Evolutionarily stable movement strategies in reaction-diffusion models with edge behavior

Gabriel Maciel[*]
Chris Cosner[†], Robert Stephen Cantrell[‡], Frithjof Lutscher[§]

February 18, 2019

Keywords: Evolution of dispersal, reaction-diffusion system, stability analysis, interface behavior

Abstract

Many types of organisms disperse through heterogeneous environments as part of their life histories. For various models of dispersal, including reaction-advection-diffusion models in continuously varying environments, it has been shown by pairwise invasibility analysis that dispersal strategies which generate an ideal free distribution are evolutionarily steady strategies (ESS, also known as evolutionarily stable strategies) and are neighborhood invader strategies (NIS). That is, populations using such strategies can both invade and resist invasion by populations using strategies that do not produce an ideal free distribution. (The ideal free distribution arises from the assumption that organisms inhabiting heterogeneous environments should move to maximize their fitness, which allows a mathematical characterization in terms of fitness equalization.) Classical reaction diffusion models assume that landscapes vary continuously. Landscape ecologists consider landscapes as mosaics of patches where individuals can make movement decisions at sharp interfaces between patches of different quality. We use a recent formulation of reaction-diffusion systems in patchy landscapes to study dispersal strategies by using methods inspired

---

[*] Department of Mathematics and Statistics, University of Ottawa, Ottawa, Canada. Current address: Universidade do Estado de Santa Catarina (UDESC), Brazil.

[†] Department of Mathematics, University of Miami, Miami, USA

[‡] Department of Mathematics, University of Miami, Miami, USA

[§] Department of Mathematics and Statistics, and Department of Biology, University of Ottawa, Ottawa, Canada

by evolutionary game theory and adaptive dynamics. Specifically, we use a version of pairwise invasibility analysis to show that in patchy environments, the behavioral strategy for movement at boundaries between different patch types that generates an ideal free distribution is both globally evolutionarily steady (ESS) and is a global neighborhood invader strategy (NIS).

1 Introduction

Organisms move in order to acquire resources and avoid dangers. In a given landscape with spatially varying habitat quality, what would be an 'optimal' strategy of movement? This question is, of course, much too general to have a meaningful answer, in particular since the meaning of 'optimal' is not easily defined. One aspect of this question has, however, generated a fascinating body of literature on the 'evolution of dispersal' that has inspired and challenged mathematicians and evolutionary biologists alike (McPeek and Holt, 1992; Lou, 2008; Cosner, 2014; Cantrell et al., 2017; Clobert et al., 2001, 2012). Here, we study the evolution of dispersal in the relatively novel framework of reaction-diffusion equations in 'patchy' landscapes (Maciel and Lutscher, 2013).

Reaction-diffusion equations have been instrumental in studying questions in theoretical spatial ecology (Cantrell and Cosner, 2003). In the simplest case, they describe the temporal dynamics of the density, $u(x,t)$, of a population from random movement and growth by

$$u_t = du_{xx} + f(x,u), \tag{1}$$

where d denotes the diffusion rate and f the spatially dependent population growth function. Hastings (1983) first studied the evolution of movement strategies ('dispersal') in this framework. He showed that a 'mutant' population can invade if and only if it has a smaller diffusion rate than the 'resident'. This conclusion that evolution in a spatially varying but temporally constant habitat favors slow dispersal rates has since been confirmed and strengthened in many respects (Dockery et al., 1998; Hutson et al., 2003; Altenberg, 2012). Intermediate or higher dispersal rates can evolve under certain conditions when movement is not just simple, random diffusion. In this case, diffusion could be space dependent and/or an additional advective or taxis term can describe directed movement. For example, when habitat variability also induces conditional dispersal, faster dispersal can be selected (McPeek and Holt, 1992; Cantrell et al., 2010; Averill et. al, 2012; Cosner, 2014). Faster dispersal can also be selected in advective environments (Lam et al., 2015; Lou and Lutscher, 2014). Directed movement towards more favorable regions combined with diffusion has also been shown to confer competitive superiority in some situations (Cantrell et al., 2006).

An approach to the evolution of dispersal that has been used extensively in recent years is a version of pairwise invasibility analysis, where one considers the dynamics of two (or more) populations that differ only in their movement strategy and asks whether one population can invade (i.e. grow from low density) and dynamically exclude the other (Dockery et al., 1998; Cantrell et al., 2010; Averill et. al, 2012; Cantrell et al., 2012a,b; Cosner, 2014; Cantrell et al., 2017). This approach is inspired by ideas from adaptive dynamics as described, for example, in Geritz et al. (1998), but does not use the full adaptive dynamics formulation, just pairwise invasibility analysis. The evolution of movement strategies can then be studied by considering a succession of such competitive interactions between an initially rare 'invader' (or 'mutant') population and a 'resident' population (Dockery et al., 1998; Cosner, 2014). The underlying assumption of this approach is that ecological processes (e.g. competition) occur on a much faster time scale than mutations. In the successive competition processes, more advantageous traits substitute others, which drives the dynamics of the evolutionary process. Some types of traits (interpreted as strategies) have special properties that make them especially relevant to understanding evolution, and are known as 'evolutionarily singular' (Geritz et al., 1998). Some terminology describing such types of traits has been developed in the contexts of adaptive dynamics and evolutionary game theory. Two such types are the evolutionarily steady strategies (ESS), which have the property that populations using them can resist invasion by populations using other strategies that are close to them, and neighborhood invader strategies (NIS) (Apaloo, 1997), which allow populations using them to invade populations using other nearby strategies. Our invasibility analysis will show that certain dispersal traits represent evolutionarily steady strategies (ESS), in the strong sense that populations using them can resist invasion by any population using another strategy, and are global neighborhood invader strategies (NIS), in the sense that a population using them can invade any population using another strategy. Another type of strategies of special interest, are convergence stable strategies, which have the property that populations using strategies closer to them can invade those using strategies that are further away. Our results do not show that the strategies are convergence stable, nor do we show directly that they cannot be evolutionary branch points. Thus, we do not show that our evolutionarily steady and neighborhood invader strategies are continuously stable in the strict sense. These are interesting topics for further study. However, since populations using the ESS/NIS strategies that we find can invade and exclude populations that do not use them, and resist invasion once established, they cannot be part of any polymorphism that includes other strategies.

While it can often be shown that evolutionarily singular strategies exist, they typically cannot be calculated analytically. There is, however, growing evidence that dispersal strategies that generate an ideal free distribution (IFD) are evolutionarily steady and are neighborhood invaders in various classes of dispersal models; see Cantrell et al. (2010); Averill et. al (2012); Cantrell et al. (2012a,b); Cosner (2014); Cantrell et al. (2017). Fretwell and Lucas (1969) introduced the IFD as a concept related to the spatial distribution of organisms. It is characterized by the feature that at equilibrium, all individuals will have equal local fitness in the sense of reproductive success in a given location, as described by the local per-capita population growth rate; see e.g., Holt and Barfield (2001); Cantrell et al. (2008); Cosner (2014). The underlying mechanism behind the concept is that if individuals are ideal in that they can assess the fitness they would have at any given location, allowing for factors such as habitat quality and crowding, and are free to move where they choose, they will move to optimize their fitness. It follows that in a population of ideal free organisms, the local fitness of individuals will be equal in all occupied habitats (since otherwise some individuals could and would move to increase their local fitness), and there should be no net movement (since once local fitness, i.e., per capita population is equalized, an individual moving to a new location would increase the crowding there, and thus would have lower fitness than it had before moving, unless two individuals just change places). Thus, an IFD is characterized as a distribution such that no individual can gain an advantage by changing its location while all others stay put. Several empirical studies have shown that approximations of IFDs naturally occur in ecological populations (Parker and Sutherland, 1986; Doncaster et al., 1997; Diffendorfer, 1998; Morris et al., 2004).

With regards to model (1), these considerations imply that when a population is at steady state and satisfies an IFD, then the local per-capita growth rate (local fitness), $f(x, u)/u$, must be identically zero across the landscape (Cantrell et al., 2008; Cosner, 2014).

Diffusion and advection have been widely used to model the dispersal of organisms, even though they have various limitations as models; see e.g., Cantrell and Cosner (2003); Cosner (2014). One issue is that if a diffusing population is introduced anywhere at any density, it will immediately have a positive density everywhere. In the present context, we envision habitats that are in principle accessible to all individuals, that is, where habitat type may affect dispersal rates but there are no complete barriers to dispersal, so this issue is not problematic. However, there are some serious problems in using classical reaction-advection-diffusion models to describe populations in heterogeneous environments. Reaction-advection-diffusion equations with spatially varying coefficients are not only difficult to study mathematically, they are also difficult to parametrize empirically. Landscape ecologists typically view heterogeneous landscapes as collections of 'patches', i.e., regions that are relatively homogeneous within but substantially different from the adjacent region. This view has the empirical advantage that it requires only one set of parameters per patch, and it allows one to include habitat preference of organisms by studying movement behavior at (or near) an edge or interface between two patch types. There is ample empirical evidence that individuals of many taxa adjust their movement behavior to habitat conditions inside patches and preferentially chose one over the other habitat at interfaces (Schultz and Crone, 2001; Schtickzelle and Baguette, 2003; Crone and Schultz, 2008; Reeve et al., 2008).

Reaction-diffusion equations in patchy landscapes were first studied by Shigesada et al. (1986), but without consideration of movement behavior at edges or interfaces between patches. Appropriate matching conditions at these interfaces were introduced by Maciel and Lutscher (2013), based on work by Ovaskainen and Cornell (2003). In the simplest case, one considers only two patches, say $\{x > x_n\}$ for patch 1 and $\{x < x_n\}$ for patch 2 with an interface at $x = x_n$. Population densities in patches i are denoted by $u_i = u_i(x,t)$. Then one has an equation of the form (1) on each patch with diffusion rate d and growth function $f(x,u)$ replaced by d_i and $f_i = f_i(u_i)$, respectively, on patch i, i.e.,

$$u_{i,t} = d_i u_{i,xx} + f_i(u_i), \quad x \in \text{patch } i.$$

At the interface, the matching conditions are

$$u_1(x_n^+,t) = k\ u_2(x_n^-,t) \tag{2}$$

$$d_1 u_{1x}(x_n^+,t) = d_2 u_{2x}(x_n^-,t), \tag{3}$$

where k is a composite parameter that encapsulates individual movement behavior as described by d_1 and d_2 and patch preference at the interface; see below. Superscripts $\pm$ denote the one-sided limits from the right and left, respectively. Equation (3) simply states the flux conservation at interfaces. We write $\alpha \in (0,1)$ for the probability that an individual at the interface chooses to move to patch 1 and $1-\alpha$ for the probability that it moves to patch 2. Then we have the explicit expression for the composite parameter k (Ovaskainen and Cornell, 2003; Maciel and Lutscher, 2013):

$$k = \frac{\alpha}{1-\alpha}\frac{d_2}{d_1}. \tag{4}$$

Thus, densities are discontinuous at interfaces in the presence of patch preference (i.e. $\alpha \neq 0.5$) and/or when diffusion rates in the two patches are different. This discontinuity and its dependence on movement behavior turns out to be crucial to basic quantities such as population persistence conditions and spread rates in periodic environment in the absence (Maciel and Lutscher, 2013) or presence (Maciel and Lutscher, 2015) of Allee effects. More recently, Maciel and Lutscher (2018) showed how different movement strategies for competing species in patchy landscapes can lead to different outcomes of the competition.

Here, we study how the outcome of competition feeds back to the evolution of movement strategies in patchy landscapes. The patch preference parameter α and diffusion rates d_1, d_2 represent the behavioural traits that we will study as strategies. Our analysis will be aimed at determining what choices of α, d_1, d_2 represent evolutionarily steady strategies. (Those will turn out to be neighbourhood invader strategies as well.) The condition for an ESS comes out in terms of the composite parameter k. To achieve the ESS values it turns out that organisms would need to assess features of the habitat types arising from f_1, f_2 such as local carrying capacities. In practice that would require certain sensory and cognitive abilities. We will not attempt to address the question of how and when organisms can obtain the information needed to achieve an ESS, although that is an interesting issue in sensory ecology; we will merely identify what information is needed in certain environments.

In this work, we consider a landscape of two patch types. We model movement and growth through reaction-diffusion equations in a patchy landscape and include interface conditions as above. We seek evolutionarily stable strategies (represented by parameters α, d_1, d_2) using two distinct approaches. First, we consider an infinite periodic landscape in which patches of the two types alternate. Our analysis of this case is based on a recently derived homogenization technique for these type of models. In the limit of small-scale heterogeneity, this technique allows us to transform a spatial problem in heterogeneous space into a problem in homogeneous space where diffusion and growth are given by certain spatial averages. The standard framework of adaptive dynamics can then be applied to the homogenized model. Secondly, we consider a landscape of only two patches. We consider this scenario as one period of the infinite landscape. We study this case by deriving the ideal-free distribution (IFD) and determining its stability analytically. We find that both approaches lead to the same result that ideal-free strategies are indeed evolutionarily stable.

2 The model

We model population dynamics of a 'resident' and an 'invader' ('mutant') in a one-dimensional patchy landscape consisting of two types of patches, denoted by Ω_i, with $i = 1, 2$. On patch (type) i, we denote by u_i the density of the resident population and by u_i' the density of the invader/mutant. In patch type i, these populations have diffusion rates d_i, d_i', respectively. For population dynamics, we assume logistic growth and Lotka-Volterra competition, and – as indicated above – we assume that resident and invader are identical with respect to these dynamics. Hence, our system of equations for $x \in \Omega_i$ is

$$u_{it} = d_i u_{ixx} + r_i u_i \left(1 - (u_i + u_i')/K_i\right), \tag{5}$$

$$u_{it}' = d_i' u_{ixx}' + r_i u_i' \left(1 - (u_i' + u_i)/K_i\right). \tag{6}$$

Parameters r_i, K_i are assumed positive.

For interface behavior, we denote by α (α') the probability with which a resident (invader) moves into the patch of type 1 at an interface. Thus, at an interface point with patch type 1 on the right and type 2 on the left, we choose the conditions in (2)-(3) with k given by (4). At points where patch type 1 is to the left and type 2 to the right, the same conditions apply with the signs denoting the directions of the one-sided limits in (2)-(3) exchanged. The matching conditions for the invader have α replaced by α' and d_i by d_i'.

For our first approach, we divide the real line into periodically alternating patches of type i with length l_i and period $l = l_1 + l_2$. Accordingly, type-1 patches may be taken as $\Omega_1 = \{x \in (nl, nl + l_1) | n \in \mathbb{Z}\}$ and type-2 patches as $\Omega_2 = \{x \in (nl + l_1, (n+1)l) | n \in \mathbb{Z}\}$. The interface points $x_n = nl$ ($n \in \mathbb{Z}$), have a type-1 patch to the right whereas the interface points $x_n = nl + l_1$ ($n \in \mathbb{Z}$) have a type-2 patch on the right.

For our second approach, we take $\Omega_1 = (0, l_1)$ and $\Omega_2 = (-l_2, 0)$. The interface $x = 0$ has the type-1 patch to the right. At the boundary points l_1 and $-l_2$, we impose no-flux conditions $u_{1x}(l_1, t) = u_{2x}(-l_2, t) = 0$. This set-up is equivalent to a periodic setting with patches of length $2l_1$ and $2l_2$ (see Appendix).

Our main goal is to determine dispersal strategies, in terms of diffusion rates and patch preference, that are evolutionarily stable. Evolutionarily stable strategies are characterized by the property that a population using such a strategy cannot be invaded by any small population using a different strategy. We present some analytical preliminaries for the second approach in the next section. In Section 4, we begin with the infinite landscape and use the homogenization approach. In

Section 5, we consider the bounded landscape and present exact stability and invasion conditions. We collect the proofs for the latter case in Section 6. We close with a discussion of biological implications.

3 Analytical preliminaries

Before we turn to the exploration of trait evolution, we need to provide some analytical results for our system of equations. This material is somewhat technical, and it is not essential to read it closely to follow most of the rest of the paper. Due to the non-standard interface conditions, several standard tools and properties of solutions of parabolic or elliptic equations are not immediately obvious for our system, e.g., the maximum and comparison principle, the existence of a dominant eigenvalue as well as existence, uniqueness and positivity of solutions in appropriate function spaces. The background theory is needed for several reasons. We need to know that the model is well posed in the sense that solutions actually exist, and stay nonnegative so they are biologically meaningful. That requires a maximum principle and some estimates for the time-dependent problem. To obtain the existence of a dominant eigenvalue, we will need to use the Krein-Rutman theorem, which is the infinite dimensional analogue of the Perron-Frobenius theorem. We need the existence of a dominant eigenvalue for the time-independent linearized model to characterize invasion fitness. To apply the Krein-Rutman theorem, we need to know that the resolvent of the linear operator describing the diffusion and patch preference in the time-independent linearized model has a compact, positive resolvent. That, in turn, requires a maximum principle and some estimates for the time-independent problem. To give a characterization of the dynamics of the time-dependent model, we need to use some ideas that underlie monotone dynamical systems theory, which require a comparison principle. All of these are standard for classical reaction-advection-diffusion models, but none of the standard results apply to models with an interface and patch preference. We provide this necessary theory here. We begin with the following maximum principle.

Proposition 3.1 *Suppose that smooth functions u_i satisfy the inequalities*

$$u_{1t} \geq d_1 u_{1xx} + a_1(x,t)u_1, \qquad 0 \leq x \leq l_1, \tag{7}$$

$$u_{2t} \geq d_2 u_{2xx} + a_2(x,t)u_2, \qquad -l_2 \leq x \leq 0, \tag{8}$$

with smooth functions a_i for $0 < t \leq T^*$ and with boundary and interface conditions

$$u_{2x}(-l_2,t) = 0 = u_{1x}(l_1,t), \quad u_1(0^+,t) = ku_2(0^-,t), \quad d_1u_{1x}(0^+,t) = d_2u_{2x}(0^-,t), \tag{9}$$

where $k > 0$. Suppose further that $u_i(x,0) \geq 0$. Then $u_i(x,t) \geq 0$ for $0 < t \leq T^*$. Furthermore, if $u_i(x,0) > 0$ for at least one i, then $u_i(x,t) > 0$ for $i = 1,2$ and $0 < t \leq T^*$.

We provide the proof in Section 6. From this proposition, we immediately obtain the following comparison principle.

Proposition 3.2 *Suppose that smooth functions $u_i = u_i(x,t)$ and $v_i = v_i(x,t)$ satisfy the inequalities*

$$u_{1t} \geq d_1u_{1xx} + f_1(u_1), \qquad 0 \leq x \leq l_1, \tag{10}$$

$$u_{2t} \geq d_2u_{2xx} + f_2(u_2), \qquad -l_2 \leq x \leq 0, \tag{11}$$

and

$$v_{1t} \leq d_1v_{1xx} + f_1(v_1), \qquad 0 \leq x \leq l_1, \tag{12}$$

$$v_{2t} \leq d_2v_{2xx} + f_2(v_2), \qquad -l_2 \leq x \leq 0, \tag{13}$$

with smooth functions f_i for $0 < t \leq T$ with boundary and interface conditions (9) for both u_i and v_i. Suppose further that $u_i(x,0) \geq v_i(x,0)$. Then $u_i(x,t) \geq v_i(x,t)$ for $0 < t \leq T$. If, in addition $u_i(x,0) > v_i(x,0)$ for at least one i, then $u_i(x,t) > v_i(x,t)$ for $0 < t \leq T$.

The proof of this proposition follows from Proposition 3.1 as in the classical case. We set $w_i = u_i - v_i$. Then

$$w_{it} \geq d_iw_{ixx} + f_i(u_i) - f_i(v_i) = d_iw_{ixx} + g_i(x)w_i,$$

where g_i are functions that depend on u_i, v_i and f_i'. By Proposition 3.1 we have $w_i \geq 0$ for $0 < t \leq T$ since $w_i(x,0) \geq 0$. The statement of strict inequality follows from the corresponding statement in Proposition 3.1. We note that the proof applies more generally when the growth functions in each patch are allowed to vary spatially, however, we are not considering this case. ■

We now move to the time-independent problem

$$
\begin{aligned}
-d_1 u_{1xx} + c_1 u_1 =& \bar{f}_1, \qquad 0 \le x \le l_1, && (14)\\
-d_2 u_{2xx} + c_2 u_2 =& \bar{f}_2, \qquad -l_2 \le x \le 0, && (15)\\
d_1 u_{1x}(0) = d_2 u_{2x}(0), \quad u_1(0) =& k u_2(0), \quad u_{1x}(l_1) = 0 = u_{2x}(-l_2), && (16)
\end{aligned}
$$

where $\bar{f}_i = \bar{f}_i(x)$ are given functions.

Proposition 3.3 *Suppose u_i solve (14)-(16) with $c_i > 0$ and $\bar{f}_i = \bar{f}_i(x) \ge 0$. Then $u_i > 0$.*

We give the proof of this proposition in Section 6.

Proposition 3.4 *Suppose $c_i > 0$. Given $\bar{f}_1 \in \mathcal{C}([0,l_1])$ and $\bar{f}_2 \in \mathcal{C}([-l_2,0])$, there is a unique solution $u_1 \in \mathcal{C}^2([0,l_1])$ and $u_2 \in \mathcal{C}^2([-l_2,0])$ of (14)-(16) with*

$$
\|u_1\|_{\mathcal{C}^2([0,l_1])} + \|u_2\|_{\mathcal{C}^2([-l_2,0])} \le C\left(\|\bar{f}_1\|_{\mathcal{C}([0,l_1])} + \|\bar{f}_2\|_{\mathcal{C}([-l_2,0])}\right). \qquad (17)
$$

The proof of this proposition can also be found in Section 6.

We are now ready to prove the existence of a principal eigenvalue that will be important in the analysis of trait evolution.

Proposition 3.5 *The operator defined by the left-hand sides of (14) and (15) and the boundary and interface conditions (16) has a principal eigenvalue with positive eigenfunction.*

The proof follows from the previous results. Assume at first that $c_i(x) > 0$. By Proposition 3.4, the solution operator of (14)-(16) defines a continuous mapping from $\mathcal{C}([0,l_1]) \times \mathcal{C}([-l_2,0])$ into $\mathcal{C}^2([0,l_1]) \times \mathcal{C}^2([-l_2,0])$, which, when combined with the compact embedding of $\mathcal{C}^2$ into $\mathcal{C}$ gives a compact mapping from $\mathcal{C}([0,l_1]) \times \mathcal{C}([-l_2,0])$ into itself. By Proposition 3.3, this mapping is strongly positive. Hence, by the Krein-Rutman theorem (Du, 2006), the operator has a positive principal eigenvalue with positive eigenfunction.

If c_i are not positive, we pick a large enough constant $q > 0$ and solve instead the problem

$$
\begin{aligned}
-d_1 u_{1xx} + (c_1 + q) u_1 =& (\lambda + q) u_1 = \tilde{\lambda} u_1, \qquad 0 \le x \le l_1, && (18)\\
-d_2 u_{2xx} + (c_2 + q) u_2 =& (\lambda + q) u_2 = \tilde{\lambda} u_2, \qquad -l_2 \le x \le 0, && (19)
\end{aligned}
$$

with conditions (16). When $c_i + q > 0$, the previous reasoning applies, and a principal eigenvalue $\tilde{\lambda} > 0$ exists. We find that $\lambda = \tilde{\lambda} - q$ need not be positive. ■

4 Homogenization analysis

Our first approach to the problem is based on homogenization. This technique is well known in multi-scale problems and frequently used in physical and engineering applications (Othmer, 1983; Bensoussan et al., 2010) but somehow less prominent in ecology (Powell and Zimmermann, 2004; Garlick et al., 2011). We give a brief heuristic explanation of the underlying ideas before we derive the details for our particular application.

Landscape attributes may vary on scales that are relatively small compared to the dispersal behavior of the species that we are interested in. For example, stands of trees and open grassland may alternate on a scale of tens or a few hundreds of meters while organisms' dispersal may occur on the scale of kilometers. In that case, one could expect that dispersal "smooths out" the different landscape characteristics so that only some appropriately averaged landscape quality determines "most" of the population dynamics. If dispersal behavior of individuals varies between the different small-scale landscape types, then one could also expect that some appropriately averaged dispersal behavior characterizes the "most important" aspects of movement on the large scale.

In our case, we expect to obtain a system of reaction-diffusion equations for two competing species (w, w') on the large scale, written as

$$w_t = \hat{D} w_{xx} + \hat{r} w(1 - c_1 w - c_2 w'), \quad w'_t = \hat{D}' w'_{xx} + \hat{r}' w'(1 - c'_1 w - c'_2 w'),$$

with "effective" diffusion coefficients $(\hat{D}, \hat{D}')$, low-density growth rates $(\hat{r}, \hat{r}')$ and intra- and inter-specific competition coefficients c_i, c'_i. These effective coefficients should be an appropriate average of the patch-wise coefficients, where "appropriate" includes movement behavior as expressed by d_i, d'_i and α, α'.

All of these ideas can be made precise by multi-scale analysis; see, e.g., Bensoussan et al. (2010). The idea is that one writes the density of the population as a function of two (or more) variables that represent the different scales. Then one expands the density in terms of a small parameter, corresponding to the ratio of the two scales, and obtains equations for each of the terms in the expansion. When solving these equations successively, the appropriate averages emerge from the equations. For example, in the case of reaction-diffusion equations with diffusion and reaction varying on a small scale only, one obtains the harmonic mean of the diffusion coefficients and the arithmetic mean of the reaction terms as the appropriate averages (Othmer, 1983). The appearance of the harmonic mean of the diffusion coefficient can be explained in terms of residence times (Powell and Zimmermann, 2004). The resulting equations on the larger scale are obviously much simpler to study because they have constant coefficients. They typically provide also a very good approximation of the two-scale model even when the scale difference is not very small (Dewhirst and Lutscher, 2009; Garlick et al., 2011).

The difficulty in applying these ideas to our model arises from the discontinuity of the densities at the interfaces. Previous theory typically assumed smooth densities. However, recent progress by Yurk and Cobbold (2018) extends this theory to include discontinuous interface conditions for a single population. The application to two competing species and stage-structured species was developed by Maciel and Lutscher (2018) and Alqawasmeh and Lutscher (2019), respectively. We now explain this set-up in more detail.

We assume that the habitat consists of two types of patches that are alternating periodically in one-dimensional space. We assume that the period is small and use it in the asymptotic expansion to obtain a spatially homogeneous averaged model on the larger scale. As described above, the two types of patches are arranged as

$$\Omega_1 = \{x \in [nl, nl + l_1] | n \in \mathbb{Z}\} \quad \text{and} \quad \Omega_2 = \{x \in [nl + l_1, (n+1)l] | n \in \mathbb{Z}\},$$

with l being the period. The dynamic equations are given by (5)-(6) and the interface conditions are as explained in Section 2.

To simplify notation, it is convenient to define a function $u(x,t)$ for $x \in \mathbb{R}$ as $u(x,t) = u_i(x,t)$ on Ω_i. Similarly, it is convenient to think of the diffusion coefficients as piecewise constant functions $d = d(x)$ with values $d(x) = d_i$ for $x \in \Omega_i$. In the same way, we can write the reaction terms as $g(x, u, u')$ according to the right-hand terms in (5)-(6). Analogous definitions can be made for $u'(x,t)$, $d'(x)$ and g'. Then formally, the densities satisfy the equations

$$u_t = [d(x)u(x,t)]_{xx} + g(x, u, u'), \qquad u'_t = [d'(x)u'(x,t)]_{xx} + g'(x, u, u'), \tag{20}$$

with the interface conditions to hold at the points of discontinuity of the coefficient functions.

Following the technique of homogenization, we choose $l = \epsilon \ll 1$ as the small parameter, and assume the existence of two distinct spatial scales, $\xi = x$ and $y = x/\epsilon$. Variable ξ is termed the "slow" and y the "fast" scale. Diffusion and growth are assumed to depend on the fast scale only, i.e., $d = d(y)$ and $g = g(y, u, u')$, while population density is a function of both scales, $u = u(\xi, y, t)$.

We deal with the equation for u first (assuming, for now, that $u' \equiv 0$). We write the population density as a series expansion in ϵ, $u(\xi, y, t) = \sum \epsilon^i u^{(i)}(\xi, y, t)$. Substituting this series into (20), we obtain equations for the different orders of ϵ, again satisfying corresponding matching conditions at interface points. Yurk and

Cobbold (2018) showed that the lowest-order term can be written as

$$u^{(0)}(\xi, y, t) = \frac{w(\xi, t)}{h(y)}, \tag{21}$$

with

$$h(y) = \begin{cases} 1 & \text{if } y \in \tilde{\Omega}_1 = \Omega_1/\epsilon, \\ k & \text{if } y \in \tilde{\Omega}_2 = \Omega_2/\epsilon. \end{cases} \tag{22}$$

Furthermore, the numerator, $w(\xi, t)$, which is a function of time and the slow scale only, solves the equation

$$w_t = \hat{l}^2 \langle d \rangle_H w_{\xi\xi} + \langle g \rangle_A, \tag{23}$$

where $\hat{l} = (l_1 + l_2)/(l_1 + l_2/k)$ is a weighted average of the patch lengths, and $\langle d \rangle_H$ and $\langle g \rangle_A$ are the harmonic mean of diffusion and arithmetic mean of growth, explicitly given by

$$\langle d \rangle_H = \left(\frac{l_1 + l_2/k}{\frac{l_1}{d_1} + \frac{l_2/k}{d_2/k^2}} \right), \quad \langle g \rangle_A = \left[\frac{l_1 g_1(w) + l_2 g_2(w/k)}{l_1 + l_2/k} \right]. \tag{24}$$

For simplicity, here we denoted $g_i(w) = g(x, w, 0)$ for $x \in \Omega_i$. (Recall that we are dealing with the case $u' = 0$ here.)

In other words, if the scale of the heterogeneity (the period) is small enough, the solution of the first equation in (20) is approximated by the leading term (21). Its time evolution is completely determined by $w(\xi, t)$, which satisfies the "homogenized equation" (23). The harmonic mean of the diffusion coefficients and the arithmetic mean of the reaction term represent the appropriate averages of the small-scale variation in movement and growth, respectively.

We now apply the homogenization procedure to the system of two equations (5)-(6) for the resident (u) and the invader (u'), as was done by Maciel and Lutscher (2018) and Alqawasmeh and Lutscher (2019). The resident leading term is given by (21), with $w(\xi, t)$ and $h(y)$ obtained from (23) and (22). The only modification arises in the average growth rate where the interaction term with the invader (u') must be considered. We thus have

$$\langle g \rangle_A = \left[\frac{l_1 g_1(w, w') + l_2 g_2(w/k, w'/k')}{l_1 + l_2/k} \right], \tag{25}$$

where $k' = d_2' \, \alpha'/[d_1'(1 - \alpha')]$ is the composite parameter that defines the density jump of the invader at an interface. We obtain similar equations for the invader substituting w', k', $g_{1,2}'$ and $d_{1,2}'$ for w, k, $g_{1,2}$ and $d_{1,2}$.

The complete homogenized equations with the averaged growth and competition coefficients written explicitly in terms of the original model parameters are

$$\begin{aligned} w_t = \hat{l}^2 \langle d \rangle_H w_{\xi\xi} + \left(\frac{r_1 l_1 + r_2 l_2/k}{l_1 + l_2/k} \right) w \Bigg[1 - \left(\frac{r_1 l_1/K_1 + r_2 l_2/(k^2 K_2)}{r_1 l_1 + r_2 l_2/k} \right) w & \\ - \left(\frac{r_1 l_1/K_1 + r_2 l_2/(k' k K_2)}{r_1 l_1 + r_2 l_2/k} \right) w' \Bigg], & \qquad (26) \\ w'_t = \hat{l'}^2 \langle d' \rangle_H w'_{\xi\xi} + \left(\frac{r_1 l_1 + r_2 l_2/k'}{l_1 + l_2/k'} \right) w' \Bigg[1 - \left(\frac{r_1 l_1/K_1 + r_2 l_2/(k'^2 K_2)}{r_1 l_1 + r_2 l_2/k'} \right) w' & \\ - \left(\frac{r_1 l_1/K_1 + r_2 l_2/(k' k K_2)}{r_1 l_1 + r_2 l_2/k'} \right) w \Bigg]. & \qquad (27) \end{aligned}$$

This system is a spatially explicit Lotka-Volterra-type competition system for w and w'. Each species diffuses according to its appropriately averaged diffusion coefficient as above. Each species has as its low-density growth rate the (weighted) arithmetic mean of the growth rates in the two patches (r_1 and r_2). The weights, l_1 and l_2/k or l_2/k', include parameters k and k', respectively, the parameters that summarize the interface behavior of individuals. If diffusion is the same in both patches ($d_1 = d_2$) and if there is no patch preference ($\alpha = \alpha' = 1/2$), then $k = k' = 1$, and we obtain the usual arithmetic means. The terms in round brackets inside the square brackets can be understood as inter- and intra-specific competition coefficients. They represent (weighted) arithmetic means of the inverse of the (scaled) carrying capacities with weights $r_1 l_1$ and $r_2 l_2/k$ (or $r_2 l_2/k'$). The scaling factor k or k' in front of K_2 indicates that the carrying capacity of a species in patch type 2 is scaled by how it uses space via interface conditions.

This homogenized competition system contains no spatial variation. In the case of spatially homogeneous Lotka-Volterra competition systems for two competitors, if one of the competitors excludes the other in the model without diffusion, then the same is true for the corresponding model with diffusion on the real line. This follows from Theorem 4.4 of Weinberger et al. (2002) after a suitable change of coordinates. Hence, we study the non-spatial dynamics of the preceding model. It turns out that the analysis is simplified by the rescaling

$$v = \left(\frac{r_1 l_1/K_1 + r_2 l_2/(k^2 K_2)}{r_1 l_1 + r_2 l_2/k} \right) w, \quad v' = \left(\frac{r_1 l_1/K_1 + r_2 l_2/(k'^2 K_2)}{r_1 l_1 + r_2 l_2/k'} \right) w', \qquad (28)$$

which leads to the system

$$\frac{dv}{dt} = r\ v\,(1 - v - av') \tag{29}$$

$$\frac{dv'}{dt} = r'\ v'\,(1 - v' - a'v)\,, \tag{30}$$

with coefficients

$$r = \left(\frac{r_1 l_1 + r_2 l_2/k}{l_1 + l_2/k}\right), \quad r' = \left(\frac{r_1 l_1 + r_2 l_2/k'}{l_1 + l_2/k'}\right), \tag{31}$$

$$a = \left(\frac{r_1 l_1/K_1 + r_2 l_2/(k'kK_2)}{r_1 l_1/K_1 + r_2 l_2/(k'^2 K_2)}\right)\left(\frac{r_1 l_1 + r_2 l_2/k'}{r_1 l_1 + r_2 l_2/k}\right), \tag{32}$$

$$a' = \left(\frac{r_1 l_1/K_1 + r_2 l_2/(k'kK_2)}{r_1 l_1/K_1 + r_2 l_2/(k^2 K_2)}\right)\left(\frac{r_1 l_1 + r_2 l_2/k}{r_1 l_1 + r_2 l_2/k'}\right). \tag{33}$$

It is straightforward to analyze the rescaled system according to the procedures of adaptive dynamics (Geritz et al., 1998). The resident-only steady state is $v = 1$, $v' = 0$. The initial invasion dynamics of a small invader population v' are governed by

$$\frac{dv'}{dt} = r'\,(1 - a')\,v', \quad \text{so that} \quad v' \propto e^{r'(1-a')t}. \tag{34}$$

We note that the competition between resident and invader depend only on the composite parameters k and k' and not on the preference and diffusion parameters individually. This fact simplifies our analysis considerably since the adaptive dynamics become a one-dimensional problem in the 'trait value' k. The invasion exponent of a potential invader with trait k' at small density in an established resident with trait k is given by

$$s(k, k') = r'\,(1 - a')\,.$$

The 'canonical equation of adaptive dynamics' (Dieckmann and Law, 1996) describes the change of the trait k over evolutionary time (T) via the selection gradient as

$$\frac{dk}{dT} = \frac{\partial s(k, k')}{\partial k'}|_{k'=k}.$$

The calculation of the selection gradient

$$\frac{\partial s}{\partial k'}|_{k'=k} = \frac{dr'}{dk'}(1 - a')|_{k'=k} - r'\frac{\partial a'}{\partial k'}|_{k'=k}$$

is simplified by the fact that $a'(k,k) = 1$. Furthermore, after a somewhat tedious calculation, we find that the sign of the selection gradient is given by

$$\text{sign}\frac{\partial s}{\partial k'}|_{k'=k} = -\text{sign}\frac{\partial a'}{\partial k'}|_{k'=k} = -\text{sign}\left(k - \frac{K_1}{K_2}\right).$$

Therefore, the evolutionarily singular strategy is given by

$$\overline{k} = \frac{K_1}{K_2}. \tag{35}$$

With this expression, we return to the nonlinear, non-spatial system (29)-(30). Species v can invade at $(0,1)$ if and only if $a < 1$. Vice versa, v' can invade at $(1,0)$ if and only if $a' < 1$. Stable coexistence is possible if and only if $a, a' < 1$. If species v uses the singular strategy (i.e., $k = \overline{k}$), one finds that $a \leq 1$ with equality only if $k' = \overline{k}$ as well. It is straightforward to see that $a' = 1$ if $k = \overline{k}$. In that case, coexistence requires $a = 1$, i.e., $k' = \overline{k}$, so that the two have to be identical. Hence, if v uses the singular strategy then it will invade all other strategies and go to fixation, i.e., dynamically exclude the other strategy. Hence, the adaptive dynamics converge to the singular strategy.

In summary, when the patches are small compared to the dispersal ability, averaging allows us to reduce the problem in the heterogeneous landscape to a homogeneous system, that, in turn, simplifies to an ODE problem for the question that we want to answer. The adaptive dynamics for the movement parameters d_i, α reduces to the dynamics of the single trait value k, the combined behavior at an interface. The result then predicts that k should evolve to $\overline{k}$. In the next section, we explain how this value of $\overline{k}$ represents an ideal-free distribution (IFD) in the full model.

5 Steady-state analysis

We now move to the steady-state analysis of the model in a heterogeneous landscape. We first identify the IFD for a single species in this case and then study its stability. Since the landscape is periodic, we study the problem on a single period only. Hence, we denote $\Omega_1 = [0, l_1]$, $\Omega_2 = [-l_2, 0]$. We have the equations

$$u_{1t} = d_1 u_{1xx} + r_1 u_1 \left(1 - u_1/K_1\right), \quad \text{in} \quad \Omega_1, \tag{36}$$

$$u_{2t} = d_2 u_{2xx} + r_2 u_2 \left(1 - u_2/K_2\right), \quad \text{in} \quad \Omega_2, \tag{37}$$

together with the boundary conditions

$$u_1(0^+,t) = \ ku_2(0^-,t), \quad d_1u_{1x}(0^+,t) = d_2u_{2x}(0^-,t), \tag{38}$$
$$u_1(l_1^-,t) = \ ku_2(-l_2^+,t), \quad u_{1x}(l_1^-,t) = 0 = u_{2x}(-l_2^+,t). \tag{39}$$

We provide a proof for the global existence of unique solutions to this system (in fact, to the system with the two competing populations) in the appendix.

We denote a candidate for an IFD by $\overline{u}_{1,2}$. As outlined in the introduction and proved by Cantrell et al. (2008); Cosner (2014), local individual fitness at the IFD must equal zero everywhere in the domain. Hence, in (36) and (37), we must have $r_i(1-\overline{u}_i/K_i) = 0$ or $\overline{u}_i = K_i$. This is the only candidate for an IFD within the set of functions that are smooth on each patch.

Consequently, at an IFD the population matches the carrying capacity in each habitat patch. The piecewise constant equilibrium $\overline{u}_{1,2}$ must also satisfy boundary conditions (38)-(39). Values of the composite parameter k for which this is possible may be termed the 'ideal-free dispersal strategies'. The only possible choice is

$$\overline{k} = \frac{K_1}{K_2}. \tag{40}$$

Hence, the ideal-free dispersal strategies are exactly the same as the ones that emerged from the adaptive dynamics approach in the previous section.

We note that because the idea-free dispersal strategy is piecewise constant, it does not depend on the diffusion coefficients d_i independently but only on their ratio through the composite parameter k, which accounts for the discontinuity at the interface.

In the following, we investigate whether $\overline{k}$ is an ESS for the non-homogenized equations as well. First, we show that a population using this strategy cannot coexist at steady state with any other population using a different strategy. Second, we will show that a population using the ideal-free strategy will be able to invade from low density any population using a different strategy.

The steady-state densities of a resident $\phi_{1,2}(x)$ and a mutant $\phi'_{1,2}(x)$ of equations (5)-(6) satisfy the equations

$$d_1\phi_{1xx} + r_1\phi_1\left(1-(\phi_1+\phi'_1)/K_1\right) = 0 \tag{41}$$
$$d_1\phi'_{1xx} + r_1\phi'_1\left(1-(\phi'_1+\phi_1)/K_1\right) = 0 \quad \text{in} \quad \Omega_1 \tag{42}$$

and

$$d_2\phi_{2xx} + r_2\phi_2\left(1 - (\phi_2 + \phi_2')/K_2\right) = 0 \tag{43}$$
$$d_2\phi'_{2xx} + r_2\phi_2'\left(1 - (\phi_2' + \phi_2)/K_2\right) = 0 \quad \text{in} \quad \Omega_2, \tag{44}$$

together with the periodic boundary conditions (38)-(39), with k and $d_{1,2}$ substituted by k' and $d'_{1,2}$ in the mutant's equations.

Theorem 5.1 *Let $\phi_{1,2}(x)$ and $\phi'_{1,2}(x)$ be positive solutions of (41)-(44) with boundary conditions as described and parameters k, k', respectively. If $k = \overline{k} = K_1/K_2$, ϕ_i and ϕ_i' are constant on Ω_i and $k' = \overline{k}$.*

Proof. Dividing (41) and (43) by ϕ_1/K_1 and ϕ_2/K_2, respectively, and integrating over space, we get

$$\begin{aligned}
&\int_0^{l_1} \left[d_1 K_1 \frac{\phi_{1xx}}{\phi_1} + r_1 K_1\left(1 - \frac{\phi_1 + \phi_1'}{K_1}\right)\right]dx \\
&\qquad + \int_{-l_2}^0 \left[d_2 K_2 \frac{\phi_{2xx}}{\phi_2} + r_2 K_2\left(1 - \frac{\phi_2 + \phi_2'}{K_2}\right)\right]dx = \\
&\int_0^{l_1} \left[d_1 K_1 \frac{(\phi_{1x})^2}{\phi_1^2} + r_1 K_1\left(1 - \frac{\phi_1 + \phi_1'}{K_1}\right)\right]dx \\
&\qquad + \int_{-l_2}^0 \left[d_2 K_2 \frac{(\phi_{2x})^2}{\phi_2^2} + r_2 K_2\left(1 - \frac{\phi_2 + \phi_2'}{K_2}\right)\right]dx = 0,
\end{aligned} \tag{45}$$

where we have integrated by parts and used boundary conditions, with $k = \overline{k}$.

Integrating equations (41) and (43) directly, we obtain

$$\begin{aligned}
&\int_0^{l_1} \left[d_1\phi_{1xx} + r_1\phi_1\left(1 - \frac{\phi_1 + \phi_1'}{K_1}\right)\right]dx + \int_{-l_2}^0 \left[d_2\phi_{2xx} + r_2\phi_2\left(1 - \frac{\phi_2 + \phi_2'}{K_2}\right)\right]dx = \\
&\int_0^{l_1} r_1\phi_1\left(1 - \frac{\phi_1 + \phi_1'}{K_1}\right)dx + \int_{-l_2}^0 r_2\phi_2\left(1 - \frac{\phi_2 + \phi_2'}{K_2}\right)dx = 0.
\end{aligned} \tag{46}$$

Similarly, the integrals of (42) and (44) result in

$$\int_0^{l_1} r_1\phi_1'\left(1 - \frac{\phi_1' + \phi_1}{K_1}\right)dx + \int_{-l_2}^0 r_2\phi_2'\left(1 - \frac{\phi_2' + \phi_2}{K_2}\right)dx = 0. \tag{47}$$

We now subtract (46) and (47) from (45) to derive the expression

$$\int_0^{l_1} d_1 K_1 \frac{(\phi_{1x})^2}{\phi_1^2} dx + \int_0^{l_1} r_1 K_1 \left(1 - \frac{\phi_1 + \phi_1'}{K_1}\right)^2 dx + \int_{-l_2}^{0} d_2 K_2 \frac{(\phi_{2x})^2}{\phi_2^2} dx + \int_{-l_2}^{0} r_2 K_2 \left(1 - \frac{\phi_2 + \phi_2'}{K_2}\right)^2 dx = 0. \quad (48)$$

Since all integrand functions are positive, they must cancel individually. We thus have

$$\phi_{ix} = 0, \quad \phi_i + \phi_i' = K_i \quad \text{in} \quad \Omega_i, \quad i = 1, 2. \quad (49)$$

In particular, ϕ_i, ϕ_i' are constant on Ω_i.

Then we can write $\phi_1 = s$ and $\phi_2 = s/\overline{k} = sK_2/K_1$. Conditions (49) then give $\phi_1' = K_1 - s$ and $\phi_2' = K_2(K_1 - s)/K_1$. Hence, $k' = \phi_1'/\phi_2' = K_1/K_2 = \overline{k}$. Thus, we have shown that if the resident employs the ideal-free movement strategy $k = \overline{k}$, a nontrivial coexistence steady state with a mutant occurs only if the mutant uses the same strategy $k' = \overline{k}$. ■

We consider now the invasibility problem of a rare population using the ideal-free movement strategy ($k = \overline{k}$) on an established population using a different strategy. We write $\phi_{1,2}'(x)$ for the steady-state density of the established population and k' for its strategy as above. It satisfies equations (42) and (44) with $\phi_i = 0$. The growth rate of a rare population in the presence of $\phi_{1,2}'$ is determined from the eigenvalue problem (Cantrell and Cosner, 2003)

$$d_1 u_{1xx} + r_1 \left(1 - \phi_1'/K_1\right) u_1 = \sigma u_1 \quad \text{in} \quad \Omega_1 \quad (50)$$

$$d_2 u_{2xx} + r_2 \left(1 - \phi_2'/K_2\right) u_2 = \sigma u_2 \quad \text{in} \quad \Omega_2, \quad (51)$$

with periodic boundary conditions as in (38)-(39) and parameter $k = \overline{k}$.

Theorem 5.2 *Under the above assumptions and conditions, we have* $\sigma > 0$.

Proof. Dividing (50) and (51) respectively by u_1/K_1 and u_2/K_2, and integrating over space, we have

$$\int_0^{l_1}\left[d_1K_1\frac{u_{1xx}}{u_1}+r_1K_1\left(1-\frac{\phi_1'}{K_1}\right)\right]dx=l_1K_1\sigma, \tag{52}$$

$$\int_{-l_2}^{0}\left[d_2K_2\frac{u_{2xx}}{u_2}+r_2K_2\left(1-\frac{\phi_2'}{K_2}\right)\right]dx=l_2K_2\sigma. \tag{53}$$

Itegrating by parts and using boundary conditions (38) and (39), with $k=\overline{k}=K_1/K_2$, we find

$$\begin{aligned}\int_0^{l_1}d_1K_1\frac{(u_{1x})^2}{{u_1}^2}dx+\int_0^{l_1}r_1K_1\left(1-\frac{\phi_1'}{K_1}\right)dx\\+\int_{-l_2}^{0}d_2K_2\frac{(u_{2x})^2}{{u_2}^2}dx+\int_{-l_2}^{0}r_2K_2\left(1-\frac{\phi_2'}{K_2}\right)dx=(l_1K_1+l_2K_2)\sigma.\end{aligned} \tag{54}$$

The steady steady $\phi_{1,2}'(x)$, in turn, is determined from the system

$$d_1\phi_{1xx}'+r_1\phi_1'\left(1-\phi_1'/K_1\right)=0,\quad \text{in}\quad \Omega_1 \tag{55}$$

$$d_2\phi_{2xx}'+r_2\phi_2'\left(1-\phi_2'/K_2\right)=0,\quad \text{in}\quad \Omega_2. \tag{56}$$

Integrating and summing the equations, we get

$$\int_0^{l_1}r_1\phi_1'\left(1-\frac{\phi_1'}{K_1}\right)dx+\int_{-l_2}^{0}r_2\phi_2'\left(1-\frac{\phi_2'}{K_2}\right)dx=0. \tag{57}$$

Finally, we subtract (57) from (54) and obtain

$$\begin{aligned}\sigma=\frac{1}{l_1K_1+l_2K_2}\Bigg[\int_0^{l_1}d_1K_1\frac{(u_{1x})^2}{{u_1}^2}dx+\int_0^{l_1}r_1K_1\left(1-\frac{\phi_1'}{K_1}\right)^2dx\\+\int_{-l_2}^{0}D_2K_2\frac{(u_{2x})^2}{{u_2}^2}dx+\int_{-l_2}^{0}r_2K_2\left(1-\frac{\phi_2'}{K_2}\right)^2dx\Bigg]\geq 0.\end{aligned} \tag{58}$$

The inequality is strict if at least one of the integrals is not zero. If the established population ϕ_i' does not use the ideal-free strategy, then at least one of the integrals involving ϕ_i' must be positive. Thus, a rare population using the ideal-free movement strategy can always invade a population using any other movement strategy at steady state. ■

Remark: With some additional analysis, it should be possible to establish that, as well as being able to invade a resident population that uses any other type of strategy, an invading population using an ideal-free strategy will actually exclude the resident population. (That is, in this setting, invasion implies fixation.) The argument would be the same as in Cantrell et al. (2010); Averill et. al (2012). The time dependent version of (41)-(44) is a competition system with the usual advection-diffusion operators replaced by diffusion operators with an interface condition for patch preference. Those operators have maximum and comparison principles by Propositions 3.1 and 3.2. Since the proof of the usual comparison principle for systems of two competitors is based on using single-equation maximum principles in a component-wise way, it would extend to the models we consider here. That would imply that the time-dependent system corresponding to (41)-(44) is monotone with respect to the usual competitive ordering. Then, by the general theory of competitive systems, if $\sigma > 0$, a small invading population using the ideal-free strategy will ultimately grow until the system approaches an equilibrium where the ideal-free population has a positive density. If, in addition, there is no coexistence equilibrium, the only equilibrium where the ideal-free competitor has a positive density will be one where the non-ideal-free competitor is not present, which would imply competitive exclusion. For examples and discussion of this aspect in the reaction-advection-diffusion case, see Cantrell et al. (2010); Averill et. al (2012); Cosner (2014).

6 Proofs

In this section, we provide the proofs of the analytical preliminaries from Section 3.

6.1 Proof of Proposition 3.1

First we note that, as in the classical maximum principle, we can assume functions a_i to be of any sign. The new variables $\tilde{u}_i = e^{\gamma t} u_i$ satisfy the same equations as u_i with a_i replaced by $a_i + \gamma$ and the same boundary and interface conditions.

Secondly, for $\varepsilon > 0$ we define

$$v_1 = u_1 + k\varepsilon e^t, \quad v_2 = u_2 + \varepsilon e^t.$$

Then $v_i(x, 0) > 0$ and v_i satisfy (7)-(8) with strict inequalities, as well as the boundary and interface conditions (9).

Now suppose that $v_i(x,t) \le 0$ for some i, some $x \in [-l_2, 0] \cup [0, l_1]$ and $t \le T$, and define

$$t_0 \colon = \sup\{0 < t \le T : \ v_i(x,t) > 0, \ x \in [-l_2, 0] \cup [0, l_1], i = 1, 2\} > 0.$$

Then there exists some $x_0 \in [-l_2, 0] \cup [0, l_1]$ with $v_1(x_0, t_0) = 0$ or $v_2(x_0, t_0) = 0$ and $v_i \geq 0$ for $t \leq t_0$, $x \in [-l_2, 0] \cup [0, l_1]$. We distinguish two cases.

If $x_0 \in (-l_2, 0)$ or $x_0 \in (0, l_1)$, then $v_{it}(x_0, t_0) \leq 0$ and $v_{ixx}(x_0, t_0) \geq 0$ but $v_{it}(x_0, t_0) - v_{ixx}(x_0, t_0) > 0$ by assumption. Hence, we have a contradiction as in the classical maximum principle.

If $x_0 = 0$, then $v_1(0, t_0) = kv_2(0, t_0) = 0$. However, since $v_i \geq 0$, for $x \in (-l_2, 0) \cup (0, l_1)$, we have $v_{1x}(0, t_0) \geq 0$ and $v_{2x}(0, t_0) \leq 0$. By the interface condition, we find $v_{1x}(0, t_0) = 0 = v_{2x}(0, t_0)$.

Now, if $v_1(x, t_0) > 0$ on some interval $(0, \delta)$ then $v_{1x}(0, t_0) > 0$, and similarly if $v_2(x, t_0) > 0$ on some interval $(-\delta, 0)$ then $v_{2x}(0, t_0) < 0$, which is a contradiction. Hence, we must have $v_1(x, t_0) = 0$ somewhere in $(0, l_1)$ and also $v_2(x, t_0) = 0$ somewhere in $(-l_2, 0)$. But then by the classical maximum principle, we have $v_1 = v_2 \equiv 0$ for $0 < t \leq t_0$. This is another contradiction.

The remaining two cases $x_0 = -l_2$ and $x_0 = l_1$ are treated in the same way. Hence, we must have $v_i > 0$. Since this argument holds for all $\varepsilon > 0$, we find $u_i \geq 0$ for $i = 1, 2$, $x \in [-l_2, 0] \cup [0, l_1]$ and $0 < t \leq T$.

Finally, assume that $u_i \not\equiv 0$ for $t = 0$ but $u_i(x_0, t_0) = 0$ for some $i = 1, 2$, $x_0 \in [-l_2, 0] \cup [0, l_1]$ and $0 < t \leq T$. Then, by the strong maximum principle $u_i \equiv 0$ for $0 < t \leq t_0$, which is a contradiction. Hence, we must have $u_i > 0$ for all $i = 1, 2$, $x \in [-l_2, 0] \cup [0, l_1]$ and $0 < t \leq T$. ■

6.2 Proof of Proposition 3.3

If u_1 has a negative minimum on $[0, l_1]$ or u_2 on $[-l_2, 0]$ then it must occur at $-l_2, 0$ or l_1 by the classical maximum principle. Suppose that a negative minimum occurs at $-l_2$. Then either $u_2 \equiv$const< 0 or $u_{2x}(-l_2) > 0$. By the boundary condition, we exclude the latter. But with the former, we find the contradiction $0 > c_2 u_2 = \bar{f}_2 \geq 0$. The same argument applies for a negative minimum at l_1.

Suppose now that the minimum occurs at $x = 0$. Then either $u_2 \equiv$const< 0 or $u_{2x}(0) < 0$. The former case cannot occur by the same argument as above; therefore the latter holds. By the interface condition, we then find $u_{1x}(0) < 0$. But this means that the minimum of u_1 must occur in $(0, l_1]$, which is impossible by the previous argument. Therefore, $u_i \geq 0$.

Now suppose that $\bar{f}_2 \geq 0$ and $\bar{f}_2 \not\equiv 0$. Then $u_2 > 0$ by the strong maximum principle. This implies $u_2(0) > 0$, so that $u_1(0) > 0$ as well. But then $u_1 > 0$ by the strong maximum principle, again. ■

6.3 Proof of Proposition 3.4

We begin with the time-independent problems (14) and (15) together with Neumann boundary conditions on each interval. The two problems decouple, so that each problem becomes an independent, regular Sturm–Liouville problem. Hence, we obtain solutions of

$$-d_i\tilde{u}_{ixx} + c_i\tilde{u}_i = \bar{f}_i$$

with Neumann boundary conditions on $[0, l_1]$ and $[-l_2, 0]$, respectively, for continuous functions $\bar{f}_i$. By the maximum principle, we have $\min \bar{f}_i < \tilde{u}_i < \max \bar{f}_i$ so that $\|\tilde{u}_i\|_\infty \leq \|\bar{f}_i\|_\infty$. From the equation, we then find $\|\tilde{u}_{ixx}\|_\infty \leq C\|\bar{f}_i\|_\infty$. We can bound the first derivative from the bound of the second derivative so that we obtain the estimates

$$\|\tilde{u}_1\|_{\mathcal{C}^2([0,l_1])} \leq C\|\bar{f}_1\|_{\mathcal{C}([0,l_1])} \quad \text{and} \quad \|\tilde{u}_2\|_{\mathcal{C}^2([-l_2,0])} \leq C\|\bar{f}_2\|_{\mathcal{C}([-l_2,0])}.$$

Secondly, we define y_1 to be the solution of

$$-d_1y_{1xx} + c_1y_1 = 0, \quad 0 \leq x \leq l_1,$$

with conditions $y_1(l_1) = 1$, and $y_{1x}(l_1) = 0$. Then we have $d_1y_{1xx}(l_1) = c_1(l_1) > 0$ by the differential equation, which leads to $y_{1x}(l_1) < 0$. Hence, near l_1 we have $y_{1x}(x) < 0$ and $y_1(x) > 1$. Suppose that $y_{1x}(x) = 0$ somewhere in $[0, l_1)$ and denote x_0 as the supremum of all such x. Then for $x > x_0$, we have $y_{1x}(x) < 0$ and $y_1(x) > 1$. Therefore, we find $y_{1xx}(x_0) \leq 0$. However, by the differential equation, $d_1y_{1xx}(x_0) = c_1(x_0)y_1(x_0) > c_1(x_0) > 0$. Hence, we have a contradiction and conclude that $y_{1x}(0) < 0$ and $y_1(0) > 0$.

The same argument goes to show that y_2 defined as the solution of

$$-d_2y_{2xx} + c_2y_2 = 0, \quad -l_2 \leq x \leq 0,$$

with conditions $y_2(-l_2) = 1$, and $y_{2x}(-l_2) = 0$ satisfies $y_2(0) > 0$ and $y_{2x}(0) > 0$.

Now we define

$$u_1 = \tilde{u}_1 + a_1y_1, \qquad u_2 = \tilde{u}_2 + a_2y_2$$

for parameters a_i. Then u_i satisfy the equations (14) and (15). Furthermore, we have $u_{1x}(l_1) = 0$ and $u_{2x}(-l_2) = 0$. We claim that we can choose the parameters in such a way as to satisfy the interface conditions in (16). The conditions can be written as

$$d_1y_{1x}(0)a_1 - d_2y_{2x}(0)a_2 = 0, \tag{59}$$

$$y_1(0)a_1 - ky_2(0)a_2 = k\tilde{u}_2(0) - \tilde{u}_1(0). \tag{60}$$

The determinant of the coefficient matrix on the left-hand side is

$$-kd_1y_{1x}(0)y_2(0) + d_2y_1(0)y_{2x}(0) > 0.$$

In particular, there is a unique solution a_i that depends on d_i, c_i, k but not on $\bar{f}_i$. Therefore the solution u_i satisfies the estimate in the statement of the proposition.■

7 Discussion

Habitat heterogeneity is ubiquitous in nature, and organisms respond to habitat variation by adjusting their movement behavior. For instance, populations can develop conditional dispersal and exhibit patch specific movement rates as well as attraction towards more favorable regions. It is critical to the persistence of the population that individuals choose ‘good’ movement strategies. These strategies are therefore traits subject to selection and evolution (Clobert et al., 2001, 2012). The theoretical study of the evolution of dispersal is well established by now (Johnson and Gaines, 1990; McPeek and Holt, 1992; Cosner, 2014), and models based on reaction-diffusion equations in particular, have generated interesting biological results as well as challenging mathematical problems (Cantrell et al., 2006; Lou, 2008; Lou and Lutscher, 2014). Our work continues this line of inquiry by using a relatively recently established reaction-diffusion model for population dynamics in a ‘patchy’ landscape.

Our model reflects a landscape-ecology point of view that separates a heterogeneous landscape into patches. Within a patch, conditions are homogeneous, at interfaces between patches, they change abruptly. Individual movement behavior depends on patch quality and may include patch preference. Our model is both, easier and more difficult than previous models. It is easier to parametrize than models with continuous variation in habitat quality and movement behavior, and it allows the straightforward inclusion of empirical results on patch preference. It is more difficult because the population density is discontinuous at an interface with a prescribed jump condition so that the classical results on existence and uniqueness of solutions cannot be applied directly. We provided a proof of the well-posedness of the equations, as well as extensions of the classical maximum- and comparison principles, and the existence of a dominant eigenvalue.

We studied the evolution of dispersal in our model within the framework of invasion analysis, a part of the adaptive dynamics framework (Geritz et al., 1998). We used two different approaches, one via homogenization and one via steady-state analysis. Somewhat surprisingly, both approaches gave the exact same results. The three movement parameters (habitat-specific movement rates and habitat preference) combine to a single effective preference parameter $k = d_2\alpha/[d_1(1-\alpha)]$ that governs the density matching at an interface and is based on a mechanistic derivation (Ovaskainen and Cornell, 2003; Maciel and Lutscher, 2013, 2015). The adaptive dynamics for the homogenized model identifies the optimal value $\overline{k} = K_1/K_2$ as the ESS and NIS of the system: We showed that a population with this strategy cannot be invaded by any other strategy but can invade every other strategy. This strategy also corresponds to an IFD in the non-homogenized model. This result fits into a growing awareness of a general principle (Cosner, 2014). It is somewhat remarkable that the optimal strategy, $\overline{k}$, does not depend on patch sizes. This observation explains why the homogenization, which is accurate for very small patch sizes, yields the same result as the non-homogenized steady-state analysis.

Because of the mechanistic interpretation of trait k, we can interpret special cases of our result. If diffusivities are equal, the evolutionarily stable habitat preference is given by $\overline{\alpha} = K_1/(K_1 + K_2)$. This intermediate preference allows the population to optimally make use of space, preventing the invasion of mutant traits. On the other hand, when there is no habitat preference, i.e. $\alpha = 0.5$, the evolutionarily stable ratio of diffusion rates is $\overline{d}_2/\overline{d}_1 = K_1/K_2$. Selected diffusion rates in a patch are then inversely proportional to the carrying capacity in that patch. Since the inverse of the diffusion rate is proportional to the 'residence index' (McNair, 1982; Turchin, 1998), our result says that the time that an individual spends in a certain patch should be proportional to the carrying capacity of that patch.

Although we have assumed an explicit expression for k based on specific movement characteristics, our analysis is quite general and is valid when other assumptions on movement are made. Potapov et al. (2014) derived a family of diffusion models by modelling random walks with transition probabilities, from one location to the next, that depend on conditions at the starting point, at the end point or at some position in between. In this formulation, transition probabilities from x to $x \pm \Delta x$ are then written as $p = p(x \pm \theta \Delta x)$, with $\theta \in [0, 1]$. Fokker-Plank (FP), $u_t = (du)_{xx}$, and Fickian, $u_t = (du_x)_x$, diffusion equations are obtained when $\theta = 0$ and $\theta = 0.5$, respectively. When $\theta = 1$ one gets the diffusion equation $u_t = (d^2(u/d)_x)_x$, which has been termed attractive dispersal (AD). In the absence of habitat preference, FP, Fickian and AD diffusion equations lead to $k = d_2/d_1$, $k = 1$ and $k = d_1/d_2$, respectively. The first case corresponds exactly to the equations we have considered (when $\alpha = 0.5$). AD results in an ESS $\overline{d}_1/\overline{d}_2 = K_1/K_2$, where diffusion is proportional to the carrying capacity in a patch. Yet for Fickian diffusion, an IFD can not be achieved as, unless $K_1 = K_2$, we always have $k \neq K_1/K_2$. Potapov et al. (2014) numerically determined that FP and AD are two candidates for ESS's within this family of diffusion models, FP being selected when diffusion is decreasing with fitness and AD being selected when diffusion increases with fitness. The FP equation is often considered to provide the best description of ecological diffusion as it aggregates individuals where movement is slow (Turchin, 1998).

Our model could be extended to study the evolution of dispersal and patch preference for more than two patches (or patch types). Perhaps the simplest extension would be to include a "corridor" between the two patches, i.e., a region that connects the two patches but in which there are no population dynamics. This extension relates to much earlier work by Slatkin (1978). Other authors have studied the evolution of dispersal among discrete habitat patches (Cantrell et al., 2017). Our work could be related to theirs by using a transition from continuous to patch models via residence times (Cobbold and Lutscher, 2014). Finally, empirical studies about patch preference and movement rates exist (e.g., Kuefler et al. (2010)) and it would be interesting to see whether or how close real populations are to the theoretical "optimal" behavior.

Acknowledgements

GAM is thankful for a postdoctoral grant from CNPq - Brazil. CC and RSC would like to acknowledge funding through the National Science Foundation of the United States (grant DMS-1514752). FL gratefully acknowledges funding from the Natural Sciences and Engineering Research Council of Canada through an individual Discovery Grant (RGPIN-2016-04759) and a Discovery Accelerator Supplement (RGPAS-2016-492872). We also thank Odo Diekmann and two anonymous reviewers for careful reading of the manuscript and helpful suggestions.

Appendix

We begin the appendix with a short remark on the relation between the infinite periodic model and the restricted two-patch model. As in Section 4, it is convenient to write the equation for a single population in the infinite periodic case as

$$u_t = [D(x)u]_{xx} + f(x,u).$$

If we assume that the two patch types have length $2l_1$ and $2l_2$, respectively, then the parameter functions D and f are periodic in x with period $2(l_1 + l_2) = 2l$, i.e., they are invariant under the translation $x \mapsto x + 2l$. Since the functions are also piecewise constant, they are also invariant under the reflection $x \mapsto 2l_1 - x$. Combining the translation invariance with the reflection invariance, we obtain a second reflection invariance under $x \mapsto 2l_1 + 2l - x$. Since the coefficient functions have these symmetry properties, every steady-state solution, u^*, has the same symmetry properties. The coefficients in the corresponding eigenvalue problem

$$\lambda\phi = [D(x)\phi]_{xx} + \partial f(x,u^*)/\partial u \; \phi$$

will then also have these symmetry properties. Therefore, the eigenfunctions will have these properties as well. The fixed points of the reflection symmetry are $x = l_1$ and $x = l_1 + l$. A smooth function with such a reflection symmetry must have zero slope that these fixed points. Hence $u_x = 0$ for $x = l_1$ and $x = -l_2$. In particular, every steady-state solution of the periodic problem is also a steady-state solution of the restricted problem, and the corresponding eigenfunctions are also eigenfunctions of the restricted problem. It is obvious that every solution of the restricted problem can be continued periodically to the real line with the symmetry conditions satisfied. A similar case was discussed in more detail for a discrete-time system of equations by Musgrave and Lutscher (2014).

In the remainder of this appendix, we show the existence and uniqueness and global boundedness of solutions to our model equations on the intervals $[-l_2, 0] \cup [0, l_1]$. Our proof is based on semi-group theory and closely follows the proof in Cosner (1987).

We consider the reaction-diffusion system

$$u_{it} = d_i u_{ixx} + (E_i - F_i u_i - G_i v_i)u_i = d_i u_{ixx} + h_i(u_i, v_i), \tag{61}$$

$$v_{it} = D_i v_{ixx} + (\tilde{E}_i - \tilde{F}_i u_i - \tilde{G}_i v_i)v_i = D_i v_{ixx} + H_i(u_i, v_i), \tag{62}$$

for $t \geq 0$ and

$$x \in \begin{cases} [0, l_1], & i = 1, \\ [-l_2, 0], & i = 2, \end{cases} \tag{63}$$

together with boundary and interface conditions

$$u_{1x}(l_1,t) = 0 = u_{2x}(-l_2,t), \quad u_1(0,t) = ku_2(0,t), \quad d_1 u_{1x}(0,t) = d_2 u_{2x}(0,t), \tag{64}$$

$$v_{1x}(l_1,t) = 0 = v_{2x}(-l_2,t), \quad v_1(0,t) = Kv_2(0,t), \quad D_1 v_{1x}(0,t) = D_2 v_{2x}(0,t). \tag{65}$$

All parameters are assumed positive. We begin by defining the appropriate function spaces.

Set-up of the problem

We cast the problem into the form of an abstract evolution equation

$$\frac{\mathrm{d}}{\mathrm{d}t}w + Aw = F(w), \tag{66}$$

where $w = (u, v)^T$ and $u = (u_1, u_2)^T$ and $v = (v_1, v_2)^T$. We define operators

$$A_u \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} -d_1 u_{1xx} \\ -d_2 u_{2xx} \end{pmatrix} \quad \text{and} \quad A_v \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} -D_1 v_{1xx} \\ -D_2 v_{2xx} \end{pmatrix}. \tag{67}$$

Then we can write

$$Aw = \begin{pmatrix} A_u + I & 0 \\ 0 & A_v + I \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -d_1 u_{1xx} + u_1 \\ -d_2 u_{2xx} + u_2 \\ -D_1 v_{1xx} + v_1 \\ -D_2 v_{2xx} + v_2 \end{pmatrix} \tag{68}$$

and $F(w) = (F_u, F_v)^T$, where

$$F(w) = \begin{pmatrix} F_u(w) \\ F_v(w) \end{pmatrix} = \begin{pmatrix} h_1(u_1, v_1) + u_1 \\ h_2(u_2, v_2) + u_2 \\ H_1(u_1, v_1) + v_1 \\ H_2(u_2, v_2) + v_2 \end{pmatrix}. \tag{69}$$

We define the following function spaces.

$$Y_u = Y_v = L^2([0, l_1]) \times L^2([-l_2, 0]), \quad \text{and} \quad Y = Y_u \times Y_v. \tag{70}$$

$$W_u = W_v = W^{2,2}([0, l_1]) \times W^{2,2}([-l_2, 0]), \quad \text{and} \quad W = W_u \times W_v. \tag{71}$$

Since we are in one space dimension, we have $W^{2,2} \hookrightarrow \mathcal{C}^1$. Hence, a function $u \in W_u$ is continuously differentiable, so that we can impose the boundary and interface conditions that we want. We therefore set $X = X_u \times X_v$ with

$$X_u = \{(u_1, u_2) \in W_u \mid u \quad \text{satisfies} \quad (64)\} \tag{72}$$

and accordingly for X_v with (64) replaced by (65).

On Y_u we define the inner product

$$\langle u, z \rangle_{Y_u} = \langle u_1, z_1 \rangle_{L^2([0,l_1])} + k \langle u_2, z_2 \rangle_{L^2([-l_2,0])} \tag{73}$$

and obtain the norm

$$\|u\|^2_{Y_u} = \|u_1\|^2_{L^2([0,l_1])} + k\|u_2\|^2_{L^2([-l_2,0])} \tag{74}$$

and similarly on Y_v with k replaced by K. Finally, we have

$$\|w\|^2_Y = \|u\|^2_{Y_u} + \|v\|^2_{Y_v}. \tag{75}$$

The linear problem

Proposition 7.1 *The linear operator A defines an analytic semigroup on Y.*

Proof. We will show that A is invertible and that the closure of the numerical range is contained in $[1,\infty)$. Then Lemma 2 in Cosner (1987) (which is a special case of Theorem V.3.2 in Kato (1966)) states that the conditions for the generation of an analytic semigroup from Part 2, Section 2 in Friedman (1969) are satisfied.

We note that since the operator A is diagonal, and since A_u and A_v are essentially identical, it is sufficient to show the two properties for A_u.

We begin by calculating the numerical range of A_u.

$$\begin{aligned}\langle A_u u, u\rangle_{Y_u} &= \int_0^{l_1}(-d_1u_{1xx}+u_1)\bar{u}_1\mathrm{d}x + k\int_{-l_2}^{0}(-d_2u_{2xx}+u_2)\bar{u}_2\mathrm{d}x\\ &= -d_1u_{1x}\bar{u}_1|_0^{l_1} + \int_0^{l_1} d_1u_{1x}\bar{u}_{1x}\mathrm{d}x - kd_2u_{2x}\bar{u}_2|_{-l_2}^{0} + \int_{-l_2}^{0} kd_2u_{2x}\bar{u}_{2x}\mathrm{d}x + \langle u,u\rangle_{Y_u}.\end{aligned}$$

(Note that $\bar{u}$ denotes the complex conjugate of the function u.)

By the boundary and interface conditions (64), the first and third term cancel. The two integral terms are non-negative, and therefore, we find

$$\begin{aligned}\langle A_u u, u\rangle_{Y_u} &= \int_0^{l_1} d_1u_{1x}\bar{u}_{1x}\mathrm{d}x + \int_{-l_2}^{0} kd_2u_{2x}\bar{u}_{2x}\mathrm{d}x + \langle u,u\rangle_{Y_u}\\ &\geq \langle u,u\rangle_{Y_u}.\end{aligned}$$

Hence, the numerical range

$$\theta(A_u) = \{\langle A_u u, u\rangle_{Y_u} \mid \|u\|_{Y_u} = 1\} \tag{76}$$

is contained in $[1,\infty)$ and so is its closure. The same is true for A_v and therefore also for A.

Secondly, we show that A_u has a bounded inverse. Consider $(\bar{f}_1,\bar{f}_2)\in Y_u$. There exist unique functions $\tilde{u}_{1,2}$ that satisfy

$$\begin{aligned}-d_1\tilde{u}_{1xx}+\tilde{u}_1 &= \bar{f}_1, \qquad x\in[0,l_1],\\ -d_2\tilde{u}_{2xx}+\tilde{u}_2 &= \bar{f}_2, \qquad x\in[-l_2,0],\end{aligned}$$

with Neumann conditions at all boundaries, i.e. $\tilde{u}_{1x}(l_1)=\tilde{u}_{1x}(0)=\tilde{u}_{2x}(0)=\tilde{u}_{2x}(-l_2)=0$. The reason is as follows. We notice that with these boundary conditions, the two equations decouple. Then each problem is an inhomogeneous boundary value problem, a special case of a regular Sturm–Liouville problem. A unique solution exists by classical methods (e.g., an explicit calculation of the Green's function). We need to estimate the norm. Classical results (e.g., Theorem 9.27 in Renardi and Rogers (2004)) give the estimate in $W^{1,2}$

$$\|\tilde{u}_1\|_{W^{1,2}([0,l_1])} \leq C\|\bar{f}_1\|_{L^2([0,l_1])}, \qquad \|\tilde{u}_2\|_{W^{1,2}([-l_2,0])} \leq C\|\bar{f}_2\|_{L^2([-l_2,0])}. \tag{77}$$

However, we need an estimate in $W^{2,2}$. We can write the equations as

$$\tilde{u}_{ixx} = \frac{1}{d_i}(\tilde{u}_i - \bar{f}_i)$$

and take norms on both sides to get

$$\|\tilde{u}_{ixx}\|_{L^2} \leq \tilde{C}(\|\tilde{u}_i\|_{L^2} + \|\bar{f}_i\|_{L^2}).$$

By the previous estimate, the right hand side can be bounded by some multiple of the L^2-norm of the data f_i alone so that we obtain the overall estimate

$$\|\tilde{u}_1\|_{W^{2,2}([0,l_1])} \leq C\|\bar{f}_1\|_{L^2([0,l_1])}, \qquad \|\tilde{u}_2\|_{W^{2,2}([-l_2,0])} \leq C\|\bar{f}_2\|_{L^2([-l_2,0])}. \tag{78}$$

We now use the same construction of functions $y_{1,2}$ in the proof of the existence of the dominant eigenvalue to obtain functions

$$u_i = \tilde{u}_i + y_i \tag{79}$$

that satisfy the differential equations and the same norm estimates as $\tilde{u}_i$ with potentially different constants.

The same construction works for A_v and therefore we have shown that A is invertible with bounded inverse. Lemma 2 in Cosner (1987) (which is a special case of Theorem V.3.2 in Kato (1966)) now states that $\mathbb{C}\backslash[1,\infty)$ is contained in the resolvent set of A and

$$\|(\lambda - A)^{-1}\| \leq \frac{1}{\operatorname{dist}(\lambda, \overline{\theta(A)})} \tag{80}$$

for all λ in the resolvent set.

Denote the distance by $d = \operatorname{dist}(\lambda, \overline{\theta(A)})$. We want to show that there exists a constant C such that

$$d \geq \frac{1+|\lambda|}{C}$$

for $\Re\lambda \leq 0$, so that from (80) we get the required estimate

$$\|(\lambda - A)^{-1}\| \leq \frac{C}{1+|\lambda|}. \tag{81}$$

On the semicircle $|\lambda|$ with $\Re\lambda \leq 0$, the function d assumes its minimum when λ is purely imaginary. Hence, it is enough to show the inequality on the imaginary line. Hence, we need to show the existence of a constant C such that

$$\sqrt{1+z^2} \geq \frac{1+z}{C}, \qquad z \geq 0.$$

The function $z \mapsto \frac{1+z^2}{(1+z)^2}$ is positive, continuous, and bounded with $f(0) = f(\infty) = 1$. Its maximum is 1 and its minimum occurs at $x = 1$. We can take C to be the inverse of the minimum of this function.

With this, we see that A satisfies the characterization to generate an analytic semigroup according to the theory developed in Friedman (1969), Part 2, Section 2. The statement is also available in Theorem 36.2 in Sell and You (2002) or in the book Pazy (1983). ∎

The nonlinear problem

We now return to the nonlinear problem (66) and prove local existence of solutions. We use the following (notation adapted) time-independent version of Lemma 3 in Cosner (1987).

Proposition 7.2 *Let A be a closed linear operator on a Banach space Y such that (81) holds. Suppose that F is a function on Y such that for some $0 < \beta < 1$ and for any $R > 0$, there exists a constant $C(R)$ such that*

$$\|F(A^{-\beta}p_1) - F(A^{-\beta}p_2)\|_Y \leq C(R)\|p_1 - p_2\|_Y \tag{82}$$

for all $p_{1,2} \in Y$ with $\|p_i\|_Y < R$. Then for any $p_0 \in \mathcal{D}(A)$ and each $R > \|A^{-\beta}p_0\|_Y$ there exists a $t^* > 0$ such that problem (66) has a unique solution in $[0, t^*]$.

From the previous section, we know that A is a closed linear operator on Y and that the norm estimate for the resolvent holds. To find an appropriate choice of β, we begin with the statement of Lemma 37.8 in Sell and You (2002).

Lemma 7.3 *Let A be a positive, sectorial operator on $L^q(\Omega, \mathbb{R}^n)$ with domain $\mathcal{D}(A) \hookrightarrow W^{m,q}$ for some $m \geq 1$. Let $0 < \beta \leq 1$. Then $\mathcal{D}(A^\beta) \mapsto W^{k,p}$ if $p \geq q$, $k \geq 0$ and $k - n/p < m\beta - n/q$.*

We apply this lemma with $n = 1$, $q = p = 2$ and $m = 2$. Then we get that $\mathcal{D}(A^{\beta}) \mapsto W^{1,2}$ for all $1/2 < \beta \leq 1$. We now fix some $\beta \in (1/2, 1)$.

We pick functions $p_{1,2} \in Y$ and set $q_i = A^{-\beta} p_i$. Since $A^{-\beta}$ maps into $\mathcal{D}(A^{\beta})$ and since by the previous lemma and our choice of β, we have the embedding into $W^{1,2}$ in each component, we see that q_i are continuous and there is a constant $C_1(R)$ such that $\|q_i\|_{\infty} \leq C_1(R)\|q_i\|_{W^{1,2}} \leq C_1(R)\|A^{-\beta}\|\|p_i\|_Y$.

For $\nu \in [0, 1]$ we define $u(\nu) = q_2 + \nu(q_1 - q_2)$. The function $\nu \mapsto F(u(\nu))$ satisfies $F(u(1)) = F(q_1)$ and $F(u(0)) = F(q_2)$. We apply the fundamental theorem of calculus and the chain rule to write

$$\begin{aligned}
\|F(A^{-\beta}p_1) - F(A^{-\beta}p_2)\|_Y &= \|F(q_1) - F(q_2)\|_Y && (83)\\
&= \|F(u(1)) - F(u(0))\|_Y && (84)\\
&= \|\int_0^1 DF(u(\nu))\frac{d}{d\nu}u(\nu)\mathrm{d}\nu\|_Y. && (85)
\end{aligned}$$

Clearly, the derivative of u is $\frac{d}{d\nu}u(\nu) = q_1 - q_2$. Furthermore, the nonlinearity of F consists of polynomials of degree at most 2 in each component. In particular, DF consists of at most linear combinations of the functions in q_i. Since q_i are bounded by the above reasoning, there is an L^{∞}-bound $C_2 = C_2(R)$ on DF for $\|p_i\| \leq R$.

Hence, we get the estimate

$$\|F(A^{-\beta}p_1) - F(A^{-\beta}p_2)\|_Y \leq C_2(R)\|q_1 - q_2\|_Y \qquad (86)$$

Therefore, the proposition applies and we obtain local existence of solutions.

Proposition 7.4 *Let* $p_0 \in \mathcal{D}(A)$ *and denote by* $w(t)$ *the unique local solution of (66) and* $w(0) = p_0$. *Then* $w(t) \in \mathcal{D}(A)$ *for all* $t \in [0, t^*]$ *and* w *as well as* $\mathrm{d}w/\mathrm{d}t$ *are strongly continuous in* $[0, t^*]$. *Furthermore,* $\mathrm{d}^2w/\mathrm{d}t^2$ *exists and is strongly continuous. Finally, if the initial condition as a function of* x *is non-negative and appropriately bounded, then so is the solution.*

Proof.
The analytic semigroup generated by A maps Y into $\mathcal{D}(A^{\beta})$ for all $\beta \geq 0$ (Theorem 37.5 in Sell and You (2002)). Therefore, the solution is in $\mathcal{D}(A)$, see also Theorem 2, in Friedman (1965). Continuity of w with respect to time follows from Theorem 2 in Friedman (1965). Higher regularity of solutions follows from the considerations following that theorem. Specifically, if the (Fréchet) derivative of $F(A^{-\beta}p)$ exists

and is Lipschitz continuous, then the solution has strong first and second derivatives and they all belong to $\mathcal{D}(A^\beta)$. Since F consists of quadratic terms, the derivative consists of linear terms and is therefore Lipschitz continuous.

To show positivity of solutions with non-negative, non-zero initial data, we apply the comparison principle (Proposition 3.2). To show the upper bounds, we proceed as follows. Function h_i is negative for $u_i > E_i/F_i$, independently of v_i. Hence, it is sufficient to show an upper bound for solutions of the u_i-equations alone.

If $kE_1/F_1 \geq E_2/F_2$, we set $u_1(x,t) = E_1/F_1$ and $u_2(x,t) = ku_1$. If $kE_1/F_1 < E_2/F_2$, we set $u_2(x,t) = E_2/F_2$ and $u_1(x,t) = u_2/k > E_1/F_1$. In either case, we find $h_i(u_i, 0) \leq 0$. Hence, we have found an upper solution and can apply the comparison principle again. ■

Proposition 7.5 *The local solutions obtained above are global solutions, i.e. they exist for $t \in [0,\infty)$.*

Proof. We pick $T_0 > 0$. As in the proof of Lemma 3 in Cosner (1987) and in the proof of Theorem 1 in Bell and Cosner (1981), we need to show that for every local solution w on $[0, T_1]$ with $T_1 \leq T_0$, there exists a constant R' such that $\|Aw\|_Y < R'$. Then we can choose $R > R'$ and apply the local existence result successively on $[0,t^*]$, $[t^*, 2t^*]$, and so on until T_0. Since T_0 was arbitrary, we have global existence.

To show the existence of the constant R', we note that $Aw = -w_t + F(w)$. Hence, we aim to estimate

$$\|-w_t + F(w)\|_Y.$$

We set

$$\begin{aligned} E(t) &= \frac{1}{2}\left(\|w\|_Y^2 + \|w_t\|_Y^2\right) \\ &= \frac{1}{2}\left(\|u\|_{Y_u}^2 + \|v\|_{Y_v}^2 + \|u_t\|_{Y_u}^2 + \|v_t\|_{Y_v}^2\right) \end{aligned} \tag{87}$$

and calculate $E'(t)$.

The first term in (87) gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}\|u\|_{Y_u}^2 = -\langle u, A_u u\rangle_{Y_u} + \langle u, F_u(u,v)\rangle_{Y_u}. \tag{88}$$

We estimate the first of these terms as we did in the calculation of the numerical range of the operator A, see Proposition 7.1. We obtain

$$-\langle u, A_u u\rangle_{Y_u} \leq -\langle u, u\rangle_{Y_u} \leq 0.$$

To estimate the second of these terms, we note that by the maximum principle, non-negative solutions (u, v) are L^∞ bounded independent of time (see previous proposition), so that the terms $E_i - F_i u_i - G_i v_i$ are also L^∞ bounded independent of time. Then we can estimate

$$\langle u, F_u(u,v)\rangle_{Y_u} = \int_0^{l_1} u_1(E_1 - F_1 u_1 - G_1 v_1)u_1 \mathrm{d}x + k\int_{-l_2}^0 u_2(E_2 - F_2 u_2 - G_2 v_2)u_2 \mathrm{d}x \le C_1 \langle u, u\rangle_{Y_u}.$$

The second term in (87) is estimated in the exact same way.

The third term in (87) consists of three terms, namely

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}\|u_t\|^2_{Y_u} = -\langle u_t, A_u u_t\rangle_{Y_u} + \langle u_t, D_u F_u(u,v)u_t\rangle_{Y_u} + \langle u_t, D_v F_u(u,v)v_t\rangle_{Y_u}. \qquad (89)$$

The first of these three terms satisfies the same estimate as the corresponding term above, i.e.

$$-\langle u_t, A_u u_t\rangle_{Y_u} \le -\langle u_t, u_t\rangle_{Y_u} \le 0.$$

The second term can be estimated in a similar way as the second term above since $D_u F_u$ consists of linear polynomials. Hence, we find

$$\langle u_t, D_u F_u(u,v)u_t\rangle_{Y_u} \le C_2 \langle u_t, u_t\rangle_{Y_u}.$$

The third term is slightly different. It is given by

$$\begin{aligned}
\langle u_t, D_v F_u(u,v)v_t\rangle_{Y_u} &= \int_0^{l_1} u_{1t}(-G_1 u_1)v_{1t}\mathrm{d}x + k\int_{-l_2}^0 u_{2t}(-G_2 u_2)v_{2t}\mathrm{d}x \\
&\le C_3\left(\int_0^{l_1} u_{1t}v_{1t}\mathrm{d}x + k\int_{-l_2}^0 u_{2t}v_{2t}\mathrm{d}x\right) \\
&\le C_3\left(\int_0^{l_1} (u_{1t}^2 + v_{1t}^2)\mathrm{d}x + k\int_{-l_2}^0 (u_{2t}^2 + v_{2t}^2)\mathrm{d}x\right) \\
&\le C_4\left(\langle u_t, u_t\rangle_{Y_u} + \langle v_t, v_t\rangle_{Y_v}\right).
\end{aligned}$$

A similar estimate holds for the v-component.

Altogether, we obtain the estimate $E'(t) \le \widehat{C}E(t)$. In particular, E can grow at most exponentially in time. In particular, $\|w\|$ and $\|w_t\|$ remain bounded for any finite time. The bound on $F(w)$ is obvious by the L^∞- bound of w. Hence, we have shown that a constant R' exists as required. ■

References

Alqawasmeh, Y. and Lutscher, F. (2019). Persistence and spread of stage-structured populations in heterogeneous landscapes. *Journal of Mathematical Biology*. DOI 10.1007/s00285-018-1317-8

Altenberg, L. (2012). Resolvent positive linear operators exhibit the reduction phenomenon. *Proceedings of the National Academy of Sciences*, 109:3705–3710.

Apaloo, J. (1997) Revisiting strategic models of evolution:the concept of neighborhood invader strategies. *Theoretical Population Biology*, 77:52-71.

Averill, I., Lou, Y, and Munther, D. (2012) On several conjectures from evolution of dispersal. *Journal of Biological Dynamics*, 6:117-130.

Bell, J. and Cosner, C. (1981). Stability properties of a model of parallel nerve fibers. *Journal of Differential Equations*, 40:303–315.

Bensoussan, A., Lions, J.-L. and Papanicolaou, G. (2010). *Asymptotic Analysis for Periodic Structures.* AMS Chelsea Publishing, Providence.

Cantrell, R. S., Cosner, C., Lou, Y., and Schreiber, S. (2017). Evolution of natal dispersal in spatially heterogeneous environments. *Mathematical Biosciences*, 283:136–144.

Cantrell, R. S. and Cosner, C. (2003). *Spatial Ecology via Reaction-Diffusion Equations.* Wiley Series in Mathematical and Computational Biology. John Wiley & Sons, Ltd., Chichester.

Cantrell, R. S., Cosner, C., and Lou, Y. (2006). Movement toward better environments and the evolution of rapid diffusion. *Mathematical Biosciences*, 204:199–214.

Cantrell, R. S., Cosner, C., and Lou, Y. (2008). Approximating the ideal free distribution via reaction-diffusion-advection equations. *Journal of Differential Equations*, 245:3687–3703.

Cantrell, R. S., Cosner, C., and Lou, Y. (2010). Evolution of dispersal and the ideal free distribution. *Mathematical Biosciences and Engineering*, 7:17–36.

Cantrell, R. S., Cosner, C., and Lou, Y. (2012a) Evolutionary stability of ideal free dispersal strategies in patchy environments. *Journal of Mathematical Biology*, 65: 943–965

Cantrell, R. S., Cosner, C., Lou, Y., and Ryan, D. (2012b) Evolutionary stability of ideal free dispersal in spatial population models with nonlocal dispersal. *Canadian Applied Mathematics Quarterly*, 20:6–38.

Clobert, J., Baguette, M., Benton, T., and Bullock, J. (2012). *Dispersal Ecology and Evolution.* Oxford University Press.

Clobert, J., Danchin, E., Dhondt, A., and Nichols, J. D. (2001). *Dispersal.* Oxford University Press.

Cobbold, C. and Lutscher, F. (2014). Mean occupancy time: linking mechanistic movement models, population dynamics and landscape ecology to population persistence. *Journal of Mathematical Biology*, 68:549–579.

Cosner, C. (1987). Existence of global solutions to a model of a myelinated nerve axon. *SIAM Journal on Mathematical Analysis*, 18(3):703–710.

Cosner, C. (2014). Reaction-diffusion-advection models for the effects and evolution of dispersal. *Discrete and continuous dynamical systems*, 35(5):1701–1745.

Crone, E. and Schultz, C. (2008). Old models explain new observations of butterfly movement at patch edges. *Ecology*, 89(7):2061–2067.

Dewhirst, S. and Lutscher, F. (2009). Dispersal in heterogeneous habitats: thresholds, spatial scales and approximate rates of spread. *Ecology*, 90:1338–1345

Dieckmann U. and Law, R. (1996). The dynamical theory of coevolution: a derivation from stochastic ecological processes. *Journal of Mathematical Biology*, 34:579–612.

Diffendorfer, J. E. (1998). Testing models of source-sink dynamics and balanced dispersal. *Oikos*, 81(3):417–433.

Dockery, J., Hutson, V., Mischaikow, K., and Pernarowski, M. (1998). The evolution of slow dispersal rates: a reaction diffusion model. *Journal of Mathematical Biology*, 37(1):61–83.

Doncaster, C. P., Clobert, J., Doligez, B., Gustafsson, L., and Danchin, E. (1997). Balanced dispersal between spatially varying local populations: an alternative to the source-sink model. *The American Naturalist*, 150(4):425–445.

Du, Y. (2006). *Order structure and topological methods in nonlinear partial differential equations. Vol 1: Maximum Principles and Applications.* World Scientific Publishing, Singapore.

Fretwell, S. and Lucas, H. (1969). On territorial behavior and other factors influencing habitat distribution in birds. *Acta Biotheoretica*, 19(1):16–36.

Friedman, A. (1965). Remarks on nonlinear parabolic equations. *Proceedings of the Symposium in Applied Mathematics*, 17:3–23.

Friedman, A. (1969). *Partial differential equations.* Holt, Rinehart and Winston.

Garlick, M. J., Powell, J. A., Hooten, M. B., and McFarlane, L. R. (2011). Homogenization of large-scale movement models in ecology. *Bulletin of Mathematical Biology*, 73(9):2088–2108.

Geritz, S., Kisdi, É., Meszéna, G., and Metz, J. (1998). Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evolutionary Ecology*, 12(1):35–57.

Hastings, A. (1983). Can spatial variation alone lead to selection for dispersal? *Theoretical Population Biology*, 24(3):244–251.

Holt, R.D. and Barfield, M. (2001) On the relationship between the ideal free distribution and the evolution of dispersal. In Clobert, J., Danchin, E., Dhondt, A., and Nichols, J.D., editors. *Dispersal.* Oxford University Press, pages 83–95.

Hutson, V., Martinez, S., Mischaikow, K., and Vickers, G. (2003). The evolution of dispersal. *Journal of Mathematical Biology*, 47:482–517.

Johnson, M. and Gaines, M. (1990). Evolution of dispersal: theoretical models and empirical tests using birds and mammals. *Annual Review of Ecology and Systematics*, 21:449–480.

Kato, T. (1966). *Perturbation theory for linear operators.* Springer.

Kuefler, D., Hudgens, B., Haddad, N.M., Morris, W.F. and Thurgate, N. (2010). The conflicting role of matrix habitats as conduits and barriers for dispersal. *Ecology*, 91:944–950.

Lam, K.-Y., Lou, Y., and Lutscher, F. (2015). Evolution of dispersal in closed advective environments. *Journal of Biological Dynamics*, 9:188–212.

Lou, Y. (2008). Some challenging mathematical problems in evolution of dispersal and population dynamics. In Friedman, A., editor, *Tutorials in Mathematical Biosciences IV: Evolution and Ecology*, chapter 5, pages 171–205. Springer, Berlin.

Lou, Y. and Lutscher, F. (2014). Evolution of dispersal in open advective environments. *Journal of Mathematical Biology*, 69:1319–1342.

Maciel, G. A. and Lutscher, F. (2013). How individual movement response to habitat edges affects population persistence and spatial spread. *The American Naturalist*, 182(1):42–52.

Maciel, G. A. and Lutscher, F. (2015). Allee effects and population spread in patchy landscapes. *Journal of Biological Dynamics*, 9(1):109–123.

Maciel, G. A. and Lutscher, F. (2018). Movement behavior determines competitive outcome and spread rates in strongly heterogeneous landscapes. *Theoretical Ecology*, 11(3):351–365.

McNair, J. (1982). Optimal giving-up times and the marginal value theorem. *The American Naturalist*, 119:511–529.

McPeek, M. A. and Holt, R. D. (1992). The evolution of dispersal in spatially and temporally varying environments. *The American Naturalist*, 140(6):1010–1027.

Morris, D. W., Diffendorfer, J. E., and Lundberg, P. (2004). Dispersal among habitats varying in fitness: reciprocating migration through ideal habitat selection. *Oikos*, 107(3):559–575.

Musgrave, J. and Lutscher, F. (2015). Integrodifference equations in patchy landscapes II: Population level consequences. *Journal of Mathematical Biology*, 69(3):617–658.

Othmer, H. (1983). A continuum model for coupled cells. *Journal of Mathematical Biology*, 17:351–369.

Ovaskainen, O. and Cornell, S. J. (2003). Biased movement at a boundary and conditional occupancy times for diffusion processes. *Journal of Applied Probability*, 40:557–580.

Parker, G. and Sutherland, W. (1986). Ideal free distributions when individuals differ in competitive ability: phenotype-limited ideal free models. *Animal Behaviour*, 34(4):1222–1242.

Pazy, A. (1983). *Semigroups of linear operators and applications to partial differential equations*. Applied Mathematical Sciences 44. Springer.

Potapov, A., Schlagel, U. E., and Lewis, M. (2014). Evolutionarily stable diffusive dispersal. *Discrete and continuous dynamical systems series B*, 19(10):3319–3340.

Powell, J. and Zimmermann, N. (2004) Multiscale analysis of active seed dispersal contributed to resolving Reid's paradox. *Ecology* 85:490–506.

Reeve, J., Cronin, J., and Haynes, K. (2008). Diffusion models for animals in complex landscapes: incorporating heterogeneity among substrates, individuals and edge behaviours. *J. Anim. Ecol.*, 77:898–904.

Renardi, M. and Rogers, R. (2004). *An Introduction to Partial Differential Equations.* Texts in Applied Mathematics 13. Springer.

Schtickzelle, N. and Baguette, M. (2003). Behavioural responses to habitat patch boundaries restrict dispersal and generate emigration-patch area relationships in fragmented landscapes. *Journal of Animal Ecology*, 72(4):533–545.

Schultz, C. and Crone, E. (2001). Edge-mediated dispersal behavior in a prairie butterfly. *Ecology*, 82(7):1879–1892.

Sell, G. and You, Y. (2002). *Dynamics of evolutionary equations.* Applied Mathematical Sciences 143. Springer.

Shigesada, N., Kawasaki, K., and Teramoto, E. (1986). Traveling periodic waves in heterogeneous environments. *Theoretical Population Biology*, 30:143–160.

Slatkin, M. (1978). On the Equilibration of Fitnesses by Natural Selection. *The American Naturalist*, 112:845–859.

Turchin, P. (1998). *Quantitative Analysis of Movement: Measuring and Modeling Population Redistribution in Animals and Plants.* Sinauer Associates.

Weinberger, H., Lewis, M., and Li, B. (2002). Analysis of linear determinacy for spread in cooperative models. *Journal of Mathematical Biology*, 45:183–218.

Yurk, B. and Cobbold, C. (2018). Homogenization techniques for population dynamics in strongly heterogeneous landscapes. *Journal of Biological Dynamics*, 12:171–193.