ON THE EVOLUTION OF DISPERSAL IN
PATCHY LANDSCAPES[*]

STEPHEN KIRKLAND[†], CHI-KWONG LI[‡], AND SEBASTIAN J. SCHREIBER[‡]

Abstract. To better understand the evolution of dispersal in spatially heterogeneous landscapes, we study difference equation models of populations that reproduce and disperse in a landscape consisting of k patches. The connectivity of the patches and costs of dispersal are determined by a $k \times k$ column substochastic matrix S, where S_{ij} represents the fraction of dispersing individuals from patch j that end up in patch i. Given S, a dispersal strategy is a $k \times 1$ vector whose ith entry gives the probability p_i that individuals disperse from patch i. If all of the p_i's are the same, then the dispersal strategy is called unconditional; otherwise it is called conditional. For two competing populations of unconditional dispersers, we prove that the slower dispersing population (i.e., the population with the smaller dispersal probability) displaces the faster dispersing population. Alternatively, for populations of conditional dispersers without any dispersal costs (i.e., S is column stochastic and all patches can support a population), we prove that there is a one parameter family of strategies that resists invasion attempts by all other strategies.

1. Introduction. Plants and animals often live in landscapes where environmental conditions vary from patch to patch. Within patches, these environmental conditions may include abiotic factors such as light, space, and nutrient availability or biotic factors such as prey, competitors, and predators. Since the fecundity and survivorship of an individual depends on these factors, an organism may decrease or increase its fitness by dispersing across the environment. Depending on their physiology and their ability to accumulate information about the environment, plants and animals can exhibit two modes of dispersals and a variety of dispersal strategies. Plants and animals can be active dispersers that move by their own energy or passive dispersers that are moved by wind, water, or other animals. Passive dispersers alter their dispersal rates by varying the likelihood of dispersing and the time spent dispersing [20]. Dispersal strategies can vary from unconditional strategies in which the probability of dispersing from a patch is independent of the local environmental conditions to conditional strategies in which the likelihood of dispersing depends on local environmental factors. Understanding how natural selection acts on these different modes and strategies of dispersal has been the focus of much theoretical work [2, 5, 8, 10, 12, 15, 16, 17]. For instance, using coupled ordinary differential equation models for populations passively dispersing between two patches, Holt [8] showed that slower dispersing populations could always invade equilibria determined by faster

dispersing populations. Hastings [5] and Dockery et al. [2] considered evolution of dispersal in continuous space using reaction diffusion equations. Dockery et al. proved that for two competing populations differing only in their diffusion constant, the population with the larger diffusion constant is excluded. In contrast, McPeek and Holt [17] using a two patch model consisting of coupled difference equations found that "dispersal between patches can be favored in spatially varying but temporally constant environment, if organisms can express conditional dispersal strategies."

In this article, we consider the evolution of conditional and unconditional dispersers for a general class of multipatch difference equations. For these difference equations, individuals in each patch disperse with some probability. When these probabilities are independent of location, the population exhibits an unconditional dispersal strategy; otherwise it exhibits a conditional dispersal strategy. For dispersing individuals, the nature of the landscape determines the likelihood S_{ji} that a disperser from patch i ends up in patch j. Unlike previous studies of the evolution of unconditional and conditional dispersal [2, 5, 8, 17], we allow for an arbitrary number of patches and place no symmetry conditions on S. For active dispersers, asymmetries in S may correspond to geographical and ecological barriers that inhibit movement from one patch to another. For passive dispersers, these asymmetries may correspond to asymmetries in the abiotic or biotic currents in which they drift.

Our main goal is to determine what types of theorems can be proved about the evolution of dispersal for this general class of difference equation models. To achieve these goals, the remainder of the article is structured as follows. In section 2, we introduce the models. Under monotonicity assumptions about the growth rates, we prove that either populations playing a single dispersal strategy go extinct for all initial conditions or approach a positive fixed point for all positive initial conditions. We also introduce models of competing populations that differ only in their dispersal ability and prove a result about invasiveness. In section 3, we prove that for two competing populations of unconditional dispersers, the slower dispersing population displaces the faster dispersing population. The proof relies heavily on proving, in section 4, monotonicity of the principal eigenvalue for a one-parameter family of nonnegative matrices. In section 5, we prove that, provided there is no cost to dispersal and all patches can support a population, there is a one-parameter family of conditional dispersal strategies that resists invasion from other types of dispersal strategies. Numerical simulations suggest that these strategies can displace all other strategies, and we prove that these strategies can weakly coexist. In section 6, we discuss our findings and suggest directions for future research.

2. The models and basic results. Consider a population exhibiting discrete reproductive and dispersal events and living in an environment consisting of k patches. The vector of population densities is given by $x = (x_1, \ldots, x_k)^T \in \mathbf{R}_+^k$, where $\mathbf{R}_+^k$ is the nonnegative cone of $\mathbf{R}^k$. To describe reproduction and survival in each patch, let $\lambda_i : \mathbf{R}_+ \to \mathbf{R}_+$ denote the per-capita growth rate of the population in the ith patch as a function of the population density in the ith patch. For these per-capita growth rates we make the following assumptions.

A1: λ_i are positive continuous decreasing functions.

A2: $\lim_{x_i \to \infty} \lambda_i(x_i) < 1$.

A3: $x_i \mapsto x_i \lambda_i(x_i)$ is increasing.

Assumption A1 corresponds to the population exhibiting increasing levels of intra specific competition or interference as population densities increase. Assumption A2 implies that at high densities the population tends to decrease in size. Assumption A3

implies that the population does not exhibit overcompensating density dependence: higher densities in the current generation yield higher densities in the next generation. Many models in the population ecology literature satisfy these three assumptions. For instance, see the Beverton–Holt model [1] in which $\lambda_i(x_i) = \frac{a_i}{1+b_i x_i}$ and the Ivlev model [14] in which $\lambda_i(x_i) = a_i(1 - \exp(-b\,x_i))$.

To describe dispersal between patches, we assume that each individual in patch i disperses with a probability p_i and S_{ji} is the probability that a dispersing individual from patch i arrives in patch j. About the matrix S we make the following assumption.

A4: S is a $k \times k$ primitive column substochastic matrix.

S can be column stochastic if all dispersing individuals migrate successfully or substochastic if some dispersing individuals experience mortality. The primitive assumption ensures that individuals (possibly after several generations) can move from any patch to any patch. S characterizes how connected the landscape is for dispersing individuals. For example, for a fully connected metapopulation, S could be the matrix whose entries all equal $\frac{1}{k}$; i.e., an individual is equally likely to end up in any patch after dispersing. Alternatively, in a landscape with a one-dimensional lattice structure with individuals able only to move to neighboring patches in one time step S is a column substochastic tridiagonal matrix that is primitive, provided it has a positive entry on the diagonal. From these p and S, the following matrix describes how the population redistributes itself across the environment in one time step:

$$S_p = I - \operatorname{diag}(p) + S \operatorname{diag}(p),$$

where $\operatorname{diag}(p)$ denotes a diagonal matrix with diagonal entries $p_1, \ldots, p_k$.

If a census of the population is taken before reproduction and after dispersal, then the dynamics of the population are given by

$$x' = S_p \Lambda(x) x =: F(x), \tag{1}$$

where x' denotes the population state in the next time step and $\Lambda(x)$ is the $k \times k$ diagonal matrix whose ith diagonal entry equals $\lambda_i(x_i)$.

Our first result characterizes the global dynamics of (1). To state this result, let $F^n(x)$ denote F composed with itself n times. Given $x, y \in \mathbf{R}_+^k$, we write $x \geq y$ if $x_i \geq y_i$ for all $1 \leq i \leq k$, $x > y$ if $x \geq y$ and $x \neq y$, and $x \gg y$ if $x_i > y_i$ for all $1 \leq i \leq k$. For a matrix A, let $r(A)$ denote the spectral radius of A.

THEOREM 2.1. *Assume that Assumptions* A1–A4 *hold and* $p \in (0,1]^k$. *If* $r(S_p \Lambda(0)) \leq 1$, *then*

$$\lim_{n \to \infty} F^n(x) = 0$$

for all $x \geq 0$. *Alternatively, if* $r(S_p \Lambda(0)) > 1$, *then there exists a fixed point* $\hat{x} \gg 0$ *for* F *such that*

$$\lim_{n \to \infty} F^n(x) = \hat{x}$$

for all $x > 0$.

Proof. Let $A(x) = S_p \Lambda(x)$. Assumptions A1, A4, and $p \gg 0$ imply that $A(x)$ is primitive for all $x \geq 0$. Assumption A3 implies that $F(x) \geq F(y)$ (resp., $F(x) > F(y)$, $F(x) \gg F(y)$) whenever $x \geq y$ (resp., $x > y$, $x \gg y$). In other words, F is a strongly monotone map.

Suppose that $r(A(0)) \leq 1$. Let $w^T \gg 0$ be a left Perron vector of $A(0)$, i.e., $r(A(0))w^T = w^T A(0)$. Define the function $L : \mathbf{R}_+^k \to \mathbf{R}_+$ by $L(x) = w^T x$. For $x > 0$, Assumption A1 implies that $w^T A(0) \gg w^T A(x)$. Hence, for any $x > 0$,

$$
\begin{aligned}
L(F(x)) &= w^T A(F(x))x \\
&= w^T A(0)x + w^T(A(F(x)) - A(0))x \\
&< r(A(0))w^T x \leq L(x).
\end{aligned}
$$

Since L is strictly decreasing along nonzero orbits of F, $L(0) = 0$, and $L(x) > 0$ for $x > 0$, it follows that $\lim_{n \to \infty} F^n(x) = 0$ for all $x \geq 0$.

Suppose $r(A(0)) > 1$. First, we show that there exists a positive fixed point $\hat{x}$. Let $v \gg 0$ be a right Perron eigenvector for $A(0)$, i.e., $A(0)v = r(A(0))v$. Since $A(0)v \gg v$, continuity of $A(x)$ implies that there exists $\epsilon > 0$ such that $A(y)y \gg y$, where $y = \epsilon v$. Since $F(x) \gg F(y)$ whenever $x \gg y$, induction implies $y \ll F(y) \ll F^2(y) \ll F^3(y) \ll \cdots$. Assumption A2 implies that the increasing sequence $F^n(y)$ is bounded. Hence, there exists $\hat{x}$ such that $\lim_{n \to \infty} F^n(y) = \hat{x}$. Continuity of F implies that $F(\hat{x}) = \hat{x}$. Second, we show that $\lim_{n \to \infty} F^n(x) = \hat{x}$ whenever $\hat{x} > x > 0$. In particular, $\hat{x}$ is a unique positive fixed point. Let w^T be the left Perron eigenvector of $A(\hat{x})$ that satisfies $w^T \hat{x} = 1$. Since $\hat{x}$ is a positive fixed point, $r(A(\hat{x})) = 1$. Define $L : \mathbf{R}_+^k \to \mathbf{R}_+$ by $L(x) = w^T x$. Let $\hat{x} > x > 0$. Then $\hat{x} > F(x) > 0$ and

$$
\begin{aligned}
L(F(x)) &= w^T A(F(x))x \\
&= w^T A(\hat{x})x + w^T(A(F(x)) - A(\hat{x}))x \\
&> r(A(\hat{x}))w^T x = L(x).
\end{aligned}
$$

Hence, $L(x), L(F(x)), L(F^2(x)), \ldots$ is a positive increasing sequence bounded above by $L(\hat{x}) = 1$. Since $L(x) < 1$ for all $x < \hat{x}$, it follows that $\lim_{n \to \infty} F^n(x) = \hat{x}$ for all $0 < x < \hat{x}$. Third, it can be shown similarly that $\lim_{n \to \infty} F^n(x) = \hat{x}$ for all $x > \hat{x}$. Fourth, consider any $x \gg 0$. Choose $\overline{x} > x$ such that $\overline{x} > \hat{x}$ and choose $\underline{x} < x$ such that $0 < \underline{x} < \hat{x}$. Since $F^n(\underline{x}) < F^n(x) < F^n(\overline{x})$ for all n and $\lim_{n \to \infty} F^n(\underline{x}) = \lim_{n \to \infty} F^n(\overline{x}) = \hat{x}$, continuity of F implies that $\lim_{n \to \infty} F^n(x) = \hat{x}$. Finally, consider any $x > 0$. Assumptions A3–A4 imply that there exists $n \geq 1$ such that $F^n(x) \gg 0$. Hence, $\lim_{n \to \infty} F^n(x) = \hat{x}$. □

To understand the evolution of dispersal, we shall consider two populations that differ only in their dispersal ability. Let $x, y \in \mathbf{R}_+^k$ denote the vector of densities of the two populations and $p, \widetilde{p}$ denote their dispersal strategies. Since the populations differ only in their dispersal abilities, their dynamics are given by

$$
\begin{aligned}
(2) \qquad x' &= S_p \Lambda(x + y)x =: G_1(x, y), \\
y' &= S_{\widetilde{p}} \Lambda(x + y)y =: G_2(x, y).
\end{aligned}
$$

From Assumption A2 it follows that (2) is dissipative i.e., there exists a compact set K such that for any $(x, y) \geq (0, 0)$, $G^n(x, y) \in K$ for n sufficiently large. Regarding the dynamics of (2) near equilibria, we need the following result about invasiveness. Since we have not assumed that $G(x, y)$ is continuously differentiable, this result does not follow immediately from the standard unstable manifold theory.

PROPOSITION 2.2. *Assume that $p, \widetilde{p} \in (0,1]^k$, S and Λ satisfy Assumptions* A1–A4, *and $r(S_p\Lambda(0)) > 1$. Let $\hat{x} \gg 0$ be the fixed point satisfying $G_1(\hat{x}, 0) = (\hat{x}, 0)$. If $r(S_{\widetilde{p}}\Lambda(\hat{x})) > 1$, then there exists a neighborhood $U \subset \mathbf{R}_+^k \times \mathbf{R}_+^k$ of $(\hat{x}, 0)$ such that for any $(x, y) \in U$ with $y > 0$, $G^n(x, y) \notin U$ for some $n \geq 1$.*

Proof. Let $A(x) = S_{\widetilde{p}}\Lambda(x)$. Assume that $r(A(\hat{x})) > 1$. Let $w^T \gg 0$ be a left Perron eigenvector of $A(\hat{x})$. Since $w^T A(\hat{x}) \gg w^T$, continuity of $x \mapsto A(x)$ implies that there exists a compact neighborhood $U \subset \mathbf{R}_+^k \times \mathbf{R}_+^k$ of $(\hat{x}, 0)$ and $c > 1$ such that $w^T A(x + y) \gg cw^T$ for all $(x, y) \in U$. Define $L : \mathbf{R}_+^k \times \mathbf{R}_+^k \to \mathbf{R}_+$ by $L(x, y) = w^T y$. Let (x, y) be in U with $y > 0$. We have $L(G(x, y)) = w^T A_{\widetilde{p}}(x + y)y > cL(x, y)$. Hence, if $(x, y), \ldots, G^n(x, y) \in U$, then $L(G^n(x, y)) > c^n w^T y$. Since U is compact and $y > 0$, it follows that there exists $n \geq 1$ such that $G^n(x, y) \notin U$. □

3. The slower unconditional disperser wins. In this section, we consider only an unconditional dispersal strategy p: a strategy that satisfies $p_1 = \cdots = p_k$ for some common value d. Equivalently, $p = d\,\mathbf{1}$, where $\mathbf{1} = (1, \ldots, 1)$. Our key result is the following theorem concerning the monotonicity of the dominant eigenvalue with respect to the parameter d.

THEOREM 3.1. *Let S be an irreducible column substochastic matrix and Λ be a diagonal matrix. If Λ is not a scalar matrix, then $d \mapsto r(((1 - d)I + dS)\Lambda)$ is decreasing on $[0, 1]$.*

The proof of Theorem 3.1 is given in section 4, where we also characterize the function $d \mapsto r(S_{d\mathbf{1}})$ when S is reducible. The following corollary follows immediately from Theorems 2.1 and 3.1.

COROLLARY 3.2. *Assume that F, S, and $\Lambda(x)$ satisfy Assumptions* A1–A4 *and $p = d\,\mathbf{1}$. Then there exists $d^* \geq 0$ such that we have the following.*

Persistence: *If $d \in [0, d^*)$, then there exists $\hat{x} \gg 0$ satisfying $\lim_{n\to\infty} F^n(x) = \hat{x}$ for all $x \gg 0$.*

Extinction: *If $d \in [d^*, 1]$, then $\lim_{n\to\infty} F^n(x) = 0$ for all $x \geq 0$.*

Moreover, $d^* = 0$ if $\max_i \lambda_i(0) \leq 1$, $d^* \in (0, 1)$ if $\max_i \lambda_i(0) > 1$ and $r(S\Lambda(0)) < 1$, and $d^* \geq 1$ if $r(S\Lambda(0)) \geq 1$.

Corollary 3.2 implies that whenever $r(S\Lambda(0)) < 1$, unconditional dispersers have a critical dispersal rate below which the population persists and above which the population is deterministically driven to extinction.

To characterize the dynamics of competing unconditional dispersers, we need an additional assumption on (2) to avoid degenerate cases. Let $v \gg 0$ be a right Perron eigenvector of S, i.e., $Sv = r(S)v$. We make the following assumption.

A5: $\Lambda(tv)$ is not a scalar matrix for any $t \geq 0$.
This assumption assures that the model exhibits a minimal amount of spatial heterogeneity in the per-capita growth rates at fixed points.

THEOREM 3.3. *Let $G = (G_1, G_2)$ satisfy Assumptions* A1–A5. *Assume that $p = d\mathbf{1}$, and $\widetilde{p} = \widetilde{d}\mathbf{1}$, where $0 < d < \widetilde{d} \leq 1$. If $r(S_p\Lambda(0)) > 1$, then for all $x > 0$ and $y \geq 0$,*

$$\lim_{n\to\infty} G^n(x, y) = (\hat{x}, 0),$$

where $\hat{x}$ is the positive fixed point of $x \mapsto G_1(x, 0)$.

Theorem 3.3 implies that the slower disperser always displaces the faster disperser. This occurs despite the fact that the faster disperser is initially able to establish itself more rapidly, as illustrated in Figures 1 and 2.
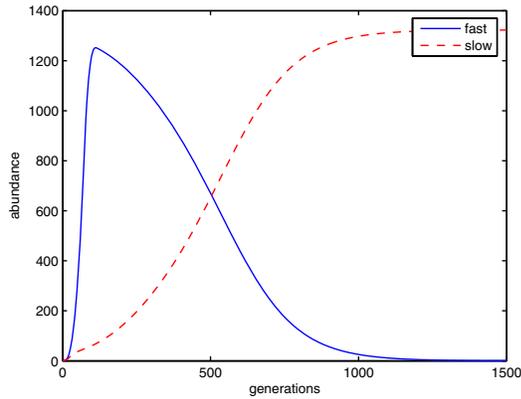
FIG. 1. *A simulation of* (2) *with* $k = 50 \times 50$ *(i.e., a two-dimensional spatial grid),* $\lambda_i(x_i) = \frac{a_i}{1+x_i}$ *with* a_i *randomly chosen from* $[1, 2]$, $d = 0.2$, $\widetilde{d} = 0.3$, *and* S *given by movement with equal likelihood to east, west, north, and south, and periodic boundary conditions. The initial condition corresponds to a density one of both populations in the center patch. The dotted and solid curves correspond to the abundances of the slower and faster dispersing populations, respectively.*

(a) slower disperser at generation 100

(b) faster disperser at generation 100

(c) slower disperser at generation 500

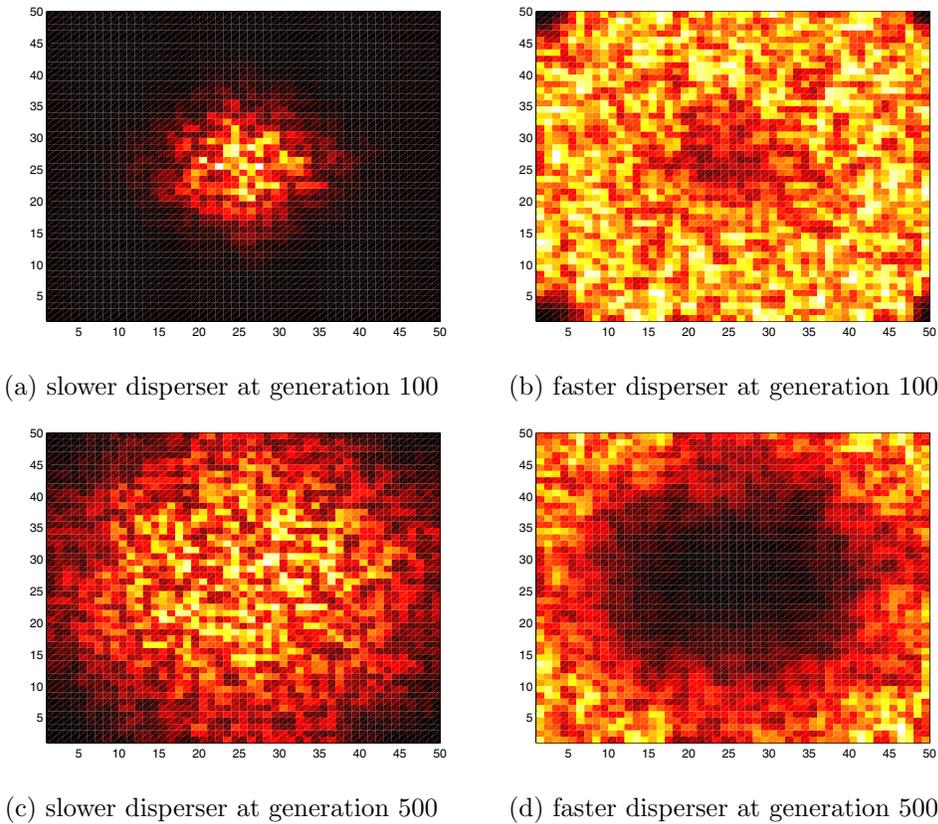(d) faster disperser at generation 500

FIG. 2. *Spatial distributions of the slower disperser in* (a) *and* (c) *and the faster dispersers in* (b) *and* (d). *The model and parameters are as in Figure* 1. *Darker (resp., lighter) shading correspond to lower (resp., higher) densities.*

Proof. The proof of this theorem relies on a result of Hsu, Smith, and Waltman [11, Theorem A] and Theorems 2.1 and 3.1. Let $A_d(x) = S_{d\mathbf{1}}(x)$. We start the proof with an important implication of Assumption A5. Suppose (x, y) satisfies $G(x, y) = (x, y)$. We claim that $\Lambda(x + y)$ is not a scalar matrix. Indeed, suppose to the contrary that $\Lambda(x + y) = tI$ for some $t > 0$. Then

$$x = S_p \Lambda(x + y)x = (1 - d)tx + dtSx,$$
$$y = S_{\widetilde{p}} \Lambda(x + y)y = (1 - \widetilde{d})ty + \widetilde{d}tSy.$$

Consequently, x and y (and hence $x + y$) are scalar multiples of v. Since this contradicts Assumption A5, $\Lambda(x + y)$ is not a scalar matrix for any fixed point (x, y) of G.

Assuming that $r(A_d(0)) > 1$, Theorem 2.1 implies that $x \mapsto G_1(x, 0)$ has a unique positive fixed point $\hat{x}$ that is globally stable. We prove the theorem in two cases. In the first case, assume that $r(A_{\widetilde{d}}(0)) > 1$. Theorem 2.1 implies that there is a unique $\hat{y} \gg 0$ such that $G(0, \hat{y}) = (0, \hat{y})$ and $\lim_{n \to \infty} G^n(0, y) = (0, \hat{y})$ whenever $y \gg 0$. To employ Theorem A in [11] we need to verify two things: G has no positive fixed point and $(0, \hat{y})$ is unstable. First, suppose to the contrary there exists $x \gg 0$ and $y \gg 0$ such that $G(x, y) = (x, y)$. Then $x = A_d(x + y)x$, $y = A_{\widetilde{d}}(x + y)y$, and $r(A_d(x + y)) = 1$. Since $\Lambda(x + y)$ is not a scalar matrix, Theorem 3.1 implies that $1 = r(A_d(x + y)) > r(A_{\widetilde{d}}(x + y)) = 1$. Hence, there can be no positive fixed point. Second, to show that $(0, \hat{y})$ is unstable, we use Theorem 3.1, which implies that $1 = r(A_{\widetilde{d}}(\hat{y})) < r(A_p(\hat{y}))$, and apply Proposition 2.2. Applying Theorem A of [11] implies that $\lim_{n \to \infty} G^n(x, y) = (\hat{x}, 0)$ whenever $x \gg 0$ and $y \gg 0$.

Suppose that $r(A_{\widetilde{d}}(0)) \leq 1$. Let $w^T \gg 0$ be a left Perron vector of $A_{\widetilde{d}}(0)$. Define the function $L : \mathbf{R}_+^k \to \mathbf{R}_+$ by $L(y) = w^T y$. Let $\pi(x, y) = y$. Since $L(G^n(x, y))$ is strictly decreasing whenever $y > 0$, $L(0) = 0$, and $L(y) > 0$ for $y > 0$, it follows that $\lim_{n \to \infty} \pi(G^n(x, y)) = 0$ for all $x \geq 0$. Hence, for any $(x, y) \in \mathbf{R}_+^k \times \mathbf{R}_+^k$, the limit points of $G^n(x, y)$ as $n \to \infty$ lie in $\mathbf{R}_+^k \times \{0\}$. By Theorem 1.8 in [18], the closure of these limit points form a connected chain recurrent set (see [18] for the definition). Since the only connected chain recurrent sets in $\mathbf{R}_+^k \times \{0\}$ are $(0, 0)$ and $(\hat{x}, 0)$, instability of $(0, 0)$ implies that $\lim_{n \to \infty} G^n(x, y) = (\hat{x}, 0)$ whenever $x > 0$. $\square$

4. Proof of Theorem 3.1. We begin with the following preliminary result.

LEMMA 4.1. *Let v and w^T be positive k-vectors so that $w^T v = 1$. Let $\mathbf{P}$ be the polytope of nonnegative matrices A such that $w^T A = w^T$ and $Av = v$. For each $A \in \mathbf{P}$, let D_A denote the diagonal matrix of column sums of A. Then*

$$\min \{w^T D_A v | A \in \mathbf{P}\} = 1.$$

A matrix $A \in \mathbf{P}$ attains the minimum value for $w^T D_A v$ if and only if $D_A = I$.

Proof. Without loss of generality, assume that $w^T = (w_1, \ldots, w_k)$ is such that $w_1 \leq \cdots \leq w_k$. Note also that if all of the entries in w^T are equal, then each matrix in $\mathbf{P}$ is a column stochastic matrix, and the statement of the lemma follows immediately. We suppose henceforth that w^T has at least two distinct entries.

Suppose that $A \in \mathbf{P}$ and that there are indices i, j, p, q satisfying the following conditions:

(3) $w_i < w_j, w_p < w_q, \quad \text{and} \quad a_{ip}, a_{jq} > 0.$

We claim that in this case, the matrix A does not satisfy

(4) $$w^T D_A v \leq w^T D_B v \quad \text{for all } B \in \mathbf{P}.$$

To see the claim, note that from (3), it follows that for sufficiently small $\epsilon > 0$, the matrix

$$\hat{A} = A + \epsilon(-e_i/w_i + e_j/w_j)(e_p/v_p - e_q/v_q)^T$$

is nonnegative, and satisfies $w^T \hat{A} = w^T$ and $\hat{A}v = v$, so that $\hat{A} \in \mathbf{P}$. Further,

$$D_{\hat{A}} = D_A + \epsilon \frac{w_j - w_i}{w_i w_j} \operatorname{diag}\left(\frac{-e_p}{v_p} + \frac{e_q}{v_q}\right),$$

so that

$$w^T D_{\hat{A}} v = w^T D_A v - \epsilon \frac{(w_j - w_i)(w_p - w_q)}{w_i w_j} < w^T D_A v.$$

Thus $w^T D_A v$ does not yield the minimum, as claimed.

Suppose the minimum entry in w is repeated a times, i.e., $w_1 = \cdots = w_a < w_{a+1}$. Partition out the first a entries of w^T, to write w^T as $[w_1 \mathbf{1}^T | \tilde{w}^T]$, and partition v conformally as

$$v = \left[\frac{\hat{v}}{\tilde{v}}\right].$$

Let $A \in \mathbf{P}$ satisfy (4). Suppose first that there are indices i and p with $1 \leq i \leq a$ and $a + 1 \leq p$, such that $a_{ip} > 0$. Since A is a minimizer, we see from the claim above that for any indices j, q with $j \geq a + 1$ and $1 \leq q \leq a$, we must have $a_{jq} = 0$. But then A has the form

$$A = \left[\begin{array}{c|c} A_1 & X \\ \hline 0 & A_2 \end{array}\right],$$

where A_1 is $a \times a$. From the facts that $w^T A = w^T$ and that the first a entries of w^T are equal and the partitioned form for A, we find that $\mathbf{1}^T A_1 = \mathbf{1}^T$. Also, $A_1 \hat{v} + X \tilde{v} = \hat{v}$, so that $\mathbf{1}^T (A_1 \hat{v} + X \tilde{v}) = \mathbf{1}^T \hat{v}$. Since $\mathbf{1}^T A_1 = \mathbf{1}^T$, we conclude that $X = 0$, a contradiction.

Consequently, we conclude that for any indices i and p with $1 \leq i \leq a$ and $a + 1 \leq p$, we must have $a_{ip} = 0$. Thus we see that A has the form

$$A = \left[\begin{array}{c|c} A_1 & 0 \\ \hline Y & A_2 \end{array}\right],$$

where A_1 is $a \times a$ and $A_1 \hat{v} = \hat{v}$. Using the fact that $w^T A = w^T$, we thus find that $w_1 \mathbf{1}^T A_1 + \tilde{w}^T Y = w_1 \mathbf{1}^T$. Hence we have $w_1 \mathbf{1}^T A_1 \hat{v} + \tilde{w}^T Y \hat{v} = w_1 \mathbf{1}^T \hat{v}$, from which we deduce that $Y = 0$.

We conclude that if $A \in \mathbf{P}$ satisfies (4), then A can be written as

$$\left[\begin{array}{c|c} A_1 & 0 \\ \hline 0 & A_2 \end{array}\right],$$

where A_1 is column stochastic. The lemma is now readily established by a deflation argument. $\square$

Our next result lends some insight into the irreducible case.

LEMMA 4.2. *Suppose that A is an irreducible nonnegative matrix, and let D_A be the diagonal matrix of column sums of A. Let Λ be a diagonal matrix such that $\Lambda \geq D_A$. For each $d \in [0,1]$ let $h(d) = r((1-d)\Lambda + dA)$. Then for any $d \in (0,1), h'(d) \leq 0$, with equality holding if and only if $\Lambda = D_A = aI$ for some $a > 0$. In that case, $h(d) = r(A) = a$ for each $d \in [0,1]$.*

Proof. Throughout, we suppose without loss of generality that $r(A) = 1$.

First, suppose that A is a primitive matrix; we claim that in this case, $h'(1) \leq 0$ with equality holding if and only if $\Lambda = D_A = I$. Let v be a right Perron vector for A. Since A is primitive, its spectral radius is a simple eigenvalue that strictly dominates the modulus of any other eigenvalue; it follows that in a sufficiently small neighborhood of 1, $h(d)$ is an eigenvalue of $(1-d)\Lambda + dA$ that is differentiable in d. For d in such a neighborhood of 1, let $w(d)^T$ be a left $h(d)$-eigenvector of $(1-d)\Lambda + dA$, normalized so that $w(d)^T v = 1$. Since $Av = v$, we have

$$h(d) = w(d)^T((1-d)\Lambda + dA)v$$
$$= (d-1)(w(d)^T(A - \Lambda)v) + w(d)^T Av$$
$$= (d-1)(1 - w(d)^T \Lambda v) + 1.$$

Since $\lim_{d \to 1} w(d)^T = w^T$, it follows that

$$\lim_{d \to 1} \frac{h(d) - h(1)}{d - 1} = \lim_{d \to 1}(1 - w(d)^T \Lambda v) = 1 - w^T \Lambda v$$
$$= -(w^T D_A v - 1) - (w^T(\Lambda - D_A)v).$$

Since $\Lambda \geq D_A$, we have $w^T(\Lambda - D_A)v \geq 0$, and by Lemma 4.1, we have $w^T D_A v - 1 \geq 0$, so certainly $h'(1) \leq 0$. Further, we see that $h'(1) = 0$ if and only if $w^T D_A v = 1$ and $w^T(\Lambda - D_A)v = 0$. It now follows from Lemma 4.1 that the former holds if and only if $D_A = I$, and since w^T and v are positive vectors, we see that the latter holds if and only if $\Lambda = D_A$. This completes the proof of the claim.

Next, suppose that A is an irreducible nonnegative matrix, and fix $d \in (0,1)$. Observe that the matrix $B = (1-d)\Lambda + dA$ is primitive and that $\Lambda \geq D_B$. For each $c \in [0,1]$, let $k(c) = r((1-c)\Lambda + cB)$, and note that $k(c) = h(cd)$. Applying the claim above to the function k, we see that $k'(1) \leq 0$, with equality holding if and only if $\Lambda = D_B = I$. But from the chain rule, we find that $k'(1) = dh'(d)$, so that $h'(d) \leq 0$, with equality if and only if $\Lambda = D_B = I$. That last condition is readily seen to be equivalent to $\Lambda = D_A = I$.

Finally, we note that if $\Lambda = D_A = I$, it is straightforward to see that for each $d \in [0,1]$, the matrix $(1-d)\Lambda + dA$ is column stochastic, so that $h(d) = 1 = r(A)$ for all such d. $\square$

The following, which evidently yields Theorem 3.1 immediately, follows from Lemma 4.2.

COROLLARY 4.3. *Suppose that A is an irreducible nonnegative matrix, and let D_A be the diagonal matrix of column sums of A. Let Λ be a diagonal matrix such that $\Lambda \geq D_A$. For each $d \in [0,1]$ let $h(d) = r((1-d)\Lambda + dA)$. Then either*

(a) *$h(d)$ is a strictly decreasing function of $d \in [0,1]$ or*

(b) *for some $a > 0, \Lambda = D_A = aI$ and $h(d) = a$ for each $d \in [0,1]$.*

We have the following generalization of Corollary 4.3.

THEOREM 4.4. *Let S be a column substochastic matrix and Λ be a diagonal matrix with positive diagonal entries. Define the function $f(d) = r(((1-d)I + dS)\Lambda)$ for $d \in [0,1]$. Then there is a $\hat{d} \in [0,1]$ such that f is strictly decreasing on $[0,\hat{d}]$ and f is constant on $[\hat{d},1]$. Specifically, let P be a permutation matrix such that*

$$
P^T S P = \begin{bmatrix} S_1 & 0 & \dots & 0 & X_1 \\ 0 & S_2 & \dots & 0 & X_2 \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & \dots & 0 & S_k & X_k \\ \hline 0 & 0 & \dots & 0 & S_{k+1} \end{bmatrix}, \quad and \quad P^T \Lambda P = \begin{bmatrix} \Lambda_1 & & & \\ & \Lambda_2 & & \\ & & \ddots & \\ & & & \Lambda_{k+1} \end{bmatrix},
$$

where (i) $P^T S P$ *and* $P^T \Lambda P$ *are partitioned conformally,* (ii) *for each $i = 1,\dots,k, S_i$ is an irreducible column stochastic matrix, and* (iii) S_{k+1} *is a column substochastic matrix such that $r(S_{k+1}) < 1$. (Note that such a permutation matrix P exists and that one part of this partitioning of $P^T S P$ may be vacuous.) Let $r(\Lambda) = \rho$. Exactly one of the following cases holds.*

(a) *For some $i = 1,\dots,k, \Lambda_i = \rho I$. In this case, $f(d) = \rho$ for all $d \in [0,1]$.*

(b) *There is an index $i_0 = 1,\dots,k$ and an $a < \rho$ such that $\Lambda_{i_0} = aI$ and in addition, for each $j = 1,\dots,k+1$, we have that either $r(S_j \Lambda_j) < a$ or $r(((1-d)I + dS_j)\Lambda_j) = a$ for all $d \in [0,1]$. In this case, there is a $\hat{d} \in (0,1)$ such that $f(d)$ is a strictly decreasing function of d for $d \in [0,\hat{d}]$, while for each $d \in [\hat{d},1], f(d) = a$.*

(c) *If $\Lambda_i \neq \rho I$ for $i = 1,\dots,k$ and there is no index i_0 and value a satisfying the hypotheses of* (b), *then $f(d)$ is strictly decreasing for $d \in [0,1]$.*

Proof. Throughout, we assume without loss of generality that $\rho = 1$. First, note that $f(d) = \max\{r(((1-d)I + dS_i)\Lambda_i) : i = 1,\dots,k+1\}$. Further, since $r(S_{k+1}) < 1$ it follows that no principal submatrix of S_{k+1} (including the entire matrix S_{k+1} itself) can have all of its column sums equal to 1; we then deduce from Corollary 4.3 that $r(((1-d)I + S_{k+1})\Lambda_{k+1})$ is strictly decreasing as a function of $d \in [0,1]$. Note further that if none of $\Lambda_1,\dots,\Lambda_k$ is a scalar matrix, then for each $i = 1,\dots,k$ the function $r(((1-d)I + dS_i)\Lambda_i)$ is strictly decreasing in d, from which we conclude that $f(d)$ is strictly decreasing.

Suppose next that for some $i = 1,\dots,k$, we have $\Lambda_i = I$. From Corollary 4.3 we see that $r(((1-d)I + dS_i)\Lambda_i) = 1$ for all $d \in [0,1]$, and we conclude readily that $f(d) = 1$ for all $d \in [0,1]$.

It remains only to consider the case that $\Lambda_i \neq I$ for $i = 1,\dots,k$ but that for one or more indices $i = 1,\dots,k, \Lambda_i$ is a scalar matrix. For concreteness, we suppose that $\Lambda_i = a_i I$ for $i = 1,\dots,j$ and that for $i = j+1,\dots,k, \Lambda_i$ is not a multiple of the identity matrix. Again without loss of generality, we can assume that $1 > a_1 \geq \cdots \geq a_j$. In this situation, we find that for each $i = 1,\dots,j, r(((1-d)I+dS_i)\Lambda_i) = a_i$, while for each $i = j+1,\dots,k+1, r(((1-d)I + dS_i)\Lambda_i)$ is a strictly decreasing function of d. It follows from the above considerations that $f(d) = \max\{a_1, r(((1-d)I + dS_{j+1})\Lambda_{j+1}),\dots,r(((1-d)I+dS_{k+1})\Lambda_{k+1})\}$.

Evidently two cases arise: either $\max\{r(S_{j+1}\Lambda_{j+1}),\dots,r(S_{k+1}\Lambda_{k+1})\} \geq a_1$ or $\max\{r(S_{j+1}\Lambda_{j+1}),\dots,r(S_{k+1}\Lambda_{k+1})\} < a_1$. In the former case we see that in fact $f(d) = \max\{r(((1-d)I+dS_{j+1})\Lambda_{j+1}),\dots,r(((1-d)I+dS_{k+1})\Lambda_{k+1})\}$ for all $d \in [0,1]$, from which we conclude that f is strictly decreasing in d. Now suppose that the latter

case holds. Since $a_1 < 1$, we see that when d is near to 0, $f(d) = \max\{r(((1-d)I + dS_{j+1})\Lambda_{j+1}), \dots, r(((1-d)I + dS_{k+1})\Lambda_{k+1})\} > a_1$. Thus, from the intermediate value theorem it follows that there is a value $\hat{d} \in (0,1)$ such that $\max\{r(((1-d)I + dS_{j+1})\Lambda_{j+1}), \dots, r(((1-d)I + dS_{k+1})\Lambda_{k+1})\} \geq a_1$ for $d \in [0,\hat{d}]$ and $\max\{r(((1-d)I + dS_{j+1})\Lambda_{j+1}), \dots, r(((1-d)I + dS_{k+1})\Lambda_{k+1})\} < a_1$ for $d \in [\hat{d}, 1]$. It now follows that $f(d)$ is strictly decreasing for $d \in [0,\hat{d}]$ and $f(d) = a_1$ for $d \in [\hat{d}, 1]$. $\qquad\square$

5. Competing conditional dispersers. In this section, we extend our study to conditional dispersers in which p need not be a constant vector. The following theorem coupled with Proposition 2.2 indicates which dispersal strategies are subject to invasion by other dispersal strategies.

THEOREM 5.1. *Assume that $\Lambda(x)$ and S satisfy Assumptions A1–A4, $p \in (0,1]^k$, and $r(S_p\Lambda(0)) > 1$. Let $\hat{x} \gg 0$ be the unique positive fixed point of F, and let $v \gg 0$ be a right Perron vector for S. Then $r(S_{\widetilde{p}}\Lambda(\hat{x})) \leq 1$ for all $\widetilde{p} \in (0,1]^k$ if and only if $\lambda_i(0) > 1$ for all i, S is column stochastic, and*

$$(5) \qquad p = t\left(\Lambda^{-1}(I)\right)^{-1} v$$

for some $t \in (0, 1/\max\{\Lambda^{-1}(I)^{-1}v\}]$. Moreover, if p is given by (5), then $\Lambda(\hat{x}) = I$.

In our proof of Theorem 5.1, we show that if either S is strictly substochastic or p is not given by (5), then there are strategies $\widetilde{p}$ arbitrarily close to p that can invade, i.e., $r(S_{\widetilde{p}}\Lambda(\hat{x})) > 1$. When S is stochastic and p is given by (5), we also show that $\Lambda(\hat{x}) = I$ and, consequently, $r(S_{\widetilde{d}}\Lambda(\hat{x})) = 1$ for all $\widetilde{p} \in [0,1]^k$. The populations playing one of these strategies exhibit an *ideal-free distribution at equilibrium* [3]; i.e., the per-capita fitness in all occupied patches are equal. Theorem 5.1 suggests the possibility that strategies of the form (5) can displace all other strategies. By [11, Theorem A] a sufficient condition for this displacement is verifying that (5) can invade any strategy $\widetilde{p}$ not given by (5) and cannot coexist at equilibrium with strategy $\widetilde{d}$. This turns out not to be true in general. For example, let $\lambda_i(x_i)$ with $i = 1, 2$ be functions such that $\lambda_1(1.2) = \lambda_2(1) = 1$, $\lambda_1(1.19) = \frac{20}{9+\sqrt{41}} \approx 1.29844$, $\lambda_2(9.52/(3 + \sqrt{41})) = \frac{10}{9+\sqrt{41}} \approx 0.642919$, where $9.52/(3 + \sqrt{41}) \approx 1.01234$, and Assumptions A1–A3 are satisfied. Define

$$S = \begin{pmatrix} 0.5 & 0.6 \\ 0.5 & 0.4 \end{pmatrix},$$

which has right Perron vector

$$v = \begin{pmatrix} 1 \\ 5/6 \end{pmatrix}.$$

Then $p = \mathbf{1}$ is a strategy of the form (5). Define

$$\widetilde{p} = \begin{pmatrix} 0.8 \\ 2/3 \end{pmatrix}.$$

The unique positive fixed point of $y \mapsto S_{\widetilde{p}}\Lambda(y)y = G(0, y)$ is by construction given by

$$\hat{y} = \begin{pmatrix} 1.19 \\ \frac{9.52}{3+\sqrt{41}} \end{pmatrix}.$$

Since a computation reveals that

$$r(S\Lambda(\widehat{y})) = 0.993735\ldots < 1 = r(S_{\widetilde{d}}\Lambda(\widehat{y})),$$

the strategy $p = \mathbf{1}$ cannot invade and displace the strategy $\widetilde{p}$. Hence, for a general $\Lambda(x)$, we cannot expect that strategies of the form (5) will displace all other strategies. However, extensive simulations with the Beverton–Holt growth functions (i.e., $\lambda_i(x_i) = \frac{a_i}{1+b_i x_i}$) suggest that the strategies given by (5) can displace any other strategy (see Figure 3). Thus we make the following conjecture.

CONJECTURE 5.1. *If $\lambda_i(x_i) = \frac{a_i}{1+b_i x_i}$, S is primitive and column stochastic, p is given by (5), $\widetilde{p}$ is not given by (5), and $r(S_p\Lambda(0)) > 1$, then*

$$\lim_{n\to\infty} G^n(x,y) = (\hat{x}, 0)$$

whenever $x > 0$.

Proof of Theorem 5.1 The key proposition (which gives us more than we need) is the following.

PROPOSITION 5.2. *Suppose that A is an irreducible nonnegative matrix with column sums c_i such that $c_1 = \min_i c_i < \max_i c_i = c_k$. If $\widetilde{A}$ is a nonnegative matrix obtained from A by changing its first column from*

$$\begin{pmatrix} a_{11} \\ \cdot \\ \vdots \\ a_{k1} \end{pmatrix} \quad to \quad \begin{pmatrix} a_{11} \\ \cdot \\ \vdots \\ a_{k1} \end{pmatrix} + \gamma \begin{pmatrix} -\sum_{i=2}^{k} a_{i1} \\ a_{21} \\ \ldots \\ a_{k1} \end{pmatrix}$$

for some positive $\gamma > 0$, then $r(A) < r(\tilde{A})$. Alternatively, if $\hat{A}$ is a nonnegative matrix obtained from A by changing its last column from

$$\begin{pmatrix} a_{1k} \\ \cdot \\ \vdots \\ a_{kk} \end{pmatrix} \quad to \quad \begin{pmatrix} a_{1k} \\ \cdot \\ \vdots \\ a_{kk} \end{pmatrix} - \gamma \begin{pmatrix} -\sum_{i=2}^{k} a_{ik} \\ a_{2k} \\ \ldots \\ a_{kk} \end{pmatrix}$$

for some $\gamma \in (0,1]$, then $r(A) > r(\hat{A})$.

Proof. Note that $c_k > r(A) > c_1$. Let w^T be the left Perron vector for A such that $w_1 = 1$, and let $\tilde{v}$ be the right Perron vector for $\tilde{A}$ normalized so that $w^T\tilde{v} = 1$. Observe that for any γ such that $\tilde{A}$ is nonnegative, $\tilde{A}$ is irreducible and, consequently, v is a positive vector. Set $W = \text{diag}\,(w_1, \ldots, w_n)$. Then WAW^{-1} has all the column sums equal to $r(A)$. Consider the first column of WAW^{-1}. We see that

$$a_{11} + \sum_{i=2}^{k} w_i a_{i1} = r(A) > c_1 = \sum_{i=1}^{k} a_{i1}.$$

Thus,

$$\sum_{i=2}^{k} w_i a_{i1} > \sum_{i=2}^{k} a_{i1}.$$

It follows that $r(\tilde{A}) = w^T\tilde{A}\tilde{v} = w^T A\tilde{v} + \gamma\tilde{v}_1(-\sum_{i=2}^{k} a_{i1} + \sum_{i=2}^{k} w_i a_{i1}) > w^T A\tilde{v} = r(A)$.
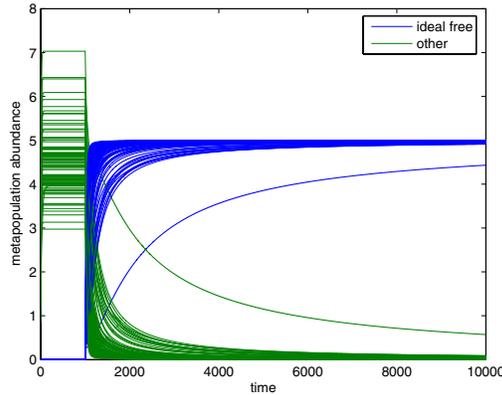
FIG. 3. *One hundred realizations of an ideal-free disperser competing against a random dispersal strategy. In the simulations, $k = 10$ and $\lambda_i(x_i) = \frac{a_i}{1+b_i x_i}$. For each simulation, the values of a_i are randomly selected from the interval $[1, 2]$, p is defined by (5), where t is randomly selected from the interval $[0, \max\left(\Lambda^{-1}(I)\right)^{-1} v]$, and $\widetilde{p}$ is randomly selected from $[0, 1]^{10}$. To normalize the local population abundances to a value of 1, in each simulation b_i is set equal to $\frac{1}{a_i - 1}$.*

A similar argument applies to the matrix $\hat{A}$ when $\gamma < 1$, while if $\gamma = 1$, we see that the first column of $\hat{A}$ is $c_k e_k$ and $r(\hat{A}) \geq c_k > r(A)$. $\square$

Now assume $p \in (0, 1]^k$, $r(S_p \Lambda(0)) > 1$, $\hat{x} \gg 0$ is the unique positive fixed point of F, and $v \gg 0$ is a right Perron vector for S. Let $A = S_p \Lambda(\hat{x})$. We begin by showing that $r(S_{\widetilde{p}} \Lambda(\hat{x})) \leq 1$ for all $\widetilde{p} \in [0, 1]^k$ implies that S is stochastic and p is given by (5). First, we show that A must have constant column sums c_i. Suppose to the contrary that there exists $1 \leq j \leq k$ such that $c_j = \max_i c_i > \min_i c_i$. Let $\widetilde{p}$ be any strategy where $\widetilde{p}_i = p_i$ for $i \neq j$ and $\widetilde{p}_j \in (0, p_j)$. Then $S_{\widetilde{p}} \Lambda(\hat{x})$ is given by replacing the jth column of A by a column which is

$$
\geq \begin{pmatrix} a_{1j} \\ . \\ \vdots \\ a_{kj} \end{pmatrix} - \gamma \begin{pmatrix} -\sum_{i=2}^{k} a_{ij} \\ a_{2j} \\ \cdots \\ a_{kj} \end{pmatrix},
$$

where $\gamma = 1 - \frac{\widetilde{p}_j}{p_j} > 0$. Proposition 5.2 implies that $r(S_{\widetilde{p}} \Lambda(\hat{x})) > r(A) = 1$, contrary to our assumption about p. Therefore A must have constant column sums $c = c_1 = \cdots = c_k$. Second, suppose to the contrary that S is substochastic. Let $\widetilde{p}$ be any strategy where $\widetilde{p}_i \in (0, p_i)$. Since S is substochastic, every column sum $S_{\widetilde{p}} \Lambda(\hat{x})$ is greater than or equal to c and at least one column sum is strictly greater than c. Hence, $r(S_{\widetilde{p}} \Lambda(\hat{x})) > r(A) = 1$, contrary to our assumption about p. Therefore, S is stochastic. Finally, since S is stochastic, it follows that $c = 1$ and $\Lambda(\hat{x}) = I$. Since $\hat{x} \gg 0$, we have $\lambda_i(0) > 1$ and $\hat{x}_i = \lambda_i^{-1}(1)$ for all i. Since $\hat{x}$ is a fixed point, we get that $\hat{x} = (I - \mathrm{diag}(p) + S\,\mathrm{diag}(p))\hat{x}$. Equivalently, $S\,\mathrm{diag}(p)\hat{x} = \mathrm{diag}(p)\hat{x}$. Hence, $\mathrm{diag}(p)\hat{x} \gg 0$ is a right Perron vector for S and p is given by (5).

Now suppose that S is stochastic and p is given by (5). Then $\Lambda(\hat{x}) = I$ and $r(S_{\widetilde{p}} \Lambda(\hat{x})) = r(S_{\widetilde{p}}) = 1$ for all $\widetilde{p} \in [0, 1]^k$. $\square$

Conjecture 5.1 suggests that for populations with Beverton–Holt local dynamics, the evolution of conditional dispersers will favor strategies on the ray defined by (5).

Hence, it is natural to ask what happens when two strategies on this ray compete against one another.

PROPOSITION 5.3. *Assume that $\Lambda(x)$ and S satisfy Assumptions* A1–A4, $\lambda_i(0) > 1$ *for all i, and S is stochastic. Let p and $\widetilde{p}$ be strategies given by* (5) *with $t = d$ and $t = \widetilde{d}$, where $0 < d < \widetilde{d} \le 1/\max\left\{\left(\Lambda^{-1}(I)\right)^{-1}v\right\}$. Then the set of fixed points of G are $(0,0)$ and*

$$L = \{(\alpha\hat{x}, (1-\alpha)\hat{x} : \alpha \in [0,1]\},$$

where $\hat{x} = \Lambda^{-1}(I)\mathbf{1}$. Moreover, if $\Lambda(x)$ is continuously differentiable with $\lambda_i'(x_i) < 0$ for all i, and $\frac{d}{dx_i}x_i\lambda_i(x_i) > 0$ for all i, then there exists a neighborhood $U \subset \mathbf{R}_+^k \times \mathbf{R}_+^k$ of L and a homeomorphism $h : [0,1] \times D \to U$ with $D = \{z \in \mathbf{R}^{2k-1} : \|z\| < 1\}$ such that $h(\alpha, 0) = \alpha\hat{x} + (1-\alpha)\hat{x}$, $h(0,D) = \{(0,y) \in U\}$, $h(1,D) = \{(x,0) \in U\}$, and $\lim_{n\to\infty} G^n(x,y) = (\alpha\hat{x}, (1-\alpha)\hat{x})$ for all $(x,y) \in h(\{\alpha\} \times D)$.

Proof. By the change of variables $x \mapsto \Lambda^{-1}(I)^{-1}\mathrm{diag}\,(v)x$, we can assume without any loss of generality that $p = d\mathbf{1}$ and $\widetilde{p} = \widetilde{d}\mathbf{1}$. Thus, a point $(x,y) > 0$ is a fixed point of G if and only if

$$((1-d)I + dS)\Lambda(x+y)x = x,$$
$$((1-\widetilde{d})I + \widetilde{d}S)\Lambda(x+y)y = y.$$

Since $r(((1-d)I + dS)\Lambda(x+y))) = r((1-\widetilde{d})I + \widetilde{d}S)\Lambda(x+y)) = 1$ and $d \ne \widetilde{d}$, Theorem 3.1 implies that $\Lambda(x+y) = I$. Therefore, (x,y) needs to satisfy $x + y = \Lambda^{-1}(I)\mathbf{1}$, $Sdx = dx$, and $S\widetilde{d}y = \widetilde{d}y$. Since S is primitive, we get that x must be a scalar multiple of y. Hence, the fixed points of G are given by $(0,0)$ and L.

Now assume that $x \mapsto \Lambda(x)$ is continuously differentiable, $\lambda_i'(x) < 0$ for all i, and $\frac{d}{dx_i}x_i\lambda_i(x_i) > 0$ for all i. We will show that L is a normally hyperbolic attractor in the sense of Hirsch, Pugh, and Shub [7]. Let $(x,y) \in L$. We have

$$DG(x,y) = \begin{pmatrix} S_d(\Lambda'(x+y)\mathrm{diag}\,(x) + \Lambda(x+y)) & S_d\Lambda'(x+y)\mathrm{diag}\,(x) \\ S_d\Lambda'(x+y)\mathrm{diag}\,(y) & S_d(\Lambda'(x+y)\mathrm{diag}\,(y) + \Lambda(x+y)) \end{pmatrix}.$$

Since $0 < \lambda_i'(x_i + y_i)(x_i + y_i) + \lambda_i(x_i + y_i) < \lambda_i'(x_i + y_i)x_i + \lambda_i(x_i + y_i)$ for all i, the diagonal blocks, $S_d(\Lambda'(x+y)\mathrm{diag}\,(x) + \Lambda(x+y))$ and $S_d(\Lambda'(x+y)\mathrm{diag}\,(y) + \Lambda(x+y))$ of $DG(x,y)$, are nonnegative primitive matrices. Since $\lambda_i'(x_i+y_i) < 0$ for all i, the off-diagonal blocks, $S_d\Lambda'(x+y)\mathrm{diag}\,(x)$ and $S_d\Lambda'(x+y)\mathrm{diag}\,(y)$, of $DG(x,y)$ are negative scalar multiples of primitive matrices. Hence, $DG(x,y)$ is a primitive matrix with respect to the competitive ordering on $\mathbf{R}_+^k \times \mathbf{R}_+^k$; i.e., $(\tilde{x}, \tilde{y}) \ge_K (x,y)$ if $\tilde{x} \ge x$ and $\tilde{y} \le y$. Since L is a line of fixed points, $DG(x,y)$ has an eigenvalue of one associated with the eigenvector $(\Lambda^{-1}(I)\mathbf{1}, -\Lambda^{-1}(I)\mathbf{1})$. The Perron–Frobenius theorem implies that all the other eigenvalues of $DG(x,y)$ are strictly less than one in absolute value. Hence, L is a normally hyperbolic one-dimensional attractor. Theorem 4.1 of [7] implies that there is a neighborhood $U \subset \mathbf{R}_+^k \times \mathbf{R}_+^k$ of L and a homeomorphism $h : [0,1] \times D \to U$ with $D = \{z \in \mathbf{R}^{2k-1} : \|z\| < 1\}$ such that $h(\alpha, 0) = \alpha\hat{x} + (1-\alpha)\hat{x}$, $h(0,D) = \{(0,y) \in U\}$, $h(1,D) = \{(x,0) \in U\}$, and $\lim_{n\to\infty} G^n(x,y) = (\alpha\hat{x}, (1-\alpha)\hat{x})$ for all $(x,y) \in h(\{\alpha\} \times D)$. $\quad\square$

Proposition 5.3 implies that once a "resident" population playing a strategy of the form (5) has established itself, a "mutant" strategy of the form (5) can invade only in a weak sense: if the mutants enter at low density, deterministically they will converge to an equilibrium with a low mutant density. After the invasion, one would

expect that demographic or environmental stochasticity would with greater likelihood result in the displacement of the mutants. Hence, once a strategy of the form (5) has established itself, it is likely to resist invasion attempts from other strategies of the form (5). Proposition 5.3 also suggests the following conjecture, which is supported by simulations using the Beverton–Holt growth function.

CONJECTURE 5.2. *Under the conditions of Proposition* 5.3, *for every* $(x, y) > 0$ *there exists* $\alpha \in [0, 1]$ *such that*

$$\lim_{n \to \infty} G^n(x, y) = (\alpha \Lambda^{-1}(I)\mathbf{1}, (1 - \alpha)\Lambda^{-1}(I)\mathbf{1}).$$

6. Discussion. For organisms that disperse unconditionally, we proved that a slower dispersing population competitively excludes a faster dispersing population. Similar results have been proven for reaction diffusion equations where the dispersal kernel is self-adjoint [2], observed in a partial analysis of two patch differential equations [8] and illustrated with simulations of two patch difference equations [17]. Our proofs apply to difference equations with an arbitrary number of patches and without any symmetry assumptions about the dispersal matrix S. Since geographical and ecological barriers often create asymmetries in the movement patterns of active dispersers and create asymmetries in abiotic and biotic currents that carry passive dispersers, accounting for these asymmetries is crucial and results in a significantly more difficult mathematical problem than the symmetric case. Theorem 3.1 provides the solution to this problem by proving for any given environmental condition (i.e., the choice of Λ and S), the principal eigenvalue for the growth dispersal matrix is a decreasing function of the dispersal rate. Hence, under all environmental conditions, populations that disperse more slowly spectrally dominate populations that disperse more quickly. Despite this spectral dominance, simulations (e.g., Figure 1) illustrate that for appropriate initial conditions, faster dispersers can be numerically dominant as they initially spread across a landscape. This initial phase of numerical dominance has empirical support in studies of northern range limits of butterflies: dispersal rates increase as species move north to newly formed favorable habitat [6]. Presumably over a long period of time, selection will favor slower dispersal rates commensurate with their ancestral rates of movement (R. Holt, *personal communication*). However, since all initial conditions do not lead to an initial phase of numerical dominance for the faster dispersers (e.g., if the initial condition is a Perron vector for the slower disperser), we still require a detailed understanding of how the local intrinsic rates of growth, the dispersal matrix, and initial conditions determine whether the faster or slower disperser is numerically dominant in the initial phase of establishment.

For conditional dispersers experiencing no dispersal costs (i.e., S is column stochastic and $\lambda_i(0) > 1$ for all i), we provide proofs that generalize previous findings in two patch models [9, 17]. We prove that all dispersal strategies outside of a one-parameter family are not evolutionarily stable: when a population adopts one of these strategies, there are nearby strategies that can invade. For populations playing strategies in this exceptional one-parameter family, the populations exhibit an ideal-free distribution at equilibrium: the per-capita growth rate is constant across the landscape [3]. Contrary to prior expectations [17], we show that are growth functions for which these ideal-free strategies cannot displace all other strategies. However, numerical simulations with the biologically plausible Beverton–Holt growth functions suggest that populations playing these ideal-free strategies can displace populations playing any other strategy. Moreover, when a population at equilibrium plays an ideal-free strategy, we prove that a population playing another ideal-free strategy cannot increase from being rare

and, consequently, is likely to be driven to extinction by stochastic forces. For populations playing these ideal-free strategies, the dispersal likelihood in a patch is inversely proportional to the equilibrium abundance in a patch. Hence, enriching one patch may result in the evolution of lower dispersal rates in that patch. Conversely, habitat degradation of a patch may result in the evolution of higher dispersal rates in that patch. These predictions about ideal-free strategies, however, have to be viewed with caution, as they are sensitive to the assumption of no dispersal costs. The inclusion of the slightest dispersal costs destroys this one-parameter family of evolutionary stable strategies and leaves only the nondispersal strategy as a candidate for an evolutionary stable strategy.

Our models make several simplifying assumptions, and relaxing these assumptions provides several mathematical problems of biological interest. Most importantly, our models do not include temporal heterogeneity, which is an important ingredient in the evolution of dispersal [17]. Temporal heterogeneity can be generated exogenously or endogenously and when combined with spatial heterogeneity can promote the evolution of faster dispersers [10, 13, 17]. For instance, Hutson, Mischaikow, and Poláčik [13] proved that a faster disperser can displace or coexist with a slower disperser for periodically forced reaction diffusion equations. Whether similar results can be proven for periodic or, more generally, random difference equations requires answering mathematically challenging questions about spectral properties of periodic and random products of nonnegative matrices. Similar challenges arise when replacing increasing growth functions with unimodal growth functions [4, 10, 19] that can generate temporal heterogeneity via periodic and chaotic population dynamics.

REFERENCES

[1] R. J. H. BEVERTON AND S. J. HOLT, *On the Dynamics of Exploited Fish Populations*, Fish. Invest. Ser. II 19, Ministry of Agriculture, Fisheries and Food, London, UK, 1957.

[2] J. DOCKERY, V. HUTSON, K. MISCHAIKOW, AND M. PERNAROWSKI, *The evolution of slow dispersal rates: a reaction diffusion model*, J. Math. Biol., 37 (1998), pp. 61–83.

[3] S. D. FRETWELL AND H. L. LUCAS, *On territorial behavior and other factors influencing patch distribution in birds*, Acta Biotheoretica, 19 (1970), pp. 16–36.

[4] W. M. GETZ, *A hypothesis regarding the abruptness of density dependence and the growth rate of populations*, Ecology, 77 (1996), pp. 2014–2026.

[5] A. HASTINGS, *Can spatial variation alone lead to selection for dispersal?*, Theor. Pop. Biol., 24 (1983), pp. 244–251.

[6] J. K. HILL, C. D. THOMAS, R. FOX, M. G. TELFER, S. G. WILLIS, J. ASHER, AND B. HUNTLEY, *Responses of butterflies to twentieth century climate warming: Implications for future ranges*, Roy. Soc. Lond. Proc. Ser. Biol. Sci., 269 (2002), pp. 2163–2171.

[7] M. W. HIRSCH, C. C. PUGH, AND M. SHUB, *Invariant Manifolds*, Springer-Verlag, Berlin, 1977.

[8] R. D. HOLT, *Population dynamics in two-patch environments: Some anomalous consequences of an optimal habitat distribution*, Theor. Pop. Biol., 28 (1985), pp. 181–208.

[9] R. D. HOLT AND M. BARFIELD, *On the relationship between the ideal-free distribution and the evolution of dispersal*, in Dispersal, J. Clobert, E. Danchin, A. Dhondt, and J. Nichols, eds., Oxford University Press, Oxford, UK, 2001, pp. 83–95.

[10] R. D. HOLT AND M. A. MCPEEK, *Chaotic population dynamics favors the evolution of dispersal*, Amer. Nat., 148 (1996), pp. 709–718.

[11] S. B. Hsu, H. L. Smith, and P. Waltman, *Competitive exclusion and coexistence for competitive systems on ordered Banach spaces*, Trans. Amer. Math. Soc., 348 (1996), pp. 4083–4094.

[12] V. Hutson, S. Martinez, K. Mischaikow, and G. T. Vickers, *The evolution of dispersal*, J. Math. Biol., 47 (2003), pp. 483–517.

[13] V. Hutson, K. Mischaikow, and P. Poláčik, *The evolution of dispersal rates in a heterogeneous time-periodic environment*, J. Math. Biol., 43 (2001), pp. 501–533.

[14] V. S. Ivlev, *Experimental Ecology of the Feeding of Fishes*, Yale University Press, New Haven, CT, 1955.

[15] M. L. Johnson and M. S. Gaines, *Evolution of dispersal: Theoretical models and empirical tests using birds and mammals*, Annu. Rev. Ecol. Syst., 21 (1990), pp. 449–480.

[16] S. A. Levin, D. Cohen, and A. Hastings, *Dispersal strategies in patchy environments*, Theoret. Pop. Biol., 26 (1984), pp. 165–191.

[17] M. A. McPeek and R. D. Holt, *The evolution of dispersal in spatially and temporally varying environments*, Amer. Nat., 6 (1992), pp. 1010–1027.

[18] K. Mischaikow, H. Smith, and H. R. Thieme, *Asymptotically autonomous semiflows: Chain recurrence and Lyapunov functions*, Trans. Amer. Math. Soc., 347 (1995), pp. 1669–1685.

[19] W. E. Ricker, *Stock and recruitment*, J. Fish. Res. Board. Can., 11 (1954), pp. 559–623.

[20] A. L. Shanks, *Mechanisms of cross-shelf dispersal of larval invertebrates and fish*, in Ecology of Marine Invertebrate Larvae, CRC, Boca Raton, FL, 1995, pp. 324–367.