



Evolution in predator–prey systems

Rick Durrett, John Mayberry*

Department of Mathematics, Cornell University, Ithaca, NY 14853, USA

Received 19 May 2009; received in revised form 22 February 2010; accepted 18 March 2010

Available online 2 April 2010

Abstract

We study the adaptive dynamics of predator–prey systems modeled by a dynamical system in which the traits of predators and prey are allowed to evolve by small mutations. When only the prey are allowed to evolve, and the size of the mutational change tends to 0, the system does not exhibit long term prey coexistence and the trait of the resident prey type converges to the solution of an ODE. When only the predators are allowed to evolve, coexistence of predators occurs. In this case, depending on the parameters being varied, we see that (i) the number of coexisting predators remains tight and the differences in traits from a reference species converge in distribution to a limit, or (ii) the number of coexisting predators tends to infinity, and we calculate the asymptotic rate at which the traits of the least and most “fit” predators in the population increase. This last result is obtained by comparison with a branching random walk killed to the left of a linear boundary and a finite branching–selection particle system.

© 2010 Elsevier B.V. All rights reserved.

MSC: primary 92D15; 92D25; secondary 60J60; 60K35

Keywords: Predator–prey; Adaptive dynamics; Coexistence; Lotka–Volterra equations; Branching random walk; Branching–selection particle system

1. Introduction

The rapidly developing field of adaptive dynamics emphasizes the combined effects of evolution and ecological interactions on population dynamics. To describe the general framework of this theory, consider a population of individuals, each associated with a trait or strategy x that characterizes its ability to survive and propagate. The current distribution of traits governs

* Corresponding author. Tel.: +1 607 255 8262; fax: +1 607 255 7149.

E-mail addresses: rtd1@cornell.edu (R. Durrett), jm858@cornell.edu (J. Mayberry).

the population dynamics by describing interactions between different individuals and their environment. Underlying this ecological process is a slower mutational process that occasionally introduces a new trait into the population. Understanding the interplay of these fast ecological and slow mutational processes is the primary objective of adaptive dynamics. Foundations of the theory were laid in the early 1990’s by Hofbauer and Sigmund [20], Metz, Nisbet, and Geritz [24], and Dieckmann, Marrow, and Law [13] and they focused on macroscopic models, i.e., ODE models describing large population limits. In their 1996 paper [12], Dieckmann and Law suggest that: “A proper mathematical theory of evolution should be dynamical... The dynamics ought to be underpinned by a microscopic theory”. A rigorous foundation for microscopic models via multi-type branching processes has now been developed (see for example Champagnat and Lambert [9] and Champagnat, Ferrière, and Méléard [8]).

In this paper, we study the dynamics of coexistence that arise as a consequence of introducing rare, small mutations into a model for predator–prey interactions. The novelty of our work lies in the establishment of the coexistence of a large number of types. This phenomenon is known as polymorphic evolution. Two other notable examples of polymorphic evolution in the adaptive dynamics literature are: (i) evolutionary branching (first described in Geritz et al. [19] and more recently studied in Champagnat and Méléard [10] from a microscopic perspective) which describes coexistence of types with diverging traits and (ii) the Tube Theorem of Geritz et al. [18, 17] where coexistence of a resident and an invading type with similar survival strategies occurs inside of a “tube” in which the sum of the invader and resident population sizes stays close to the former resident attractor. The situation that we encounter more closely resembles the second scenario: types with very similar traits can coexist. In our model, this is due to the fact that interspecies competition (competition with individuals of different types) has less of an effect than intraspecies competition (competition with individuals of the same type).

Since our focus will be on the dynamics of the random process of types that emerges from our underlying mutational process, we shall take a macroscopic perspective of population dynamics, using a Lotka–Volterra system of ODE’s to describe predator–prey interactions. In particular, we will suppose that if we have M prey types and N predator types, then the densities are governed by the ODE’s

$$\begin{aligned} \frac{du_i}{dt} &= u_i \left(\beta_i \left(1 - \sum_k u_k \right) - 1 - \sum_j \alpha_{i,j} v_j \right) \\ \frac{dv_j}{dt} &= v_j \left(\sum_i \alpha_{i,j} u_i - \delta_j - v_j \right) \end{aligned} \tag{1.1}$$

where the u_i , $1 \leq i \leq M$, are the densities of the prey, and the v_j , $1 \leq j \leq N$, are the densities of the predators. Our main interest is in the effect of small mutations in the resident types on the equilibrium behavior of (1.1). While the co-evolutionary case in which both predator and prey are allowed to vary is certainly of interest and can lead to exotic behavior (see, for example, Dieckmann et al. [13] and Dercole et al. [11]), we will here only consider the two cases of fixed predator/evolving prey and evolving predator/fixed prey. Such examples are also of interest and have been studied in laboratory experiments (see, for example, Jones and Ellner [22]).

Following the usual approach in adaptive dynamics, we shall assume that mutations take place on a much slower time scale than the population dynamics reaching equilibrium. In particular, suppose that we are considering predator evolution and we currently have k predator types (which we call the resident types) and one prey type coexisting in equilibrium. We introduce a small

density of a new type of predator, called the mutant type, with traits chosen according to a specified mutation distribution, and let the densities evolve according to (1.1) until a possibly new equilibrium is reached before introducing the next mutant type into the population. By traits, we mean the parameters in (1.1) (birth, death, and consumption rates) that characterize each individual’s ability to survive and propagate. We then repeat this process, using only those predators which could coexist at the previous step. In this way, we obtain a Markov chain of resident types with transitions determined by the outcomes of the ecological interactions between the previous residents and the mutant. Once we have introduced some preliminary results for the ODE, we will formulate this process more precisely.

Our evolutionary algorithm leads to a variety of different scenarios depending on the underlying parameters being varied.

- (a) If we have a single, fixed predator and allow mutations in the prey’s traits (α, β) , then prolonged coexistence of prey does not occur and the traits of the resident prey converge to the solution of an ODE (see Theorem 1).
- (b) If we have a single, fixed prey and allow the consumption rate α of predators to evolve, then coexistence of predators occurs, but the number of coexisting predators remains tight and the differences of the parameters from a reference type converge in distribution to a limit (see Theorem 2).
- (c) If we have a single, fixed prey and allow the death rate δ of predators to evolve, the number of coexisting predators tends to infinity and we can calculate the speed at which the traits of the least and most fit predators in the population increase (see Theorem 3).

In all three cases, our results are more mathematically interesting than biologically relevant since in (b), for example, the consumption rate of all predators currently present in the population increases without bound.

The remainder of this section is dedicated to statements of these results and some conjectures for future research. Proofs of the three results in (a)–(c) above will be contained in Sections 2–5.

1.1. Prey evolution

We begin with the case in which we have a single predator with fixed death rate $\delta > 0$ and we allow prey types to evolve. Throughout the remainder of this subsection when we refer to (1.1), we shall always assume that $N = 1$ and let v denote the density of our single predator. Prey types are characterized by their two-dimensional trait vectors $y = (\alpha, \beta) \in \mathbb{R}^2$ and we say that prey types y_1, \dots, y_M can coexist with the predator if whenever we run (1.1) started from positive initial densities, $v(0), u_i(0) > 0$, the densities $v(t), u_i(t)$ remain bounded away from zero for all time. Our first step is to discuss criteria for coexistence. We use the notation $u = (u_1, u_2, \dots, u_M)$ and

$$\Gamma_{M,1} = \{(u, v) \in \mathbb{R}^{M+1} : v, u_i \geq 0, \sum u_i \leq 1\}$$

$$\Gamma_{M,1}^{J,+} = \{(u, v) \in \Gamma_{M,1} : v, u_i > 0, \forall i \in J\}$$

for any $J \subset \{1, 2, \dots, M\}$. If $J = \{1, \dots, M\}$, we simply write $\Gamma_{M,1}^+ = \Gamma_{M,1}^{J,+}$. Note that $\Gamma_{M,1}$ is invariant under (1.1). Here and elsewhere, we shall use $|\cdot|$ to denote the cardinality of a finite set, the absolute value of a real number, and the Lebesgue measure of a set of real numbers depending on the context.

Proposition 1.1. For all prey $y_i \in \mathbb{R}^2, i \leq M$, with different birth rates, (1.1) has an explicitly calculable equilibrium $\sigma = (\sigma_1, \dots, \sigma_M, \sigma_{M+1}) \in \Gamma_{M,1}$ which is globally attracting on $\Gamma_{M,1}^+$.

Furthermore, if $J_\sigma = \{i \leq M : \sigma_i > 0\}$, then $|J_\sigma| \leq 2$ and σ is globally attracting on $\Gamma_{M,1}^{J_\sigma,+}$ as well.

Proposition 1.1 follows from Lemmas 2.1 and 2.2 in Section 2. Since we need explicit formulas for the equilibria of (1.1), we will need to re-work some standard results on Lotka–Volterra equations (see, for example, Chapters 13 and 15 of Hofbauer and Sigmund [21] and Chapter 3 in Takeuchi [26]) to prove our results.

We are now ready to describe the Prey Evolutionary Process (Prey EP). This process is a continuous time Markov jump process which keeps track of the current resident prey types in the population. Proposition 1.1 tells us that we will never have more than two coexisting prey types so at time t , the state of the Prey EP is $\mathbf{Y}(t) = (Y_1(t), Y_2(t)) \in \mathbb{R}^2 \times \mathbb{R}^2$. For initial conditions, we set $Y_2(0) = (0, 0)$, i.e., the second prey species is initially absent, and choose any $Y_1(0) = (\alpha(0), \beta(0))$ satisfying

$$Y_1(0) \in \mathcal{V} \equiv \left\{ (\alpha, \beta) \in \mathbb{R}_+^2 : \beta > \frac{\alpha}{\alpha - \delta} > 1 \right\}.$$

The reason for this choice of $Y_1(0)$ is that if $M = 1$, the globally attracting equilibrium described in Proposition 1.1 satisfies $\sigma_1, \sigma_2 > 0$ if and only if the prey type has traits in \mathcal{V} (see (2.3) in Section 2). As long as $Y_1(t) \neq (0, 0)$, mutational events occur at rate 1 and after a mutation, the transitions for $\mathbf{Y}(t)$ are determined by the following procedure. We pick one of the non-zero $Y_i(t-)$, $i = 1, 2$, at random and choose $Y_{new} = (\alpha_{new}, \beta_{new})$ uniformly from $B_\varepsilon(Y_i(t-))$, the ball of radius ε around $Y_i(t-)$. If $Y_2(t-) \neq (0, 0)$, let σ be the equilibrium obtained in Proposition 1.1 when $M = 3$ and the prey have traits $y_1, y_2, y_3 = Y_1(t-), Y_2(t-), Y_{new}$. If $Y_2(t-) = (0, 0)$, then let σ be the equilibrium obtained in Proposition 1.1 when $M = 2$ and the prey have traits $y_1, y_2 = Y_1(t-), Y_{new}$. Note that since the probability of inserting a mutant with the same birth rate as one of the residents is 0, we do not have to worry about the exceptional case where Proposition 1.1 does not apply. If $|J_\sigma| = 2$, then we set $Y_1(t)$ and $Y_2(t)$ equal to the parameter values of the two prey with positive equilibrium densities. If $|J_\sigma| = 1$, then we set $Y_1(t)$ equal to the parameter values of the single prey with positive equilibrium density and take $Y_2(t) = (0, 0)$. If $|J_\sigma| = 0$, we set $Y_1(t), Y_2(t) = (0, 0)$ and the process enters an absorbing state. We say that the population is monomorphic when $Y_2(t) = (0, 0)$ and refer to the events where $Y_2(t)$ jumps from $(0, 0)$ as coexistence events.

Our first main result says that in the small mutation limit, the population is essentially monomorphic.

Theorem 1. *Let $T > 0$. As $\varepsilon \rightarrow 0$, $Y_1^\varepsilon(t) \equiv Y_1(t/\varepsilon) \rightarrow y_1(t)$ in probability uniformly on $[0, T]$. $y_1(t)$ is the unique solution to the ODE*

$$\frac{dy_1}{dt} = \frac{2}{3\pi} \mathcal{N}(y_1(t)) \tag{1.2}$$

with initial conditions $y_1(0) = Y_1(0)$ where $\mathcal{N}(\cdot)$ is explicitly calculable; see (2.4). Furthermore, if we let $Y_2^\varepsilon(t) = Y_2(t/\varepsilon)$ and for $t \leq T$, define

$$N_t^\varepsilon \equiv |\{s \leq t : Y_2^\varepsilon(s-) = 0, Y_2^\varepsilon(s) \neq 0\}|$$

as the number of times Y_2 jumps from 0 before time t/ε , then as $\varepsilon \rightarrow 0$,

$$N^\varepsilon \Rightarrow N$$

where N is a non-homogeneous Poisson process on $[0, T]$ and “ \Rightarrow ” denotes convergence in distribution.

We will prove **Theorem 1** in Section 2. The proof reveals that $|\{t \leq T : Y_2^\varepsilon(t \wedge \tau) > 0\}| \rightarrow 0$ (see (2.8)) which justifies our earlier claim that prolonged coexistence of prey does not occur. The constant on the right hand side of (1.2) is EY^+ when (X, Y) is chosen at random from the ball of radius 1 and appears due to our choice of mutation distribution. (1.2) is essentially a special case of the “Canonical Equation of Adaptive Dynamics”; see (6.2) in Dieckmann and Law [12] or (1) in Champagnat and Lambert [9]. In words, evolution proceeds in the direction of fastest increase in fitness. We do not have an explicitly defined fitness, but the infinitesimal drift in the traits is perpendicular to the region of values that cannot invade the resident.

The limiting ODE in **Theorem 1** is not biologically sensible because the prey birth rate increases without bound. This could be remedied by restricting the permitted values of (α, β) to a curve, or making them functions of other underlying parameters (see for example Dieckmann et al. [13], where traits depend on the “body size” of the predator and prey). However, since our main interest in including **Theorem 1** is as a contrast to the results on predator evolution below, we do not here endeavor to carry through the details of this more realistic scenario.

1.2. Predator evolution

We now consider the case where we have a single prey with fixed birth rate $\beta > 1$ and density u ($M = 1$ in (1.1)), but we allow our predators to evolve. Predators are characterized by their two-dimensional trait vector $x = (\alpha, \delta) \in \mathbb{R}_+^2$. As in the previous section, the first step is to develop a criterion for determining coexistence of multiple predators. The next result, which is proved in Section 3, tells us that this can be done by checking a simple algebraic condition. In order to state the result, we define the *characteristic ratio* of the predator $x = (\alpha, \delta)$ as $\ell = \alpha/\delta$ and use the notation

$$\begin{aligned} \Gamma_{1,N} &= \{(u, v_1, \dots, v_N) \in \mathbb{R}^{N+1} : 0 \leq u \leq 1, v_i \geq 0\} \\ \Gamma_{1,N}^+ &= \{(u, v_1, \dots, v_N) \in \mathbb{R}^{N+1} : 0 < u \leq 1, v_i > 0\}. \end{aligned}$$

Note that $\Gamma_{1,N}$ is invariant for (1.1).

Proposition 1.2. *For any $N \geq 1$, (1.1) has a unique, globally attracting equilibrium σ for initial densities in $\Gamma_{1,N}^+$. This equilibrium has the following characterization. Suppose that predators x_1, \dots, x_N are ordered by increasing characteristic ratios. Then the globally attracting equilibrium has a positive i th component if and only if $i = 1, \dots, m$ where $m \leq N$ is the largest value of k satisfying the condition*

$$\sum_{j=1}^k \alpha_j^2 (\ell_k - \ell_j) < r - \beta \ell_k \tag{1.3}$$

and $r = \beta - 1 > 0$ is the intrinsic growth rate of the prey.

We prove **Proposition 1.2** in Section 3. The definition for the Predator Evolution Process (Predator EP) is similar to the definition of the Prey EP except that the state space is now $(\mathbb{R}^2)^{\mathbb{N}}$ as there is no limit on the number of predators that can coexist. The state of the process at time t is

$$\mathbf{X}(t) = ((\alpha_1(t), \delta_1(t)), (\alpha_2(t), \delta_2(t)), \dots).$$

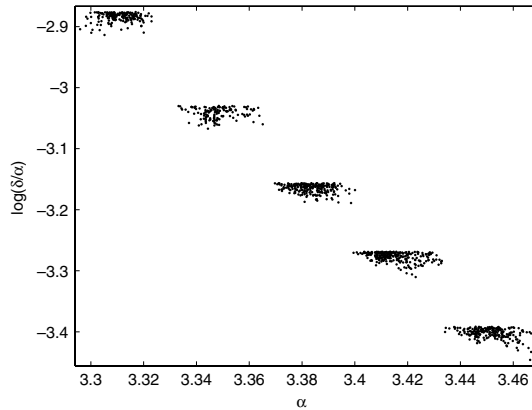


Fig. 1. The five clusters, from upper left to lower right, are the characteristics of the coexisting predators in a sample run of the Predator EP after 10^4 , 1.25×10^4 , 1.5×10^4 , 1.75×10^4 and 2×10^4 mutations have occurred. The consumption rates, α , of all coexisting predators are plotted on the x-axis and the corresponding values of $\log \ell = \log(\delta/\alpha)$ are plotted on the y-axis. Parameters: $r = 1$, $\alpha(0) = 3$, $\delta(0) = 0.45$, $\varepsilon = 0.01$.

For initial conditions, we choose $(\alpha_1(0), \delta_1(0))$ so that $\alpha_1(0)r/\beta > \delta_1(0)$ and set $\alpha_k(0), \beta_k(0) = 0$ for all $k \geq 2$. Our choice of $\alpha_1(0), \delta_1(0)$ guarantees that the initial predator can coexist with the prey (see Section 3). As long as $(a_1(t), \beta_1(t)) \neq (0, 0)$, mutations occur at rate 1 and if a mutation occurs at time t , transitions are described by the following rules. Define $N_t = \max\{i : (\alpha_i(t-), \beta_i(t-)) \neq (0, 0)\}$, choose one of $(\alpha_i(t-), \delta_i(t-))$, $i \leq N_t$, at random, and introduce a new mutant predator with traits $\alpha_{new} = \alpha_i + \varepsilon U_1$ and $\delta_{new} = \delta_i e^{\varepsilon U_2}$ where $U_1, U_2 \sim \text{Uniform}[-1, 1]$. We then order the predators $(\alpha_1(t-), \beta_2(t-)), \dots, (\alpha_{N_t}(t-), \beta_{N_t}(t-)), (\alpha_{new}, \beta_{new})$ by increasing characteristic ratios and call the ordered traits x_1, \dots, x_{N_t+1} . Applying Proposition 1.2 with predators x_1, \dots, x_{N_t+1} , we obtain a globally attracting equilibrium σ with $\sigma_i > 0$ if and only if $i \leq m$ for some $m \leq N_t + 1$. We then set $(\alpha_i(t), \beta_i(t)) = x_i$ for all $i \leq m$ and set $(\alpha_i(t), \beta_i(t)) = (0, 0)$ for all $i > m$.

Simulations suggest that we see a growing cloud of coexisting predators with some limiting shape and all predators have consumption rates α going off to infinity and $\log(\ell)$ going to $-\infty$ (see Figs. 1 and 2). We were not able to analyze the two-dimensional system so we will specialize to the two cases where only α or δ varies and the other remains fixed.

1.3. The alpha predator evolution process

In this section, we assume that $\delta = 1$ remains fixed and allow for mutations in α . Note that the characteristic ratio of a predator is now $1/\alpha$, so ordering predators by increasing characteristic ratio is equivalent to ordering predators by decreasing consumption rate. We shall prove our results for a discrete time version of the Predator EP in which the n th mutation occurs at time n . Once this is done, it is straightforward to generalize the result to continuous time.

To more precisely describe the version that we study, suppose that at time 0, we have a single predator $\alpha(0)$ which can coexist with the prey. If at time n , we have N_n predators $\alpha_1(n), \dots, \alpha_{N_n}(n)$ in decreasing order that satisfy (1.3) with $k = N_n$, then at time $n + 1$ we choose one of the $\alpha_j(n)$, $j \leq N_n$, at random, introduce a mutant with trait $\alpha_{new} = \alpha_j(n) + \varepsilon U$, $U \sim \text{Uniform}[-1, 1]$, and then use (1.3) to decide on the state of the process at time $n + 1$ in the same manner as we did for the Predator EP. We shall refer to this process as the Alpha Predator Evolution Process (APEP) and use the following notation throughout the remainder of

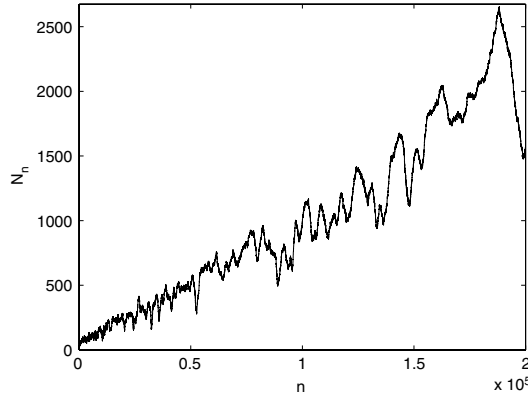


Fig. 2. Plot of N_n = number of coexisting species in the population after the n th mutation has occurred in the Predator EP from Fig. 1.

the section:

- N_n = number of coexisting predators at time n .
- $\alpha_j(n)$ = j th-largest α amongst all coexisting predators at time n for $j \leq N_n$.
- $\alpha_{\min}(n) = \alpha_{N_n}(n)$.
- $\alpha_j(n) = \alpha_{\min}(n)$ if $j > N_n$.
- $d_j(n) = \alpha_j(n) - \alpha_{\min}(n)$ = differences between predator consumption rates.
- $\Delta_n = (d_1(n), d_2(n), \dots)$.

Setting $\alpha_j(n) = \alpha_{\min}(n)$ for $j > N_n$ is done for convenience so that $d_j = 0$ for all $j \geq N_n$.

Substituting $\delta = 1$, the condition (1.3) for coexistence of $\alpha_1, \dots, \alpha_N$ simplifies to

$$\sum_{j=1}^N \frac{\alpha_j}{\alpha_N} (\alpha_j - \alpha_N) < r - \frac{\beta}{\alpha_N}. \tag{1.4}$$

Since $\alpha_j/\alpha_N > 1$, this implies that all the differences $\alpha_j - \alpha_N$ must be $< r$, so we define $\mathcal{S} := [0, r]^{\mathbb{N}}$ and let $\|\cdot\|_{TV}$ denote the total variation norm on $\mathcal{M}_1(\mathcal{S})$, the space of probability measures on \mathcal{S} . We denote by P^α the law of the APEP started from an initial predator with trait α .

Theorem 2. N_n is tight and $\alpha_{\min}(n) \rightarrow \infty$ a.s. as $n \rightarrow \infty$. In addition, there exists a measure π_ε on \mathcal{S} and a constant $a_\varepsilon > 0$ such that

$$\|P^{\alpha(0)}(\Delta_n \in \cdot) - \pi_\varepsilon(\cdot)\|_{TV} \rightarrow 0$$

and $\alpha_{\min}(n)/n \rightarrow a_\varepsilon > 0$ as $n \rightarrow \infty$.

This result is proved in Section 4. The reason for the difference from Theorem 1 is that mutant types with traits similar to those of the resident type can always invade because predators only suffer density dependent killings from their own type. In the case of prey evolution, this is not the case since interspecies and intraspecies competition affect the prey equally.

The key to the proof of Theorem 2 is the observation that as $\alpha_N \rightarrow \infty$, the condition (1.4) becomes

$$\sum_{j=1}^N (\alpha_j - \alpha_N) < r$$

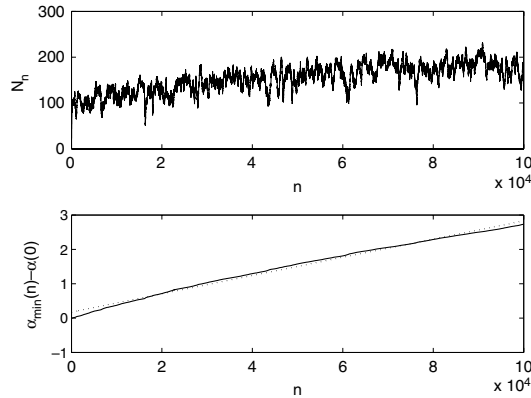


Fig. 3. Top: Plot of the number of coexisting predators at time n in the APEP with $\varepsilon = 0.01$, $r = 1$, $\alpha(0) = 3$. Bottom: Plot of the change in $\alpha_{\min}(n)$ for the same simulation.

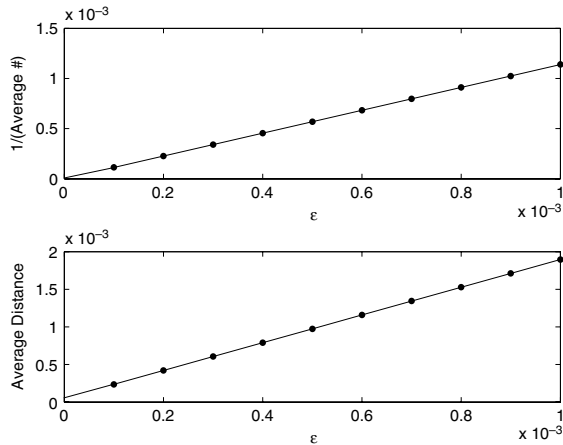


Fig. 4. The top panel shows the inverse of the average number of species and the bottom panel shows the average maximum distance between α 's as a function of ε for the APEP. Here, we have run one simulation for each value of $\varepsilon = 0.001, 0.002, \dots, 0.01$ with $r = 1$, $\alpha(0) = 3$ and then averaged out the results of each simulation over the last 25,000 time steps to obtain the values for the plotted points. The solid lines are the corresponding least square lines. It appears that the number of coexisting species is $O(1/\varepsilon)$ and the maximum distance between coexisting types is $O(\varepsilon)$ as $\varepsilon \rightarrow 0$.

and we can show that the differences Δ_n are asymptotically a positive recurrent Harris chain with stationary distribution π_ε . A coupling argument shows that the non-homogeneous chain also converges to π_ε . The linear growth of α_{\min} then follows from a standard result on functionals of positive recurrent Markov chains.

Fig. 3 illustrates the tightness of N_n and linear growth of α_{\min} . Fig. 4 suggests that as the size of the perturbation $\varepsilon \rightarrow 0$, the spacing between traits is $O(\varepsilon)$, and the number of coexisting types is $O(1/\varepsilon)$. We believe that if one converts the rescaled spacings Δ_n/ε into a measure by assigning each one mass ε then as $\varepsilon \rightarrow 0$, the distribution of this measure under π_ε converges to a deterministic limit in which the density of particles is roughly, but not exactly, exponential; see Fig. 5.

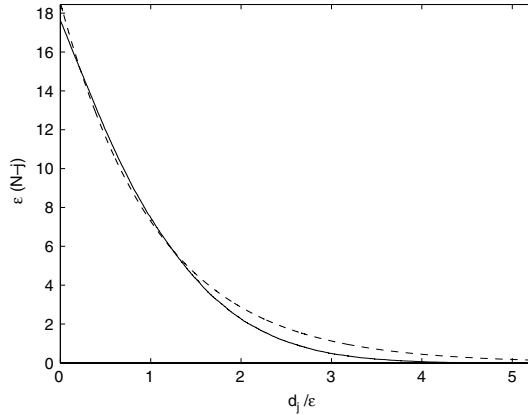


Fig. 5. Plot of the distribution of predator types for a single run of the APEP at time $n = 50,000$ with $\varepsilon = 0.001$ and $r = 1$. The solid line connects the points $(d_j/\varepsilon, \varepsilon(N_n - j))$, $j = 1, 2, \dots, N_n = 17626$. The dashed line gives an exponential approximation.

1.4. The delta predator evolution process

The final evolution process we consider is the Delta Predator Evolution Process (DPEP). The DPEP is defined in continuous time and follows the same rules as the Predator EP except that all predators have fixed $\alpha = 1$ and we only allow for mutations in δ . For convenience, we assume that $\varepsilon = 1$ and define $X_j(t) = -\log \delta_j(t)$. We also set $N_t =$ the number of coexisting predators at time t . Note that since $\delta_1(t) < \delta_2(t) < \dots < \delta_{N_t}(t)$ by definition of the DPEP, we always have $X_1(t) > X_2(t) > \dots > X_{N_t}(t)$ and so $X_{\max}(t) = X_1(t)$ and $X_{\min}(t) = X_{N_t}(t)$ give the traits of the most and least fit predators, respectively. Furthermore, (1.3) implies that

$$e^{-X_{N_t}(t)} \left(\beta + \sum_{j=1}^{N_t} 1 - \exp(-[X_j(t) - X_{N_t}(t)]) \right) < r \tag{1.5}$$

for all $t \geq 0$.

To get started in the analysis of the DPEP, our first step in Section 5 is to prove a simple result which already shows that the behavior is very different from the APEP.

Lemma 1.1. *As $t \rightarrow \infty$, $N_t \rightarrow \infty$ a.s.*

Our next result describes the asymptotic rates at which the smallest and largest predator traits increase. In what follows, we let S_t be a random walk starting at 0 that takes jumps at rate 1 uniform on $[-1, 1]$. The theory of large deviations tells us that

$$\Lambda(x) = \lim_{t \rightarrow \infty} \frac{1}{t} \log P(S_t > xt)$$

exists and can be calculated in terms of the moment generating function of S_t .

Theorem 3. *$X_{\max}(t)/t \rightarrow a$ and $X_{\min}(t)/t \rightarrow b$ a.s. as $t \rightarrow \infty$ where $a \approx 0.9053$ and $b \approx 0.5667$ satisfy the equations*

$$\Lambda(a) = -1, \quad \Lambda(b) = -1 + b. \tag{1.6}$$

Furthermore, we have $\liminf(1/t) \log N_t \geq b$ a.s.

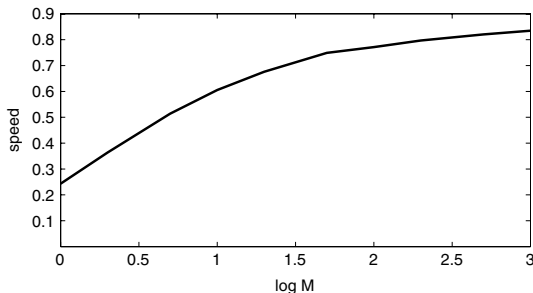


Fig. 6. Graph of the speed a_M versus $\log M$ showing slow convergence to the limit $a \approx 0.9053$ for the finite branching–selection particle system Y^M defined in Section 1.4.

We will prove Theorem 3 in Section 5. To explain why it is true, it is convenient to adopt the perspective that the different predator types correspond to particles and their traits correspond to positions on the real line. Let Z_t be the branching random walk in which particles give birth at rate 1 and their offspring are displaced by an amount uniform on $[-1, 1]$. A result of Biggins [3] implies that r_t , the position of the rightmost particle at time t in the branching random walk, has $r_t/t \rightarrow a$ and

$$\frac{1}{t} \log Z_t([xt, \infty)) \rightarrow 1 + \Lambda(x)$$

for $0 \leq x < a$ so $(1/t) \log Z_t([bt, \infty)) \rightarrow b$. Since we can construct Z_t in such a way that all particles in $X(t)$ are in Z_t , we must have $\limsup X_{\max}(t)/t \leq a$ a.s. The definition of b and an argument by contradiction using (1.5) give the upper bound $\limsup_{t \rightarrow \infty} X_{\min}(t)/t \leq b$ for the speed of the leftmost particle.

To bound $\liminf_{t \rightarrow \infty} X_{\max}(t)/t$, we consider the following branching–selection particle system: at any time t , we have M particles with positions $Y_1^M(t) > \dots > Y_M^M(t)$, all giving birth at rate 1. Whenever a new particle is born, we reorder and delete the leftmost particle. Using techniques from the study of the APEP in Section 4, we could show that $Y_1^M(t)/t \rightarrow a_M$, but instead we complete the proof of the first result by showing

$$\lim_{M \rightarrow \infty} \liminf_{t \rightarrow \infty} Y_1^M(t)/t = a. \tag{1.7}$$

Nina Gantert has pointed out to us that Berard and Guere [5] have recently proved

$$a - a_M \sim C(\log(M))^{-2}$$

for a related discrete time model in which all M particles split into two and only the rightmost M are kept. This confirms a slow rate of convergence, which was predicted much earlier by Brunet and Derrida [7], and which we observed in our numerical attempts to verify the limit in (1.7); see Fig. 6.

To bound $\liminf_{t \rightarrow \infty} X_{\min}(t)/t$, we study the branching random walk with killing at $-K + \gamma t$. Our result given in Lemma 5.4 is a cousin of a result of Kesten [23] for branching Brownian motion on $[0, \infty)$ where during its lifetime, each particle moves according to Brownian motion with drift $\mu < 0$ and variance σ^2 , all particles die at rate c and give birth to a mean m number of offspring upon death with particles killed when they hit 0. Kesten’s result states that the system has positive probability of survival when $\mu < \mu_0 = (2\sigma^2 c(m - 1))^{1/2}$ (Theorem 1.1, (1.6)), and in this supercritical case, if we start with one particle at x , then for every interval I ,

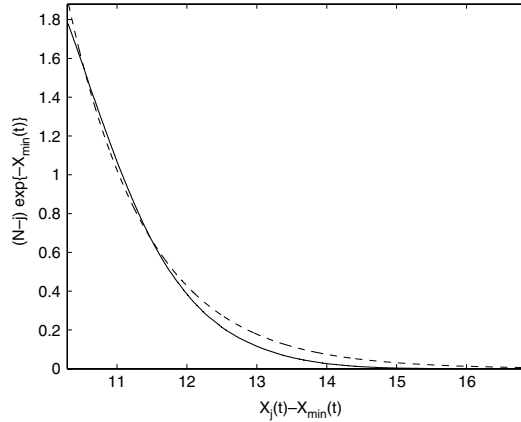


Fig. 7. Plot of the distribution of predator types for a single run of the DPEP with $r = 1$ at time $t \approx 20.25$ (after $n = 50,000$ insertions). The solid line shows the point $(X_j(t) - X_{\min}(t), (N_t - j)e^{-X_{\min}(t)})$, $j = 1, \dots, N_t = 25467$. The dashed line gives an exponential approximation.

$Z_t(I)/E_x Z_t(I) \rightarrow W$ a.s for some finite random variables W (Theorem 1.1 (1.5)). However, Kesten’s efforts are concentrated on the exotic behavior in the critical case $\mu = \mu_0$, and he says “so far we have only an ugly and complicated proof of the growth results in the supercritical case, and we shall therefore not prove Theorem 1.1”. In Section 5, we show that using ideas of Biggins [3] it is easy to prove results for $(1/t) \log Z_t([ct, \infty))$.

The result $\limsup_{t \rightarrow \infty} X_{\min}(t)/t \leq b$ implies that if T is large and we start the branching random walk with one particle at $X_{\max}(T)$ at time T then all of the particles in the branching random walk with killing at $(b + \varepsilon)t$ are present in the $X_i(t)$. If $X_{\min}(t)$ is too far to the left then we would contradict (1.5). The last part of the proof suggests that most particles are near $X_{\min}(t)$. Simulations (see Fig. 7) further suggest that:

Conjecture. *If we put mass $\exp(-X_{\min}(t))$ at $X_i(t) - X_{\min}(t)$ then this measure converges to a deterministic limit, which again is roughly but not exactly exponential.*

However, proving this seems to be a difficult problem. Recently, Durrett and Remenik [15] have proved convergence of the toy model to the solution of a free boundary problem as $M \rightarrow \infty$.

The final conclusion $\liminf_{t \rightarrow \infty} (\log N_t)/t \geq b$ follows from the result for $X_{\min}(t)$ and the proof of Lemma 1.1. Since the result comes from replacing (1.5) by $e^{-X_N(t)}(\beta + N) < r$, it seems unlikely that b is the right constant, but finding the right constant would require proving the conjecture.

The proof of $\liminf_{t \rightarrow \infty} X_{\max}(t)/t \geq a$ also leads to the following result regarding the limiting behavior of the most fit predator in the APEP as the mutation radius $\varepsilon \rightarrow 0$.

Corollary 1. *If we run the APEP in continuous time and let*

$$a_\varepsilon^* = \lim_{t \rightarrow \infty} \alpha_{\max}^\varepsilon(t)/t,$$

then $\lim_{\varepsilon \rightarrow 0} a_\varepsilon^* = a$.

2. Prey evolution

In this section we will prove Theorem 1. We first establish Proposition 1.1 as a consequence of Lemmas 2.1 and 2.2 below. Note that these results only cover $M \leq 3$, but the proof of Lemma 2.2

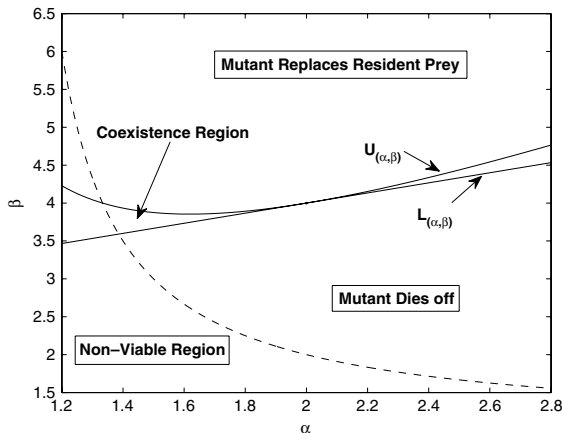


Fig. 8. Plot of the invadability curves for $(\alpha_1, \beta_1) = (2, 4)$. For the predator, we set $\delta = 1$. The dashed line shows the boundary of the viable region.

can easily be extended to show that coexistence for $M \geq 4$ is not possible. Our results rely heavily on the notion of invadability (see Durrett [14]) and results on Lotka–Volterra systems (see Chapter 13 and 15 of Hofbauer and Sigmund [21]) which we shall quote as needed. We shall also make use of the notation introduced in Section 1.1 and assume throughout this section that $N = 1$ in (1.1). Therefore, we have a single predator, whose density we shall denote by v , with fixed death rate $\delta > 0$.

2.1. Prey ODE results

Let $y_1 = (\alpha_1, \beta_1) \in \mathbb{R}_+^2$. To determine when the predator and a prey with trait y_1 can coexist, we note that if $\beta_1 > 1$ then in the absence of predators the prey reach an equilibrium density

$$\sigma_1^0(y_1) = (\beta_1 - 1)/\beta_1. \tag{2.1}$$

If the prey are in equilibrium then the predators can increase when v is small if

$$\alpha_1 \sigma_1^0(y_1) - \delta > 0. \tag{2.2}$$

Using the formula for $\sigma_1^0(y_1)$, we see that this holds if and only if $\alpha_1 > \delta$ and

$$\beta_1 > \frac{\alpha_1}{\alpha_1 - \delta} > 1. \tag{2.3}$$

We call this set of (α_1, β_1) the *viable region* for prey and label it \mathcal{V} . See Fig. 8 for an example.

Algebra shows that when (2.3) occurs, there is a predator–prey equilibrium $\sigma^1(y_1)$ with coordinates

$$\sigma_1^1(y_1) = \frac{(\beta_1 - 1) + \alpha_1 \delta}{\beta_1 + \alpha_1^2}, \quad \sigma_2^1(y_1) = \frac{(\beta_1 - 1)\alpha_1 - \beta_1 \delta}{\beta_1 + \alpha_1^2}.$$

A second prey type with trait $y_2 = (\alpha_2, \beta_2)$ can invade the first prey and the predator in equilibrium when

$$\beta_2(1 - \sigma_1^1(y_1)) - 1 - \alpha_2 \sigma_2^1(y_1) > 0.$$

By interchanging the roles of y_1 and y_2 we get the condition for the first prey to invade the second prey and predator in equilibrium. If both prey traits are viable and the two invadability conditions hold, then it is shown in Section 7.1 of [14] that there is coexistence in the ODE, i.e., the three densities stay bounded away from 0. That the densities actually converge to a positive equilibrium in this case is the result of Lemma 2.1 below.

Following [14], we use \succ for “invades” (prey j can invade prey $1, \dots, j - 1$ in equilibrium if its density will increase whenever $1, \dots, j - 1$ are in equilibrium with the predator and a small initial density of j 's is introduced). Using the new notation and defining

$$F(y_1, y_2) = \beta_2(1 - \sigma_1^1(y_1)) - 1 - \alpha_2\sigma_2^1(y_1),$$

we have $2 \succ 1$ if and only if

$$y_2 \in \{y : F(y_1, y) > 0\} =: \mathcal{L}_{y_1}$$

and $1 \succ 2$ if and only if

$$y_2 \in \{y : F(y, y_1) > 0\} =: \mathcal{U}_{y_1}.$$

We call the boundary curves $L_{y_1} \equiv \{y : F(y_1, y) = 0\}$ and $U_{y_1} \equiv \{y : F(y, y_1) = 0\}$ the invadability curves and note that we have the formulas

$$(\alpha, \beta) \in U_{y_1} \Leftrightarrow \beta = g(y_1, \alpha)$$

$$(\alpha, \beta) \in L_{y_1} \Leftrightarrow \beta = h(y_1, \alpha)$$

where

$$g(y_1, \alpha) = \frac{(\beta_1 - 1)\alpha^2 + (\alpha_1 - \beta_1\delta)\alpha + \beta_1}{1 + \alpha_1(\alpha - \delta)}$$

and

$$h(y_1, \alpha) = \frac{\alpha\sigma_2^1(y_1) + 1}{1 - \sigma_1^1(y_1)}$$

are well defined provided $y_1, (\alpha, \beta) \in \mathcal{V}$. Calculus shows that the curve U_{y_1} is tangent to the curve L_{y_1} at y_1 and we let $\mathcal{N}(y_1)$ denote the corresponding unit normal vector:

$$\mathcal{N}(y_1) = c(-\sigma_2^1(y_1), 1 - \sigma_1^1(y_1)) \tag{2.4}$$

where c is chosen to make the length 1. The situation is depicted in Fig. 8. Lemma 2.1 describes the set of all possible ecological outcomes based on this splitting of type space.

Lemma 2.1. *Let $y_1, y_2 \in \mathcal{V}$ with $\beta_1 \neq \beta_2$ and suppose that $(u_1(0), u_2(0), v(0)) \in \Gamma_{2,1}^+$. Then one of the following must be true:*

- (a) $y_2 \in \mathcal{L}_{y_1} \cap \mathcal{U}_{y_1}$ in which case the solution to (1.1) converges to a unique positive equilibrium $\sigma^2(y_1, y_2) = (\sigma_1^2(y_1, y_2), \sigma_2^2(y_1, y_2), \sigma_3^2(y_1, y_2)) \in \Gamma_{2,1}^+$.
- (b) $y_2 \in \mathcal{L}_{y_1}$, but $y_2 \notin \mathcal{U}_{y_1}$ in which case the solution to (1.1) converges to the equilibrium $(0, \sigma_1^1(y_2), \sigma_2^1(y_2))$.
- (c) $y_2 \in \mathcal{U}_{y_1}$, but $y_2 \notin \mathcal{L}_{y_1}$ in which case the solution to (1.1) converges to the equilibrium $(\sigma_1^1(y_1), 0, \sigma_2^1(y_1))$.

Since the probability of inserting a mutant with the same birth rate as the resident is zero, the condition $\beta_1 \neq \beta_2$ does not impose any additional restrictions and saves us the headache of dealing with a scenario in which we have an infinite number of equilibria.

Before beginning the proof, we note that setting $u_3 = v$, we can rewrite (1.1) in Lotka–Volterra form as

$$\frac{du_i}{dt} = u_i(r_i + (Au)_i)$$

where $r_i = \beta_i - 1, i = 1, 2, r_3 = -\delta$, and

$$A = \begin{bmatrix} -\beta_1 & -\beta_1 & -\alpha_1 \\ -\beta_2 & -\beta_2 & -\alpha_2 \\ \alpha_1 & \alpha_2 & -1 \end{bmatrix}.$$

If $f_i(u) = du_i/dt$, then the components of the Jacobian are given by

$$\frac{df_i(u)}{du_j} = \delta_{ij}(r_i + (Au)_i) + u_i A_{ij}. \tag{2.5}$$

Let J_p denote the value of the Jacobian matrix $J = (df_i/du_j)$ evaluated at p . Following [21] (see pages 155, 159 therein), we shall say that an equilibrium point p for (1.1) is *regular* if $\det(J_p) \neq 0$ and *saturated* if $r_i + (Ap)_i \leq 0$ for all i . Note that if we have an equilibrium $p \in \Gamma_{2,1}^+$, then p is trivially saturated since in this case, $r_i + (Ap)_i = 0$ for all i . More generally, an equilibrium point σ with $\sigma_i = 0$ for all $i \in I$ is saturated if it cannot be invaded by types $i \in I$. In particular, $(\sigma_1^1(y_1), 0, \sigma_2^1(y_1))$ and $(0, \sigma_1^1(y_2), \sigma_2^1(y_2))$ are saturated exactly when $y_2 \notin \mathcal{L}_{y_1}$ and $y_2 \notin \mathcal{U}_{y_1}$, respectively. The assumptions $y_1, y_2 \in \mathcal{V}$ and $\beta_1, \beta_2 > 1$ imply that $\sigma^0(y_1), \sigma^0(y_2)$, and the origin are never saturated. It is easy to see that since $\beta_1 \neq \beta_2$, there can be no other possible equilibria $\sigma \notin \Gamma_{2,1}^+$.

Let

$$ind(p) = sign(\det(J_p))$$

denote the *index* of a regular equilibrium p . The index theorem for Lotka–Volterra equations (13.4.4 in [21]) tells us that if all saturated equilibria p are regular, we must always have

$$\sum_{p:p \text{ saturated}} ind(p) = (-1)^3 = -1. \tag{2.6}$$

The key to the proof will be showing that (1.1) has a unique saturated fixed point in all three cases (a)–(c). However, since it is not always true that a unique saturated fixed point is globally attracting (see page 195 in [21]), we need to work a little bit harder to get the result. To ease notation, we shall let F_i denote the face in $\Gamma_{2,1}$ on which $u_i = 0$ and $E_{i,j}$ denote the edge where $u_i = u_j = 0$.

Proof. Suppose first that we are in case (a) so that $(\sigma_1^1(y_1), 0, \sigma_2^1(y_1))$ and $(0, \sigma_1^1(y_2), \sigma_2^1(y_2))$ are not saturated. Then Theorem 7.1 in [14] implies that we have coexistence and hence by Theorem 13.3.1 and 13.5.2 in [21], (1.1) has a unique regular equilibrium $\sigma^2 \in \Gamma_{2,1}^+$. To show that it is globally attracting, we will show that all eigenvalues of J_{σ^2} have negative real parts. The conclusion that σ^2 is globally attracting on $\Gamma_{2,1}^+$ then follows by Theorem 15.3.1 in [21] (A is Volterra–Lyapunov stable with $d_i = \sigma_i^2$). The Routh–Hurwitz (R–H) conditions (see pages 702–703 of Murray [25] for the version used here or Anagnost and Desoer [1] for an elementary proof) tell us that all eigenvalues of J_{σ^2} will have negative real parts if (1) $\text{trace}(J_{\sigma^2}) < 0$, (2) $\det(J_{\sigma^2}) < 0$, and (3)

$$\det(J_{\sigma^2}) > \text{trace}(J_{\sigma^2})\Sigma_2$$

where Σ_2 is the sum of the 2×2 principal minors of J_{σ^2} . But since $r + A\sigma^2 = 0$, substituting σ^2 into (2.5) yields

$$J_{\sigma^2} = \begin{bmatrix} -\sigma_1^2\beta_1 & -\sigma_1^2\beta_1 & -\sigma_1^2\alpha_1 \\ -\sigma_2^2\beta_2 & -\sigma_2^2\beta_2 & -\sigma_2^2\alpha_2 \\ \sigma_3^2\alpha_1 & \sigma_3^2\alpha_2 & -\sigma_3^2 \end{bmatrix}.$$

so the first R–H condition is obvious and the third follows from a simple algebraic calculation. The second condition follows from (2.6) since σ^2 is the unique saturated equilibrium point for (1.1) and is regular.

Suppose now we are in case (b) so that $(0, \sigma_1^1(y_2), \sigma_2^1(y_2))$ is saturated, but $(\sigma_1^1(y_1), 0, \sigma_2^1(y_1))$ is not saturated. If we let $\sigma = (0, \sigma_1, \sigma_2) = (0, \sigma_1^1(y_2), \sigma_2^1(y_2))$, then

$$J_{\sigma} = \begin{bmatrix} r_1 + (A\sigma)_1 & 0 & 0 \\ -\sigma_1\beta_2 & -\sigma_1\beta_2 & -\sigma_1\alpha_2 \\ \sigma_2\alpha_1 & \sigma_2\alpha_2 & -\sigma_2\gamma \end{bmatrix}.$$

The assumptions in (b) imply that $2 > 1, 3$ and $1 \neq 2, 3$ so we must have $r_1 + (A\sigma)_1 < 0$. Therefore,

$$\det(J_{\sigma}) = (r_1 + (A\sigma)_1)(\beta_2\gamma + \alpha_2^2)\sigma_1\sigma_2 < 0$$

implying that σ is regular. If coexistence was possible, then as in the proof of (a), we would have a regular equilibrium $\rho \in \Gamma_{2,1}^+$ and (1.1) would have exactly two regular, saturated equilibria, violating (2.6). Therefore we know that $u_i(t) \rightarrow 0$ as $t \rightarrow \infty$ for some $i = 1, 2, 3$. But since we have the invadability conditions, $1 > 0, 3 > 1, 3 > 2, 2 > 1, 3$, and $1 \neq 2, 3$, the proof of Theorem 7.1 in [14] tells us that there exists a repelling function for the set

$$F \equiv F_3 \cup E_{1,2} \cup F_2$$

and therefore we know that (1.1) must leave $\Gamma_{2,1}^+$ through $F_1 \setminus F$ on which σ is globally attracting (Lemma 5.0 in [14]). The proof of (c) is identical after interchanging the roles of y_1 and y_2 . \square

Our next result describes the possible outcomes of adding a new species when two are already coexisting.

Lemma 2.2. *Let $y_1, y_2, y_3 \in \mathcal{V}$, $y_2 \in \mathcal{L}_{y_1} \cap \mathcal{U}_{y_1}$, $\beta_1 \neq \beta_2$, and*

$$(u_1(0), u_2(0), u_3(0), v(0)) \in \Gamma_{3,1}^+.$$

Then one of the following must be true:

- (a) $y_3 \in \mathcal{L}_{y_1} \cap \mathcal{L}_{y_2}$ and $y_3 \notin \mathcal{U}_{y_1} \cup \mathcal{U}_{y_2}$ in which case the solution to (1.1) converges to $(0, 0, \sigma_1^1(y_3), \sigma_2^1(y_3))$.
- (b) $y_3 \in \mathcal{U}_{y_1} \cap \mathcal{U}_{y_2}$ and $y_3 \notin \mathcal{L}_{y_1} \cup \mathcal{L}_{y_2}$ in which case the solution to (1.1) converges to $(\sigma_1^2(y_1, y_2), \sigma_2^2(y_1, y_2), 0, \sigma_3^2(y_1, y_2))$.
- (c) Neither (a) nor (b) is satisfied in which case either $u_1(t) \rightarrow 0$ or $u_2(t) \rightarrow 0$.

Proof. It is easy to check that if we have $M = 3$ in (1.1), relabel the predator’s density as u_4 and write (1.1) in Lotka–Volterra form, then $\det(A) = 0$ so there can be no coexistence of the three prey types by Theorem 13.5.2 in [21]. Therefore, at least one of the types dies out. Which one(s) can be sorted out using the same idea as in the proof of Lemma 2.1. \square

To see why we do not need to be concerned with case (c), we note that because of the tangency of U_{y_1} and L_{y_1} , the chance of inserting a mutant which can coexist with a current resident prey is always of order

$$\varepsilon^{-2} \int_0^\varepsilon x^2 dx = O(\varepsilon).$$

But then the probability of inserting a third prey which can coexist with two resident prey is also $O(\varepsilon)$ and therefore, we must wait $O(1/\varepsilon^2)$ time steps until the first time we encounter the situation in Lemma 2.2, (c). Of course, we will still get convergence to an equilibrium that can be determined as in Lemma 2.1 provided $\beta_3 \neq \beta_1, \beta_2$.

2.2. Proof of Theorem 1

Throughout this section, we shall use $C = C_T$ to denote a positive constant which depends on T and may change from line to line. Write

$$F^\varepsilon(y_1) = \frac{1}{2\pi\varepsilon^2} \int_{-\varepsilon}^\varepsilon f(y_1, x) dx$$

where $f(y_1, \alpha) = g(y_1, \alpha) - h(y_1, \alpha) \geq 0$. We also define

$$p^\varepsilon(t) = \begin{cases} F^\varepsilon(Y_1^\varepsilon(t)) & \text{if } Y_2^\varepsilon(t) = 0 \\ \max(F^\varepsilon(Y_1^\varepsilon(t)), F^\varepsilon(Y_2^\varepsilon(t))) & \text{if } Y_2^\varepsilon(t) \neq 0. \end{cases}$$

For $Y_1^\varepsilon(t), Y_2^\varepsilon(t) \in \mathcal{V}$, p^ε gives the probability of inserting a new type that can coexist when only one resident type is present and an upper bound on the probability that a new type can coexist with one of the resident types when two resident types are present. Note that the tangency of $g(y_1, \cdot), h(y_1, \cdot)$ and Taylor’s theorem imply that

$$\varepsilon^{-1} F^\varepsilon(y_1) \rightarrow c\phi(y_1) \tag{2.7}$$

for any $y_1 \in \mathcal{V}$ and some constant $c > 0$ where

$$\phi(y_1) = \frac{\partial^2(g - h)}{\partial \alpha^2}(y_1, \alpha_1)$$

is a continuous function of y_1 . Therefore, if $Y_1^\varepsilon(t), Y_2^\varepsilon(t) \in K$, K bounded, then there exists $C_K > 0$ such that $p^\varepsilon(t) \leq C_K \varepsilon$.

Choose an open, bounded set $K_1 \subset \mathcal{V}$ and a compact set $K_2 \subset K_1$ such that $y_1(t) \in K_2$ for all $t \leq T$. The existence of K_1 is guaranteed since on the boundary of the viable region, $\beta = \alpha/(\alpha - \delta)$ and so the slope of U_{y_1} at y_1 is 0 implying that $\mathcal{N}(y_1)$ points straight up, and it is impossible for $y_1(t)$ to leave the viable region. Let $\rho > 0$ be small enough that $K_2 + \rho \subset K_1$ and define

$$\tau = \inf\{t : \mathbf{Y}^\varepsilon(t) \notin (K_2 + \rho) \times (K_2 + \rho)\}$$

as the first time $Y_1^\varepsilon(t)$ or $Y_2^\varepsilon(t)$ leaves $K_2 + \rho$. Then $p^\varepsilon(t \wedge \tau) \leq C\varepsilon, \forall t \leq T$. Since mutations occur at rate $1/\varepsilon$, it follows that the expected number of times before $T \wedge \tau$ that there is one prey type and an inserted type coexists is $\leq C$. If two prey types coexist, Lemma 2.2 implies that with probability $\geq 1 - C\varepsilon$, the next time a new type is inserted in $\mathcal{L}_{y_1} \cap \mathcal{L}_{y_2}$, it will replace the two coexisting types. Therefore, if ε is small, the amount of time during which two types coexist is approximately exponential with mean 2ε and so

$$|\{t \leq T : Y_2^\varepsilon(t \wedge \tau) > 0\}| \rightarrow 0 \tag{2.8}$$

a.s. as $\varepsilon \rightarrow 0$ and hence, we can ignore these isolated episodes when studying the evolution of $Y_1^\varepsilon(t \wedge \tau)$.

When there is no coexistence, mutations in the direction of L_{y_1} leave the resident type unchanged and mutations in the direction of $\mathcal{N}(y_1)$ replace the resident, so the infinitesimal mean of $Y_1^\varepsilon(t \wedge \tau)$ is given by

$$b(y_1) = \frac{2}{3\pi} \mathcal{N}(y_1) \tag{2.9}$$

where the $2/3\pi$ comes from the fact that if we choose a point at random from the upper half of the ball of radius 1 in the (α, β) plane, then the β component has density $(4/\pi)\sqrt{1 - \beta^2}$ and hence mean

$$\frac{4}{\pi} \int_0^1 \beta \sqrt{1 - \beta^2} dy = \frac{4}{3\pi}.$$

(2.9) then follows on noting that choices from the half of the ball above L_{y_1} occur with probability 1/2. It is clear from the scaling that the entries in the infinitesimal covariance are of order ε and therefore, the infinitesimal mean and covariance of $Y_1^\varepsilon(\cdot \wedge \tau)$ converge to $b(y_1)$ and $a(y_1) = 0$ respectively. Since b is Lipschitz continuous, the martingale problem for (a, b) is well posed, so convergence of $Y_1^\varepsilon(\cdot \wedge \tau)$ to y_1 follows from Theorem 7.4.1 in Ethier and Kurtz [16]. But then we can choose ρ small enough that $P(\tau \leq T) \rightarrow 0$ and we obtain (1.2).

It remains to prove that

$$N_t^\varepsilon = |\{s \leq t : Y_2^\varepsilon(s-) = 0, Y_2^\varepsilon(s) \neq 0\}|$$

converges to a non-homogeneous Poisson process. Since $F^\varepsilon(Y_1^\varepsilon(t))$ gives the jump probabilities for N_t^ε when $Y_2^\varepsilon(t) = 0$, the compensator for N_t^ε is given by

$$A_t^\varepsilon = \int_0^t \mathbf{1}_{\{Y_2^\varepsilon(s)=0\}} \varepsilon^{-1} F^\varepsilon(Y_1^\varepsilon(s)) ds.$$

(1.2), (2.7) and (2.8) then imply that

$$A_t^\varepsilon \rightarrow m(t) \equiv \int_0^t c\phi(y_1(s)) ds. \tag{2.10}$$

$m(t)$ is continuous and deterministic so we conclude from Theorem 1 in Brown [6] that $N^\varepsilon \Rightarrow N$ where N is a non-homogeneous Poisson process with mean function $m(t)$. \square

3. Multiple-predator ODE facts

The goal of this section is the derivation of Lemmas 3.1 and 3.2 which together imply Proposition 1.2. The first result gives the algebraic condition for existence of positive equilibrium densities and the second proves convergence to equilibrium. See Section 1.2 for relevant notation.

In the absence of predators, the prey have equilibrium density $\sigma^0 = r/\beta$ where $r = \beta - 1 > 0$ by assumption. Suppose we wish to find a positive equilibrium $\sigma^k = (\sigma_0^k, \sigma_1^k, \dots, \sigma_k^k)$ on the face

$$\Gamma_{1,k} = \{v \in \Gamma_{1,N} : v_{k+1} = \dots = v_N = 0\}.$$

Then, solving the equations $\alpha_j \sigma_0^k - \delta_j - \sigma_j^k = 0$ for $\sigma_j^k, j = 1, \dots, k$ we obtain

$$\sigma_j^k = \sigma_j^k(x_1, \dots, x_k) = \alpha_j \sigma_0^k - \delta_j$$

and substituting these expressions into the equation $r - \beta\sigma_0^k - \sum_{j=1}^k \alpha_j \sigma_j^k = 0$ yields

$$r - \beta\sigma_0^k = \sum_{j=1}^k \alpha_j^2 \sigma_0^k - \sum_{j=1}^k \alpha_j \delta_j.$$

We can conclude that

$$\sigma_0^k = \sigma_0^k(x_1, \dots, x_k) = \frac{r + \sum_{i=1}^k \alpha_i \delta_i}{\beta + \sum_{i=1}^k \alpha_i^2} > 0$$

for all $k \geq 0$. To determine when $\sigma_j^k > 0$, $1 \leq j \leq k$, write $S_k = \sum_{i=1}^k \alpha_i^2$ and

$$\alpha_j \sigma_0^k = \frac{\alpha_j r + \alpha_j^2 \delta_j + \alpha_j \sum_{i \neq j} \alpha_i \delta_i}{\beta + S_k}.$$

Adding $\delta_j - \delta_j(\beta + S_k)/(\beta + S_k)$, the above is

$$\begin{aligned} & \alpha_j r - \beta \delta_j + \sum_{i \neq j} (\alpha_j \alpha_i \delta_i - \alpha_i^2 \delta_j) \\ = & \delta_j + \frac{\alpha_j r - \beta \delta_j + \sum_{i \neq j} (\alpha_j \alpha_i \delta_i - \alpha_i^2 \delta_j)}{\beta + S_k} \\ = & \delta_j + \frac{(\beta + \sum_{i \neq j} \alpha_i^2)(\alpha_j \sigma_0^{k-1}(x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_k) - \delta_j)}{\beta + S_k} \end{aligned}$$

since $(\beta + \sum_{i \neq j} \alpha_i^2)\sigma_0^{k-1} = r + \sum_{i \neq j} \alpha_i \delta_i$. From this it follows that σ_j^k will be positive if and only if

$$\sigma_0^{k-1}(x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_k) > \ell_j. \tag{3.1}$$

where $\ell_j = \delta_j/\alpha_j$ is the characteristic ratio for predator x_j .

Lemma 3.1. *Suppose x_1, \dots, x_k are ordered by increasing characteristic ratio. Then $\sigma_k^k > 0$ if and only if*

$$\beta \ell_k + \sum_{j=1}^k \alpha_j^2 (\ell_k - \ell_j) < r \tag{3.2}$$

and if (3.2) is satisfied, then $\sigma_j^k > 0$ for all $j \leq k$.

Proof. (3.1) implies that $\sigma_k^k > 0$ if and only if

$$\ell_k < \sigma_0^{k-1}(x_1, \dots, x_{k-1}) = \frac{r + \sum_{j=1}^{k-1} \alpha_j^2 \ell_j}{\beta + \sum_{j=1}^{k-1} \alpha_j^2} \tag{3.3}$$

where we have used the definition of $\ell_j = \delta_j/\alpha_j$ on the right. Multiplying both sides by the denominator of the right and then rearranging terms, we obtain (3.2) (since $\ell_k - \ell_k = 0$). Now suppose that (3.2) holds. Then since $\ell_k > \ell_j$, for all $j = 1, \dots, k - 1$, (3.2) also holds if we replace ℓ_k by ℓ_j , $j < k$, on the left and reversing the algebra used to derive (3.2) from (3.3) shows that this is equivalent to $\sigma_j^k > 0$, proving the result. \square

Lemma 3.2. *Suppose we have a collection of predators x_1, \dots, x_N ordered by increasing ℓ and let $k \leq N$ be the largest integer for which (3.2) is satisfied (with the convention that $k = 0$ if (3.2) fails for all $k \leq N$). Then $\sigma = (\sigma_0^k, \sigma_1^k, \dots, \sigma_k^k, 0, \dots, 0)$ is a globally attracting fixed point on $\Gamma_{1,N}^+$ with Lyapunov function*

$$V(u, v_1, \dots, v_N) = u - \sigma_0^k \log u + \sum_{i=1}^k (v_i - \sigma_i^k \log v_i) + \sum_{i=k+1}^N v_i.$$

Proof. Differentiating V yields

$$\begin{aligned} \frac{dV}{dt} &= (u - \sigma_0^k)(r - \beta u - \sum_{i=1}^k \alpha_i v_i - \sum_{i=k+1}^N \alpha_i v_i) \\ &\quad + \sum_{i=1}^k (v_i - \sigma_i^k)(-\delta_i - v_i + \alpha_i u) + \sum_{i=k+1}^N v_i(-\delta_i - v_i + \alpha_i u) \\ &= -\beta(u - \sigma_0^k)^2 - \sum_{i=1}^k (v_i - \sigma_i^k)^2 - \sum_{i=k+1}^N v_i(\delta_i - \alpha_i \sigma_0^k) - \sum_{i=k+1}^N v_i^2. \end{aligned}$$

If $(u, v_1, \dots, v_N) = \sigma$, this expression is 0 and otherwise it is < 0 since Lemma 3.1 and (3.1) imply that

$$\ell_i \geq \ell_{k+1} > \sigma_0^k$$

for all $i \geq k + 1$ so all terms on the left are negative. \square

4. Proof of Theorem 2

In this section, we prove Theorem 2 for the APEP as defined in Section 1.3 and use the notation defined there. We also define the Markov chain $Y_n = (\alpha_{\min}(n), \Delta_n)$.

If $\delta_j = 1$ for all j , (3.2) with $k = N$ can be rewritten as

$$\sum_{j=1}^N \frac{\alpha_j}{\alpha_N} (\alpha_j - \alpha_N) < r - \frac{\beta}{\alpha_N}. \tag{4.1}$$

Our first step is to show:

Lemma 4.1. *The sequence N_n is tight.*

Proof. Define the sets $A^m = [0, r]^m \times \{0\}^{\mathbb{N}}$, for $m \in \mathbb{N}$. Then $\Delta_n \in A^m$ if and only if $N_n \leq m$. Let $M = M(r, \varepsilon) = \lceil \frac{4r}{\varepsilon} \rceil$ be the smallest integer $> 4r/\varepsilon$ and suppose that $Y_n = y \in \mathbb{R}^+ \times \mathcal{S}$. From (4.1), at most M of the $\alpha_j(n)$'s can be $\geq \alpha_{\min}(n) + \varepsilon/4$. With probability at least $1/4^M$, the next M mutants will be inserted to the right of $\alpha_{\min}(n) + \varepsilon/2$. But then none of the predators

to the left $\alpha_{\min}(n) + \varepsilon/4$ can be in the coexisting set at time $n + M$ because otherwise, by the definition of the APEP, any predator with $\alpha > \alpha_{\min}(n) + \varepsilon/2$ would also be in the set, and since there are at least M such predators,

$$\sum_{j=1}^{\infty} d_j(n + M) > M(\varepsilon/2 - \varepsilon/4) > r$$

contradicting (4.1). Therefore, we have the uniform lower bound

$$P(\Delta_{n+M} \in A^{2M} | Y_n = y) \geq 4^{-M} \tag{4.2}$$

which holds for all $y \in \mathbb{R}^+ \times \mathcal{S}$. Since this bound is uniform in y , tightness follows. \square

Lemma 4.2. *As $n \rightarrow \infty$, the marginal transition probabilities for Δ_n :*

$$p_\alpha(\Delta, \cdot) := P(\Delta_{n+1} \in \cdot | Y_n = (\alpha, \Delta))$$

converge in total variation to the transition probabilities for a time homogeneous Markov chain X_n with transition probabilities $p(\Delta, \cdot)$.

Proof. Suppose that $Y_n = (\alpha, \Delta)$. Since $0 \leq \alpha_j(n) - \alpha_{\min}(n) \leq r$ for all $n \geq 1$, we can see that as $\alpha \rightarrow \infty$, (4.1) simplifies to

$$\sum_{j=1}^N d_j(n) < r. \tag{4.3}$$

This implies that, in the limit, the differences evolve according to the following algorithm: pick a species $1 \leq k \leq N_n$ at random, insert a random mutation in $(d_k(n) - \varepsilon, d_k(n) + \varepsilon)$, and then modify the algorithm in Proposition 1.2 to use (4.3) instead of (4.1) with the rule that we shift the differences before calculating the sum if the new insertion is left of 0. \square

Our next result concerns the limiting behavior of X_n . Writing x instead of Δ for the vector of differences, we set

$$p(x, A) = P(X_{n+1} \in A | X_n = x).$$

Lemma 4.3. *X_n is a positive recurrent, Harris chain and hence, has a unique stationary distribution π .*

Proof. Following the arguments in Athreya and Ney [2], it suffices to show that there exists a “regenerative” set $A \subset \mathcal{S}$ satisfying:

- (C1) $P^x(\tau_A < \infty) = 1$ for all $x \in \mathcal{S}$ where τ_A is hitting time of A .
- (C2) There exists a probability measure ρ on A , $\lambda > 0$, and $\kappa \in \mathbb{N}$ such that $p^\kappa(x, B) \geq \lambda\rho(B)$ for all $x \in A$, $B \subset A$.

The same calculation that led to (4.2) shows that A^{2M} satisfies the condition in (C1), but (C2) may not hold for this set. We therefore define a set G (for good) that will be reached from A^{2M} with probability 1 and satisfies (C2). To this end, let

$$\kappa = 1 + \sup \left\{ k : \sum_{j=1}^k j = \frac{k(k+1)}{2} < 2r/\varepsilon \right\}$$

and choose η small enough that

$$\sum_{j=1}^k j(\varepsilon/2 + \eta) < r. \tag{4.4}$$

Let $G = \{d_i - d_{i+1} \in (\varepsilon/2, \varepsilon/2 + \eta) \text{ for } i < \kappa \text{ and } d_i = 0 \text{ for } i \geq \kappa\}$. In other words, $d \in G$ corresponds to κ types coexisting with α 's that have spacings between $\varepsilon/2$ and $\varepsilon/2 + \eta$ units apart.

The first step in showing that (C1) and (C2) hold for $A = G$ is to show that if $X_0 = x \in A^{2M}$, then we can get to A in κ steps by the following path: first, we choose d_1 (the predator with the largest values of α) as our mutating predator at time 1 (which happens with probability at least $(2M)^{-1}$) and then choose a mutant type g_1 in $(d_1 + \varepsilon/2, d_1 + \varepsilon/2 + \eta)$ (which happens with probability $\eta/(2\varepsilon)$). At the next time step, we choose g_1 as our mutating type (which happens with probability at least $(2M + 1)^{-1}$) and then mutate to $g_2 \in (g_1 + \varepsilon/2, g_1 + \varepsilon/2 + \eta)$. If we continue for κ steps, then each $g_j, 1 \leq j \leq \kappa$, will be at least as big as $d_1 + j\varepsilon/2$, so by (4.3), no member of the coexisting set at time 0 will remain at time κ . Furthermore, by (4.4), the shifted set $d'_j = g_{\kappa-j+1} - g_1, 1 \leq j \leq \kappa$, will satisfy (4.3) and therefore, $X_\kappa \in G$. It is clear from the construction that we have

$$p^\kappa(x, G) \geq \left(\frac{\eta}{2\varepsilon(2M + \kappa)}\right)^\kappa. \tag{4.5}$$

To prove (C2) holds, we first consider cylinder sets of the form $B = \{d_i - d_{i+1} \in B_i \subset (\varepsilon/2, \varepsilon/2 + \eta) \text{ for } i < \kappa \text{ and } d_i = 0 \text{ for } i \geq \kappa\}$. Then if $x \in G$, taking the same path as led to (4.5) yields the lower bound

$$p^\kappa(x, B) \geq \frac{|B_1| \cdots |B_{\kappa-1}|}{(2\varepsilon)^{\kappa-1}} \left(\frac{1}{2M + \kappa}\right)^\kappa. \tag{4.6}$$

If we let ρ = the Lebesgue measure on G normalized to be a probability and recall that the Radon–Nikodym derivative $dp^\kappa(x, \cdot)/d\rho(\cdot)$ evaluated at a general measurable set B can be written as the limit of $p^\kappa(x, B_k)/\rho(B_k)$ where B_k is a sequence of cylinder sets, (C2) follows.

To check positive recurrence, we let τ_A be the first-hitting time of our regenerative set G . (4.2) and (4.5) tell us that there is a positive constant $\theta = \theta(r, \varepsilon)$ such that

$$p^{2M+\kappa}(x, G) \geq \theta > 0$$

for any $x \in \mathcal{S}$. Therefore, we have $E^x(\tau_A) \leq (2M + \kappa)/\theta < \infty$, completing the proof. \square

The construction in the previous lemma also yields:

Lemma 4.4. $\alpha_{\min}(n) \rightarrow \infty$ a.s. as $n \rightarrow \infty$.

Proof. We can modify the construction in the previous lemma to show that there exist constants $K, J \geq 1, \rho > 0$ such that

$$P(\alpha_1((n + 1)K) - \alpha_1(nK) \geq J\varepsilon/2 | Y_{nK} = y) \geq \rho$$

for any $y \in \mathbb{R}^+ \times \mathcal{S}$ and $n \geq 0$. Therefore, $\alpha_1(n) \rightarrow \infty$ a.s. by the Borel–Cantelli Lemma and the result follows since $\alpha_1(n) - \alpha_{\min}(n) < r$. \square

Theorem 4. As $n \rightarrow \infty, \|P^\alpha(\Delta_n \in \cdot) - \pi(\cdot)\|_{TV} \rightarrow 0$ for any initial $\alpha \in \mathbb{R}^+$.

Proof. It suffices to prove the result for the subsequences $n = m\kappa + j$ for $0 \leq j < \kappa$, but then by using the Markov property at time j , it is enough to prove the result for $n = m\kappa$ and a general initial distribution. To prepare for the proof, recall that one can modify the state space of a Harris recurrent Markov chain to have a point ζ that corresponds to being distributed on the set A according to ρ with the exact position being independent of the past.

To prove the result, we will construct a process $(\tilde{X}_n, \tilde{\Delta}_n)$ on $\mathcal{S} \times \mathcal{S}$ such that the marginal law of $\tilde{\Delta}_n$ is the law of $\Delta_{n\kappa}$, the marginal distribution of \tilde{X}_n is π for all n , and $\mathbb{P}(\tilde{X}_n \neq \tilde{\Delta}_n) \rightarrow 0$ as $n \rightarrow \infty$. Let U_1, U_2, \dots and V_1, V_2, \dots be independent and uniform on $[0, 1]$. To begin, let $q_\alpha(x, \cdot) \equiv P(\Delta_\kappa \in \cdot | \Delta_0 = (\alpha, x))$ and

$$q(x, \cdot) \equiv \lim_{\alpha \rightarrow \infty} q_\alpha(x, \cdot) = p^\kappa(x, \cdot)$$

by Lemma 4.2. Define the function $J_n : \mathcal{S} \times [0, 1] \rightarrow \mathcal{S}$ by

$$P(J_n(x, U_n) \in B) = q_{\alpha_{\min}(n\kappa)}(x, B).$$

Since $q_{\alpha_{\min}(n\kappa)}(x, \cdot) \in \mathcal{M}_1(\mathcal{S})$ and \mathcal{S} is a separable metric space, defining J_n is possible by Theorem 3.2 in Billingsley [4]. Suppose that \tilde{X}_n has distribution π , define $Z_{n+1} = J_n(\tilde{X}_n, U_n)$ and

$$\mu_n(A) \equiv P(Z_n \in A | \alpha_{\min}(n\kappa)) = \int q_{\alpha_{\min}(n\kappa)}(x, A) \pi(dx),$$

and let $(\tilde{X}_{n+1}, Z_{n+1})$ be a maximal coupling of (X_n, Z_n) such that

$$\mathbb{P}(\tilde{X}_{n+1} \neq Z_{n+1}) = \|\mu_n - \pi\|_{TV}$$

(see, for example Thorisson [27]). Then from the definition of μ_n and $(\tilde{X}_{n+1}, Z_{n+1})$ we have

$$\begin{aligned} \eta_{n+1} &\equiv \mathbb{P}(\tilde{X}_{n+1} \neq Z_{n+1}) \\ &= \left\| \int q_{\alpha_{\min}(n\kappa)}(x, \cdot) \pi(dx) - \int q(x, \cdot) \pi(dx) \right\|_{TV} \rightarrow 0 \end{aligned}$$

as $n \rightarrow \infty$ by Lemmas 4.2 and 4.4.

When $\{\tilde{\Delta}_n = \tilde{X}_n\}$, we set $\tilde{\Delta}_{n+1} = J_n(\tilde{X}_n, U_n) = Z_{n+1}$, so

$$P(\tilde{X}_{n+1} \neq \tilde{\Delta}_{n+1}, \tilde{X}_n = \tilde{\Delta}_n) \leq \eta_{n+1}.$$

On $\{\tilde{X}_n \neq \tilde{\Delta}_n\}$, we take $\tilde{\Delta}_{n+1} = J_n(\tilde{X}_n, V_n)$. (4.6) implies that $q(x, \zeta) \geq \lambda$, so it follows from Lemma 4.2 that if $\alpha_{\min}(n\kappa) \geq \alpha_0$ then $q(\alpha_{\min}(n\kappa), x, \zeta) \geq \lambda/2$, and we have

$$P(\tilde{X}_{n+1} = \tilde{\Delta}_{n+1} | \tilde{X}_n \neq \tilde{\Delta}_n) > \lambda/2$$

so if $\zeta_n = P(\tilde{X}_n \neq \tilde{\Delta}_n)$, then

$$\zeta_{n+1} \leq (1 - \lambda/2)\zeta_n + \eta_{n+1}.$$

Iterating yields the inequality

$$\zeta_{n+1} \leq \sum_{i=1}^{n+1} (1 - \lambda/2)^{n+1-i} \eta_i. \tag{4.7}$$

Since $|1 - \lambda/2| < 1$ and $\eta_n \rightarrow 0$, the right hand side of (4.7) must also go to zero which yields

$$\|P(\Delta_{n\kappa} \in \cdot) - \pi(\cdot)\|_{TV} \leq P(\tilde{X}_n \neq \tilde{\Delta}_n) = \eta_n \rightarrow 0$$

completing the proof. \square

It remains to prove the result on the linear growth of $\alpha_{\min}(n)$. Since $\alpha_j - \alpha_{\min} \leq r$, it suffices to establish this for α_{\max} . To do this, we look at the chains $Z_n = (X_n, U_n, V_n)$ with U_n uniform on $[0, 1]$ giving the index $k = \lceil N_n U_n \rceil$ of the value to be mutated, and V_n independently uniform on $[-\varepsilon, \varepsilon]$ giving the change in the value due to mutation. It is clear that the distribution of Z_n will converge in distribution to the product measure $\tilde{\pi} = \pi \times \text{uniform}[0, 1] \times \text{uniform}[-\varepsilon, \varepsilon]$, so if we let $f(Z_n) = \alpha_{\max}(n) - \alpha_{\max}(n - 1)$ be the amount shifted at the n th step, then f is non-negative and bounded above by ε so the strong law for functionals of Markov chains implies

$$\frac{\alpha_{\max}(n) - \alpha_{\max}(0)}{n} = \frac{1}{n} \sum_{m=1}^n f(Z_m) \rightarrow \int f(x)\tilde{\pi}(dx) = \bar{\alpha}. \tag{4.8}$$

Since $f > 0$ with positive probability, $\bar{\alpha} > 0$. To extend this result to the real chain, let $(\tilde{X}_n, \tilde{\Delta}_n)$ be the coupled chain from the proof of Theorem 4 and define $D_n = 1$ if $\tilde{X}_n \neq \tilde{\Delta}_n$ and $D_n = 0$ otherwise. From the proof of Theorem 4, we can dominate D_n by a Markov chain B_n that has

$$\begin{aligned} P(B_{n+1} = 1 | B_n = 0) &= \eta_{n+1} \\ P(B_{n+1} = 0 | B_n = 1) &= \frac{\lambda}{2} \end{aligned}$$

i.e., we can define the two processes on the same space such that $B_n \geq D_n$ for all n . Coupling B_n with a homogeneous chain B_n^ρ that has $P(B_{n+1}^\rho = 1 | B_n^\rho = 0) = \rho$, $P(B_{n+1}^\rho = 0 | B_n^\rho = 1) = \lambda/2$, and stationary distribution π^ρ with $\pi^\rho(1) = \rho/(\rho + \lambda/2)$, and recalling that $\eta_n \rightarrow 0$, it follows that

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{m=1}^n D_m \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{m=1}^n B_m \leq \frac{\rho}{\rho + \lambda/2}.$$

Since this holds for any $\rho > 0$, we must have

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{m=1}^n D_m = 0$$

and the desired result now follows from (4.8) and the fact that $0 \leq \alpha_{\max}(n) - \alpha_{\max}(n - 1) \leq 1$ for all $n \geq 0$.

5. Proof of Theorem 3

In this section, we prove Theorem 2 for the DPEP as defined in Section 1.3 and use the notation defined there. Since one of the keys to deriving our results will be comparison with a branching random walk, we continue adopting the perspective that $X_j(t)$ refers to the position of particle j on the positive half-line. Note that if we set $\alpha_j = 1$, $k = N$ in (3.2), we obtain the condition for coexistence:

$$\delta_N \left(\beta + \sum_{j=1}^N \left(1 - \frac{\delta_j}{\delta_N} \right) \right) < r. \tag{5.1}$$

Proof of Lemma 1.1. Let $X_i(t) = -\log(\delta_i(t))$ and $X_1(t) > \dots > X_M(t)$ be the rightmost M particles at this time. It should be clear from (5.1) that if

$$e^{-X_M(T)}(\beta + M) < r \tag{5.2}$$

then we will have $N_t \geq M$ for $t \geq T$. Let $y = -\log(r/(\beta + M))$. The rightmost particle is increasing in t . Since the number of particles changes by ≤ 1 each time and $\sum_{m=1}^{\infty} 1/m = \infty$, the rightmost particle gives birth to the right of its current position plus $1/2$ infinitely many times. Thus at some time T , we will have a point $\geq y + M$. Since $|X_i(t) - X_{i+1}(t)| \leq 1$ and points are only erased when (5.1) fails, (5.2) follows. \square

5.1. Asymptotics for X_{max}

For the remainder of the paper, we let Z_t be a branching random walk started from one particle at 0, in which particles give birth at rate 1 and displacements are uniform on $[-1, 1]$. It is well known that the mean measure

$$EZ_t(A) = e^t P(S_t \in A) \tag{5.3}$$

where S_t is a continuous time random walk that jumps at rate 1 and takes step uniform on $[-1, 1]$. If we let $\phi(\theta) = (e^\theta - e^{-\theta})/2\theta$ be the moment generating function for the displacements, then

$$Ee^{\theta S_t} = \sum_{n=0}^{\infty} e^{-t} \frac{t^n}{n!} \phi^n(\theta) = \exp(t(\phi(\theta) - 1)).$$

Chebyshev’s inequality implies that if $\theta > 0$,

$$P(S_t > xt) \leq \exp(-t(\theta x - \phi(\theta) + 1)) \tag{5.4}$$

and standard large deviations results imply that for $x \geq 0$,

$$\frac{1}{t} \log P(S_t > xt) \rightarrow \Lambda(x) = -\left(\sup_{\theta>0} \{\theta x - \phi(\theta)\} + 1\right) \tag{5.5}$$

where $\Lambda(0) = 0$ and Λ is strictly decreasing on $[0, \infty)$.

Biggins [3], Theorem 2, shows that the rightmost particle in the branching random walk $Z_{max}(t)/t \rightarrow a$ a.s. where a , defined in (1.6), is the smallest $x > 0$ such that $\Lambda(x) \leq -1$. Since the particles $X_i(t)$ in our evolution model are a subset of those in the branching random walk, we have

$$\limsup_{t \rightarrow \infty} X_1(t)/t \leq a.$$

The remainder of this section is dedicated to the proof of the lower bound

$$\liminf_{t \rightarrow \infty} X_1(t)/t \geq a. \tag{5.6}$$

By Lemma 1.1, we know that there exists some time T such that $N_t \geq M$ for $t \geq T$. By the proof of Lemma 1.1, we can take T to be the first time $e^{-X_M(T)}(\beta + M) < r$, which is a stopping time, so the future behavior of the process is not affected.

Lemma 5.1. *If we start the toy model at time T with positions equal to the rightmost M particles at this time $X_1(T) > \dots > X_M(T)$, then the $X_i(t)$ and $Y_i^M(t)$ can be defined on the same space such that $X_i(t) \geq Y_i^M(t)$ for all $1 \leq i \leq M$ and $t \geq T$.*

Proof. Couple the birth times of $X_i(t)$ and $Y_i^M(t)$ and the displacements of their offspring. Recall that if a birth from $X_k(t)$ with $k > M$ lands to the right of some $X_i(t)$, $i \leq M$, we renumber

the X_i and put them in decreasing order. Births of particles from $X_k(t)$ for $k > M$ may cause the X 's to get ahead of the Y 's, but coupled births for $i \leq M$ cause the vectors of X 's and Y 's to move in parallel and the desired comparison follows. \square

For our next comparison consider the branching random walk started with one particle at $Y_1^M(0)$. Let T_k be the time of the k th birth, with $T_0 = 0$, and for $t \in [T_{k-1}, T_k)$ let $\zeta_1^k(t) > \zeta_2^k(t) > \dots > \zeta_k^k(t)$ be the locations of the particles present.

Lemma 5.2. *We can couple the branching random walk and the toy model so that for $t \in [T_{k-1}, T_k)$, $Y_j^M(t) \geq \zeta_j^k(t)$ for $1 \leq j \leq k$ and $k < M$.*

Proof. Couple the birth times of $\zeta_j^k(t)$ and $Y_j^M(t)$ for $j \leq k$ and $t \in (T_{k-1}, T_k]$, i.e., there will be no births in (T_{k-1}, T_k) and the same particle will give birth at time T_k . Births of particles from $Y_j(t)$ for $j > k$ may cause the Y 's to get ahead of the ζ 's, but coupled births for $j \leq k$ cause the vectors of ζ 's and Y 's to move in parallel. \square

Lemma 5.3. *Let B_M be the time of the M th birth in the branching random walk.*

$$\liminf_{t \rightarrow \infty} \frac{Y_1^M(t)}{t} \geq \frac{EZ_{\max}(B_M)}{EB_M} \rightarrow a \quad \text{as } M \rightarrow \infty$$

Proof. Let $T_{k,1} = T_k$ where the T_k are as in Lemma 5.2 and for $j > 1$, let $T_{k,j}$, $k \leq M$, denote the time of the k th birth in a BRW started with a single particle at $Y_1^M(T_{M,j-1})$ at time $T_{M,j-1}$ and let $\zeta_{M,j}$ denote the position of the rightmost particle at time $T_{M,j}$. Repeatedly applying the comparison in Lemma 5.2 yields

$$\begin{aligned} \frac{Y_1^M(t)}{t} &\geq \frac{\sum_{j: T_{M,j} \leq t} (Y_1^M(T_{M,j+1}) - Y_1^M(T_{M,j}))}{t} \\ &\geq \frac{\sum_{j: T_{M,j} \leq t} (\zeta_{M,j+1} - Y_1^M(T_{M,j}))}{t}. \end{aligned}$$

But the time intervals $T_{M,j+1} - T_{M,j}$ are iid with mean EB_M so the first part of the result follows from the renewal theorem. To prove the second part, we note that Biggins' result implies

$$Z_{\max}(B_M)/B_M \rightarrow a \quad \text{almost surely.}$$

Since $B_M = \xi_1 + \dots + \xi_M$ where the ξ_i are independent exponentials with mean $1/i$, it is easy to see that $B_M/EB_M \rightarrow 1$, so

$$Z_{\max}(B_M)/EB_M \rightarrow a \quad \text{almost surely.}$$

Therefore, the result will follow from the dominated convergence theorem if we can show that

$$E \left(\sup \frac{Z_{\max}(B_M)}{EB_M} \right) < \infty.$$

By the Cauchy–Schwarz inequality, it suffices to show

$$E \left(\sup_{t \geq 1} \frac{Z_{\max}(t)}{t} \right)^2 < \infty \tag{5.7}$$

and

$$E \left(\sup \frac{B_M}{EB_M} \right)^2 < \infty. \tag{5.8}$$

To prove (5.7), we note that (5.3) and (5.4) imply that

$$P(Z_{\max}(t) > xt) \leq e^{t(1+\Lambda(x))}$$

and since Λ is concave with $\Lambda(0) = 0$ and $\Lambda(a) = -1$ with $a < 1$, it follows that for $x \geq 1$,

$$P(Z_{\max}(t) > xt) \leq e^{t(1-x)}.$$

Now if $Z_{\max}(t)/t > 2x$ for some t , then since $Z_{\max}(t)$ is non-decreasing, we must have $Z_{\max}(s)/s > x$ for some $s \in [t, t + 1]$ and therefore, summing over all integers t from 1 to ∞ , we see that if $x > 2$,

$$P \left(\sup_{t \geq 1} Z_{\max}(t)/t > 2x \right) \leq e^{1-x}$$

which proves (5.7). To prove (5.8), we note that $EB_M = \sum_{i=1}^M 1/i$ and

$$E \exp(\theta B_M) = \prod_{i=1}^M \frac{1}{1 - \theta/i}$$

for $0 < \theta < 1$, so by Chebyshev’s inequality,

$$P(B_M > yEB_M) \leq \exp \left(-\theta y \sum_{i=1}^M \frac{1}{i} - \sum_{i=1}^M \log(1 - \theta/i) \right).$$

Taking $\theta = 1/2$ and choosing c so that $\log(1 - x) \geq -x - cx^2$ when $0 < x < 1/2$, we have

$$\begin{aligned} P(B_M > yEB_M) &\leq \exp \left(\sum_{i=1}^M \frac{1}{2i} (1 - y) + \frac{c}{4i^2} \right) \\ &\leq C \exp \left(\frac{1 - y}{2} \log(M + 1) \right) = C \cdot (M + 1)^{(1-y)/2}. \end{aligned}$$

Therefore if $y > 3$,

$$\sum_{M=2}^{\infty} (M + 1)^{(1-y)/2} \leq \int_2^{\infty} x^{(1-y)/2} dy = \frac{2^{(3-y)/2}}{(y - 3)/2}$$

which yields (5.8), completing the proof. \square

(5.6) follows from Lemmas 5.1 and 5.3 which completes the proof that the speed of the rightmost particle is a . We shall complete the proof of Theorem 3 in the next section by showing the speed of the leftmost particle is b , but first we pause to prove Corollary 1.

Proof of Corollary 1. Suppose we choose ε small enough that $\varepsilon M(M - 1)/2 < r$. Using the coupling in Lemma 5.2 we can use the particles $\zeta_j^k, j \leq k \leq M$, from the branching random walk started at X_{max} to get a lower bound on the rightmost $k \leq M$ particles in the predator evolution with fixed δ . An induction argument shows that the spacings between the corresponding particles in the predator evolution are $\leq \varepsilon$ at all times. Since we have assumed $\varepsilon \sum_{j=1}^{M-1} j < r$, the rightmost $k \leq M$ particles are never killed. The remainder of the proof is the same as before. \square

5.2. Asymptotics for X_{min}

In order to get the speed of the leftmost particle, we will need the following result on a branching random walk with killing which is an adaptation of Biggins [3], Theorems 1 and 2, which proves this result without killing.

Lemma 5.4. *Let $Z_t(\gamma, A)$ denote the number of particles in A under a branching random walk with birth rate 1, displacements uniform on $[-1, 1]$, killing to the left of $-K + \gamma t$, and started with one particle at 0. Then for any $c > \gamma$ on the set of non-extinction*

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log Z_t(\gamma, [ct, \infty)) = I(c) \tag{5.9}$$

where $I(c) = 1 + \Lambda(c)$, and the probability of extinction tends to 0 as $K \rightarrow \infty$.

Proof. Theorem 2 in Biggins along with (5.3) and (5.5) yields (5.9) in the case of no killing and since $Z_t(\gamma, [ct, \infty)) \subset Z_t([ct, \infty))$, we get the upper bound in (5.9). To get the lower bound, we recall that to prove the corresponding lower bound for the process without killing, Biggins lets Z_{m+1}^k be the points at time $(m + 1)k$ that are at least kc units to the right of their ancestor in Z_m^k at time mk . $|Z_m^k|$ is a branching process with offspring distribution $|Z_1^k|$ so $(|Z_m^k|)^{1/m} \rightarrow E|Z_1^k|$ on the non-extinction set. Combining (5.3) and (5.5) implies $(1/k) \log E|Z_1^k| \rightarrow I(c)$ which yields the desired lower bound.

To extend this construction to the process with killing, let \bar{Z}_{m+1}^k be the points at time $(m + 1)k$ that are at least kc units to the right of their ancestor in \bar{Z}_m^k at time mk and are not killed by going to the left of $-K + \gamma t$ of $mk \leq t \leq (m + 1)k$. By construction, all points in \bar{Z}_m^k are $\geq cmk$ and we have chosen $\gamma < c$, so for large m , the killing has little effect and on the set of non-extinction we have

$$\frac{1}{m} \log |\bar{Z}_m^k| \rightarrow \log E|Z_1^k|.$$

Using (5.3) and (5.5) again gives the desired lower bound. \square

With this result in hand, we can complete the:

Proof of $X_{min}(t)/t \rightarrow b$. When $X_{min}(t)$ increases we must have

$$N_t e^{-X_{min}(t)} \geq r.$$

Since the particles in X are a subset of the particles in the branching random walk, it follows that if $X_{min}(t) \geq (b + \varepsilon)t$,

$$N_t e^{-X_{min}(t)} \leq Z_t([(b + \varepsilon)t, \infty)) e^{-(b+\varepsilon)t} \rightarrow 0$$

as $t \rightarrow \infty$ since $I(c) < c$ for all $c > b$. Therefore, $\limsup X_{min}(t)/t \leq b$ a.s.

To prove that $\liminf X_{min}(t)/t \geq b$ a.s., let $c \in (b, a)$ and $\varepsilon > 0$. Choose K large enough that the probability of extinction in the branching random walk with killing at $-k + bt$ is less than ε for all $k \geq K$ and then take T large enough that $X_1(t) \geq ct$ for all $t \geq T$ (which is possible since $\lim X_1(t)/t = a$) and that $bT > K$. Suppose that $X_{min}(t) \leq (b - \rho)t$ for some $\rho > 0$. Then by comparing with a branching random walk with killing at $-X_1(T) + bt$, we have

$$\begin{aligned} F(t) &:= e^{-X_{min}(t)} \sum_{j=1}^{N_t} (1 - e^{-X_j(t)/X_{min}(t)}) \\ &\geq e^{-(b-\varepsilon)t} (1 - e^{-(c-b+\varepsilon)t}) Z_t(b, [ct, \infty)). \end{aligned} \tag{5.10}$$

But on the non-extinction set (which has probability at least $1 - \varepsilon$), we have

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log[e^{-(b-\varepsilon)t} (1 - e^{-(c-b+\varepsilon)t}) Z_t(\gamma, [ct, \infty))] = I(c) - b + \varepsilon \rightarrow \varepsilon > 0$$

as $c \downarrow b$ and therefore, we must have $X_{\min}(t) > (b - \varepsilon)t$ eventually or there would exist a sequence of points $t_i \rightarrow \infty$ for which $F(t_i) \rightarrow \infty$, contradicting (1.5). Therefore, $P(\liminf X_{\min}(t)/t < b) < \varepsilon$ and since ε is arbitrary, this proves the result.

To conclude that $\liminf_{t \rightarrow \infty} (\log N_t)/t \geq b$ a.s., note that if $\varepsilon > 0$ then for large times there are at least $\exp((I(c) - \varepsilon)t)$ points of X to the left of ct . Picking c close to b and ε small gives the desired result. \square

Acknowledgements

The first author was partially supported by NSF grant DMS 0704996 from the probability program. The second author was partially supported by NSF RTG grant DMS 0739164.

References

- [1] J. Anagnost, C. Desoer, An elementary proof of the Routh–Hurwitz stability criterion, *Circuits Systems Signal Process* 10 (1) (1991) 101–114.
- [2] K. Athreya, P.E. Ney, A new approach to the limit theory of recurrent Markov chains, *Trans. Amer. Math. Soc.* 245 (1978) 493–501.
- [3] J.D. Biggins, Chernoff’s theorem in branching random walk, *J. Appl. Probab.* 14 (1977) 630–636.
- [4] P. Billingsley, *Weak Convergence of Measures*, SIAM, Philadelphia, PA, 1971.
- [5] J. Berard, J. Gouere, Brunet–Derrida behavior of branching–selection particle systems on the line, 2008. [arXiv:0811.2782v1](https://arxiv.org/abs/0811.2782v1).
- [6] T. Brown, A martingale approach to the Poisson convergence of simple point processes, *Ann. Probab.* 6 (1978) 615–628.
- [7] E. Brunet, B. Derrida, Shift in the velocity of a front due to a cutoff, *Phys. Rev. E* 56 (1997) 2597–2604.
- [8] N. Champagnat, R. Ferrière, S. Méléard, From individual stochastic processes to macroscopic models in adaptive evolution, *Stoch. Models* 24 (1) (2008) 2–44. Suppl.
- [9] N. Champagnat, A. Lambert, Evolution of discrete populations and the canonical diffusion of adaptive dynamics, *Ann. Appl. Probab.* 17 (2007) 102–155.
- [10] N. Champagnat, S. Méléard, Polymorphic evolution sequence and evolutionary branching, *Probab. Theory Related Fields* (2010) (in press).
- [11] F. Dercole, J.O. Irisson, S. Rinaldi, Bifurcation analysis of a predator–prey coevolution model, *SIAM J. Appl. Math.* 63 (2002) 1378–1391.
- [12] U. Dieckmann, R. Law, The dynamical theory of coevolution: a derivation from stochastic ecological processes, *J. Math. Biol.* 3 (1996) 579–612.
- [13] U. Dieckmann, P. Marrow, R. Law, Evolutionary cycling in predator–prey interactions: population dynamics and the Red Queen, *J. Theor. Biol.* 176 (1995) 91–102.
- [14] R. Durrett, Mutual invadability implies coexistence in spatial models, *Mem. Amer. Math. Soc.* 156 (740) (2002).
- [15] R. Durrett, D. Remenik, Brunet–Derrida particle systems, free boundary problems and Wiener–Hopf equations, *Ann. Probab.* (2010) (in press).
- [16] S. Ethier, T. Kurtz, *Markov Processes: Characterization and Convergence*, John Wiley and Sons, 1986.
- [17] S.A.H. Geritz, Resident–invader dynamics and the coexistence of similar strategies, *J. Math. Biol.* 50 (2005) 67–82.
- [18] S.A.H. Geritz, M. Gyllenberg, F.J.A. Jacobs, K. Parvinen, Invasion dynamics and attractor inheritance, *J. Math. Biol.* 44 (2002) 548–560.
- [19] S.A.H. Geritz, J.A.J. Metz, E. Kisdi, G. Meszina, Dynamics of Adaptation and Evolutionary Branching, *Phys. Rev. Lett.* 78 (1997) 2024–2027.
- [20] J. Hofbauer, K. Sigmund, Adaptive dynamics and evolutionary stability, *Appl. Math. Lett.* 3 (4) (1990) 75–79.
- [21] J. Hofbauer, K. Sigmund, *Evolutionary Games and Replicator Dynamics*, Cambridge University Press, Cambridge, United Kingdom, 1998.

- [22] L.E. Jones, S.P. Ellner, Effects of rapid prey evolution on predator–prey cycles, *J. Math. Biol.* 55 (2007) 541–573.
- [23] H. Kesten, Branching Brownian motion with absorption, *Stoch. Proc. Appl.* 7 (1978) 9–47.
- [24] J.A.J. Metz, R.M. Nisbet, S.A.H. Geritz, How should we define fitness for general ecological scenarios?, *Trends Ecol. Evol.* 7 (1992) 198–202.
- [25] J.D. Murray, *Mathematical Biology*, Springer-Verlag, Berlin, 1989.
- [26] Y. Takeuchi, *Global Dynamical Properties of Lotka–Volterra Systems*, World Scientific Publishing Co., Singapore, 1996.
- [27] H. Thorisson, On maximal and distributional coupling, *Ann. Appl. Probab.* 44 (1986) 873–876.