, R]$. Moreover, Theorem 5 of Cox and Durrett (1995) implies that the leftmost-zero, $\ell'_t = \inf\{x : \zeta_t(x) = 0\}$, is recurrent, so

$$P(\tau_{t_1}^2 < \infty) = 1 \quad \text{and} \quad P(\ell_t \text{ visits } [1, R] \text{ i.o.}) = 1.$$

To conclude, we use the argument described above: if $r_{t_2} - \ell_{t_2} < M$ then $r_{t_2} < R + M$ and the number of possible configurations ξ_{t_2} is at most 2^{R+M} . In particular, there exists an $\varepsilon_2 > 0$ such that the probability that from time t_2 the process r_t decreases to 0 is greater than or equal to ε_2 . In conclusion, each time ℓ_t visits the set $[1, R]$, the process r_t returns to 0 with probability at least $\varepsilon_2(1 - \varepsilon_0) > 0$. The Borel–Cantelli lemma then implies the result.

5. Proofs of Theorems 5 and 6

Proof of Theorem 5. The key of the proof is duality. Recall that if $\beta < 1$, the dual process performs coalescing branching random walks in which each particle gives birth to a new particle at rate $1 - \beta$. We denote by $X_s^{(x,t)} = \max \hat{\xi}_s^{(x,t)}$ the right edge of the dual process. If we start the process from the initial configuration $\xi_0(x) = \mathbf{1}_{\{y \leq 0\}}(x)$, the duality property

$$\xi_t(x) = 0 \iff \xi_0(z) = 0 \quad \text{for some } z \in \hat{\xi}_t^{(x,t)}$$

implies that $\xi_t(x) = 1$ if and only if $X_t^{(x,t)} > 0$. Moreover, in the simplest nearest neighbor case, the right edge process $X_s^{(x,t)}$ is easy to describe:

1. if $X_s^{(x,t)}$ meets a δ , then it is forced to move along the corresponding δ -arrow;
2. if a particle at $X_s^{(x,t)}$ branches, then $X_s^{(x,t)}$ moves only if it takes it further from 0.

In conclusion, if $x > 0$ then $X_s^{(x,t)}$ is a continuous-time random walk which makes transitions

$$X_s^{(x,t)} \rightarrow \begin{cases} X_s^{(x,t)} + 1 & \text{at rate } 1, \\ X_s^{(x,t)} - 1 & \text{at rate } \beta, \end{cases}$$

and with 0 as an absorbing state. Let u_x be the probability that $X_s^{(x,t)} = 0$ at some time $s \geq 0$. By decomposing according to whether $X_s^{(x,t)}$ first jumps on the left or on the right gives

$$u_x = \frac{\beta}{\beta + 1} u_{x-1} + \frac{1}{\beta + 1} u_{x+1} \quad \text{for all } x \geq 1.$$

This implies that $u_{x+1} - u_x$ is a geometrical sequence with parameter β so that

$$u_x - u_0 = \sum_{k=0}^{x-1} \beta^k (u_1 - u_0) = \frac{1 - \beta^x}{1 - \beta} (u_1 - u_0).$$

Now, since $X_s^{(x,t)}$ drifts to the right, $\lim_{x \rightarrow \infty} u_x = 0$. Moreover, $u_0 = 1$ so it follows, from the previous equation, that $u_1 = \beta$. In conclusion, $u_x = \beta^x$ and

$$\mathbb{P}(r_t \geq x) = \mathbb{P}(X_t^{(x,t)} = 0) \leq \mathbb{P}(X_s^{(x,t)} = 0 \text{ at some time } s \geq 0) = \beta^x.$$

This completes the proof.

Proof of Theorem 6. We denote by τ the last time the rightmost-one process r_t returns to 0. Since $\beta > 1$, there exist $C < \infty$ and $\gamma > 0$ such that

$$\mathbb{P}(\tau > \varepsilon t) \leq C \exp(-\gamma \varepsilon t).$$

In particular, it suffices to prove the result for the continuous-time random walk Y_t starting at $Y_0 = 0$ which makes transitions

$$Y_t \rightarrow \begin{cases} Y_t + 1 & \text{at rate } \beta, \\ Y_t - 1 & \text{at rate } 1. \end{cases}$$

Let Z_n be the discrete-time version of Y_t , namely the Bernoulli random walk which jumps on the right with probability $\beta/(1 + \beta)$ and on the left with probability $1/(1 + \beta)$. Then, Proposition 3 of Spitzer (1976, Section 1.5) implies that

$$\mathbb{P}\left(\left|\frac{Z_n}{n} - \frac{\beta - 1}{\beta + 1}\right| \geq \varepsilon\right) \leq 2e^{-\gamma \varepsilon^2 n}$$

for any $n \geq 1$ and $\varepsilon > 0$. To deduce the analogous result for the continuous-time version Y_t , we observe that large-deviation results for the Poisson distribution imply that the probability of more than $(\beta + 1 + \varepsilon)t$ or less than $(\beta + 1 - \varepsilon)t$ jumps by time t is smaller than $Ce^{-\gamma t}$ for some $C, \gamma \in (0, \infty)$. It follows that

$$\mathbb{P}(|Y_t - (\beta - 1)t| \geq \varepsilon t) \leq Ce^{-\gamma \varepsilon^2 t}$$

for appropriate $C < \infty$ and $\gamma > 0$.

References

- [1] BRAMSON, M. AND GRIFFEATH, D. (1980). On the Williams–Bjerknes tumor growth model. II. *Math. Proc. Camb. Philos. Soc.* **88**, 339–357.
- [2] BRAMSON, M. AND GRIFFEATH, D. (1981). On the Williams–Bjerknes tumor growth model. I. *Ann. Prob.* **9**, 173–185.
- [3] COX, J. T. AND DURRETT, R. (1995). Hybrid zones and voter model interfaces. *Bernoulli* **1**, 343–370.
- [4] DURRETT, R. (1988). *Lecture Notes on Particle Systems and Percolation*. Wadsworth, Pacific Grove, CA.
- [5] ELLSTRAND, N. C., PRENTICE, H. C. AND HANCOCK, J. F. (1999). Gene flow and introgression from domesticated plants into their wild relatives. *Ann. Rev. Ecol. Syst.* **30**, 539–563.
- [6] HARRIS, T. E. (1972). Nearest neighbor Markov interaction processes on multidimensional lattices. *Adv. Math.* **9**, 66–89.

- [7] HAYGOOD, R., IVES, A. R. AND ANDOW, D. A. (2003). Consequences of recurrent gene flow from crops to wild relatives. *Proc. R. Soc. London B* **270**, 1879–1886.
- [8] HOLLEY, R. A. AND LIGGETT T. M. (1975). Ergodic theorems for weakly interacting systems and the voter model. *Ann. Prob.* **3**, 643–663.
- [9] HUXEL, G. R. (1999). Rapid displacement of native species by invasive species: effects of hybridization. *Biol. Conserv.* **89**, 143–152.
- [10] KARLIN, S. AND TAYLOR, H. M. (1975). *A First Course in Stochastic Processes*, 2nd edn. Academic Press, New York.
- [11] SPITZER, F. (1976). *Principles of Random Walks*, 2nd edn. Springer, New York.
- [12] WATRUD, L. S. *et al.* (2004). Evidence of landscape-level, pollen-mediated gene flow from genetically modified creeping bentgrass with *CP4 EPSPS* as a marker. *Proc. Nat. Acad. Sci. USA* **101**, 14533–14538.
- [13] WOLF, D. E., TAKEBAYASHI, N. AND RIESEBERG, L. H. (2001). Predicting the risk of extinction through hybridization. *Conserv. Biol.* **15**, 1039–1053.