Mathematical Biology

Resident-invader dynamics of similar strategies in fluctuating environments

Yuhua Cai[1] · Stefan A. H. Geritz[1]

Abstract

We study resident-invader dynamics in fluctuating environments when the invader and the resident have close but distinct strategies. First we focus on a class of continuous-time models of unstructured populations of multi-dimensional strategies, which incorporates environmental feedback and environmental stochasticity. Then we generalize our results to a class of structured population models. We classify the generic population dynamical outcomes of an invasion event when the resident population in a given environment is non-growing on the long-run and stochastically persistent. Our approach is based on the series expansion of a model with respect to the small strategy difference, and on the analysis of a stochastic fast-slow system induced by time-scale separation. Theoretical and numerical analyses show that the total size of the resident and invader population varies stochastically and dramatically in time, while the relative size of the invader population changes slowly and asymptotically in time. Thereby the classification is based on the asymptotic behavior of the relative population size, and which is shown to be fully determined by invasion criteria (i.e., without having to study the full generic dynamical system). Our results extend and generalize previous results for a stable resident equilibrium (particularly, Geritz in J Math Biol 50(1):67–82, 2005; Dercole and Geritz in J Theor Biol 394:231-254, 2016) to non-equilibrium resident population dynamics as well as resident dynamics with stochastic (or deterministic) drivers.

✉ Yuhua Cai
yuhua.cai@helsinki.fi

Stefan A. H. Geritz
stefan.geritz@helsinki.fi

1   Department of Mathematics and Statistics, University of Helsinki, PO Box 68, 00014 Helsinki, Finland

1 Introduction

Two important issues in the framework of adaptive dynamics are which mutant strategies can invade a population of given resident strategies, and what would be the population dynamical outcomes of an invasion event. The long-term growth rate of a newly arrived and initially rare mutant with strategy y in the environment generated by a population of resident strategy x (i.e., invasion fitness of mutant y in resident x) determines whether an invasion event may occur or not (Metz et al. 1992, 1996; Dieckmann and Law 1996; Geritz et al. 1997, 1998). If y can invade x but not *vice versa*, does it mean that y will eventually take over x and becomes the new resident? In general, this need not happen as examples of unprotected coexistence and the "resident strikes back" phenomenon show (see e.g., Doebeli 1998; Parvinen 1999; Mylius and Diekmann 2001; Dercole et al. 2002). However, if y is close to, but not identical to x, Geritz (2005) and Dercole and Geritz (2016) have shown that invasion without back-invasion generically implies substitution, and mutual invasion generically leads to coexistence in a broad class of population models. They found that the generic population dynamical outcomes of an invasion event of similar strategies are determined by invasion criteria alone, i.e., without having to study the full generic dynamical system.

The focus of Geritz (2005) and Dercole and Geritz (2016) is on resident-invader dynamics of similar strategies in a constant environment, providing that the resident population in a large ecological community is at a hyperbolic attracting steady state. Recently, there has been considerable interest in questions related to evolution of phenotypic diversity in fluctuating environments (Kussell and Leibler 2005; Kisdi and Liu 2006; Geritz et al. 2007; Schreiber 2012b; Wakano and Iwasa 2013; Ripa and Dieckmann 2013; Melbinger and Vergassola 2015; Sæther and Engen 2015; Ferris and Best 2019). The purpose of this paper is to extend and generalize the results of Geritz (2005) and Dercole and Geritz (2016) to environmental fluctuations due to nonequilibrium population dynamics (e.g., cycle, quasiperiodic trajectory and chaos) or environmental stochasticity. This requires (i) considering an explicit model that takes account of environmental feedback and environmental stochasticity, and (ii) specifying the population dynamics of residents in fluctuating environments.

A population affects its environment, and the environment in turn affects the population. Such an environmental feedback loop characterizes the interaction of populations with their environments and plays a central role in their ecological and evolutionary dynamics (Metz et al. 1996, 2008; Meszéna and Metz 1999; Kisdi and Geritz 2016; Lion 2018b). Examples of feedback variables are: interacting species (e.g., predator, prey and food); different classes in a population who is structured by age, sex, spatial location, etc; and physical variables (e.g., temperature and humidity) which have been shown that they may provide a feedback between ecological communities and local physical patterns in some concrete examples (see the Introduction of Benaïm and Schreiber (2019) and references therein). These feedback variables need to be modelled with their own dynamics. Environmental stochasticity, however, generally refers to effects of fluctuations in external factors (which can be biotic or abiotic) on model parameters (Nisbet and Gurney 1982; Kliemann 1983; Chesson 1986; Tuljapurkar 1990). These external factors drive the dynamics of the aforementioned feedback

variables but not *vice versa*. This is what differentiates environmental stochasticity from environmental feedback.

A population of a resident strategy in a fluctuating environment generally requires to be non-growing on the long-run and stochastically persistent. The non-growing property means that the long-term growth rate of the resident population in a given environment is zero. Stochastic persistence is an important concept to characterize population dynamics in fluctuating environments, which captures the property that if initially present, even if only in small size, then it will be present over arbitrary long periods of time in a stochastic sense. There are serval different (not necessarily equivalent) definitions of stochastic persistence (see e.g., a review in Schreiber 2012a). In this paper, we use the definition of Schreiber et al. (2011) and Benaïm (2018) which asserts that the probability of a population being near extinction is arbitrarily small.

The rest of this paper is organized as follows. In Sect. 2, we focus on the analysis of resident-invader dynamics of a class of unstructured population models. In Sect. 2.1, a class of polymorphic unstructured population models for multi-dimensional strategies and an explicit formulation of environmental feedback and environmental stochasticity is presented. In Sect. 2.2, we review some basic concepts used in this paper. In Sect. 2.3, we specify when a population of a given strategy is a resident. Section 2.4 gives the definition of invasion fitness. In Sect. 2.5, we present the basic analysis of invasion dynamics of similar strategies. Particularly, in Sects. 2.5.2 and 2.5.3, we classify the generic population dynamical outcomes of an invasion event. In Sect. 3, we illustrate how to generalize the results of the unstructured population models to a class of structured population models. In Sect. 4, we apply our results to examples of evolving bacteria in a chemostat, Lotka-Volterra competition, structured SIRS epidemic dynamics, and the evolution of timidity of the prey in a prey-predator model. The first two examples are designed to highlight different ways that our results can be used. The third example is designed to illustrate how a concrete structured population model can be reformulated into our framework. The last example is designed to demonstrate that our results are applicable to the evolution with the non-equilibrium resident dynamics. In Sect. 5, we conclude with discussing how our results relate to the existing literature and what possible generalizations are. Proofs are given in the "Appendix".

Before proceeding with the analysis, we introduce some default notations used in this paper. Let $\mathcal{X}$ be the strategy space, $\mathcal{N}$ be the space of non-negative population sizes, $\mathcal{E}$ be the space of feedback variables and Θ be the space of external factors. Each of these spaces is assumed to be a subset of a normed vector space. $f^{(i)}$ is the ith derivative of function f with respect to its first argument, and $f^{(i,j)}$ is the ith derivative of function f with respect to its first argument and the jth derivative with respect to its second argument. The superscript $\top$ means transpose of a vector or a matrix.

2 Unstructured population models

2.1 The polymorphic population model

Consider a population of strategies $x_1, \ldots, x_k \in \mathcal{X}$ whose sizes at time t are given by $n_{1,t}, \ldots, n_{k,t} \in \mathcal{N}$, respectively. The population of $x_1, \ldots, x_k$ interact with $e_t \in \mathcal{E}$ that describes the feedback environment. Generally, e_t may correspond to interacting species, different classes of the population, physical factors, or any combination thereof. $n_{1,t}, \ldots, n_{k,t}$ as well as e_t may be influenced by a stochastic driver $\theta_t \in \Theta$ that describes external factors. The per-capita growth rate f of individuals with strategy x_i depends on the current environmental condition (e_t, θ_t). Under these considerations, the dynamics of the polymorphic unstructured population are given by

$$\dot{n}_{i,t} = f(x_i, e_t, \theta_t) n_{i,t}, \quad i = 1, \ldots, k \tag{2.1a}$$

where the dot denotes differentiation with respect to time variable t. Notice that for a given environmental condition (e_t, θ_t), (2.1a) is linear in the population size. All the nonlinearity of (2.1a) comes from how the environment e_t depends on the population sizes.

Before introducing the formulation of the dynamics of e_t, let us look at how the environment e_t depends on the population sizes in the following two deterministic models.

Example 1 Lehtinen and Geritz (2019) studied the evolution of timidity in a prey species whose predator has cannibalistic tendencies, in which the population dynamics is given by

$$\text{(prey } i) \qquad \dot{n}_{i,t} = n_{i,t}^F g\left(\sum_j n_{j,t}^F\right) - \mu n_{i,t} - \beta n_{i,t}^F p_t^S,$$
$$\text{(adult predator) } \dot{p}_t = \tfrac{1}{T} J_t - \delta p_t,$$

where the reproduction of prey i is limited by competition (of resources, territories, breeding sites, etc.) among the foraging prey $n_{i,t}^F = \frac{n_{i,t}}{1 + x_i p_t}$ with the per-capita birth rate $g\left(\sum_j n_{j,t}^F\right) = a - c \sum_j n_{j,t}^F$, a searching adult predator $p_t^S = \frac{p_t}{1 + \beta h \sum_j n_{j,t}^F}$ captures the foraging prey at the rate β, the maturation of juvenile predators $J_t = \frac{\gamma \beta p_t^S \sum_j n_{j,t}^F}{\sigma + (1 - \lambda) \alpha p_t^S}$ with the mean time T depends on the cannibalistic pressure they experience throughout their juvenile period, and the per-capita death rate of the adult predator and the predation-independent per-capita death rate of the prey are δ and μ, respectively. The prey strategies differ only in their level of timidity $x_i \in \mathbb{R}_+ := \mathcal{X}$. Other parameters are all positive constant. Let

$$\text{(adult predators) } e_{1,t} = p_t,$$
$$\text{(foraging preys) } e_{2,t} = \sum_j n_{j,t}^F = \sum_j \frac{1}{1 + x_j e_{1,t}} n_{j,t},$$

then

$$\dot{e}_{1,t} = \sum_j \frac{\gamma \left(1 + \beta h e_{2,t}\right) \frac{\beta e_{1,t}}{(1+\beta h e_{2,t})(1+x_j e_{1,t})}}{T\left(\sigma(1 + \beta h e_{2,t}) + (1 - \lambda)\alpha e_{1,t}\right)} n_{i,t} - \delta e_{1,t}.$$

Since $e_{1,t}$ is an implicit function of the prey sizes and $e_{2,t}$ is an explicit function of $e_{1,t}$, the environmental variable $e_{2,t}$ becomes nonlinear in the prey sizes. Thus, the per-capita population growth rate f of prey type i in the nonlinear environment $(e_{1,t}, e_{2,t}) := e_t$ is given by

$$f(x_i, e_t) = \frac{g(e_{2,t})}{1 + x_i e_{1,t}} - \mu - \frac{\beta e_{1,t}}{(1 + x_i e_{1,t})(1 + \beta h e_{2,t})}.$$

Example 2 Dercole and Rinaldi (2002) and Dercole (2003) studied the evolution of cannibalistic strategies, in which the interactions between k cannibalistic consumer sub-populations with sizes $n_{i,t}$ and strategies $x_i \in \mathbb{R}_+ := \mathcal{X}$, $i = 1, \ldots, k$, are described by

$$\begin{aligned}
\text{(type } i) \quad \dot{n}_{i,t} = n_{i,t} \Bigg(&\gamma \frac{a_0(x_i)n_0 + \sum_j a(x_i, x_j)n_{j,t}}{1 + h_0(x_i)a_0(x_i)n_0 + \sum_j h(x_i)a(x_i, x_j)n_{j,t}} \\
&- \sum_j \frac{a(x_j, x_i)n_{j,t}}{1 + h_0(x_j)a_0(x_j)n_0 + \sum_\ell h(x_j)a(x_j, x_\ell)n_{\ell,t}} \\
&- c \sum_j n_{j,t} \Bigg),
\end{aligned}$$

where the conversion factor γ, the competition coefficient c, and the size of a common resource n_0 are assumed to be positive constant, while the attack rate a (and a_0) and the handling time h (and h_0) depend upon the strategies. In the bracket of the equation, the first term is reproduction through the harvesting of the common resource and cannibalism toward the sub-population j, the second term is mortality due to cannibalism, and the last term is mortality due to competition. For a given strategy x_i, let

$$\begin{aligned}
\text{(harvested food)} \quad & e_{21,t}(x_i) = \sum_j a(x_i, x_j)n_{j,t}, \\
\text{(handling times)} \quad & e_{22,t}(x_i) = \sum_j h(x_i)a(x_i, x_j)n_{j,t}, \\
\text{(cannibalism)} \quad & e_{23,t}(x_i) = \sum_j \frac{a(x_j, x_i)}{1 + h_0(x_j)a_0(x_j)n_0 + e_{22,t}(x_j)}n_{j,t}, \\
\text{(competition)} \quad & e_{24,t} = \sum_j cn_{j,t}.
\end{aligned}$$

Here the environmental variables $e_{21,t}$, $e_{22,t}$ and $e_{24,t}$ are all linear in the population sizes, but the environmental variable $e_{23,t}$ is a nonlinear and explicit function of $e_{22,t}$ which leads to the nonlinearity of $e_{23,t}$ in the population sizes. Thus, the per-capita population growth rate f of sub-population i with a given strategy x_i in the nonlinear

environment $(e_{21,t}, e_{22,t}, e_{23,t}, e_{24,t}) := e_t$ is given by

$$f(x_i, e_t) = \gamma \frac{a_0(x_i)n_0 + e_{21,t}(x_i)}{1 + h_0(x_i)a_0(x_i) + e_{22,t}(x_i)} - e_{23,t}(x_i) - e_{24,t}$$

(N.B., $e_{21,t}$, $e_{22,t}$ and $e_{23,t}$ are functions of strategy x_i and the strategy of sub-population type i has infinitely many choices, which causes an infinite-dimensional environment e_t to appear. We will provide more interpretations for the case of infinite-dimensional environment below).

Following Diekmann et al. (2001) and Geritz (2005), the environment e_t is assumed to be a linear function of the population sizes

$$e_t = \sum_j I(x_j)n_{j,t},$$

where $I(x_j)$ describes the environmental impact of a single individual with strategy x_j. However, this assumption excludes many interesting models with nonlinear environments such as the previous two examples. In Example 1, the predator size necessarily is an environmental variable for the preys, and the fraction of time individual prey spend foraging depends on the adult predators. In Example 2, the harvested food, the total handling times, and the compete-caused and cannibalize-caused death are environmental variables for the sub-population i, in which the cannibalize-caused death depends on the total handling times. We note that different environmental variables may have different forms even in the same model. For instance, in Example 1, $e_{1,t}$ is govern by a differential equation but $e_{2,t}$ is a function of $e_{1,t}$. For these reasons, we describe the environment by implicit (differential) equations.

In this paper, we consider an explicit formulation of the dynamics of the feedback environment $e_t = (e_{1,t}, e_{2,t})$ as illustrated in Fig. 1, which is given by

$$
\begin{aligned}
\dot{e}_{1,t} &= G_1(e_t, \theta_t) + \sum_j H_1(x_j, e_t, \theta_t)n_{j,t}, \\
e_{2,t} &= G_2(e_t, \theta_t) + \sum_j H_2(x_j, e_t, \theta_t)n_{j,t}.
\end{aligned}
\tag{2.1b}
$$

The dynamics of $e_{1,t}$ is given by a differential equation, while $e_{2,t}$ is an implicit function of the current state of the system. The functions G_1 and G_2 describe the intrinsic dynamics of the virgin environment (the environment unaffected by the population of $x_1, \ldots, x_k$). Given the current environment, the functions H_1 and H_2 describe the environmental impact of a single individual with a given strategy. Consequently, the second terms of (2.1b) gives the total environmental impact of the entire population of $x_1, \ldots, x_k$. These functions all involve the current environmental condition (e_t, θ_t). Here we would like to highlight that $e_{2,t}$ is assumed to be unique defined by the implicit equation. As Example 2 and "Appendix A" show, H_2 can be a function of $e_{2,t}$, but $e_{2,t}$ is still unique defined. Alternatively, by allowing e_t to be infinite-dimensional, we can also incorporate it into (2.1b) (Geritz 2005). For instance, in Example 2 and the

model in Sect. 4.2, where the feedback variables at any time are a function of strategy and where $f(x_i, e_t, \theta_t)$ involves the evaluation of e_t in x_i (ref to "Appendix A" and Sect. 4.2, respectively).

To consider external factors, we assume that $\{\theta_t\}_{t \geq 0}$ is a continuous time Markov process which is defining on an underlying probability space and taking value in Θ. From the definition of Markov process, the future state is independent of the past and depends only the present, which is a natural consideration in reality and contains a large number of ecologically interpretable factors. There are many ways to construct a Markov process, e.g., a pure jump process and a process generated by stochastic differential equations. In this paper, we assume that the process $\{\theta_t\}_{t \geq 0}$ is governed by the following stochastic differential equation:

$$\dot{\theta}_t = A(\theta_t) + B(\theta_t)\dot{W}_t, \tag{2.1c}$$

where $\dot{W}_t$ is a white noise, namely W_t is a Brownian motion defined on a complete probability space $(\Omega, \mathscr{F}, \mathbb{P})$ with a filtration $\{\mathscr{F}_t\}_{t \geq 0}$ satisfying the usual conditions (i.e., it is right continuous and increasing while $\mathscr{F}_0$ contains all $\mathbb{P}$-null sets). (2.1c) can be calculated using one of two methods: Itô or Stratonovich, but in fact our results do not depend on the calculation methods used.

Model (2.1) is an extension of the deterministic population models used in Geritz (2005), Dercole and Rinaldi (2008) and Dercole and Geritz (2016), which is intuitive and interpretable in biology and covers a large class of ecological models with an explicit formulation of environmental feedback and environmental stochasticity. Meanwhile, model (2.1) can be viewed as a revision of the stochastic population model used in Kliemann (1983), but now takes account of strategies and feedback variables. From the application point of view, any complicated considerations of biological mechanisms can be formulated into functions f, G_1, H_1, G_2, H_2, A and B, and a certain level of smoothness is assumed to make them mathematically meaningful.

Regarding (2.1), we make the following assumptions:

A1 The Markov process $\{\theta_t\}_{t \geq 0}$ is ergodic with invariant probability measure ν on Θ.
A2 The partial derivatives $f^{(i,j)}$, $G_1^{(i)}$, $H_1^{(i,j)}$, $G_2^{(i)}$ and $H_2^{(i,j)}$ exit and are locally Lipschitz continuous and measurable in (e, θ) for all strategies and for all nonnegative integers i, $j \leq 2$. And functions A and B are locally Lipschitz continuous and measurable in θ.

Assumption **A1** indicates the process $\{\theta_t\}_{t \geq 0}$ is completely realized before we start to observe the dynamics of $(n_{1,t}, \ldots, n_{k,t}, e_t)$. Both meanings of ergodicity and invariance are given in Sect. 2.2. Process $\{n_{1,t}, \ldots, n_{k,t}, e_t\}_{t \geq 0}$ is usually not Markovian. By Assumption **A1**, however, process $\{n_{1,t}, \ldots, n_{k,t}, e_t, \theta_t\}_{t \geq 0}$ is Markovian under weak assumptions (ref to Arnold and Kliemann (1983, Lemma 2.1)). Thereby it allows us to utilize the excellent works in the theory of Markov processes for the analysis of the process $\{n_{1,t}, \ldots, n_{k,t}, e_t, \theta_t\}_{t \geq 0}$. Assumption **A2** is a technical requirement to guarantee the existence of chain rule and justify the series expansion.

2.2 Some basic concepts

In order to state our results, we review some basic concepts in Markov process and ergodic theory.

Throughout Sect. 2, let

$$Z_t = (n_{1,t}, \ldots, n_{k,t}, e_t, \theta_t),$$

which lies in space

$$\mathcal{Z} = \mathcal{N}^k \times \mathcal{E} \times \Theta.$$

For any Borel set $C \subset \mathcal{Z}$, let

$$\mathbb{P}_z\{Z_t \in C\} = \mathbb{P}\{Z_t \in C | Z_0 = z\}$$

be the probability of Z_t in C given $Z_0 = z \in \mathcal{Z}$. Let $\{\mathcal{T}_t\}_{t \geq 0}$ be the associated semigroup of Markov process $\{Z_t\}_{t \geq 0}$ on $\mathcal{Z}$ defined by conditional expectations:

$$\mathcal{T}_t h(z) = \mathbb{E}_z\big[h(Z_t)\big] = \mathbb{E}\big[h(Z_t)|Z_0 = z\big],$$

for bounded and measurable function $h : \mathcal{Z} \mapsto \mathbb{R}$. Assumption **A2** implies that the Markov semigroup $\{\mathcal{T}_t\}_{t \geq 0}$ is $\mathscr{C}_b(\mathcal{Z})$-*Feller* (Benaïm 2018), i.e., operator $\mathcal{T}_t$ takes bounded continuous functions h on $\mathcal{Z}$ to bounded continuous functions $\mathcal{T}_t h$ on $\mathcal{Z}$ for all $t \geq 0$. A probability measure μ on $\mathcal{Z}$ is *invariant* for $\{Z_t\}_{t \geq 0}$ (or $\{\mathcal{T}_t\}_{t \geq 0}$) if

$$\int_{\mathcal{Z}} \mathcal{T}_t h(z) \mu(dz) = \int_{\mathcal{Z}} h(z) \mu(dz),$$

for all bounded and measurable functions $h : \mathcal{Z} \mapsto \mathbb{R}$ and all $t \geq 0$. An invariant probability measure μ is *ergodic* if it can not be written as a non-trivial convex combination of other distinct invariant probability measures.

If Z_t initially follows the distribution of an ergodic probability measure μ, then Birkhoff's ergodic theorem tells us that, in simple terms, the time average of a function of the process along the trajectories exists with probability one and equals the space average. More precisely, for all measurable functions $h : \mathcal{Z} \mapsto \mathbb{R}$ with $\int_{\mathcal{Z}} |h(z)| \mu(dz) < +\infty$,

$$\lim_{t \to +\infty} \frac{1}{t} \int_0^t h(Z_s) ds = \int_{\mathcal{Z}} h(z) \mu(dz)$$

with probability one.

2.3 Resident and resident system

From (2.1a) with initial value $Z_0 = z \in \mathcal{Z}$, a straightforward calculation shows that the species x_i tends to increase in its population size if

$$\lim_{t \to +\infty} \frac{1}{t} \int_0^t f(x_i, e_s, \theta_s)ds > 0,$$

and tends to decrease in its population size if

$$\lim_{t \to +\infty} \frac{1}{t} \int_0^t f(x_i, e_s, \theta_s)ds < 0,$$

provided that the limit exists. Let μ be an invariant probability measure for process $\{Z_t\}_{t \geq 0}$. From Birkhoff's ergodic theorem,

$$\lim_{t \to +\infty} \frac{1}{t} \int_0^t f(x_i, e_s, \theta_s)ds = \lambda_{x_i}(\mu)$$

with probability one and for μ-almost every z, where $\lambda_{x_i}(\mu)$ is the *per-capita growth rate of species x_i with respect to μ*:

$$\lambda_{x_i}(\mu) = \int_{\mathcal{Z}} f(x_i, e, \theta)\mu(de, d\theta).$$

We now introduce two notations for further analysis. We define the *extinction set* of a population of strategies $x_1, \ldots, x_k$

$$\mathcal{Z}_0 = \left\{ z \in \mathcal{Z} : \min_i n_i = 0 \right\}$$

is the set which at least one species is absent, and the *η-neighborhood of the extinction set*

$$\mathcal{Z}_\eta = \left\{ z \in \mathcal{Z} : \min_i n_i \leq \eta \right\}$$

is the set which at least one species has a size less than η.

Next, we specify when a population of strategies $x_1, \ldots, x_k$ is a resident.

Definition 1 A population of strategies $x_1, \ldots, x_k$ with corresponding size $\{n_{1,t}, \ldots, n_{k,t}\}_{t \geq 0}$ in the environment $\{e_t, \theta_t\}_{t \geq 0}$ is called a *resident* with respect to μ if μ is an invariant probability measure for process $\{Z_t\}_{t \geq 0}$ satisfying $\mu(\mathcal{Z} \setminus \mathcal{Z}_0) = 1$. Meanwhile,

  (i) *(non-growing)* $\lambda_{x_i}(\mu) = 0$ for every $i \in \{1, \ldots, k\}$, and
  (ii) *(stochastically persistent)* for all $\varepsilon > 0$, there exists a $\eta > 0$ such that $\mu(\mathcal{Z}_\eta) \leq \varepsilon$.
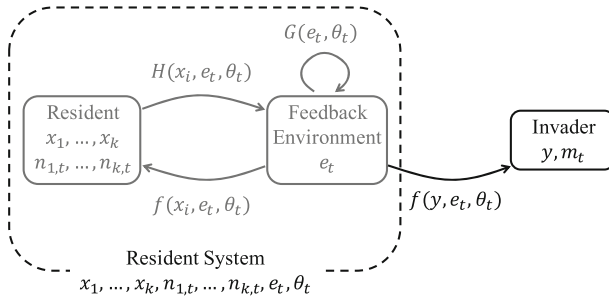
Fig. 1 Diagram of the interaction between the resident population and the feedback environment and the one-directional impact on the invader population

Definition 2 The multi-tuple $(x_1, \ldots, x_k, n_{1,t}, \ldots, n_{k,t}, e_t, \theta_t)$ is called a *resident system* with respect to μ if μ is an invariant probability measure for process $\{Z_t\}_{t \geq 0}$ with properties satisfying Definition 1.

In Definition 1, the supports of required invariant probability measures should exclude the extinction set $\mathcal{Z}_0$. The first reason is that it is the technical requirement to guarantee the properties of *non-growing* and *stochastically persistence*. The second reason is from a probabilistic perspective that if a population initially follows an invariant probability measure μ with $\mu(\mathcal{Z}_0) \neq 0$, then the extinction of the population is a non-zero probabilistic event on the long-term. From Definition 1, a resident is non-growing on the long-run and stochastically persistent, where the concept of persistence asserts that the probability of population size being near the extinction state is very small. In this paper, a resident (resp. a resident system) is characterized by an invariant probability measure defined on the state space. Actually, there might be more than one invariant probability measures, but we focus our attention on one of them. Each adequate invariant probability measure corresponds to a resident (resp. a resident system).

2.4 Invasion fitness

Assume that $(x_1, \ldots, x_k, n_{1,t}, \ldots, n_{k,t}, e_t, \theta_t)$ is a resident system with an invariant probability measure μ satisfying Definition 1, which further determines the ecological environment for a mutant type $y \in \mathcal{X}$ with initially infinitesimal size $m_t \in \mathcal{N}$. Thus, the population dynamics of a newly arrived and initially rare mutant y is described by the following linear differential equation

$$\dot{m}_t = f(y, e_t, \theta_t) m_t. \tag{2.2}$$

Notice that (2.2) holds as long as the mutant population remains rare. The tuple (y, m_t) is referred to an *invader*. Figure 1 shows the relationship of the invader (y, m_t) and the resident system $(x_1, \ldots, x_k, n_{1,t}, \ldots, n_{k,t}, e_t, \theta_t)$ where arrows with special indexes illustrate different impacts or contributions between them.

Formally, the *invasion fitness of mutant y in resident* $(x_1, \ldots, x_k)$ is given by

$$S_{x_1,\ldots,x_k}(y) := \lim_{t \to +\infty} \frac{\log m_t}{t} = \lim_{t \to +\infty} \frac{1}{t} \int_0^t f(y, e_s, \theta_s) ds = \lambda_y(\mu) \quad (2.3)$$

with probability one and for μ-almost every z and for every positive size m_0 (Metz et al. 1992; Ferriere and Gatto 1995). The invader y dies out if $S_{x_1,\ldots,x_k}(y) < 0$ and can spread if $S_{x_1,\ldots,x_k}(y) > 0$. From Definition 1, if $y = x_i$ for any $i \in \{1, \ldots, k\}$, then we have

$$S_{x_1,\ldots,x_k}(x_i) = \lambda_{x_i}(\mu) = 0 \quad (2.4)$$

which is so-called *the principle of selective neutrality of residents* in the framework of adaptive dynamics.

2.5 Population dynamics of similar strategies

Consider now a population of two similar strategies (i.e., two close but distinct strategies) $x_1, x_2 \in \mathcal{X}$ with corresponding sizes $n_{1,t}, n_{2,t} \in \mathcal{N}$ at time t. Without loss of generality, let strategies

$$x_1 = x + \epsilon\xi_1, \ x_2 = x + \epsilon\xi_2,$$

where $\xi_1 \neq \xi_2$ with $\|\xi_2 - \xi_1\| < +\infty$ and where small $\epsilon > 0$ (N.B., here ϵ is different from the ε used in Definition 1) and where x is a reference strategy. Biologically, ϵ can be viewed as the intensity of mutation, and ξ_i is the direction of mutation with respect to the reference strategy x. Model (2.1) applied to the case of $k = 2$ we further have a dimorphic population model. Obviously, $n_{1,t}, n_{2,t}$ and e_t are functions of x_1 and/or x_2. Thus they are functions of ϵ, which are denoted by $n_{1,t}^\epsilon, n_{2,t}^\epsilon$ and e_t^ϵ, respectively.

To study the dimorphic population dynamics of x_1 and x_2 in fluctuating environments, following Meszéna et al. (2005) and Dercole and Geritz (2016), we introduce the total population size N_t^ϵ of x_1 and x_2 at time t and the relative population size P_t^ϵ of x_2 at time t, i.e.,

$$N_t^\epsilon = n_{1,t}^\epsilon + n_{2,t}^\epsilon, \ P_t^\epsilon = \frac{n_{2,t}^\epsilon}{N_t^\epsilon}.$$

Under these assumptions, we arrive at the following equations:

$$\dot{N}_t^\epsilon = \left(f(x + \epsilon\xi_1, e_t^\epsilon, \theta_t)(1 - P_t^\epsilon) + f(x + \epsilon\xi_2, e_t^\epsilon, \theta_t)P_t^\epsilon \right) N_t^\epsilon, \quad (2.5a)$$

$$\dot{P}_t^\epsilon = P_t^\epsilon(1 - P_t^\epsilon)\left(f(x + \epsilon\xi_2, e_t^\epsilon, \theta_t) - f(x + \epsilon\xi_1, e_t^\epsilon, \theta_t) \right), \quad (2.5b)$$

$$\dot{e}_{1,t}^\epsilon = G_1(e_t^\epsilon, \theta_t)$$
$$\quad + H_1(x + \epsilon\xi_1, e_t^\epsilon, \theta_t)(1 - P_t^\epsilon)N_t^\epsilon + H_1(x + \epsilon\xi_2, e_t^\epsilon, \theta_t)P_t^\epsilon N_t^\epsilon, \quad (2.5c)$$

$$e_{2,t}^\epsilon = G_2(e_t^\epsilon, \theta_t)$$

$$+ H_2(x + \epsilon\xi_1, e_t^\epsilon, \theta_t)(1 - P_t^\epsilon)N_t^\epsilon + H_2(x + \epsilon\xi_2, e_t^\epsilon, \theta_t)P_t^\epsilon N_t^\epsilon, \quad (2.5d)$$

$$\dot{\theta}_t = A(\theta_t) + B(\theta_t)\dot{W}_t. \quad (2.5e)$$

Next, we unfold the dimorphic population dynamics of x_1 and x_2 based on the series expansion of model (2.5) in powers of ϵ, and on time-scale separation. In the following analysis, we will see that $(N_t^\epsilon, e_t^\epsilon, \theta_t)$ are fast variables with dynamics acted on t-timescale and P_t^ϵ is a slow variable with dynamics expounded on $\epsilon^i t$-timescales for $i = 1, 2, \ldots$.

2.5.1 Fast dynamics

For $\epsilon = 0$, we get

$$\begin{aligned}
\dot{N}_t^0 &= f(x, e_t^0, \theta_t)N_t^0, \\
\dot{P}_t^0 &= 0, \\
\dot{e}_{1,t}^0 &= G_1(e_t^0, \theta_t) + H_1(x, e_t^0, \theta_t)N_t^0, \\
e_{2,t}^0 &= G_2(e_t^0, \theta_t) + H_2(x, e_t^0, \theta_t)N_t^0, \\
\dot{\theta}_t &= A(\theta_t) + B(\theta_t)\dot{W}_t.
\end{aligned} \quad (2.6)$$

A comparison of (2.6) and the dynamics of the reference monomorphic system (x, n_t, e_t, θ_t) shows that (N_t^0, e_t^0, θ_t) has the same dynamics as (n_t, e_t, θ_t). Assume that (x, n_t, e_t, θ_t) is a resident system with associated invariant probability measures μ. Then, in the limit $\epsilon = 0$, $(N_t^\epsilon, e_t^\epsilon, \theta_t)$ enter the dynamics of the observables through μ, and P_t^ϵ becomes irrelevant. Thus, (2.5) is a stochastic fast-slow system with fast variables $(N_t^\epsilon, e_t^\epsilon, \theta_t)$ and slow variable P_t^ϵ.

For positive but small ϵ, the trajectories of $(N_t^\epsilon, P_t^\epsilon, e_t^\epsilon, \theta_t)$ can be viewed as a small variation of that of $(N_t^0, P_t^0, e_t^0, \theta_t)$ on t-timescale. Under some mild assumptions, the dynamics of (2.6) are arbitrarily good approximation of the dynamics of (2.5) in terms of conditional expectations when $\epsilon \to 0$ (ref to "Appendix B").

2.5.2 Slow dynamics on ϵt-timescale

To study the dynamics of the relative population size on slow timescales, we have to take account of the high-order terms in the series expansion.

Consider the first-order term in the series expansion of the right side of (2.5b) and let $t_1 = \epsilon t$, we have

$$\dot{P}_{t_1}^\epsilon = P_{t_1}^\epsilon(1 - P_{t_1}^\epsilon)f^{(1,0)}(x, e_{t_1/\epsilon}, \theta_{t_1/\epsilon})^\top(\xi_2 - \xi_1) + \mathcal{O}(\epsilon). \quad (2.7)$$

Here the term $e_{t_1/\epsilon}^0$ in the series expansion has been replaced by $e_{t_1/\epsilon}$ because they have the same dynamics on t-timescale (see the interpretation in Sect. 2.5.1). Notice that $f^{(1,0)}(x, e, \theta)^\top$ becomes a row vector if x is a multi-dimensional strategy. Since (x, n_t, e_t, θ_t) is assumed to be a resident system with associated invariant probability

measures μ satisfying $\mu(\mathcal{Z} \setminus \mathcal{Z}_0) = 1$, we can apply Birkhoff's ergodic theorem to an observation $f^{(1,0)}$ of (x, n_t, e_t, θ_t), i.e., we have

$$\lim_{T \to +\infty} \frac{1}{T} \int_0^T f^{(1,0)}(x, e_s, \theta_s)^\top ds = \int_{\mathcal{Z} \setminus \mathcal{Z}_0} f^{(1,0)}(x, e, \theta)^\top \mu(de, d\theta) \tag{2.8}$$
$$= \partial_y \mathcal{S}_x(y)|_{y=x}$$

with probability one and for μ-almost every $Z_0 = z \in \mathcal{Z} \setminus \mathcal{Z}_0$. Here $\partial_y \mathcal{S}_x(y)|_{y=x}$ is so called the *selection gradient* at the reference strategy x. A strategy x is an *evolutionarily singular strategy* if

$$\partial_y \mathcal{S}_x(y)|_{y=x} = 0.$$

(ref to Metz et al. 1996; Geritz et al. 1997, 1998, 1999).

Throughout this section, assuming that

$$\partial_y \mathcal{S}_x(y)|_{y=x}(\xi_2 - \xi_1) \neq 0.$$

Consider the following deterministic system

$$\dot{\bar{P}}_{t_1} = \bar{P}_{t_1}(1 - \bar{P}_{t_1})\partial_y \mathcal{S}_x(y)|_{y=x}(\xi_2 - \xi_1). \tag{2.9}$$

From the averaging principle for fast-slow systems with stochastic and stationary fast dynamics, solutions of (2.7) are well approximated by solutions of (2.9) on every finite time interval when $\epsilon \to 0$, where (2.9) is so-called the *averaged system* with respect to (2.7) (ref to Freidlin and Wentzell (2012, Theorem 1.3 of Chapter 2 and Theorem 2.1 of Chapter 7), or Khas'minskii 1966; Arnold 2001; Kifer 2001; Liu and Krstic 2012). In the averaged system (2.9), there are exactly two equilibria:

$$0, \ 1,$$

with stabilities dominated by the signs of $\partial_y \mathcal{S}_x(y)|_{y=x}(\xi_2 - \xi_1)$. Figure 2a gives all generic cases. Under the stability of the averaged system (2.9), we show the property of solutions of (2.7) in arbitrarily long time intervals.

Lemma 1 *Let* $P_0^\epsilon = \bar{P}_0 = p$ *and* $Z_0 = z$. *Assume that for a small* $\sigma > 0$,

$$\int_{\mathcal{Z} \setminus \mathcal{Z}_0} |f(y, e, \theta)|^2 \mu(de, d\theta) < +\infty \text{ for all } y \in \left\{ \tilde{x} \in \mathcal{X} : \|\tilde{x} - x\| \leq \sigma \right\},$$

and $\lim_{t_1 \to +\infty} \bar{P}_{t_1} = p^*$. *Then for every* $\delta \in (0, 1)$, *there exists a constant* $T_\delta > 0$ *such that the solution of (2.7) satisfies*

$$\mathbb{P}_{(p,z)} \left\{ \lim_{\epsilon \to 0} \sup_{t_1 \geq T_\delta} |P_{t_1}^\epsilon - p^*| \leq \delta \right\} = 1 \tag{2.10}$$

for all $p \in [0, 1]$ and for μ-almost every $z \in \mathcal{Z} \setminus \mathcal{Z}_0$, where $p^* = 0$ or 1.

In words, for $\epsilon > 0$ sufficiently small and at some later time, the trajectories of $P_{t_1}^\epsilon$ will remain δ-close to the stable equilibrium of the averaged system (2.9) for arbitrarily long time with probability one.

Remark 1 In the hypotheses of Lemma 1, the boundedness ensures that the invasion fitness $\mathcal{S}_x(y)$ for a mutant strategy y close to the reference resident strategy x and the selection gradient $\partial_y \mathcal{S}_x(y)|_{y=x}$ are well defined, and it further guarantees the approximation of solutions of (2.7) and (2.9) in arbitrarily long time intervals.

Expansions of invasion fitnesses $\mathcal{S}_{x_1}(x_2)$ and $\mathcal{S}_{x_2}(x_1)$ up to the first order terms in ϵ give

$$\mathcal{S}_{x_1}(x_2) = -\mathcal{S}_{x_2}(x_1) = \epsilon \partial_y \mathcal{S}_x(y)|_{y=x}(\xi_2 - \xi_1). \tag{2.11}$$

Using (2.11), a stochastic analogy of the *invasion implies fixation theorem* of Geritz (2005, Proposition 1) immediately follows from Lemma 1, which is called as *invasion implies substitution theorem* in the present paper.

Theorem 1 *("Invasion implies substitution theorem") Assume that the hypotheses of Lemma 1 are satisfied. Let the strategy pair (x_1, x_2) and the reference strategy x are such that $\partial_y \mathcal{S}_x(y)|_{y=x}(\xi_2 - \xi_1) \neq 0$.*

(i) If $\mathcal{S}_{x_1}(x_2) > 0$ and $\mathcal{S}_{x_2}(x_1) < 0$, then (2.10) holds for $p^* = 1$.
(ii) If $\mathcal{S}_{x_1}(x_2) < 0$ and $\mathcal{S}_{x_2}(x_1) > 0$, then (2.10) holds for $p^* = 0$.

For $\mathcal{X} \subset \mathbb{R}$, provided that $\partial_y \mathcal{S}_x(y)|_{y=x} \neq 0$ (i.e., x isn't an evolutionarily singular strategy), the dimorphic population dynamics of x_1 and x_2 is completely determined by the invasion fitnesses. Since there are only two equilibria that are at the boundary for the averaged system (2.9), invasion implies substitution in terms that the dynamics of the relative population size satisfies (2.10). In other words, similar strategies cannot coexist when both of them are away from an evolutionarily singular strategy.

For $\mathcal{X} \subset \mathbb{R}^d$ with $d \geq 2$, geometrically, the deviation $\xi_2 - \xi_1$ (which is equivalent to $x_2 - x_1$) shall be away from the null-set of the linear form $\alpha \rightarrow \partial_y \mathcal{S}_x(y)|_{y=x}\alpha$ for all $\alpha \neq 0$ and $\partial_y \mathcal{S}_x(y)|_{y=x}$ non-vanishes, so that (2.9) is non-degenerate (i.e., the right side isn't equal to zero). Consequently, in a multi-dimensional strategy space, invasion implies substitution provided that not only the two strategies are away from an evolutionarily singular strategy, but also their deviation isn't orthogonal to the selection gradient at the reference strategy.

The averaged system (2.9) gives a geometric interpretation of the population dynamics of similar strategies in terms of their deviation and the selection gradient at a reference strategy, which indicates the sufficient condition of substitution of similar strategies. Thus, on t_1-timescale, any invader that can spread in a given resident system will ultimately oust the resident.

2.5.3 Slow dynamics on $\epsilon^2 t$-timescale

Once $\partial_y \mathcal{S}_x(y)|_{y=x}(\xi_2 - \xi_1) = 0$, the right side of the averaged system (2.9) becomes zero. Then the high-order term $\mathcal{O}(\epsilon)$ in (2.7) plays a role on slower timescales.

Consider the second-order term in the series expansion of the right side of (2.5b) and let $t_2 = \epsilon^2 t$, we have the associated averaged system:

$$\dot{\bar{P}}_{t_2} = \bar{P}_{t_2}(1 - \bar{P}_{t_2})\bigg((\xi_2 + \xi_1)^\top (\mathcal{C}_{22} + \mathcal{C}_{21})(\xi_2 - \xi_1)$$
$$+ \Big(\frac{1}{2} - \bar{P}_{t_2}\Big)(\xi_2 - \xi_1)^\top (\mathcal{C}_{22} + \mathcal{C}_{11})(\xi_2 - \xi_1)\bigg) \qquad (2.12)$$

with matrixes

$$\mathcal{C}_{11} = \mathcal{C}_{11}^\top = \frac{1}{2}\partial_{xx}\mathcal{S}_x(y)|_{y=x}, \ \mathcal{C}_{12} = \mathcal{C}_{21}^\top = \frac{1}{2}\partial_{xy}\mathcal{S}_x(y)|_{y=x},$$
$$\mathcal{C}_{21} = \mathcal{C}_{12}^\top = \frac{1}{2}\partial_{yx}\mathcal{S}_x(y)|_{y=x}, \ \mathcal{C}_{22} = \mathcal{C}_{22}^\top = \frac{1}{2}\partial_{yy}\mathcal{S}_x(y)|_{y=x}. \qquad (2.13)$$

For the reader's convenience, we give the detailed derivation of (2.12) in "Appendix D". Notice that $\{\mathcal{C}_{ij}\}_{i,j=1,2}$ are matrices if x is a multi-dimensional strategy.

Throughout this section, assuming that

$$(\xi_2 + \xi_1)^\top (\mathcal{C}_{22} + \mathcal{C}_{21})(\xi_2 - \xi_1) \text{ and } (\xi_2 - \xi_1)^\top (\mathcal{C}_{22} + \mathcal{C}_{11})(\xi_2 - \xi_1)$$

don't vanish at the same time. If $(\xi_2 - \xi_1)^\top (\mathcal{C}_{22} + \mathcal{C}_{11})(\xi_2 - \xi_1) \neq 0$, then the averaged system (2.12) has an interior equilibrium. In case of $(\xi_2 - \xi_1)^\top (\mathcal{C}_{22} + \mathcal{C}_{11})(\xi_2 - \xi_1) = 0$, only boundary equilibria exist. Figure 2b and c conclude all possible equilibria and their stability.

From the series expansions of invasion fitnesses $\mathcal{S}_{x_1}(x_2)$ and $\mathcal{S}_{x_2}(x_1)$, the second-order terms in ϵ give

$$\mathcal{S}_{x_1}(x_2)(1 - \bar{P}_{t_2}) - \mathcal{S}_{x_2}(x_1)\bar{P}_{t_2}$$
$$= \epsilon^2\Big(\xi_2^\top \mathcal{C}_{22}\xi_2 - \xi_1^\top \mathcal{C}_{22}\xi_1 + 2\big((1 - \bar{P}_{t_2})\xi_1 + \bar{P}_{t_2}\xi_2\big)^\top \mathcal{C}_{21}(\xi_2 - \xi_1)\Big)$$
$$= \epsilon^2\bigg((\xi_2 + \xi_1)^\top (\mathcal{C}_{22} + \mathcal{C}_{21})(\xi_2 - \xi_1) \qquad (2.14)$$
$$+ \Big(\frac{1}{2} - \bar{P}_{t_2}\Big)(\xi_2 - \xi_1)^\top (\mathcal{C}_{22} + \mathcal{C}_{11})(\xi_2 - \xi_1)\bigg),$$

where the derivation of the last equality can be found in "Appendix D". Using (2.14), the averaged system (2.12) have at most three equilibria in $[0, 1]$:

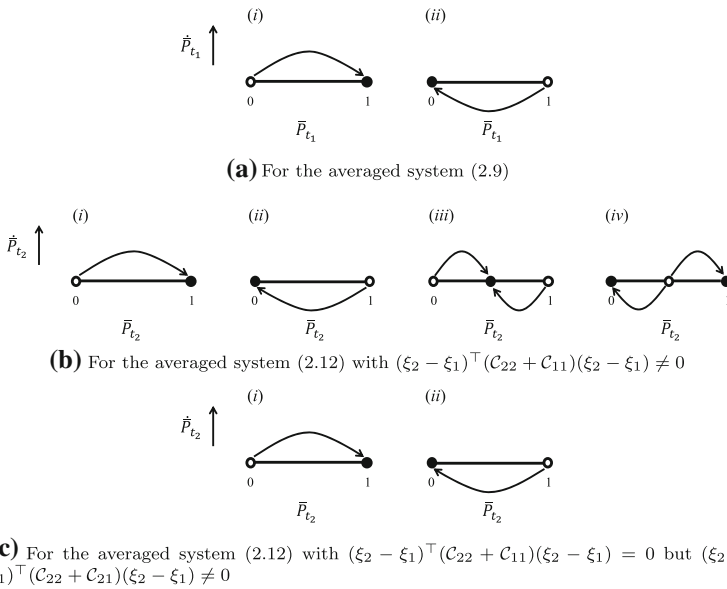$$0, \ \frac{\mathcal{S}_{x_1}(x_2)}{\mathcal{S}_{x_1}(x_2) + \mathcal{S}_{x_2}(x_1)}, \ 1,$$

$\dot{\bar{P}}_{t_1}$ (i) (ii)

$\bar{P}_{t_1}$ $\bar{P}_{t_1}$

(a) For the averaged system (2.9)

$\dot{\bar{P}}_{t_2}$ (i) (ii) (iii) (iv)

$\bar{P}_{t_2}$ $\bar{P}_{t_2}$ $\bar{P}_{t_2}$ $\bar{P}_{t_2}$

(b) For the averaged system (2.12) with $(\xi_2 - \xi_1)^\top (\mathcal{C}_{22} + \mathcal{C}_{11})(\xi_2 - \xi_1) \neq 0$

$\dot{\bar{P}}_{t_2}$ (i) (ii)

$\bar{P}_{t_2}$ $\bar{P}_{t_2}$

(c) For the averaged system (2.12) with $(\xi_2 - \xi_1)^\top (\mathcal{C}_{22} + \mathcal{C}_{11})(\xi_2 - \xi_1) = 0$ but $(\xi_2 + \xi_1)^\top (\mathcal{C}_{22} + \mathcal{C}_{21})(\xi_2 - \xi_1) \neq 0$

Fig. 2 Equilibria and their stability of the associated averaged systems on different timescales, where "∘" denotes an unstable equilibrium and "•" denotes a stable equilibrium

with stabilities dominated by the combination of signs of $\mathcal{S}_{x_1}(x_2)$ and $\mathcal{S}_{x_2}(x_1)$. Under the stability of the averaged system (2.12), we have a similar consequence as Lemma 1.

Lemma 2 *Let* $P_0^\epsilon = \bar{P}_0 = p$ *and* $Z_0 = z$. *Assume that for a small* $\sigma > 0$,

$$\int_{\mathcal{Z} \setminus \mathcal{Z}_0} |f(y, e, \theta)|^2 \mu(de, d\theta) < +\infty \text{ for all } y \in \{\tilde{x} \in \mathcal{X} : \|\tilde{x} - x\| \leq \sigma\}.$$

Then for every $\delta \in (0, 1)$, *there exists a constant* $T_\delta > 0$ *such that*

(i) *if the averaged system* (2.12) *has a unique stable equilibrium, then the solution of* (2.12) *satisfies*

$$\mathbb{P}_{(p,z)} \left\{ \lim_{\epsilon \to 0} \sup_{t_2 \geq T_\delta} |P_{t_2}^\epsilon - p^*| \leq \delta \right\} = 1 \tag{2.15}$$

for all $p \in [0, 1]$ *and for* μ-*almost every* $z \in \mathcal{Z} \setminus \mathcal{Z}_0$, *where* $p^* = 0$, *or* $\frac{\mathcal{S}_{x_1}(x_2)}{\mathcal{S}_{x_1}(x_2) + \mathcal{S}_{x_2}(x_1)}$ *or 1;*

(ii) *if the averaged system* (2.12) *has two stable equilibria (i.e., 0 and 1), then we have*

$$q_{(p,z)}^0 + q_{(p,z)}^1 = 1 \tag{2.16}$$

for all $p \in [0, 1]$ and for μ-almost every $z \in \mathcal{Z} \setminus \mathcal{Z}_0$, where probabilities

$$q^0_{(p,z)} = \mathbb{P}_{(p,z)} \left\{ \lim_{\epsilon \to 0} \sup_{t_2 \geq T_\delta} |P^\epsilon_{t_2} - 0| \leq \delta \right\},$$

$$q^1_{(p,z)} = \mathbb{P}_{(p,z)} \left\{ \lim_{\epsilon \to 0} \sup_{t_2 \geq T_\delta} |P^\epsilon_{t_2} - 1| \leq \delta \right\}.$$

In words, for $\epsilon > 0$ sufficiently small and at some later time, the trajectories of $P^\epsilon_{t_2}$ will remain δ-close to the attractor (either a unique stable equilibrium or the bistable equilibria) of the averaged system (2.12) for arbitrarily long time with probability one.

Remark 2 In the hypotheses of Lemma 2, the boundedness ensures that (2.13) are well defined, and it further guarantees the approximation of solutions of $P^\epsilon_{t_2}$ and $\bar{P}_{t_2}$ in arbitrarily long time intervals. The boundedness is a technical requirement for Lemmas 1 and 2, which meets by many models.

Let $(\xi_2 - \xi_1)^\top (\mathcal{C}_{22} + \mathcal{C}_{11})(\xi_2 - \xi_1) \neq 0$. Under the stability of the averaged system (2.12) with equivalent form (2.14), a stochastic analogy of the *classification theorem* of Geritz (2005, Proposition 2) immediately follows from Lemma 2.

Theorem 2 ("Classification theorem") *Assume that the hypotheses of Lemma 2 are satisfied. Let the strategy pair (x_1, x_2) and the reference strategy x are such that $\partial_y \mathcal{S}_x(y)|_{y=x}(\xi_2 - \xi_1) = 0$ but $(\xi_2 - \xi_1)^\top (\mathcal{C}_{22} + \mathcal{C}_{11})(\xi_2 - \xi_1) \neq 0$.*

  (i) If $\mathcal{S}_{x_1}(x_2) > 0$ and $\mathcal{S}_{x_2}(x_1) < 0$, then (2.15) holds for $p^* = 1$.
  (ii) If $\mathcal{S}_{x_1}(x_2) < 0$ and $\mathcal{S}_{x_2}(x_1) > 0$, then (2.15) holds for $p^* = 0$.
  (iii) If $\mathcal{S}_{x_1}(x_2) > 0$ and $\mathcal{S}_{x_2}(x_1) > 0$, then (2.15) holds for $p^* = \frac{\mathcal{S}_{x_1}(x_2)}{\mathcal{S}_{x_1}(x_2) + \mathcal{S}_{x_2}(x_1)}$.
 *(iv) If $\mathcal{S}_{x_1}(x_2) < 0$ and $\mathcal{S}_{x_2}(x_1) < 0$, then (2.16) holds.*

Let $(\xi_2 - \xi_1)^\top (\mathcal{C}_{22} + \mathcal{C}_{11})(\xi_2 - \xi_1) = 0$. Then the averaged system (2.12) only has boundary equilibria. From the assertion (i) of Lemma 2, we have a special invasion-substitution theorem.

Theorem 3 ("Special invasion-substitution theorem") *Assume that the hypotheses of Lemma 2 are satisfied. Let the strategy pair (x_1, x_2) and the reference strategy x are such that $\partial_y \mathcal{S}_x(y)|_{y=x}(\xi_2 - \xi_1) = 0$ and $(\xi_2 - \xi_1)^\top (\mathcal{C}_{22} + \mathcal{C}_{11})(\xi_2 - \xi_1) = 0$ but $(\xi_2 + \xi_1)^\top (\mathcal{C}_{22} + \mathcal{C}_{21})(\xi_2 - \xi_1) \neq 0$.*

  (i) If $\mathcal{S}_{x_1}(x_2) > 0$ and $\mathcal{S}_{x_2}(x_1) < 0$, then (2.15) for $p^* = 1$.
  (ii) If $\mathcal{S}_{x_1}(x_2) < 0$ and $\mathcal{S}_{x_2}(x_1) > 0$, then (2.15) for $p^* = 0$.

For $\mathcal{X} \subset \mathbb{R}$, Theorem 2 corresponds to the invasion dynamics of similar strategies close to a generic evolutionarily singular strategy that satisfies $\mathcal{C}_{22} + \mathcal{C}_{11} \neq 0$, while Theorem 3 corresponds to the invasion dynamics of similar strategies with $\xi_1 + \xi_2 \neq 0$ close to an evolutionarily singular strategy satisfying $\mathcal{C}_{22} + \mathcal{C}_{11} = 0$ and $\mathcal{C}_{22} + \mathcal{C}_{21} \neq 0$. In the neighborhood of a generic evolutionarily singular strategy, the invasion dynamics is essentially "Lotka-Volterra"—dominance of one strategy, coexistence of two strategies, and mutual exclusion of two strategies—in sense that the relative

Fig. 3 The geometric interpretation of the coexistence condition in a one-dimensional (or a one-dimensional parameterization of a multi-dimensional) strategy space. The "+" in the Pairwise Invasibility Plot (PIP) indicates the area where the mutant can invade the resident, and the "−" indicates the area where the mutant cannot invade
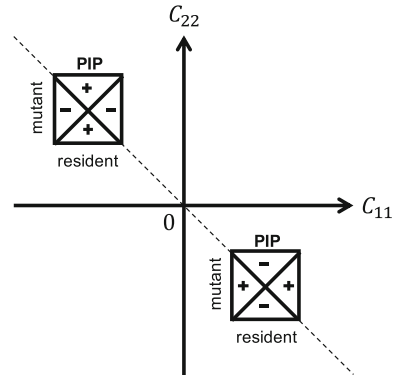
population size remains δ-close to the attractor of the averaged system for arbitrarily long time with probability one. In the neighborhood of an evolutionarily singular strategy satisfying $C_{22} + C_{11} = 0$ and $C_{22} + C_{21} \neq 0$, only substitution happens for the invasion dynamics of any two similar but non-opposite strategies (i.e., $\xi_1 + \xi_2 \neq 0$). In Fig. 3, the dashed line indicates the scenario $C_{22} + C_{11} = 0$ associated with the local configuration of the Pairwise Invasibility Plot (Matsuda 1985; van Tienderen and de Jong 1986). Invasion implies substitution if (C_{11}, C_{22}) is on the dash line. Coexistence of similar strategies may occur if (C_{11}, C_{22}) is away from the dash line.

For $\mathcal{X} \subset \mathbb{R}^d$ with $d \geq 2$, the hypothesis $(\xi_2 - \xi_1)^\top (C_{22} + C_{11})(\xi_2 - \xi_1) \neq 0$ in Theorem 2 means that the deviation $\xi_2 - \xi_1$ shall be away from the null-set of the quadratic form $\alpha \mapsto \alpha^\top (C_{22} + C_{11})\alpha$ for all $\alpha \neq 0$. Apart from $\partial_y \mathcal{S}_x(y)|_{y=x}(\xi_2 - \xi_1) = 0$, Theorem 3 requires that $(\xi_2 - \xi_1)^\top (C_{22} + C_{11})(\xi_2 - \xi_1) = 0$ and $(\xi_2 + \xi_1)^\top (C_{22} + C_{21})(\xi_2 - \xi_1) \neq 0$. The correspondingly geometric interpretation is complicate. However, if let $\xi_1 = 0$ (i.e., $x_1 = x$), the two hypotheses are equivalent to $\xi_2^\top (C_{22} + C_{11})\xi_2 = 0$ and $\xi_2^\top (C_{22} + C_{21})\xi_2 \neq 0$. From a conservation law $C_{11} + C_{22} + C_{21} + C_{12} = 0$ (ref to (d.9) in "Appendix D"), we have $C_{22} + C_{21} = \frac{1}{2}(C_{22} - C_{11}) + \frac{1}{2}(C_{21} - C_{12})$ where $C_{22} - C_{11}$ is symmetric and $C_{21} - C_{12}$ is skew-symmetric. Then by the property of skew-symmetric matrices, $\alpha^\top (C_{21} - C_{12})\alpha = 0$ for all α, we obtain that $\xi_2^\top (C_{22} + C_{21})\xi_2 = \frac{1}{2}\xi_2^\top (C_{22} - C_{11})\xi_2$. Thus, if $\xi_2 \neq 0$ is such that $\xi_2^\top (C_{22} + C_{11})\xi_2 = 0$ but $\xi_2^\top C_{11}\xi_2 \neq 0$ and $\xi_2^\top C_{22}\xi_2 \neq 0$, then $\xi_2^\top (C_{22} + C_{21})\xi_2 \neq 0$ is readily satisfied.

Remark 3 Once $(\xi_2 - \xi_1)^\top (C_{22} + C_{11})(\xi_2 - \xi_1) = 0$ and $(\xi_2 + \xi_1)^\top (C_{22} + C_{21})(\xi_2 - \xi_1) = 0$, the higher order terms in ϵ of the series expansion are imperative to reveal the invasion dynamics of similar strategies. The degeneracy of a higher degree opens the possibility of unprotected coexistence for two strategies (ref to Priklopil (2012) and Dercole and Geritz (2016) for population models with point equilibria).

3 Structured population models

In this section, we illustrate how to generalize the results of the unstructured population models to a class of structured population models.

Consider a structured population of strategies $x_1, \ldots, x_k \in \mathcal{X}$ with corresponding sizes $n_{1,t}, \ldots, n_{k,t} \in \mathcal{N}$ at time t where $n_{i,t} = (n_{i,t}^1, \ldots, n_{i,t}^\ell)^\top$ for $\ell \geq 2$ and for $i = 1, \ldots, k$. Under the same considerations in Sect. 2.1, our polymorphic structured population model is

$$
\begin{aligned}
\dot{n}_{i,t} &= F(x_i, e_t, \theta_t) n_{i,t}, \quad i = 1, \ldots, k \\
\dot{e}_{1,t} &= G_1(e_t, \theta_t) + \sum_j H_1(x_j, e_t, \theta_t) n_{j,t}, \\
e_{2,t} &= G_2(e_t, \theta_t) + \sum_j H_2(x_j, e_t, \theta_t) n_{j,t}, \\
\dot{\theta}_t &= A(\theta_t) + B(\theta_t) \dot{W}_t,
\end{aligned}
\tag{3.1}
$$

where $F(x_i, e_t, \theta_t)$ is a $\ell \times \ell$ matrix whose entries are transition rates between different states of individuals with strategy x_i and depending on the current environmental condition (e_t, θ_t). Model (3.1) is an extension of the deterministic models used in Durinx et al. (2008) and Priklopil and Lehmann (2019), but now with an explicit formulation of environmental feedback and environmental stochasticity.

To place (3.1) into our framework, let $\|n_{i,t}\|_1$ be the total population size of x_i, and let $v_{i,t}$ be the proportion of each component in the population of x_i, i.e., for $i = 1, \ldots, k$,

$$
\|n_{i,t}\|_1 = \sum_j n_{i,t}^j, \quad v_{i,t} = (v_{i,t}^1, \cdots, v_{i,t}^\ell)^\top = \left(\frac{n_{i,t}^1}{\|n_{i,t}\|_1}, \cdots, \frac{n_{i,t}^\ell}{\|n_{i,t}\|_1} \right)^\top.
$$

From (3.1), we obtain that for $i = 1, \ldots, k$,

$$
\begin{aligned}
\|n_{i,t}\|_1^\cdot &= \mathbb{1}^\top F(x_i, e_t, \theta_t) v_{i,t} \|n_{i,t}\|_1, \\
\dot{v}_{i,t} &= \big(F(x_i, e_t, \theta_t) - \mathbb{1}^\top F(x_i, e_t, \theta_t) v_{i,t} \mathbb{I} \big) v_{i,t}, \\
\dot{e}_{1,t} &= G_1(e_t, \theta_t) + \sum_j H_1(x_j, e_t, \theta_t) v_{j,t} \|n_{j,t}\|_1, \\
e_{2,t} &= G_2(e_t, \theta_t) + \sum_j H_2(x_j, e_t, \theta_t) v_{j,t} \|n_{j,t}\|_1, \\
\dot{\theta}_t &= A(\theta_t) + B(\theta_t) \dot{W}_t,
\end{aligned}
\tag{3.2}
$$

where $\mathbb{1}^\top = (1, \ldots 1)$ and $\mathbb{I}$ is the identity matrix. Although (3.2) cannot be written in the form of (2.1), the results in Sect. 2 still can be generalized to model (3.2) where population structures $v_{1,t}, \ldots, v_{k,t}$ are viewed as auxiliary variables that lie in the space $\Delta = \{ v \in [0, 1]^\ell : \|v\|_1 = 1 \}$. Throughout this section, denote $Z_t = \big(\|n_{1,t}\|_1, \ldots, \|n_{k,t}\|_1, v_{1,t}, \ldots, v_{k,t}, e_t, \theta_t \big)$. The state space of the dynamics of Z_t is $\mathcal{Z} = \mathbb{R}_+^k \times \Delta^k \times \mathcal{E} \times \Theta$, the extinction set is $\mathcal{Z}_0 = \{ z \in \mathcal{Z} : \min_i \|n_i\|_1 = 0 \}$, and the η-neighborhood of the extinction set is $\mathcal{Z}_\eta = \{ z \in \mathcal{Z} : \min_i \|n_i\|_1 \leq \eta \}$. Like Assumption **A2**, the partial derivative $F^{(i,j)}$ requires to exit and to be locally Lipschitz

continuous and measurable in (e, θ) for all strategies and for all nonnegative integers $i, j \leq 2$.

Next, we focus on the analysis of dimorphic population dynamics of similar strategies. Like the discussion in Sect. 2.5, let $x_1 = x + \epsilon \xi_1$, $x_2 = x + \epsilon \xi_2 \in \mathcal{X}$ with corresponding population sizes $n_{1,t}^\epsilon, n_{2,t}^\epsilon \in \mathcal{N}$ and population structures $v_{1,t}^\epsilon, v_{2,t}^\epsilon \in \Delta$ at time t. Let $(x, \|n_t\|_1, v_t, e_t, \theta_t)$ is the reference monomorphic resident system with invariant probability measure μ. We now are interested in the dynamics of the total population size N_t^ϵ of x_1 and x_2 and the relative population size P_t^ϵ of x_2, where

$$N_t^\epsilon = \|n_{1,t}^\epsilon\|_1 + \|n_{2,t}^\epsilon\|_1, \quad P_t^\epsilon = \frac{\|n_{2,t}^\epsilon\|_1}{N_t^\epsilon}.$$

Expanding (3.2) around $\epsilon = 0$ (i.e., $x_1 = x_2 = x$) and then taking $\epsilon = 0$, we have

$$\dot{N}_t^0 = \mathbb{1}^\top F(x, e_t^0, \theta_t)\big(v_{2,t}^0 P_t^0 + v_{1,t}^0(1 - P_t^0)\big)N_t^0, \tag{3.3a}$$

$$\dot{P}_t^0 = P_t^0(1 - P_t^0)\mathbb{1}^\top F(x, e_t^0, \theta_t)(v_{2,t}^0 - v_{1,t}^0). \tag{3.3b}$$

Recalling the analysis in Sect. 2.5.1, the difference here is that P_t^0 is not longer a constant on t-timescale. Instead, P_t^0 evolves with N_t^0 on t-timescale. The following we derive that (i) the dynamics of the monomorphic system $(\|n_t\|_1, v_t, e_t, \theta_t)$ still represents the dynamics of $(N_t^\epsilon, v_{i,t}^\epsilon, e_t^\epsilon, \theta_t)$ on t-timescale when $\epsilon = 0$ for $i = 1, 2$, and (ii) P_t^ϵ evolves on t-timescale but eventually returns to its initial state when $\epsilon = 0$.

Denote

$$\bar{v}_t^0 = v_{2,t}^0 P_t^0 + v_{1,t}^0(1 - P_t^0),$$

which is the averaged proportion of population structures $v_{1,t}^0$ and $v_{2,t}^0$ at time t. From (3.3a) and the dynamics of $(v_{1,t}^0, v_{2,t}^0, e_t^0)$, we have

$$\begin{aligned}
\dot{N}_t^0 &= \mathbb{1}^\top F(x, e_t^0, \theta_t)\bar{v}_t^0 N_t^0, \\
\dot{\bar{v}}_t^0 &= \big(F(x, e_t^0, \theta_t) - \mathbb{1}^\top F(x, e_t^0, \theta_t)\bar{v}_t^0\mathbb{1}\big)\bar{v}_t^0, \\
\dot{e}_{1,t}^0 &= G_1(e_t^0, \theta_t) + H_1(x, e_t^0, \theta_t)\bar{v}_t^0 N_t^0, \\
\dot{e}_{2,t}^0 &= G_2(e_t^0, \theta_t) + H_2(x, e_t^0, \theta_t)\bar{v}_t^0 N_t^0, \\
\dot{\theta}_t &= A(\theta_t) + B(\theta_t)\dot{W}_t.
\end{aligned} \tag{3.4}$$

It shows that $(N_t^0, \bar{v}_t^0, e_t^0, \theta_t)$ satisfies the same equations as $(\|n_t\|_1, v_t, e_t, \theta_t)$. In words, the dynamics of $(\|n_t\|_1, v_t, e_t, \theta_t)$ represents that of $(N_t^\epsilon, \bar{v}_t^\epsilon, e_t^\epsilon, \theta_t)$ on t-timescale when $\epsilon = 0$. Turning to the dynamics of $v_{1,t}^0$ and $v_{2,t}^0$,

$$\dot{v}_{i,t}^0 = \big(F(x, e_t^0, \theta_t) - \mathbb{1}^\top F(x, e_t^0, \theta_t)v_{i,t}^0\mathbb{1}\big)v_{i,t}^0, \quad i = 1, 2$$

with e_t^0 calculated from (3.4), we see that $v_{1,t}^0$ and $v_{2,t}^0$ satisfy the same equation as $\bar{v}_t^0$ (or, equivalently, v_t). Then $v_{1,t}^0$ and $v_{2,t}^0$ have the same dynamics as $\bar{v}_t^0$ (or, equivalently,

v_t). Thus, we claim that $(N_t^\epsilon, v_{i,t}^\epsilon, e_t^\epsilon, \theta_t)$ has the same dynamics as $\left(\|n_t\|_1, v_t, e_t, \theta_t\right)$ on t-timescale when $\epsilon = 0$ for $i = 1, 2$.

Now we return to (3.3b), a straightforward calculation shows that

$$P_t^0 = \left(1 + \frac{1-p}{p} \exp\left(\int_0^t \mathbb{1}^\top F(x, e_\tau^0, \theta_\tau)(v_{2,\tau}^0 - v_{1,\tau}^0)d\tau\right)\right)^{-1}$$

provided that the initial state is $(n, v_1, v_2, e, \theta, p)$. Using Birkhoff's ergodic theorem and the principle of selective neutrality of residents (i.e., zero growth rate of the reference resident x), we obtain

$$\lim_{t \to +\infty} P_t^0 = p$$

with probability one and for all $p \in [0, 1]$ and for μ-almost every (n, v_1, v_2, e, θ) (N.B., initial states v_1 and v_2 do not have to be identical). It means that P_t^0 will sufficiently close to its initial state some time later with probability one. Thus, we need consider the dynamics of P_t^ϵ on slow timescales to revel the dimorphic population dynamics of x_1 and x_2.

The remaining generalization of our theorems to model (3.2) can be done by using of arguments similar to Sects. 2.5.2 and 2.5.3.

4 Applications

To illustrate the utility of our theorems, we apply them to four concrete examples. The first example we consider the evolving bacteria in a chemostat model. We will see that substitution is the unique outcome after an invasion event. The second example we consider a Lotka-Volterra competition model. Theorems 1, 2 and 3 are applied to predict the population dynamical outcomes of an invasion event when the strategy pairs satisfy the corresponding hypotheses of these theorems. The third example we consider a structured SIRS epidemic model, which is designed to illustrate how a concrete structured population model can be reformulated into our framework. In the last example, we consider the evolution of timidity of the prey in a predator-prey model, which is designed to demonstrate that our results are applicable to the evolution with the non-equilibrium resident dynamics.

In these four examples, we start from a monomorphic population with a certain strategy. To verify the existence of an invariant probability measure for a resident system, one might consider the tightness of one of the following two families of measures:

(a) the statistics associated with a single realization of process $\{Z_t\}_{t \geq 0}$, i.e. measure

$$\Pi_t^z(\cdot) = \frac{1}{t} \int_0^t \delta_{Z_s}(\cdot)ds$$

with a Dirac measure δ_{Z_s} at Z_s (ref to the argument presented in Schreiber et al. (2011) or Benaïm (2018));

(b) the statistics associated with transition probabilities $\{\mathcal{P}_t\}_{t\geq 0}$ of process $\{Z_t\}_{t\geq 0}$, i.e. measure

$$\pi_t^z(\cdot) = \frac{1}{t}\int_0^t \mathcal{P}_s(z,\cdot)ds$$

which is a common tool in the literature on Markov process (see e.g. Arnold and Kliemann 1983, Proposition 3.15).

Different ways to study the existence of invariant probability measures give different interpretations for the asymptotic behavior of process $\{Z_t\}_{t\geq 0}$. Assume that the limit exists as time goes to infinity, the (a) gives the long-term frequency of a typical realization of process $\{Z_t\}_{t\geq 0}$ visiting a particular configuration. Since a transition probability corresponds to the frequency of observing a particular event across many realizations of process $\{Z_t\}_{t\geq 0}$, the (b) gives the long-term frequency of probabilities for process visiting a particular configuration. Following the strategies of proofs of Benaïm (2018, Theorem 4.4), Schreiber et al. (2011) and Arnold and Kliemann (1983, Proposition 3.15), one can show that all weak* limit points of Π_t^z and π_t^z are candidate μ of Definition 1.

In each example, we present the simulated trajectories of (i) the resident and invader population in the corresponding phase plane, (ii) the total population size N_t, and (iii) the relative population size P_t of the invader. The numerical analysis shows that N_t varies stochastically and dramatically in time, while P_t changes slowly and asymptotically in time. Since we don't perform the complete time-scale separation in numerical simulations, one can find that there are some small fluctuations remain in the simulated trajectories of P_t in Figs. 4, 6, 8, and 10 for different examples, respectively. With complete time-scale separation (i.e., $\epsilon \to 0$), the trajectories of P_t will be smooth.

4.1 Chemostat model

Consider the evolving bacteria with strategy $x = (\beta, \gamma, \delta) \in \mathbb{R}_+^3 := \mathcal{X}$ and concentration n_t at time t in a chemostat model, in which the dynamics is given by

$$\dot{R}_t = D(R_t^{in} - R_t) - \beta R_t n_t,$$
$$\dot{n}_t = \gamma \beta R_t n_t - (\delta + D)n_t,$$

$$(4.1)$$

where R_t is the concentration of nutrient at time t. Let R_t^{in} be the time-varying concentration of nutrient at the input. Positive constants D, β, γ and δ are dilution rate, nutrient uptake rate, conversion efficiency and death rate of bacteria, respectively. In this model, we assume

$$R_t^{in} = \rho_1 - \rho_2 \frac{2\theta_t}{1 + \theta_t^2}$$

with constants $\rho_1 > \rho_2 > 0$ and $\{\theta_t\}_{t \geq 0}$ is the stationary Ornstein-Uhlenbeck process generated by a linear stochastic differential equation

$$\dot{\theta}_t = -a\theta_t + b\dot{W}_t, \quad \theta_0 = 0, \tag{4.2}$$

with positive constants a and b. Then R_t^{in} takes value in the interval $[\rho_1 - \rho_2, \rho_1 + \rho_2]$ for all time t and tends to peaks around $\rho_1 \pm \rho_2$, which is a random switching scenario. Different ways that an unbounded noise induces bounded fluctuations on model parameters refer to d'Onofrio (2013) and Caraballo and Han (2016).

Denote

$$\begin{aligned}
e_t &= (e_{1,t}, e_{2,t}) = (R_t, 0), \\
f(x, e_t, \theta_t) &= \gamma \beta e_{1,t} - \delta - D, \\
G_1(e_t, \theta_t) &= D(R_t^{in} - e_{1,t}), \\
H_1(x, e_t, \theta_t) &= \beta e_{1,t}.
\end{aligned}$$

Then (4.1) can be written as the general form (2.1). From the boundedness of R_t^{in}, one can show that $R_t \leq \rho_1 + \rho_2$ and $n_t \leq \frac{\gamma\beta(\rho_1+\rho_2)}{\delta+D}$ for t sufficiently large. Thus, the state space of the dynamics of (n_t, e_t, θ_t) is $\mathcal{Z} = \left[0, \frac{\gamma\beta(\rho_1+\rho_2)}{\delta+D}\right] \times [0, \rho_1 + \rho_2] \times \{0\} \times \mathbb{R}$. The partial derivatives $f^{(i,j)}$, $G_1^{(i,j)}$ and $H_1^{(i,j)}$ exist for all strategies $x \in \mathcal{X}$ and for all nonnegative integers $i, j \leq 2$, which satisfy Assumption **A2**.

For bacteria x, its virgin environment R_t^{vir} is generated by

$$\dot{R}_t^{vir} = D(R_t^{in} - R_t^{vir}).$$

The boundedness of R_t^{in} and the Feller property ensure the existence of invariant probability measures for process $\{R_t^{vir}, \theta_t\}_{t \geq 0}$ with support $[\rho_1 - \rho_2, \rho_1 + \rho_2] \times \mathbb{R} \subset \mathcal{Z}_0$ and expectation $\mathbb{E}[R^{vir}] = \mathbb{E}[R^{in}] = \rho_1$. Bacteria x can invade the virgin environment R^{vir} if

$$\gamma\beta\mathbb{E}[R^{vir}] - \delta - D = \gamma\beta\rho_1 - \delta - D > 0. \tag{4.3}$$

Likewise, the boundedness of (n_t, R_t) and the Feller property ensure the existence of invariant probability measures for process $\{n_t, R_t, \theta_t\}_{t \geq 0}$, denoted by μ. One hand, by Poincaré recurrence theorem and the boundedness of n_t, we obtain that $\lambda_x(\mu) = 0$ for all μ with $\mu(\mathcal{Z} \setminus \mathcal{Z}_0) = 1$ (ref to Schreiber et al. 2011, assertion (iii) of Proposition 1). On the other hand, since $\mathcal{Z} \setminus \mathcal{Z}_0$ and $\mathcal{Z}_0$ are invariant, there exists $\alpha \in (0, 1]$ such that $\mu = (1 - \alpha)\mu_0 + \alpha\mu_1$ where μ_0 is an invariant probability measure with $\mu_0(\mathcal{Z}_0) = 1$ and μ_1 is an invariant probability measure with $\mu_1(\mathcal{Z} \setminus \mathcal{Z}_0) = 1$. Notice that $\lambda_x(\mu) \leq 0$ due to the boundedness of n_t. From this fact and $\lambda_x(\mu_1) = 0$, it follows that $(1 - \alpha)\lambda_x(\mu_0) \leq 0$. By (4.3) (i.e., $\lambda_x(\mu_0) > 0$), we obtain that $\alpha = 1$. Thus, all μ satisfy $\mu(\mathcal{Z} \setminus \mathcal{Z}_0) = 1$, provided that (4.3) holds. Using (4.3) and arguments similar to the proof of Benaïm (2018, Theorem 4.4), we further can show that for all $\varepsilon > 0$, there exists a $\eta > 0$ such that $\mu(\mathcal{Z}_\eta) \leq \varepsilon$ for every μ with $\mu(\mathcal{Z} \setminus \mathcal{Z}_0) = 1$.
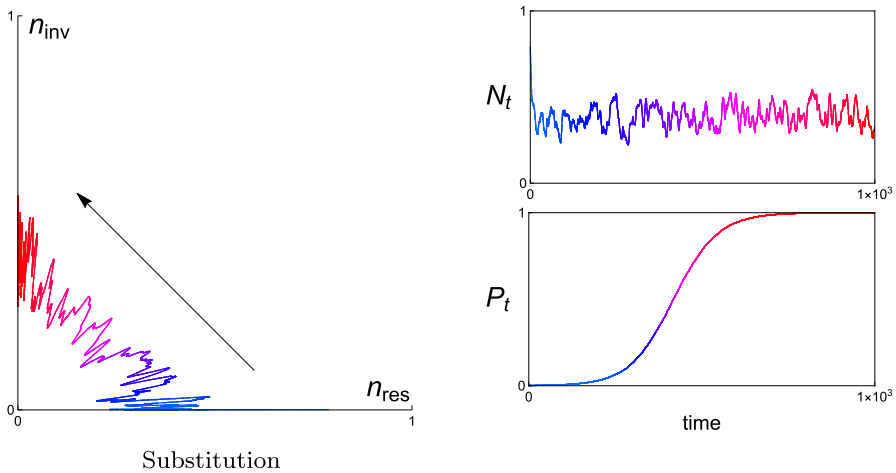
Fig. 4 A simulated population trajectory in the phase plane and associated dynamics of total concentration $N_t = n_{res,t} + n_{inv,t}$ and relative concentration $P_t = n_{inv,t}/N_t$ for a resident bacteria with strategy $x = (2, 0.8, 0.095)$ and an invader bacteria with strategy $y = (1.99, 0.85, 0.09)$. The corresponding invasion fitnesses $\mathcal{S}_x(y) = 0.0161516 > 0$ and $\mathcal{S}_y(x) = -0.0102779 < 0$. Parameter values $D = 0.1$, $\rho_1 = 1, \rho_2 = 0.5, a = b = 1$

From Definition 1, we thus claim that bacteria x is non-growing on the long-run and stochastically persistent.

Once bacteria x becomes a resident, the invasion fitness of an initially rare mutant $y = (\beta', \gamma', \delta')$ in resident x is

$$\mathcal{S}_x(y) = \gamma' \beta' \mathbb{E}[R] - \delta' - D.$$

For similar strategies (x, y) satisfying $\partial_y \mathcal{S}_x(y)|_{y=x}(y - x) \neq 0$, Theorem 1 can be employed to predict the invasion dynamics. Instead we apply Theorems 2 or 3 if $\partial_y \mathcal{S}_x(y)|_{y=x}(y - x) = 0$. However, we will see that substitution is the unique outcome after an invasion event whatever strategy pairs are. In fact, from (2.4), we get that $\mathbb{E}[R] = \frac{\delta+D}{\gamma\beta}$. Hence, invasion happens if and only if $\frac{\delta'+D}{\gamma'\beta'} < \frac{\delta+D}{\gamma\beta}$. It further implies that for all $x, y \in \mathcal{X}, \mathcal{S}_y(x) < 0$ if $\mathcal{S}_x(y) > 0$.

Figure 4 shows how a specific population trajectory develops from n_{res}-axis to n_{inv}-axis in the phase plane provided that $x = (2, 0.8, 0.095)$ and $y = (1.99, 0.85, 0.09)$ with $\mathcal{S}_x(y) > 0$ and $\mathcal{S}_y(x) < 0$. The initially rare bacteria y will actually take over the bacteria x and becomes the new resident. The total concentration N_t of the resident x and the invader y varies stochastically and dramatically in time. However, the relative concentration P_t of the invader y changes slowly in time and asymptotically closes to 1. The dynamics of P_t illustrates that the resident x is substituted by the invader y eventually.

4.2 Lotka–Volterra competition model

Consider a stochastic Lotka-Volterra competition model

$$\dot{n}_{i,t} = \tilde{r}_t(x_i) n_{i,t} \left(1 - \frac{\sum_j \alpha(x_i, x_j) n_{j,t}}{\widetilde{K}_t(x_i)} \right), \quad i = 1, \ldots, k \tag{4.4}$$

where $n_{i,t}$ is the population size of strategy $x_i \in \mathbb{R} := \mathcal{X}$. Assume that the intrinsic growth rate and the carrying capacity of the i-th species are influenced by an external factor θ_t such that they fluctuate around strategy-related values. For simplicity, let

$$\tilde{r}_t(x_i) = r(x_i) \exp(\rho_1 \theta_t) = \exp(-d_1 x_i^2 + \rho_1 \theta_t),$$
$$\widetilde{K}_t(x_i) = K(x_i) \exp(\rho_2 \theta_t) = \exp(-d_2 x_i^2 + \rho_2 \theta_t)$$

where d_1, d_2, ρ_1 and ρ_2 are scaling parameters and where the $\{\theta_t\}_{t \geq 0}$ is a stationary Ornstein-Uhlenbeck process generated by (4.2). Then the expected growth rate and the expected carrying capacity of the i-th species are $r(x_i) \exp\left(\frac{\rho_1^2 b^2}{2a}\right)$ and $K(x_i) \exp\left(\frac{\rho_2^2 b^2}{2a}\right)$, respectively. Let the competitive coefficient between strategy x_i and x_j be of the form

$$\alpha(x_i, x_j)$$
$$= \left(1 - (x_j - x_i)(c_0 + c_1 x_i^2 + c_2 x_i x_j + c_3 x_j^2) \exp(c_4 x_j^2)\right) \exp\left(c_5(x_i^2 - x_j^2)\right),$$

where $\alpha(x_i, x_i) = 1$ for all i.

Let $\mathcal{E}$ be the set of all functions $e : \mathcal{X} \mapsto \mathbb{R}^2$ of the form

$$e_t(\cdot) = (e_{1,t}, e_{2,t})(\cdot) = \left(0, \sum_j \alpha(\cdot, x_j) n_{j,t}\right)$$

for $n_{j,t} \geq 0$ and $x_j \in \mathcal{X}$. Since the competition is modelled as a direct interaction, we can define the per-capita environmental impact $H_2 : \mathcal{X} \times \mathcal{E} \times \mathbb{R} \mapsto \mathbb{R}$ by

$$H_2(x_j, e_t, \theta)(\cdot) = \alpha(\cdot, x_j)$$

which is the competitive impact of a single individual with strategy x_j on its competitor with any strategy. Since the competitor's strategy has infinitely many choices, $e_{2,t}$ has infinite dimensions. In addition, define the map $G_2 : \mathcal{E} \times \mathbb{R} \mapsto \mathbb{R}$ by

$$G_2(e_t, \theta) = 0.$$

Further, define the map $f : \mathcal{X} \times \mathcal{E} \times \mathbb{R} \mapsto \mathbb{R}$ by

$$f(x_i, e_t(x_i), \theta) = \tilde{r}(x_i) \left(1 - \frac{e_{2,t}(x_i)}{\widetilde{K}(x_i)} \right).$$
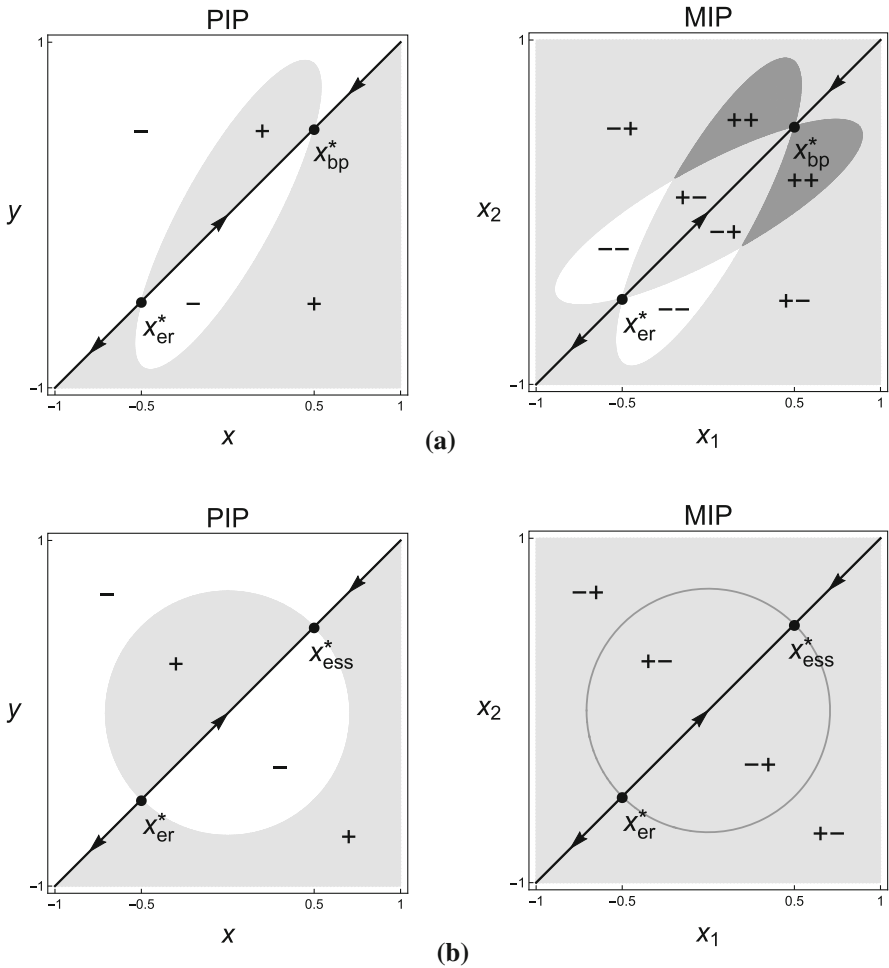
Fig. 5 Pairwise invasibility plot (PIP) and corresponding Mutual invasibility plot (MIP) for the invasion fitness (4.5) with parameter values $d_1 = 1$, $d_2 = 0.5$, $\rho_1 = 1$, $\rho_2 = 0.8$, $a = b = 1$, and $(c_0, c_1, c_2, c_3, c_4, c_5) =$ **(a)** $(1, -11, 11, -4, 1, 0.5)$ and **(b)** $(1, -2, 0, -2, 1, 0.5)$. White area in PIP: y cannot invade x; light gray area in PIP: y can invade x; white area in MIP: x_1 and x_2 cannot invade each other; light gray area in MIP: one of x_1 and x_2 can invader the other but not *vice versa*; gray area in MIP: x_1 and x_2 can invade each other; "+" and "−": signs of the invasion fitness for given strategy pairs; arrows: the direction of monomorphic evolution

Thus, (4.4) can be written as the general form (2.1). Furthermore, the state space of the dynamics of $(n_{1,t}, \ldots, n_{k,t}, e_t, \theta_t)$ is $\mathcal{Z} = \mathbb{R}_+^k \times \mathcal{E} \times \mathbb{R}$. The partial derivatives $f^{(i,j)}$, $G_2^{(i,j)}$ and $H_2^{(i,j)}$ exist for all strategies and for all nonnegative integers i, $j \leq 2$, which satisfy Assumption **A2**.

We now start from a monomorphic population of strategy x with size n_t at time t. Similar to the previous example, one can show that species x successfully establishes with population dynamics satisfying Definition 1. The only difference that should be

pointed out is that the population is unbounded in this example. Instead the population is ultimately bounded in mean (Miyahara 1972, 1973), i.e., $\lim \sup_{t \to +\infty} \mathbb{E}[n_t] \leq \mathbb{E}[\lim \sup_{t \to +\infty} n_t] = K(x)\mathbb{E}[\lim \sup_{t \to +\infty} \exp(\rho_2 \theta_t)] = K(x) < +\infty$ for all $x \in \mathcal{X}$. Further, this stochastic boundedness combined with the Feller property guarantees the existence of invariant probability measures for process $\{n_t, \theta_t\}_{t \geq 0}$. Now, from (4.4), the invasion fitness of an initially rare mutant y in resident x is given by

$$\mathcal{S}_x(y) = r(y) \exp\left(\frac{\rho_1^2 \sigma^2}{2a}\right)\left(1 - \frac{\alpha(y, x)K(x)}{K(y)}\right). \tag{4.5}$$

Here we have used the principle of selective neutrality of residents (i.e., $\mathcal{S}_x(x) = 0$ for all $x \in \mathcal{X}$) to derive (4.5).

Figure 5a shows the PIP and the MIP corresponding to (4.5) with parameters $(c_0, c_1, c_2, c_3, c_4, c_5) = (1, -11, 11, -4, 1, 0.5)$. The features of these two plots are similar to that of Kisdi et al. (2001) which also studies the evolutionary of strategy x but for a deterministic Lotka-Volterra model with $\tilde{r}_t(x) = \widetilde{K}_t(x) = 1$ for all $x \in \mathcal{X}$ and for all $t \geq 0$. Away from evolutionarily singular strategies x_{er}^* (evolutionary repeller) and x_{bp}^* (evolutionary branching point), it follows from Theorem 1 that every successful invasion of mutant y in a sufficiently small neighborhood of resident x will takeover the population. Further, directional evolution proceeding by successive invasions and substitutions with small phenotypic effect leads to an evolutionary branching point x_{bp}^* provided that the initial strategy value is above x_{er}^*, otherwise is towards low-valued strategies if the initial strategy value is below x_{er}^*. Both x_{er}^* and x_{bp}^* satisfy $\mathcal{C}_{22} + \mathcal{C}_{11} \neq 0$, so that we can apply Theorem 2 to all strategies in a neighborhood of x_{er}^* or x_{bp}^*. Having approached x_{bp}^*, the population undergoes evolutionary branching that gives rise to two district strategies coexistence. Figure 6a gives a simulated population trajectory in the phase plane for the coexistence of strategies $x = 0.48$ and $y = 0.51$ with corresponding invasion fitnesses $\mathcal{S}_x(y) > 0$ and $\mathcal{S}_y(x) > 0$. The dynamics of P_t shows that the relative population size of the invader y asymptotically tends to an interior value of $(0, 1)$.

To find the case of mutual exclusion, we choose two strategies lie in the white area of the MIP in Fig. 5a and close to the x_{er}^*. Figure 6b gives two simulated population trajectories in the phase plane for strategies $x = -0.53$ and $y = -0.47$ with corresponding invasion fitnesses $\mathcal{S}_x(y) < 0$ and $\mathcal{S}_y(x) < 0$, provided that initial population states are identical. In the phase plane, we see that one simulated population trajectory eventually reaches n_{res}-axis, while the other trajectory eventually reaches n_{inv}-axis. The corresponding dynamics of P_t shows a simulated trajectory asymptotically tends to 0 but the other one asymptotically tends to 1.

Figure 5b shows the PIP and the MIP corresponding to (4.5) with parameters $(c_0, c_1, c_2, c_3, c_4, c_5) = (1, -2, 0, -2, 1, 0.5)$. Now the lower singularity is still an evolutionary repeller, while the upper singularity becomes an evolutionarily stable strategy. What's more, the associated $\mathcal{C}_{22} + \mathcal{C}_{11} = 0$ but $\mathcal{C}_{22} \neq 0$ and $\mathcal{C}_{11} \neq 0$ for both two evolutionarily singular strategies x_{er}^* and x_{ess}^*. Therefore, we apply Theorem 3 to all strategies in a neighborhood of x_{er}^* or x_{ess}^*, in which substitution is the unique outcome of an invasion event. Figure 6c gives a simulated population trajectory in the phase plane for strategies $x = 0.55$ and $y = 0.51$ with corresponding invasion
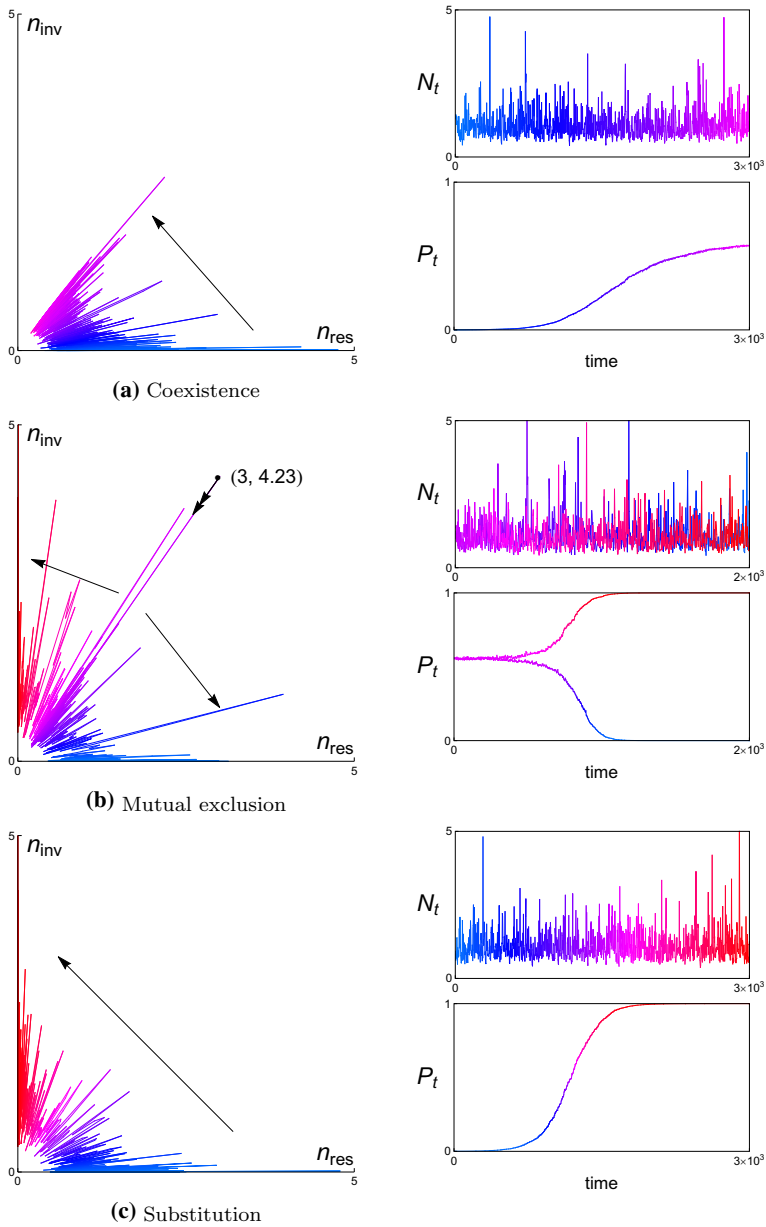
(a) Coexistence

(b) Mutual exclusion

(c) Substitution

Fig. 6 Simulated population trajectories in the phase plane and associated dynamics of the total population size $N_t = n_{res,t} + n_{inv,t}$ and the relative population size $P_t = n_{inv,t}/N_t$ for strategy pairs **(a)** $(x, y) = (0.48, 0.51)$ with corresponding invasion fitnesses $\mathcal{S}_x(y) = 0.00454545 > 0$ and $\mathcal{S}_y(x) = 0.00346302 > 0$, **(b)** $(x, y) = (-0.53, -0.47)$ with corresponding invasion fitnesses $\mathcal{S}_x(y) = -0.0179815 < 0$ and $\mathcal{S}_y(x) = -0.0143759 < 0$, and **(c)** $(x, y) = (0.55, 0.51)$ with corresponding invasion fitnesses $\mathcal{S}_x(y) = 0.0064304 > 0$ and $\mathcal{S}_y(x) = -0.0064304 < 0$, where other parameter values see Fig. 5

finesses $\mathcal{S}_x(y) > 0$ and $\mathcal{S}_y(x) < 0$. The dynamics of P_t verifies that the invader y substitutes the resident x eventually.

4.3 A structured SIRS epidemic model

Consider a structured SIRS epidemic model for the evolution of viral type. For a given moment t and $i = 1, \ldots, k$, let S_t be the numbers of susceptible individuals, $I_{i,t}$ be the numbers of individual infected by a virus $x_i \in \mathbb{R}_+ := \mathcal{X}$, and $R_{i,t}$ be the number of recovered individual whose is infected by the virus x_i. Let M_t be the total number of susceptible, infected and recovered individuals at time t, i.e.,

$$M_t = S_t + \sum_j I_{j,t} + \sum_j R_{j,t}.$$

Assume that all individuals are born free of the disease at a constant rate $\Lambda > 0$, and all individuals naturally die at a density-dependent rate δM_t with constant $\delta > 0$. In a well-mixed population, the probability of a contact being with an infected individual of x_i is given by $\frac{I_{i,t}}{M_t}$. The probability of that contact giving rise to an infection of x_i is given by $\tilde{\beta}_t(x_i)$, which is commonly called transmission rate. Infectious individuals of x_i recover at a rate $\tilde{\gamma}_t(x_i)$ and disease-caused die at a rate $\tilde{\alpha}_t(x_i)$. Recovered individuals of x_i lose their protection against a reinfection and become susceptible again at a rate $\tilde{\zeta}_t(x_i)$. Here $\tilde{\beta}_t(x_i)$, $\tilde{\gamma}_t(x_i)$, $\tilde{\alpha}_t(x_i)$ and $\tilde{\zeta}_t(x_i)$ are all positive, time-varying and virus-dependent. For simplicity, we assume that the transmission rate fluctuates around a virus-dependent value in the way: $\tilde{\beta}_t(x_i) = \beta(x_i) \exp(\rho_1 \theta_t)$ with $\beta(x_i) > 0$ for all $x_i \in \mathcal{X}$, where the $\{\theta_t\}_{t \geq 0}$ is a stationary Ornstein-Uhlenbeck process generated by (4.2) and ρ_1 is a scaling parameter to measure the effect of θ_t on the transmission rate. Likewise, $\tilde{\gamma}_t(x_i)$, $\tilde{\alpha}_t(x_i)$ and $\tilde{\zeta}_t(x_i)$ fluctuate in time in the same way as $\tilde{\beta}_t(x_i)$ but the scaling parameters are different, i.e., $\tilde{\gamma}_t(x_i) = \gamma(x_i) \exp(\rho_2 \theta_t)$, $\tilde{\alpha}_t(x_i) = \alpha(x_i) \exp(\rho_3 \theta_t)$ and $\tilde{\zeta}_t(x_i) = \zeta(x_i) \exp(\rho_4 \theta_t)$ with $\gamma(x_i), \alpha(x_i), \zeta(x_i) > 0$ for all $x_i \in \mathcal{X}$. Assume further that one disease-free individual is infected with a certain virus, it cannot be infected by any other virus. Under these assumptions, a structured SIRS epidemic model is given by the following differential equations:

$$\dot{S}_t = \Lambda M_t - \sum_j \tilde{\beta}_t(x_j) \frac{I_{j,t}}{M_t} S_t - \delta M_t S_t + \sum_j \tilde{\zeta}_t(x_j) R_{j,t},$$

$$\begin{pmatrix} \dot{I}_{i,t} \\ \dot{R}_{i,t} \end{pmatrix} = \underbrace{\begin{pmatrix} \tilde{\beta}_t(x_i) \frac{S_t}{M_t} - \tilde{\gamma}_t(x_i) - \tilde{\alpha}_t(x_i) - \delta M_t & 0 \\ \tilde{\gamma}_t(x_i) & -\tilde{\zeta}_t(x_i) - \delta M_t \end{pmatrix}}_{:= F(x_i, S_t, M_t, \theta_t)} \tag{4.6}$$

$$\times \begin{pmatrix} I_{i,t} \\ R_{i,t} \end{pmatrix}, \quad i = 1, \ldots, k.$$
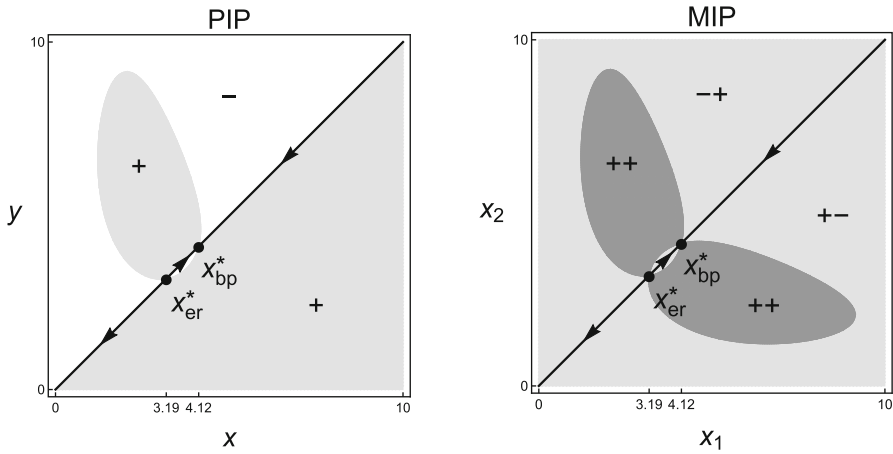
Fig. 7 Pairwise Invasibility Plot (PIP) and corresponding Mutual Invasibility Plot (MIP) for the small-noise approximation of the invasion fitness (4.9) for $\alpha(x) = x$, $\beta(x) = 5.65 + \frac{2x^2}{6+0.1x^2}$, $\gamma(x) = 0.3$ and $\zeta(x) = \exp(-0.5x^2)$. We further have $\mathcal{S}_x(y) = (\beta(y) - \beta(x))\left(\frac{\widehat{S}}{\widehat{M}} + \frac{1}{2}g(\widehat{S}, \widehat{M}, \hat{\theta})\mathrm{Cov}_{(S,M,\theta)}(0)\right) - (\alpha(y) - \alpha(x)) + \mathcal{O}(\rho_1^2)$, where $\widehat{S}$ and $\widehat{M}$ are positive equilibrium values of the deterministic resident system, and $\hat{\theta}$ is the expectation of the process $\{\theta_t\}_{t\geq 0}$, and $g(\widehat{S}, \widehat{M}, \hat{\theta})$ is the second derivative of $\exp(\rho_1\theta)\frac{S}{M}$ with respect to (S, M, θ) evaluated at $(\widehat{S}, \widehat{M}, \hat{\theta})$, and $\mathrm{Cov}_{(S,M,\theta)}(0)$ is the covariance matrix in the components of (S, M, θ) at the stationary distribution of the monomorphic resident system. Notice that $\widehat{S}$, $\widehat{M}$ and $\mathrm{Cov}_{(S,M,\theta)}(0)$ are functions of the resident strategy x. For the interpretations of local areas, markers and arrows in the PIP and the MIP see Fig. 5. Parameter values $\Lambda = 3$, $\delta = 0.2$, $a = 2$, $b = 1$, $\rho_1 = 0.1$ and $\rho_2 = \rho_3 = \rho_4 = 0$

Following Sect. 3, let

$$n_{i,t} = (n_{i,t}^1, n_{i,t}^2)^\top = (I_{i,t}, R_{i,t})^\top, \ v_{i,t} = (v_{i,t}^1, v_{i,t}^2)^\top = \left(\frac{n_{i,t}^1}{\|n_{i,t}\|_1}, \frac{n_{i,t}^2}{\|n_{i,t}\|_1}\right)^\top,$$

where the state space of $v_{i,t}$ is $\Delta = \{v \in [0, 1]^2 : \|v\|_1 = 1\}$. Then the dynamics of (4.6) are equivalent to

$$\|n_{i,t}\|_1^\cdot = \mathbb{1}^\top F(x_i, S_t, M_t, \theta_t)v_{i,t}\|n_{i,t}\|_1,$$
$$\dot{v}_{i,t} = \left(F(x_i, S_t, M_t, \theta_t) - \mathbb{1}^\top F(x_i, S_t, M_t, \theta_t)v_{i,t}\mathbb{I}\right)v_{i,t},$$
$$\dot{S}_t = \Lambda M_t - \sum_j \tilde{\beta}_t(x_j)\frac{S_t}{M_t}v_{j,t}^1\|n_{j,t}\|_1 - \delta M_t S_t + \sum_j \tilde{\zeta}_t(x_j)v_{j,t}^2\|n_{j,t}\|_1,$$
$$M_t = S_t + \sum_j \|n_{j,t}\|_1,$$

which can be written as the form specified in (3.2) with

$$
\begin{aligned}
e_t &= (e_{1,t}, e_{2,t}) = (S_t, M_t),\\
G_1(e_t, \theta_t) &= (\Lambda - \delta e_{1,t})e_{2,t},\\
H_1(x_j, e_t, \theta_t) &= \left(-\tilde{\beta}_t(x_j)\frac{e_{1,t}}{e_{2,t}},\ \tilde{\zeta}_t(x_j)\right),\\
G_2(e_t, \theta_t) &= e_{1,t},\\
H_2(x_j, e_t, \theta_t) &= 1.
\end{aligned}
$$

It is easy to show that $M_t \leq \frac{\Lambda}{\delta}$ for t sufficiently large. Thus, the state space of the dynamics of $Z_t = \big(\|n_{1,t}\|_1, \ldots, \|n_{k,t}\|_1, v_{1,t}, \ldots, v_{k,t}, e_t, \theta_t\big)$ is $\mathcal{Z} = \big[0, \frac{\Lambda}{\delta}\big]^k \times \Delta^k \times \big[0, \frac{\Lambda}{\delta}\big]^2 \times \mathbb{R}$, the extinction set $\mathcal{Z}_0 = \big\{z \in \mathcal{Z} : \min_i \|n_i\|_1 = 0\big\}$, and the η-neighborhood of the extinction set $\mathcal{Z}_\eta = \big\{z \in \mathcal{Z} : \min_i \|n_i\|_1 \leq \eta\big\}$. The partial derivatives $F^{(i,j)}$, $G_1^{(i,j)}$, $H_1^{(i,j)}$, $G_2^{(i,j)}$ and $H_2^{(i,j)}$ exist and are locally Lipschitz continuous and measurable in (e, θ) for all strategies and for all nonnegative integers $i, j \leq 2$.

We now start from a monomorphic population of viral type x with the total population size $\|n_t\|_1$ and the population structure v_t. To proceeding the further analysis, we need to know when a given viral type x becomes a resident, i.e., it successfully establishes in the virgin environment. The dynamics on the boundary of $\big(\|n_t\|_1, S_t, M_t\big)$-space globally converges to a unique stable equilibrium $\big(0, \frac{\Lambda}{\delta}, \frac{\Lambda}{\delta}\big)$, provided that $S_0 \neq 0$. Hence, the ergodic measure on the boundary of $\big(\|n_t\|_1, S_t, M_t\big)$-space is the Dirac measure at $\big(0, \frac{\Lambda}{\delta}, \frac{\Lambda}{\delta}\big)$. Substitution of this stable equilibrium into the equation v_t gives

$$
\dot{v}_t^1 = \underbrace{\left(\tilde{\beta}_t(x) - \tilde{\gamma}_t(x) - \tilde{\alpha}_t(x) - \Lambda - \big(\tilde{\beta}_t(x) - \tilde{\alpha}_t(x) - \Lambda\big)v_t^1 + \big(\tilde{\zeta}_t(x) + \Lambda\big)v_t^2\right)}_{:= h(v_t^1, v_t^2, \theta_t)} v_t^1
$$

$$
v_t^2 = 1 - v_t^1. \tag{4.7}
$$

Providing that the expectation

$$
\mathbb{E}\big[\tilde{\beta}(x) - \tilde{\gamma}(x) - \tilde{\alpha}(x) - \Lambda\big] > 0, \tag{4.8}
$$

the dynamics of (4.7) consist of a Dirac measure at the unstable equilibrium $v = (0, 1)$ and invariant probability measures $\hat{v}$ supported on $\Delta \setminus \{(0, 1)\}$. Let μ be the product of the Dirac measure at $\big(\|n\|_1, S, M\big) = \big(0, \frac{\Lambda}{\delta}, \frac{\Lambda}{\delta}\big)$ and the $\hat{v}$. Then $\lambda_x(\mu) = \mathbb{E}\big[\big(\tilde{\beta}(x)\frac{S}{M} - \tilde{\alpha}(x) - \delta M\big)v^1 - \big(\tilde{\zeta}(x) + \delta M\big)v^2\big]$ corresponds to the per-capita growth rate of $\|n_t\|_1$ with respect to μ. From the dynamics of v_t^1 we have that with probability one $\lim_{t \to +\infty} \frac{1}{t} \log v_t^1 = \lim_{t \to +\infty} \frac{1}{t} \int_0^t h(v_s^1, v_s^2, \theta_s)ds = \mathbb{E}\big[h(v^1, v^2, \theta)\big] = 0$ for $(v_0^1, v_0^2) \in \Delta \setminus \{(0, 1)\}$, which implies that $\lambda_x(\mu) = \mathbb{E}\big[\tilde{\beta}(x) - \tilde{\gamma}(x) - \tilde{\alpha}(x) - \Lambda\big]$. Thus, if (4.8) holds, a similar argument as the first example implies that the population of viral type x is non-growing on the long-run and stochastically persistent.

(a) Substitution
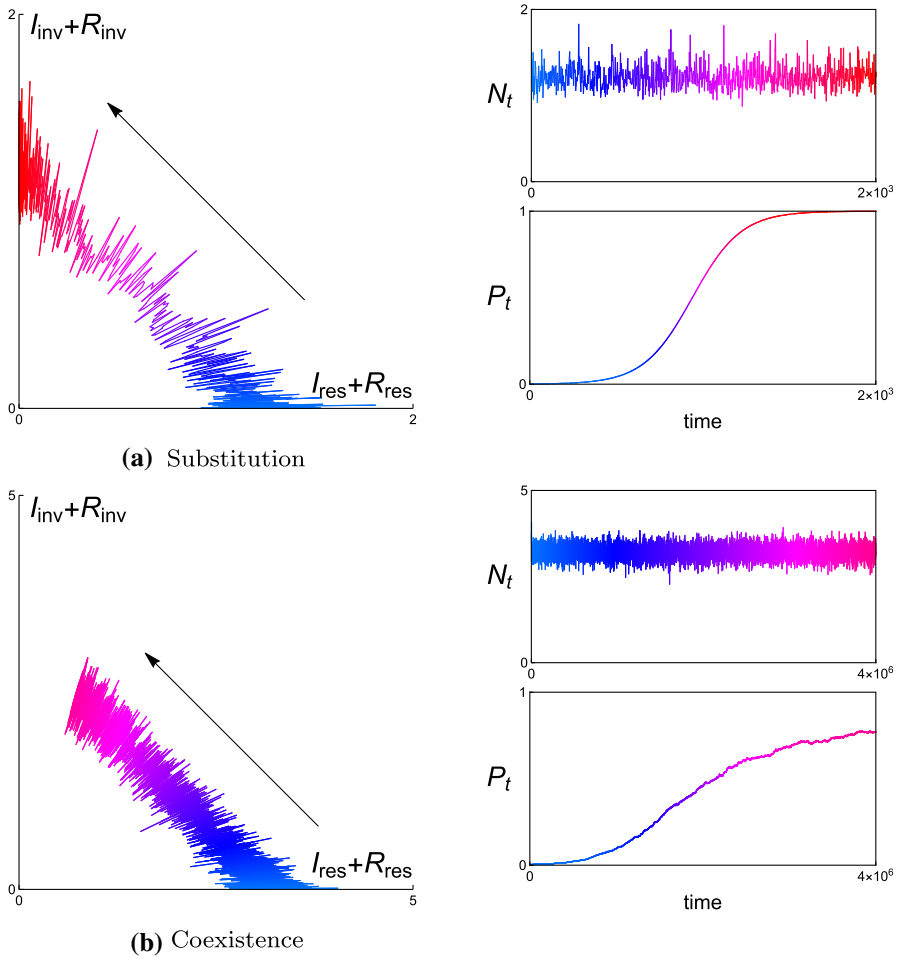

(b) Coexistence

Fig. 8 Simulated population trajectories in the phase plane and associated dynamics of the total population size $N_t = (I_{res,t} + R_{res,t}) + (I_{inv,t} + R_{inv,t})$ and the relative population size $P_t = (I_{inv,t} + R_{inv,t})/N_t$ for strategy pairs **(a)** $(x, y) = (8.7, 8.68)$ with corresponding invasion fitnesses $\mathcal{S}_x(y) \approx 0.00708633 > 0$ and $\mathcal{S}_y(x) \approx -0.00709391 < 0$, and **(b)** $(x, y) = (4.1158, 4.1238)$ with corresponding invasion fitnesses $\mathcal{S}_x(y) \approx 2.27275 \times 10^{-6} > 0$ and $\mathcal{S}_y(x) \approx 2.27315 \times 10^{-6} > 0$, where parameter values refer to Fig. 7

Once virus x becomes a resident, the invasion fitness of an initially rare mutant y in resident x is given by

$$\mathcal{S}_x(y) = \mathbb{E}\left[\left(\tilde{\beta}(y) - \tilde{\beta}(x)\right)\frac{S}{M} - \left(\tilde{\gamma}(y) - \tilde{\gamma}(x)\right) - \left(\tilde{\alpha}(y) - \tilde{\alpha}(x)\right)\right]. \qquad (4.9)$$

Here we have used the principle of selective neutrality of residents (i.e., $\mathcal{S}_x(x) = 0$ for all $x \in \mathcal{X}$) to derive (4.9). Generally, there isn't an explicit expression of (4.9) in terms of strategies x and y. Thus the numerical PIP and MIP are based on a small-noise approximation of (4.9) (see e.g., Vilar and Rubi 2018).

Figure 7 shows the PIP and the MIP corresponding to a small-noise approximation of (4.9) (ref to the legend). Away from the two evolutionarily singular strategies x_{er}^* and x_{bp}^*, every successful invasion of mutant y in a sufficiently small neighborhood of resident x will takeover the population. Figure 8a gives a simulated population trajectory in the phase plane for strategies $x = 8.7$ and $y = 8.68$ with corresponding invasion fitnesses $\mathcal{S}_x(y) > 0$ and $\mathcal{S}_y(x) < 0$. The dynamics of P_t shows that the invader y will eventually oust the resident x and become the new resident. Close to the two evolutionarily singular strategies, Figure 8b gives a simulated population trajectory in the phase plane for strategies $x = 4.1158$ and $y = 4.1238$ with corresponding invasion fitnesses $\mathcal{S}_x(y) > 0$ and $\mathcal{S}_y(x) > 0$. The dynamics of P_t shows that the relative population size asymptotically tends to an interior value of $(0, 1)$, which implies that the two viral types can coexist.

4.4 A prey-predator model: continuation of Example 1

In order to show that the deterministic model of Example 1 belongs to the general class of models considered in this paper, let

$$G_1(e_t) = -\delta e_{1,t},$$

$$H_1(x_j, e_t) = \frac{\gamma\big(1 + \beta h e_{2,t}\big)^{\frac{\beta e_{1,t}}{(1+\beta h e_{2,t})(1+x_j e_{1,t})}}}{T\big(\sigma(1 + \beta h e_{2,t}) + (1-\lambda)\alpha e_{1,t}\big)},$$

$$G_2(e_t) = 0,$$

$$H_2(x_j, e_t) = \frac{1}{1 + x_j e_{1,t}},$$

then the deterministic model of Example 1 can be written as the general form (2.1). The partial derivatives $f^{(i,j)}$, $G_1^{(i,j)}$, $H_1^{(i,j)}$, $G_2^{(i,j)}$ and $H_2^{(i,j)}$ exist for all strategies and for all nonnegative integers $i, j \leq 2$, which satisfy Assumption **A2**.

The evolutionary dynamics of the deterministic model of Example 1 has been well studied by Lehtinen and Geritz (2019). We are interesting in the evolution of timidity of the prey with the non-equilibrium resident dynamics. When only a single prey type x is present, the system has a stable interior equilibrium at $x = 1$. Decreasing x destabilises the equilibrium through a supercritical Hopf bifurcation at $x_{\text{Hopf}} = 0.6289$, after which periodic attractors are present (ref to Lehtinen and Geritz (2019, Fig. 3B)). In a periodic resident environment set by a single resident type x, the invasion fitness of an initially rare mutant y is given by

$$\mathcal{S}_x(y) = \frac{1}{\tau(x)} \int_0^{\tau(x)} f(y, e_t)dt, \tag{4.10}$$

where $\tau(x)$ is the period (ref to Metz et al. 1992).

Figure 9 shows the PIP and the MIP corresponding to (4.10) with the same parameters as Lehtinen and Geritz (2019, Fig. 4), in which the monomorphic resident environment is an equilibrium environment if $x \geq x_{\text{Hopf}}$ and becomes a periodic
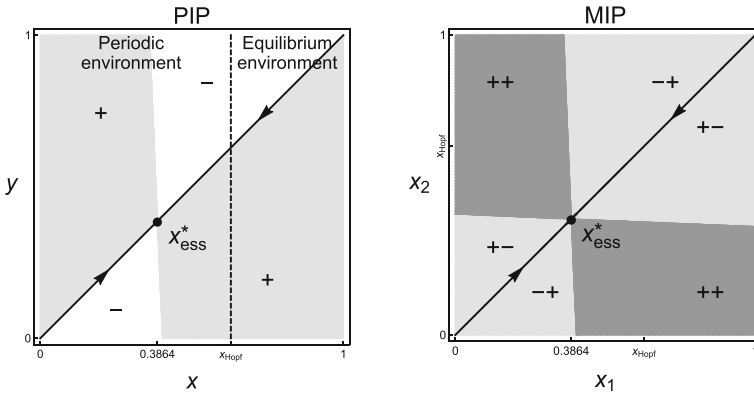
Fig. 9 Pairwise Invasibility Plot (PIP) and corresponding Mutual Invasibility Plot (MIP) of the invasion fitness (4.10). For the interpretations of local areas, markers and arrows in the PIP and the MIP see Fig. 5. Parameter values $a = c = 2$, $\mu = \delta = h = T = 1$, $\gamma = 3$, $\lambda = 0.6$, $\sigma = 0.7$, $\alpha = 0$, and $\beta = 6$. Further, the Hopf bifurcation point $x_{\text{Hopf}} = 0.6289$

environment if $x < x_{\text{Hopf}}$. In the periodically resident environment, away from the evolutionarily singular strategy x^*_{ess}, it follows from Theorem 1 that successful invasion of mutant y in resident x will takeover the population. Figure 10a gives a simulated population trajectory in the phase plane for strategies $x = 0.27$ and $y = 0.3$ with corresponding invasion fitnesses $\mathcal{S}_x(y) > 0$ and $\mathcal{S}_y(x) < 0$. The total population size of $N_t = n_{res,t} + n_{inv,t}$ varies significantly in time. However, the relative population size $P_t = \frac{n_{inv,t}}{N_t}$ changes slowly in time and asymptotically increases from 0 to 1. The dynamics of P_t implies that the invader y will oust the resident x and becomes a new resident. Figure 10c shows the periodic attractor of the monomorphic model with the new resident strategy $x = 0.3$. In the neighborhood of x^*_{ess}, we can apply Theorem 2 to predict the population dynamical outcomes of an invasion event. If both strategies of the resident and the mutant are in the gray area of MIP, it follows from Theorem 2 that the resident and the mutant can coexist eventually (notice that all coexistence strategies are not evolutionarily stable). Figure 10b gives a simulated population trajectory in the phase plane for the coexistence strategies $(x, y) = (0.3764, 0.4314)$ with corresponding invasion fitnesses $\mathcal{S}_x(y) > 0$ and $\mathcal{S}_y(x) > 0$. The total population size N_t still fluctuates significantly in time. but the relative population size P_t slowly and asymptotically tends to an interior value of $(0, 1)$. The dynamics of P_t verifies that the invader y eventually coexists with the resident x. Figure 10d shows the periodic attractor of the dimorphic model with coexistence strategies $(x_1, x_2) = (0.3764, 0.4314)$.

5 Discussion

Our main result is the complete classification of generic population dynamical outcomes of resident-invader dynamics in fluctuating environments when the invader and the resident have similar strategies. The outcomes are essentially "Lotka-Volterra": (i) one strategy ousts the other if the one can invade a population of the other but not the

(a) Substitution
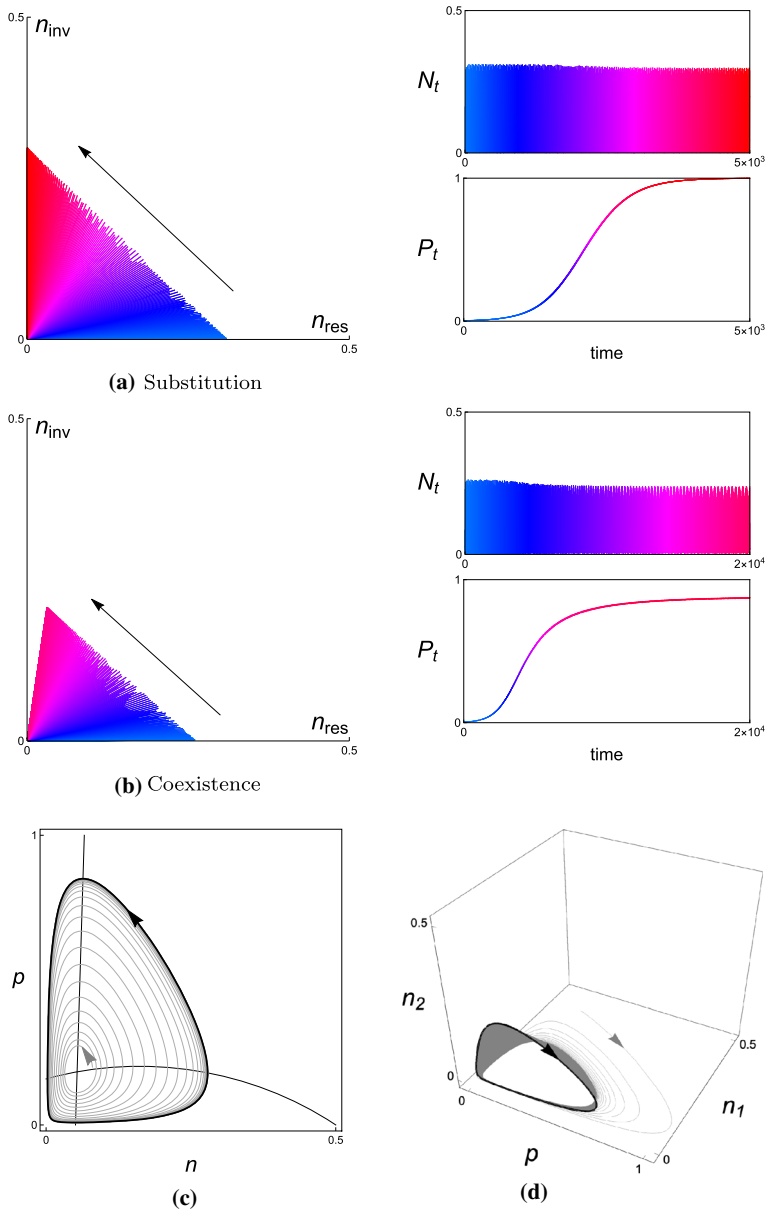

(b) Coexistence

(c)

(d)

Fig. 10 Simulated population trajectories in the phase plane and associated dynamics of the total population size $N_t = n_{res,t} + n_{inv,t}$ and the relative population size $P_t = n_{inv,t}/N_t$ for strategy pairs **(a)** $(x, y) = (0.27, 0.3)$ with corresponding invasion fitnesses $\mathcal{S}_x(y) \approx 0.002141 > 0$ and $\mathcal{S}_y(x) \approx -0.001671 < 0$, and **(b)** $(x, y) = (0.3764, 0.4314)$ with corresponding invasion fitnesses $\mathcal{S}_x(y) \approx 0.000277 > 0$ and $\mathcal{S}_y(x) \approx 0.001496 > 0$. **(c)** the periodic attractor of the monomorphic model with resident strategy $x = 0.3$. **(d)** the periodic attractor of the dimorphic model with coexistence strategies $(x_1, x_2) = (0.3764, 0.4314)$. Parameter values are the same as Fig. 9

other way around, i.e., invasion implies substitution; (ii) the two strategies coexist if they can mutual invade; (iii) they mutually exclude one another if neither can invade a population of the other. Which of the four Lotka-Volterra-type outcomes occurs depends only on the signs of the invasion fitnesses. In a one-dimensional strategy space or a one-dimensional parameterization of a multi-dimensional strategy space (see e.g., Kisdi 2015), invasion implies substitution away from an evolutionarily singular strategy, while all four Lotka-Volterra-type outcomes are possible close to an evolutionarily singular strategy. In a multi-dimensional strategy space, however, the situation is generally complicated because all four Lotka-Volterra-type outcomes may occur also away from an evolutionarily singular strategy.

We extend and generalize previous results of resident-invader dynamics of similar strategies to models incorporating (i) explicit feedback environments with their own dynamics, (ii) scalar-valued as well as vector-valued strategies, (iii) unstructured as well as structured populations, (iv) monomorphic as well as polymorphic resident populations, and (v) non-equilibrium resident population dynamics as well as resident dynamics with stochastic (or deterministic) drivers. Although we show all results for models with a monomorphic resident population (i.e., single resident phenotype), the generalization of them to polymorphic resident populations (i.e., multiple resident phenotypes) is straight forward because of the way we modelled the environment feedback loop. Arbitrarily polymorphic resident populations can be accounted for by treating the corresponding population sizes of the extra resident phenotypes as environmental feedback variables. In the next paragraphs we focus on the differences between our work and previous studies.

For a class of unstructured population models with point equilibria, Geritz (2005) has shown that the resident-invader dynamics generically behaves like the Lotka-Volterra competition model by using of the Poincaré-Bendixson theorem. The results hold for vector-valued strategies but were proved only for a feedback environment that is given as an explicit linear function of the resident and invader population sizes. However, environmental feedback variables such as resources, predators, or competitors are often given implicitly (e.g., by differential equations) as opposed to explicitly. Our implicit representation is more general than the explicit and liner environmental feedback used in Geritz (2005).

Dercole and Geritz (2016) has the same aims and results as Geritz (2005), but it is more general in some aspects but less general in others, and it uses a basically different mathematical approach, i.e., time-scale arguments. The results only apply to unstructured population models with point equilibria and scalar-valued strategies, but the formulation of environmental feedback is very general—probably as general as in our approach (in their formulation, they allow the resident and the invader to interact with finitely many other populations whose corresponding sizes are packed in the feedback variable e_t. The dynamics of e_t is govern by a differential equation, in which the growth rate is an implicit function of the resident, the invader, and the environmental variable itself. Given four structural properties of the growth rate of the e_t as well as the per-capita growth rates of the resident and the invader, their formulation generalizes the law of mass action (see Dercole 2016). The generalization takes into account that pairwise interactions can depend on the concomitant activities of the encountered individuals, which leads a nonlinear density dependence in the

per-capita growth rates of the resident and the invader). The paper also studies some mathematically degenerate cases which confirm the results of Priklopil (2012) on unprotected coexistence of two strategies near an evolutionarily singular strategy.

The "invasion implies substitution" outcome for unstructured population models with point equilibria also can be found in Dercole and Rinaldi (2008, Section 3.4) and Oba and Kigami (2018).

Our paper is the first to study resident-invader dynamics in fluctuating environments. We study the resident-invader dynamics as a stochastic fast-slow system where the total population size of the resident and the invader is the fast variable and the relative population size of the invader is the slow variable. We show that trajectories of the slow variable on slow timescale are well approximated by that of an associated averaged system, and the stability of the averaged system depends on invasion criteria alone. From these results, Theorems 1, 2 and 3 give the complete classification of generic population dynamical outcomes of an invasion event.

For a class of structured population models with point equilibria, recently, Priklopil and Lehmann (2019) have shown that invasion implies substitution for ecological communities with finite-class-structured populations of scalar-valued strategies. Their approach is based on the analysis of a weighted average of the relative invader population sizes in each class, where the weighting coefficient is the class-specific reproductive value (see also Lion 2018a). However, it is not clear how their method can be used to give a full classification of the generic population dynamical outcomes of an invasion event. Moreover, it is not clear how their approach can be applied to fluctuating environments.

Cantrell et al. (2017) has extended the *tube theorem* of Geritz et al. (2002) and the *invasion implies fixation theorem* of Geritz (2005) to a class of reaction-diffusion models for understanding evolution of dispersal in space. Their focus is on a dimorphic system (i.e., one resident and one invader) of infinite-dimensional structured populations of scalar-valued strategies.

When populations are in a fluctuating environment, we illustrate how to generalize the results of unstructured population models to a class of structured population models, but only for finite-class-structured populations. The extension of our results to populations with a more general structure (e.g., size distributions, continuous age and spatial diffusion) we leave for future work.

As mentioned in the Introduction, there are serval different (not necessarily equivalent) definitions of stochastic persistence for a resident population. Most of them can be summed up from the "ensemble" perspective in terms of transition probabilities (see e.g., Chesson 1982; Chesson and Ellner 1989; Li and Mao 2009) or the "typical trajectory" perspective in terms of empirical measures (i.e., how typical sample trajectories of the population process are distributed in time, see e.g., Benaïm and Schreiber 2009, 2019; Schreiber et al. 2011; Roth and Schreiber 2014; Benaïm 2018) for the long-term population dynamics. Different definitions often give different interpretations of the population dynamics. Extending our results to models with different kinds of resident dynamics would be useful for the applicability of the theory of adaptive dynamics. Our definition of resident persistence follows Schreiber et al. (2011) and Benaïm (2018) is a rather weak one. This means that our results also apply to more restrictive definitions.

Appendix A Rewrite the environmental variable of Example 2 as the general form (2.1b)

Still denote the environmental variable by e_t. Let $\mathcal{E}$ be the set of all functions $e : \mathcal{X} \mapsto \mathbb{R}_+^2$ of the form

$$
\begin{aligned}
e_t(\cdot) &= (e_{1,t}, e_{2,t})(\cdot) \\
&= \big(e_{1,t}, (e_{21,t}, e_{22,t}, e_{23,t}, e_{24,t})\big)(\cdot) \\
&= \bigg(0, \bigg(\sum_j a(\cdot, x_j)n_{j,t}, \ \sum_j h(\cdot)a(\cdot, x_j)n_{j,t}, \\
&\qquad \sum_j \frac{a(x_j, \cdot)}{1 + h_0(x_j)a_0(x_j)n_0 + e_{22,t}(x_j)}n_{j,t}, \ \sum_j cn_{j,t}\bigg)\bigg)
\end{aligned}
$$

for $n_{j,t} \geq 0$ and $x_j \in \mathcal{X}$. Then we can define the per-capita environmental impact $H_2 : \mathcal{X} \times \mathcal{E} \mapsto \mathbb{R}$ by

$$
H_2(x_j, e_t)(\cdot) = \bigg(a(\cdot, x_j), \ h(\cdot)a(\cdot, x_j), \ \frac{a(x_j, \cdot)}{1 + h_0(x_j)a_0(x_j)n_0 + e_{22,t}(x_j)}, \ c\bigg).
$$

Since the strategy of sub-population i has infinitely many choices, $e_{2,t}$ has infinite dimensions. In addition, define the map $G_2 : \mathcal{E} \times \mathbb{R} \mapsto \mathbb{R}$ by

$$
G_2(e_t, \theta) = (0, 0, 0, 0).
$$

Hence, e_t gets the general form (2.1b).

Appendix B Approximation on *t*-timescale when $\epsilon \to 0$

Writing (2.5) into the following form:

$$\dot{\widetilde{Z}}_t^\epsilon = \Phi(x + \epsilon\xi_1, x + \epsilon\xi_2, \widetilde{Z}_t^\epsilon),$$

where

$$\widetilde{Z}_t^\epsilon = (N_t^\epsilon, P_t^\epsilon, e_t^\epsilon, \theta_t) \in \mathcal{N} \times [0,1] \times \mathcal{E} \times \Theta := \widetilde{\mathcal{Z}}.$$

For $\epsilon = 0$, we have $\widetilde{Z}_t^0 = (N_t^0, P_t^0, e_t^0, \theta_t) \in \widetilde{\mathcal{Z}}$. Then (2.6) can be written as the same form, i.e.,

$$\dot{\widetilde{Z}}_t^0 = \Phi(x, x, \widetilde{Z}_t^0).$$

From Assumption **A1**, $\{\widetilde{Z}_t^\epsilon\}_{t \geq 0}$ is a Markov process for any given ϵ (ref to Arnold and Kliemann (1983, Lemma 2.1)). Let $\{\widetilde{\mathcal{T}}_t^\epsilon\}_{t\geq0}$ and $\{\widetilde{\mathcal{T}}_t^0\}_{t\geq0}$ are the associated semigroup for Markov process $\{\widetilde{Z}_t^\epsilon\}_{t\geq0}$ with $\epsilon > 0$ and for Markov process $\{\widetilde{Z}_t^0\}_{t\geq0}$ when $\epsilon = 0$, respectively. In addition, the function Φ requires to satisfy the following assumption:

A3 The partial derivatives $\Phi^{(i,j,k)}$ exit and are locally Lipschitz continuous and measurable in z for all strategies and all ϵ and for all $i, j, k \in \{0,1\}$.

For positive but small ϵ, $\widetilde{Z}_t^\epsilon$ is sufficiently closed to $\widetilde{Z}_t^0$ in terms of conditional expectations as the following lemma shows.

Lemma 3 *Let $\widetilde{Z}_0^\epsilon = \widetilde{Z}_0^0 = z$. Then, for all bounded and measurable function $h : \widetilde{\mathcal{Z}} \mapsto \mathbb{R}$, we have*

$$\lim_{\epsilon \to 0} \widetilde{\mathcal{T}}_t^\epsilon h(z) = \widetilde{\mathcal{T}}_t^0 h(z)$$

for all $t \geq 0$, and is uniformly in $z \in \widetilde{\mathcal{Z}}$.

Proof For all $t \geq 0$

$$\widetilde{Z}_t^\epsilon - \widetilde{Z}_t^0 = \int_0^t \left(\Phi(x + \epsilon\xi_1, x + \epsilon\xi_2, \widetilde{Z}_s^\epsilon) - \Phi(x + \epsilon\xi_1, x + \epsilon\xi_2, \widetilde{Z}_s^0) \right) ds$$

$$+ \int_0^t \left(\Phi(x + \epsilon\xi_1, x + \epsilon\xi_2, \widetilde{Z}_s^0) - \Phi(x + \epsilon\xi_1, x, \widetilde{Z}_s^0) \right) ds$$

$$+ \int_0^t \left(\Phi(x + \epsilon\xi_1, x, \widetilde{Z}_s^0) - \Phi(x, x, \widetilde{Z}_s^0) \right) ds.$$

Let

$$\varphi_z(t) = \mathbb{E}_z \left[\| \widetilde{Z}_t^\epsilon - \widetilde{Z}_t^0 \|^2 \right].$$

Using the Cauchy–Schwartz inequality we have

$$
\begin{aligned}
\varphi_z(t) \le 3\mathbb{E}_z \Bigg[& \left\| \int_0^t \left(\Phi(x + \epsilon\xi_1, x + \epsilon\xi_2, \widetilde{Z}_s^\epsilon) - \Phi(x + \epsilon\xi_1, x + \epsilon\xi_2, \widetilde{Z}_s^0) \right) ds \right\|^2 \\
& + \left\| \int_0^t \left(\Phi(x + \epsilon\xi_1, x + \epsilon\xi_2, \widetilde{Z}_s^0) - \Phi(x + \epsilon\xi_1, x, \widetilde{Z}_s^0) \right) ds \right\|^2 \\
& + \left\| \int_0^t \left(\Phi(x + \epsilon\xi_1, x, \widetilde{Z}_s^0) - \Phi(x, x, \widetilde{Z}_s^0) \right) ds \right\|^2 \Bigg] \\
\le 3t \int_0^t & \mathbb{E}_z \Big[\left\| \Phi(x + \epsilon\xi_1, x + \epsilon\xi_2, \widetilde{Z}_s^\epsilon) - \Phi(x + \epsilon\xi_1, x + \epsilon\xi_2, \widetilde{Z}_s^0) \right\|^2 \Big] ds \\
+ 3t \int_0^t & \mathbb{E}_z \Big[\left\| \Phi(x + \epsilon\xi_1, x + \epsilon\xi_2, \widetilde{Z}_s^0) - \Phi(x + \epsilon\xi_1, x, \widetilde{Z}_s^0) \right\|^2 \Big] ds \\
+ 3t \int_0^t & \mathbb{E}_z \Big[\left\| \Phi(x + \epsilon\xi_1, x, \widetilde{Z}_s^0) - \Phi(x, x, \widetilde{Z}_s^0) \right\|^2 \Big] ds.
\end{aligned}
$$

By Assumption **A3**, there exist some positive constants K_0, K_1 and K_2 such that

$$
\varphi_z(t) \le 3t K_0 \int_0^t \varphi_z(s) ds + 3t^2 (K_1\xi_1 + K_2\xi_2)\epsilon^2.
$$

Then, by the Gronwall's lemma, it follows that for all $t \in [0, T]$

$$
\varphi_z(t) \le \epsilon^2 \widetilde{K}_1 \exp(t\widetilde{K}_0),
$$

where $\widetilde{K}_0 = 3T K_0$ and $\widetilde{K}_1 = 3T^2(K_1\xi_1 + K_2\xi_2)$.

Let $h : \widetilde{\mathcal{Z}} \mapsto \mathbb{R}$ is a bounded and measurable function. By uniform continuity, there exists $\delta > 0$ such that for all $z_1, z_2 \in \widetilde{\mathcal{Z}}$ with $\|z_1 - z_2\| \le \delta$, we have $|h(z_1) - h(z_2)| \le \epsilon$. Using the Markov's inequality, we obtain that for $t \in [0, T]$

$$
\begin{aligned}
\left| \widetilde{T}_t^\epsilon h(z) - \widetilde{T}_t^0 h(z) \right| &= \left| \mathbb{E}_z \big[h(\widetilde{Z}_t^\epsilon) \big] - \mathbb{E}_z \big[h(\widetilde{Z}_t^0) \big] \right| \\
&\le 2\|h\| \mathbb{P}_z \big\{ \|\widetilde{Z}_t^\epsilon - \widetilde{Z}_t^0\| \ge \delta \big\} + \epsilon \\
&\le 2\|h\| \delta^{-1} \mathbb{E}_z \big[\|\widetilde{Z}_t^\epsilon - \widetilde{Z}_t^0\| \big] + \epsilon \\
&\le 2\|h\| \delta^{-1} \sqrt{\varphi_z(t)} + \epsilon \\
&\le \epsilon \Big(2|h|\delta^{-1} \sqrt{\widetilde{K}_1 \exp(t\widetilde{K}_0)} + 1 \Big).
\end{aligned}
$$

This inequality holds for any $T > 0$, we thus get the desired result for all $t \ge 0$.
    The proof is complete.   □

Appendix C Proof of Lemma 1

From now on, we fix initial values $P_0^\epsilon = \bar{P}_0 = p$ and $Z_0 = z$. The proof of Lemma 1 will be separated into the following three steps.

Firstly, we show that the trajectory of $\bar{P}_{t_1}$ is sufficiently close to that of $P_{t_1}^\epsilon$ for $\epsilon > 0$ small in every finite time interval.

Lemma 4 *Assume that for a small $\sigma > 0$, $\int_{\mathcal{Z} \setminus \mathcal{Z}_0} |f(y, e, \theta)|^2 \mu(de, d\theta) < +\infty$ for all $y \in \{\tilde{x} \in \mathcal{X} : \|\tilde{x} - x\| \leq \sigma\}$. Then for every $T > 0$, the solutions of (2.7) and (2.9) satisfy*

$$\mathbb{P}_{(p,z)} \left\{ \lim_{\epsilon \to 0} \sup_{0 \leq t_1 \leq T} |P_{t_1}^\epsilon - \bar{P}_{t_1}| = 0 \right\} = 1 \qquad \text{(c.1)}$$

for all $p \in [0, 1]$ and for μ-almost every $z \in \mathcal{Z} \setminus \mathcal{Z}_0$.

Proof Denote

$$h(z) = f^{(1,0)}(x, e, \theta)^\top (\xi_2 - \xi_1)$$

where $z = (n, e, \theta)$. Let $\bar{h}$ be the expectation of $h(z)$ with respect to μ, i.e.,

$$\bar{h} = \int_{\mathcal{Z} \setminus \mathcal{Z}_0} h(z) \mu(dz) = \partial_y \mathcal{S}_x(y)|_{y=x} (\xi_2 - \xi_1).$$

For all $t_1 \geq 0$, ignoring the term $\mathcal{O}(\epsilon)$ in (2.7),

$$
\begin{aligned}
P_{t_1}^\epsilon - \bar{P}_{t_1} &= \int_0^{t_1} \left(P_s^\epsilon (1 - P_s^\epsilon) h(Z_{s/\epsilon}) - \bar{P}_s (1 - \bar{P}_s) \bar{h} \right) ds \\
&= \int_0^{t_1} \left(P_s^\epsilon (1 - P_s^\epsilon) h(Z_{s/\epsilon}) - \bar{P}_s (1 - \bar{P}_s) h(Z_{s/\epsilon}) \right) ds \\
&\quad + \int_0^{t_1} \left(\bar{P}_s (1 - \bar{P}_s) h(Z_{s/\epsilon}) - \bar{P}_s (1 - \bar{P}_s) \bar{h} \right) ds.
\end{aligned}
$$

Fixed $T > 0$, using the Cauchy–Schwartz inequality we have

$$
\begin{aligned}
\sup_{0 \leq t_1 \leq T} |P_{t_1}^\epsilon - \bar{P}_{t_1}|^2 &\leq \sup_{0 \leq t_1 \leq T} 2 \left| \int_0^{t_1} (P_s^\epsilon - \bar{P}_s)(1 - P_s^\epsilon - \bar{P}_s) h(Z_{s/\epsilon}) ds \right|^2 \\
&\quad + \sup_{0 \leq t_1 \leq T} 2 \left| \int_0^{t_1} \bar{P}_s (1 - \bar{P}_s) (h(Z_{s/\epsilon}) - \bar{h}) ds \right|^2 \\
&\leq \sup_{0 \leq t_1 \leq T} 2 \int_0^{t_1} |P_s^\epsilon - \bar{P}_s|^2 ds \int_0^T |h(Z_{s/\epsilon})|^2 ds \\
&\quad + \sup_{0 \leq t_1 \leq T} 2 \left| \int_0^{t_1} \bar{P}_s (1 - \bar{P}_s) (h(Z_{s/\epsilon}) - \bar{h}) ds \right|^2.
\end{aligned}
$$

Denote

$$\alpha_\epsilon = 2 \int_0^T |h(Z_{s/\epsilon})|^2 ds = 2T \frac{1}{T/\epsilon} \int_0^{T/\epsilon} |h(Z_s)|^2 ds,$$

$$\beta_\epsilon = \sup_{0 \le t_1 \le T} 2 \left| \int_0^{t_1} \bar{P}_s (1 - \bar{P}_s)(h(Z_{s/\epsilon}) - \bar{h}) ds \right|^2.$$

By the Gronwall's lemma, it follows that

$$\sup_{0 \le t_1 \le T} |P_{t_1}^\epsilon - \bar{P}_{t_1}|^2 \le \beta_\epsilon \exp(\alpha_\epsilon).$$

We now focus on the limit of α_ϵ and β_ϵ as $\epsilon \to 0$, respectively. By Birkhoff's ergodic theorem, we obtain that for all $p \in [0, 1]$ and for μ-almost every $z \in \mathcal{Z} \setminus \mathcal{Z}_0$

$$\mathbb{P}_{(p,z)} \left\{ \lim_{\epsilon \to 0} \alpha_\epsilon = 2T \int_{\mathcal{Z} \setminus \mathcal{Z}_0} |h(z)|^2 \mu(dz) < +\infty \right\} = 1. \tag{c.2}$$

Here, we have used that $\int_{\mathcal{Z} \setminus \mathcal{Z}_0} |f^{(1,0)}(x, e, \theta)^\top (\xi_2 - \xi_1)|^2 \mu(de, d\theta) < +\infty$ due to the hypothesis of the lemma and

$$\left| f^{(1,0)}(x, e, \theta)^\top (\xi_2 - \xi_1) \right|^2 \le K_\sigma \left| f^{(1,0)}(x, e, \theta)^\top (y - x) \right|^2$$

$$= K_\sigma \left| f(y, e, \theta) - f(x, e, \theta) - \mathcal{O}(\|y - x\|^2) \right|^2$$

$$\le 3 K_\sigma |f(y, e, \theta)|^2 + 3 K_\sigma |f(x, e, \theta)|^2 + \mathcal{O}(\|y - x\|^4)$$

where the positive constant K_σ satisfies $\xi_2 - \xi_1 = \sqrt{K_\sigma}(y - x)$ for $y \in \{\tilde{x} \in \mathcal{X} : \|\tilde{x} - x\| \le \sigma\}$. In fact, for given ξ_1, ξ_2 and x, we can find a σ to be such that the hypothesis of the lemma is satisfied, and then determine y and K_σ.

To estimate the term β_ϵ, we divide $[0, s]$ into some intervals depending of a given size $\triangle$ and define a function $\bar{P}_s^\triangle$ by

$$\bar{P}_s^\triangle = \bar{P}_{[\frac{s}{\triangle}]\triangle}$$

where $[\frac{s}{\triangle}]$ is the integer part of $\frac{s}{\triangle}$. Then, using the Cauchy–Schwartz inequality we have

$$\beta_\epsilon = \sup_{0 \le t_1 \le T} 2 \left| \int_0^{t_1} \left(\bar{P}_s (1 - \bar{P}_s) h(Z_{s/\epsilon}) - \bar{P}_s^\triangle (1 - \bar{P}_s^\triangle) h(Z_{s/\epsilon}) \right. \right.$$

$$+ \bar{P}_s^\triangle (1 - \bar{P}_s^\triangle) h(Z_{s/\epsilon}) - \bar{P}_s^\triangle (1 - \bar{P}_s^\triangle) \bar{h}$$

$$\left. \left. + \bar{P}_s^\triangle (1 - \bar{P}_s^\triangle) \bar{h} - \bar{P}_s (1 - \bar{P}_s) \bar{h} \right) ds \right|^2$$

$$\le \sup_{0 \le t_1 \le T} 6 \left| \int_0^{t_1} \left(\bar{P}_s (1 - \bar{P}_s) h(Z_{s/\epsilon}) - \bar{P}_s^\triangle (1 - \bar{P}_s^\triangle) h(Z_{s/\epsilon}) \right) ds \right|^2$$

$$+ \sup_{0 \le t_1 \le T} 6 \left| \int_0^{t_1} \left(\bar{P}_s^\Delta (1 - \bar{P}_s^\Delta) h(Z_{s/\epsilon}) - \bar{P}_s^\Delta (1 - \bar{P}_s^\Delta) \bar{h} \right) ds \right|^2$$

$$+ \sup_{0 \le t_1 \le T} 6 \left| \int_0^{t_1} \left(\bar{P}_s^\Delta (1 - \bar{P}_s^\Delta) \bar{h} - \bar{P}_s (1 - \bar{P}_s) \bar{h} \right) ds \right|^2$$

$$\le \sup_{0 \le t_1 \le T} 6 \int_0^{t_1} \left| (\bar{P}_s - \bar{P}_s^\Delta)(1 - \bar{P}_s - \bar{P}_s^\Delta) \right|^2 ds \int_0^T \left| h(Z_{s/\epsilon}) \right|^2 ds$$

$$+ \sup_{0 \le t_1 \le T} 6 \left| \int_0^{t_1} \bar{P}_s^\Delta (1 - \bar{P}_s^\Delta) \left(h(Z_{s/\epsilon}) - \bar{h} \right) ds \right|^2$$

$$+ \sup_{0 \le t_1 \le T} 6 \bar{h}^2 T \int_0^T \left| (\bar{P}_s^\Delta - \bar{P}_s)(1 - \bar{P}_s^\Delta - \bar{P}_s) \right|^2 ds$$

$$\le \sup_{0 \le t_1 \le T} 6 \int_0^{t_1} |\bar{P}_s - \bar{P}_s^\Delta|^2 ds \int_0^T \left| h(Z_{s/\epsilon}) \right|^2 ds$$

$$+ \sup_{0 \le t_1 \le T} 6 \left| \int_0^{t_1} \bar{P}_s^\Delta (1 - \bar{P}_s^\Delta) \left(h(Z_{s/\epsilon}) - \bar{h} \right) ds \right|^2$$

$$+ \sup_{0 \le t_1 \le T} 6 \bar{h}^2 T \int_0^T |\bar{P}_s^\Delta - \bar{P}_s|^2 ds$$

$$\le (3\alpha_\epsilon T + 6\bar{h}^2 T^2) \sup_{0 \le s \le T} |\bar{P}_s - \bar{P}_s^\Delta|^2$$

$$+ \sup_{0 \le t_1 \le T} 6 \left| \int_0^{t_1} \bar{P}_s^\Delta (1 - \bar{P}_s^\Delta) \left(h(Z_{s/\epsilon}) - \bar{h} \right) ds \right|^2.$$

Following the proof of Liu and Krstic (2012, (4.110) and (4.111) in pp. 73–74), we immediately have, for all $p \in [0, 1]$ and for μ-almost every $z \in \mathcal{Z} \setminus \mathcal{Z}_0$,

$$\mathbb{P}_{(p,z)} \left\{ \lim_{\epsilon \to 0} \sup_{0 \le t_1 \le T} 6 \left| \int_0^{t_1} \bar{P}_s^\Delta (1 - \bar{P}_s^\Delta) \left(h(Z_{s/\epsilon}) - \bar{h} \right) ds \right|^2 = 0 \right\} = 1.$$

By this limit and

$$\lim_{\Delta \to 0} \sup_{0 \le s \le T} |\bar{P}_s - \bar{P}_s^\Delta|^2 = 0,$$

we obtain that for all $p \in [0, 1]$ and for μ-almost every $z \in \mathcal{Z} \setminus \mathcal{Z}_0$

$$\mathbb{P}_{(p,z)} \left\{ \lim_{\epsilon \to 0} \beta_\epsilon = 0 \right\} = 1. \tag{c.3}$$

Hence, by (c.2) and (c.3), for all $p \in [0, 1]$ and for μ-almost every $z \in \mathcal{Z} \setminus \mathcal{Z}_0$

$$\mathbb{P}_{(p,z)} \left\{ \lim_{\epsilon \to 0} \sup_{0 \le t_1 \le T} |P_{t_1}^\epsilon - \bar{P}_{t_1}|^2 = 0 \right\} = 1$$

which implies the desired result of the lemma.

The proof is complete.   □

Secondly, following the proof of Liu and Krstic (2012, the assertion (i) of Theorem 4.1, see pp. 65–66), the finite-time approximation (c.1) can be extended to arbitrarily long time intervals as the following lemma shows. For the sake of self-containedness, we give the proof.

Lemma 5 *Assume that the hypothesis of Lemma* 4 *is satisfied. Then for all* $\delta \in (0, 1)$, *the solutions of* (2.7) *and* (2.9) *satisfy*

$$\mathbb{P}_{(p,z)}\left\{ \lim_{\epsilon \to 0} \inf \left\{ t_1 \ge 0 : |P_{t_1}^{\epsilon} - \bar{P}_{t_1}| > \delta \right\} = +\infty \right\} = 1 \qquad \text{(c.4)}$$

for all $p \in [0, 1]$ *and for* μ-*almost every* $z \in \mathcal{Z} \setminus \mathcal{Z}_0$.

Proof Define a set $\widetilde{\Omega} = \left\{ \omega : \lim_{\epsilon \to 0} \sup_{0 \le t_1 \le T} |P_{t_1}^{\epsilon}(\omega) - \bar{P}_{t_1}| = 0, \ \forall \, T \in \mathbb{N} \right\}$. From Lemma 4, it follows that $\mathbb{P}\{\widetilde{\Omega}\} = 1$. Let stoping time $\tau_{\delta}^{\epsilon} = \inf \left\{ t_1 \ge 0 : |P_{t_1}^{\epsilon} - \bar{P}_{t_1}| > \delta \right\}$ for $\delta \in (0, 1)$ and $\epsilon > 0$. Notice that $P_0^{\epsilon} - \bar{P}_0 = 0$. By the continuity of the trajectories, $\tau_{\delta}^{\epsilon} \in (0, +\infty]$ for all δ and ϵ. Typically, $|P_{\tau_{\delta}^{\epsilon}}^{\epsilon} - \bar{P}_{\tau_{\delta}^{\epsilon}}| = \delta$ provided that $\tau_{\delta}^{\epsilon} < +\infty$.

For fixed $\delta \in (0, 1)$ and any $\omega \in \widetilde{\Omega}$, we get that for any $T > 0$, there exists $\tilde{\epsilon}(\omega, \delta, T) > 0$ such that $\tau_{\delta}^{\epsilon}(\omega) > T$ for any $0 < \epsilon < \tilde{\epsilon}(\omega, \delta, T)$. It further implies that $\lim_{\epsilon \to 0} \tau_{\delta}^{\epsilon}(\omega) = +\infty$. From the fact $\mathbb{P}(\widetilde{\Omega}) = 1$, it follows that $\mathbb{P}\left\{ \lim_{\epsilon \to 0} \tau_{\delta}^{\epsilon} = +\infty \right\} = 1$.

The proof is complete.   □

Finally, from the averaged system (2.9), we have

$$\lim_{t_1 \to +\infty} \bar{P}_{t_1} = p^*$$

with $p^* = 0$ or 1, which is determined by the sign of $\partial_y \mathcal{S}_x(y)|_{y=x} (\xi_2 - \xi_1)$. It means that for every $\delta \in (0, 1)$, there exists a constant $T_{\delta} > 0$ such that

$$\sup_{t_1 \ge T_{\delta}} |\bar{P}_{t_1} - p^*| < \frac{\delta}{2}.$$

From this, the event

$$\left\{ \sup_{t_1 \ge T_{\delta}} |P_{t_1}^{\epsilon} - p^*| > \delta \right\} = \bigcup_{t_1 \ge T_{\delta}} \left\{ |P_{t_1}^{\epsilon} - \bar{P}_{t_1} + \bar{P}_{t_1} - p^*| > \delta \right\}$$

$$\subseteq \bigcup_{t_1 \ge T_{\delta}} \left\{ |P_{t_1}^{\epsilon} - \bar{P}_{t_1}| > \frac{\delta}{2} \right\}$$

$$\subseteq \bigcup_{t_1 \ge 0} \left\{ |P_{t_1}^{\epsilon} - \bar{P}_{t_1}| > \frac{\delta}{2} \right\}$$

$$= \left\{ \sup_{t_1 \geq 0} |P_{t_1}^\epsilon - \bar{P}_{t_1}| > \frac{\delta}{2} \right\},$$

and then the stopping time

$$\inf \left\{ t_1 \geq T_\delta : |P_{t_1}^\epsilon - p^*| > \delta \right\} \geq \inf \left\{ t_1 \geq 0 : |P_{t_1}^\epsilon - \bar{P}_{t_1}| > \frac{\delta}{2} \right\}.$$

Thus, it follows from (c.4) that

$$\mathbb{P}_{(p,z)} \left\{ \lim_{\epsilon \to 0} \inf \left\{ t_1 \geq T_\delta : |P_{t_1}^\epsilon - p^*| > \delta \right\} = +\infty \right\} = 1$$

which is equivalent to (2.10).

The proof of Lemma 1 is complete.

Appendix D Derivation of (2.12)

On t_2-timescale, (2.7) becomes

$$\dot{P}_{t_2}^\epsilon = \frac{1}{2} P_{t_2}^\epsilon (1 - P_{t_2}^\epsilon) \big(\xi_2^\top f^{(2,0)}(x, e_{t_2/\epsilon^2}, \theta_{t_2/\epsilon^2}) \xi_2 - \xi_1^\top f^{(2,0)}(x, e_{t_2/\epsilon^2}, \theta_{t_2/\epsilon^2}) \xi_1$$

$$+ 2 e_{t_2/\epsilon^2}^{(1)\top} f^{(1,1)}(x, e_{t_2/\epsilon^2}, \theta_{t_2/\epsilon^2}) (\xi_2 - \xi_1) \big) + \mathcal{O}(\epsilon) \qquad \text{(d.1)}$$

$$:= P_{t_2}^\epsilon (1 - P_{t_2}^\epsilon) \phi(x, e_{t_2/\epsilon^2}, \theta_{t_2/\epsilon^2}, e_{t_2/\epsilon^2}^{(1)}) + \mathcal{O}(\epsilon).$$

The displacement of the trajectory $P_{t_2}^\epsilon$ starting from p over a small time $\triangle$ is

$$P_\triangle^\epsilon - p = \int_0^\triangle \big(P_s^\epsilon (1 - P_s^\epsilon) \phi(x, e_{s/\epsilon^2}, \theta_{s/\epsilon^2}, e_{s/\epsilon^2}^{(1)}) + \mathcal{O}(\epsilon) \big) ds$$

$$= \int_0^\triangle \big(p(1 - p) \phi(x, e_{s/\epsilon^2}, \theta_{s/\epsilon^2}, e_{s/\epsilon^2}^{(1)}) + \mathcal{O}(\epsilon) \big) ds$$

$$+ \int_0^\triangle \big(P_s^\epsilon (1 - P_s^\epsilon) - p(1 - p) \big) \phi(x, e_{s/\epsilon^2}, \theta_{s/\epsilon^2}, e_{s/\epsilon^2}^{(1)}) ds$$

$$= \triangle \left(\frac{\epsilon^2}{\triangle} \int_0^{\frac{\triangle}{\epsilon^2}} p(1 - p) \phi(x, e_s, \theta_s, e_s^{(1)}) ds \right) + \mathcal{O}(\epsilon, \triangle),$$

where the high order term $\mathcal{O}(\epsilon, \triangle)$ converges to zero as $\triangle \to 0$ and $\epsilon \to 0$. Following from a standard argument as Freidlin and Wentzell (2012, Chapter 7), we claim that the solutions of $P_{t_2}^\epsilon$ converges to that of $\bar{P}_{t_2}$ of the following averaged system

$$\dot{\bar{P}}_{t_2} = \bar{P}_{t_2}(1 - \bar{P}_{t_2}) \langle \phi(x, e_s, \theta_s, e_s^{(1)}) \rangle_{+\infty}, \quad \bar{P}_0 = p, \qquad \text{(d.2)}$$

on every finite time interval as $\epsilon \to 0$, where $\langle \cdot \rangle_{+\infty}$ means the time average for $t_2/\epsilon^2 \to +\infty$ as $\epsilon \to 0$. This result further can be extended to arbitrarily long time intervals in the same manner as in Lemmas 4 and 5.

To arrive the averaged system (2.12), we need the time evolution of $e_t^{(1)}$ in fast time t. From now on, we will use the simplified symbols (N_t, e_t) to replace (N_t^0, e_t^0) used in Sect. 2.5.1. Consider the series expansion of model (2.5) with $P_t = p$ for all t, the first-order terms in ϵ of (2.5a), (2.5c) and (2.5d) are

$$
\begin{aligned}
\dot{N}_t^{(1)} &= f N_t^{(1)} + \big(e_t^{(1)\top} f^{(0,1)} + \big((1-p)\xi_1 + p\xi_2 \big)^\top f^{(1,0)} \big) N_t, \\
\dot{e}_{1,t}^{(1)} &= e_t^{(1)\top} G_1^{(1)} + H_1 N_t^{(1)} + \big(e_t^{(1)\top} H_1^{(0,1)} + \big((1-p)\xi_1 + p\xi_2 \big)^\top H_1^{(1,0)} \big) N_t, \\
\dot{e}_{2,t}^{(1)} &= e_t^{(1)\top} G_2^{(1)} + H_2 N_t^{(1)} + \big(e_t^{(1)\top} H_2^{(0,1)} + \big((1-p)\xi_1 + p\xi_2 \big)^\top H_2^{(1,0)} \big) N_t, \\
\dot{\theta}_t &= A(\theta_t) + B(\theta_t) \dot{W}_t,
\end{aligned}
$$

where functions f, $f^{(0,1)}$, $f^{(1,0)}$, $G_i^{(1)}$, H_i, $H_i^{(0,1)}$ and $H_i^{(1,0)}$ are all evaluated at (x, e_t, θ_t), and where (N_t, e_t, θ_t) are calculated from (2.6) that is independent of p. Generally, it is impossible to obtain an explicit expression for $e_t^{(1)}$ by solving the above equations. Instead, we focus on an equivalent relation which will be employed to derive the averaged system (2.12).

Consider the second-order terms in ϵ of the series expansions of $\mathcal{S}_{x_1}(x_2)$ and $\mathcal{S}_{x_1}(x_1)$, we have

$$
\begin{aligned}
\lim_{\epsilon \to 0} \frac{\mathcal{S}_{x_1}(x_2)}{\epsilon^2} &= \lim_{\epsilon \to 0} \frac{\mathcal{S}_{x+\epsilon\xi_1}(x+\epsilon\xi_2) - \mathcal{S}_{x+\epsilon\xi_1}(x+\epsilon\xi_1)}{\epsilon^2} \\
&= \frac{1}{2} \big\langle \xi_2^\top f^{(2,0)} \xi_2 - \xi_1^\top f^{(2,0)} \xi_1 + 2 \hat{e}_t^{(1)\top} f^{(1,1)} (\xi_2 - \xi_1) \big\rangle_{+\infty}
\end{aligned}
\tag{d.3}
$$

where $\hat{e}_t^{(1)}$ is calculated from

$$
\begin{aligned}
\dot{\hat{N}}_t^{(1)} &= f \hat{N}_t^{(1)} + \big(\hat{e}_t^{(1)\top} f^{(0,1)} + \xi_2^\top f^{(1,0)} \big) N_t, \\
\dot{\hat{e}}_{1,t}^{(1)} &= \hat{e}_t^{(1)\top} G_1^{(1)} + H_1 \hat{N}_t^{(1)} + \big(\hat{e}_t^{(1)\top} H_1^{(0,1)} + \xi_2^\top H_1^{(1,0)} \big) N_t, \\
\hat{e}_{2,t}^{(1)} &= \hat{e}_t^{(1)\top} G_2^{(1)} + H_2 \hat{N}_t^{(1)} + \big(\hat{e}_t^{(1)\top} H_2^{(0,1)} + \xi_2^\top H_2^{(1,0)} \big) N_t, \\
\dot{\theta}_t &= A(\theta_t) + B(\theta_t) \dot{W}_t.
\end{aligned}
\tag{d.4}
$$

Similarly, consider the second-order terms in ϵ of the series expansions of $\mathcal{S}_{x_2}(x_1)$ and $\mathcal{S}_{x_2}(x_2)$, we have

$$
\begin{aligned}
\lim_{\epsilon \to 0} \frac{\mathcal{S}_{x_2}(x_1)}{\epsilon^2} &= \lim_{\epsilon \to 0} \frac{\mathcal{S}_{x+\epsilon\xi_2}(x+\epsilon\xi_1) - \mathcal{S}_{x+\epsilon\xi_2}(x+\epsilon\xi_2)}{\epsilon^2} \\
&= -\frac{1}{2} \big\langle \xi_2^\top f^{(2,0)} \xi_2 - \xi_1^\top f^{(2,0)} \xi_1 + 2 \check{e}_t^{(1)\top} f^{(1,1)} (\xi_2 - \xi_1) \big\rangle_{+\infty}
\end{aligned}
\tag{d.5}
$$

where $\check{e}_t^{(1)}$ is calculated from

$$
\begin{aligned}
\dot{\check{N}}_t^{(1)} &= f\check{N}_t^{(1)} + \big(\check{e}_t^{(1)\top} f^{(0,1)} + \xi_1^\top f^{(1,0)}\big)N_t, \\
\dot{\check{e}}_{1,t}^{(1)} &= \check{e}_t^{(1)\top} G_1^{(1)} + H_1 \check{N}_t^{(1)} + \big(\check{e}_t^{(1)\top} H_1^{(0,1)} + \xi_1^\top H_1^{(1,0)}\big)N_t, \\
\check{e}_{2,t}^{(1)} &= \check{e}_t^{(1)\top} G_2^{(1)} + H_2 \check{N}_t^{(1)} + \big(\check{e}_t^{(1)\top} H_2^{(0,1)} + \xi_1^\top H_2^{(1,0)}\big)N_t, \\
\dot{\theta}_t &= A(\theta_t) + B(\theta_t)\dot{W}_t.
\end{aligned}
\tag{d.6}
$$

Denote

$$
\widetilde{N}_t^{(1)} = (1-p)\hat{N}_t^{(1)} + p\check{N}_t^{(1)}, \quad \tilde{e}_t^{(1)} = (1-p)\hat{e}_t^{(1)} + p\check{e}_t^{(1)}.
$$

By (d.3) and (d.5), we get

$$
\begin{aligned}
&\lim_{\epsilon \to 0} \frac{(1-p)\mathcal{S}_{x_1}(x_2) - p\mathcal{S}_{x_2}(x_1)}{\epsilon^2} \\
&= \frac{1}{2}\big\langle \xi_2^\top f^{(2,0)} \xi_2 - \xi_1^\top f^{(2,0)} \xi_1 + 2\tilde{e}_t^{(1)\top} f^{(1,1)}(\xi_2 - \xi_1) \big\rangle_{+\infty}.
\end{aligned}
\tag{d.7}
$$

It follows from (d.4) and (d.6) that the dynamics of $\widetilde{N}_t^{(1)}$ and $\tilde{e}_t^{(1)}$ are given by

$$
\begin{aligned}
\dot{\widetilde{N}}_t^{(1)} &= f\widetilde{N}_t^{(1)} + \Big(\tilde{e}_t^{(1)\top} f^{(0,1)} + \big((1-p)\xi_1 + p\xi_2\big)^\top f^{(1,0)}\Big)N_t, \\
\dot{\tilde{e}}_{1,t}^{(1)} &= \tilde{e}_t^{(1)\top} G_1^{(1)} + H_1 \widetilde{N}_t^{(1)} + \Big(\tilde{e}_t^{(1)\top} H_1^{(0,1)} + \big((1-p)\xi_1 + p\xi_2\big)^\top H_1^{(1,0)}\Big)N_t, \\
\tilde{e}_{2,t}^{(1)} &= \tilde{e}_t^{(1)\top} G_2^{(1)} + H_2 \widetilde{N}_t^{(1)} + \Big(\tilde{e}_t^{(1)\top} H_2^{(0,1)} + \big((1-p)\xi_1 + p\xi_2\big)^\top H_2^{(1,0)}\Big)N_t, \\
\dot{\theta}_t &= A(\theta_t) + B(\theta_t)\dot{W}_t,
\end{aligned}
$$

which shows that $(\widetilde{N}_t^{(1)}, \tilde{e}_t^{(1)})$ has the same dynamics as $(N_t^{(1)}, e_t^{(1)})$.

Since by (d.1), (d.2) and (d.7) we obtain that

$$
\frac{\dot{\bar{P}}_{t_2}}{\bar{P}_{t_2}(1-\bar{P}_{t_2})} = \lim_{\epsilon \to 0} \frac{(1-\bar{P}_{t_2})\mathcal{S}_{x_1}(x_2) - \bar{P}_{t_2}\mathcal{S}_{x_2}(x_1)}{\epsilon^2},
\tag{d.8}
$$

indicating that now we can express the dynamics of $\bar{P}_{t_2}$ in terms of the expansions of the two invasion fitnesses. From $\mathcal{S}_x(x) = 0$ applied to the second-order terms in its series expansion we have

$$
\mathcal{C}_{11} + \mathcal{C}_{22} + \mathcal{C}_{12} + \mathcal{C}_{21} = 0,
\tag{d.9}
$$

where $\{\mathcal{C}_{ij}\}_{i,j \in \{1,2\}}$ are the second derivatives of invasion fitness (ref to (2.13)). Using the second-order terms in ϵ of the series expansions of $\mathcal{S}_{x_1}(x_2)$ and $\mathcal{S}_{x_2}(x_1)$ and

perusing (d.9), (d.8) is equivalent to

$$
\begin{aligned}
\frac{\dot{\bar{P}}_{t_2}}{\bar{P}_{t_2}(1 - \bar{P}_{t_2})} &= (1 - \bar{P}_{t_2})(\xi_1^\top C_{11}\xi_1 + \xi_2^\top C_{22}\xi_2 + 2\xi_1^\top C_{21}\xi_2) \\
&\quad - \bar{P}_{t_2}(\xi_2^\top C_{11}\xi_2 + \xi_1^\top C_{22}\xi_1 + 2\xi_2^\top C_{21}\xi_1) \\
&= (1 - \bar{P}_{t_2})(\xi_2^\top C_{22}\xi_2 - \xi_1^\top C_{22}\xi_1 + 2\xi_1^\top C_{21}(\xi_2 - \xi_1)) \\
&\quad - \bar{P}_{t_2}(-\xi_2^\top C_{22}\xi_2 + \xi_1^\top C_{22}\xi_1 - 2\xi_2^\top C_{21}(\xi_2 - \xi_1)) \\
&= \xi_2^\top C_{22}\xi_2 - \xi_1^\top C_{22}\xi_1 + 2((1 - \bar{P}_{t_2})\xi_1 + \bar{P}_{t_2}\xi_2)^\top C_{21}(\xi_2 - \xi_1).
\end{aligned}
$$

To complete the derivation of (2.12), the right side of the last equality needs to be simplified via the following three steps: firstly,

$$
\xi_2^\top C_{22}\xi_2 - \xi_1^\top C_{22}\xi_1 = (\xi_2 + \xi_1)^\top C_{22}(\xi_2 - \xi_1);
$$

secondly,

$$
\begin{aligned}
&2((1 - \bar{P}_{t_2})\xi_1 + \bar{P}_{t_2}\xi_2)^\top C_{21}(\xi_2 - \xi_1) \\
&= ((1 - \bar{P}_{t_2})\xi_1 + \bar{P}_{t_2}\xi_2)^\top C_{21}(\xi_2 - \xi_1) + (\xi_2 - \xi_1)^\top C_{12}((1 - \bar{P}_{t_2})\xi_1 + \bar{P}_{t_2}\xi_2) \\
&= \xi_1^\top C_{21}(\xi_2 - \xi_1) + (\xi_2 - \xi_1)^\top C_{12}\xi_1 - \bar{P}_{t_2}(\xi_2 - \xi_1)^\top (C_{22} + C_{11})(\xi_2 - \xi_1);
\end{aligned}
$$

finally,

$$
\begin{aligned}
&\xi_1^\top C_{21}(\xi_2 - \xi_1) + (\xi_2 - \xi_1)^\top C_{12}\xi_1 \\
&= (\xi_2 + \xi_1)^\top C_{21}(\xi_2 - \xi_1) - (\xi_2 - \xi_1)^\top C_{21}(\xi_2 - \xi_1) \\
&= (\xi_2 + \xi_1)^\top C_{21}(\xi_2 - \xi_1) + \frac{1}{2}(\xi_2 - \xi_1)^\top (C_{22} + C_{11})(\xi_2 - \xi_1).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\frac{\dot{\bar{P}}_{t_2}}{\bar{P}_{t_2}(1 - \bar{P}_{t_2})} &= (\xi_2 + \xi_1)^\top (C_{22} + C_{21})(\xi_2 - \xi_1) \\
&\quad + \left(\frac{1}{2} - \bar{P}_{t_2}\right)(\xi_2 - \xi_1)^\top (C_{22} + C_{11})(\xi_2 - \xi_1)
\end{aligned}
$$

which is the desired result.

Appendix E Proof of Lemma 2

The first assertion can be proven in the same manner as in Lemma 1, in which the associated averaged system becomes (2.12). The details are left to the reader.

Next, we focus on the proof of the second assertion. Let $P_0^\epsilon = \bar{P}_0 = p$ and $Z_0 = z$. Denote

$$p_0^* = 0, \quad p_{1/2}^* = \frac{\mathcal{S}_{x_1}(x_2)}{\mathcal{S}_{x_1}(x_2) + \mathcal{S}_{x_2}(x_1)}, \quad p_1^* = 1.$$

One hand, from the averaged system (2.12), the bistability indicates that $p_{1/2}^* \in (0, 1)$ is a saddle node such that

$$\lim_{t_2 \to +\infty} \bar{P}_{t_2} = p_0^* \quad \text{for all } p \in [0, p_{1/2}^*),$$

and

$$\lim_{t_2 \to +\infty} \bar{P}_{t_2} = p_1^* \quad \text{for all } p \in (p_{1/2}^*, 1].$$

On the other hand, on t_2-timescale, we also have the similar result as Lemma 5, i.e., for all $\delta \in (0, 1)$

$$\mathbb{P}_{(p,z)} \left\{ \lim_{\epsilon \to 0} \inf \left\{ t_2 \geq 0 : |P_{t_2}^\epsilon - \bar{P}_{t_2}| > \delta \right\} = +\infty \right\} = 1 \qquad \text{(e.1)}$$

for all $p \in [0, 1] \setminus \{p_{1/2}^*\}$ and for μ-almost every $z \in \mathcal{Z} \setminus \mathcal{Z}_0$.

Let stopping times

$$\tau_{0,\delta}^\epsilon := \inf \left\{ t_2 \geq T_\delta : |P_{t_2}^\epsilon - p_0^*| > \delta \right\},$$
$$\tau_{1,\delta}^\epsilon := \inf \left\{ t_2 \geq T_\delta : |P_{t_2}^\epsilon - p_1^*| > \delta \right\}.$$

For any $\varepsilon > 0$, by (e.1), the continuity of trajectories and the Markov property, there is $\tilde{\delta} \in \left(0, \min\{p_{1/2}^*, 1 - p_{1/2}^*\}\right)$ such that for every $\delta \in (0, \tilde{\delta})$, there exists $T_\delta > 0$ such that

$$\tilde{q}_{(p,z)}^0 := \mathbb{P}_{(p,z)} \left\{ \lim_{\epsilon \to 0} \tau_{0,\delta}^\epsilon = +\infty \right\} \geq 1 - \varepsilon \quad \text{for all } p \in [0, \delta)$$

and

$$\tilde{q}_{(p,z)}^1 := \mathbb{P}_{(p,z)} \left\{ \lim_{\epsilon \to 0} \tau_{1,\delta}^\epsilon = +\infty \right\} \geq 1 - \varepsilon \quad \text{for all } p \in (1 - \delta, 1].$$

We now show that $\tilde{q}_{(p,z)}^0 + \tilde{q}_{(p,z)}^1 = 1$ for all $p \in [0, 1] \setminus \{p_{1/2}^*\}$ and for μ-almost every $\mathcal{Z} \setminus \mathcal{Z}_0$. For this we only need to prove that the interval $[\delta, 1 - \delta]$ is transient, i.e., almost every trajectory of $P_{t_2}^\epsilon$ will escape from $[\delta, 1 - \delta]$ after some finite times later. In fact, by (e.1), there exist $\tilde{T}_\delta > 0$ and $\tilde{\epsilon} > 0$ such that $\mathbb{P}_{(p,z)}\{|P_{t_2}^\epsilon - \bar{P}_{t_2}| \leq \frac{\delta}{2}$ for all $0 < \epsilon \leq \tilde{\epsilon}$ and $t_2 \geq \tilde{T}_\delta\} \geq 1 - \varepsilon$. Hence, for any $t_2 \geq \tilde{T}_\delta$ and $p \in (\delta, 1 - \delta)$, the event

$$\left\{ P_{t_2}^\epsilon \notin [\delta, 1 - \delta] \right\} = \left\{ P_{t_2}^\epsilon - \bar{P}_{t_2} + \bar{P}_{t_2} \notin [\delta, 1 - \delta] \right\}$$

$$= \left\{ \bar{P}_{t_2} \notin \left[\delta - (P_{t_2}^\epsilon - \bar{P}_{t_2}), \, 1 - \delta - (P_{t_2}^\epsilon - \bar{P}_{t_2}) \right] \right\}$$
$$\supseteq \left\{ \bar{P}_{t_2} \notin \left[\frac{\delta}{2}, \, 1 - \frac{\delta}{2} \right] \right\},$$

and then the stopping time

$$\inf \left\{ t_2 \geq 0 : P_{t_2}^\epsilon \notin [\delta, 1 - \delta] \right\} \leq \inf \left\{ t_2 \geq \widetilde{T}_\delta : P_{t_2}^\epsilon \notin [\delta, 1 - \delta] \right\}$$
$$\leq \inf \left\{ t_2 \geq \widetilde{T}_\delta : \bar{P}_{t_2} \notin \left[\frac{\delta}{2}, \, 1 - \frac{\delta}{2} \right] \right\},$$

which implies that

$$\mathbb{P}_{(p,z)} \left\{ \liminf_{\epsilon \to 0} \left\{ t_2 \geq 0 : P_{t_2}^\epsilon \notin [\delta, 1 - \delta] \right\} < +\infty \right\} \geq 1 - \varepsilon.$$

It follows from the Markov property that

$$\mathbb{P}_{(p,z)} \left\{ \lim_{\epsilon \to 0} \tau_{0,\delta}^\epsilon = +\infty \ \text{ or } \ \lim_{\epsilon \to 0} \tau_{1,\delta}^\epsilon = +\infty \right\} \geq 1 - \varepsilon$$

for all $p \in [0, 1] \setminus \{p_{1/2}^*\}$ and for μ-almost every $\mathcal{Z} \setminus \mathcal{Z}_0$. Since the ε is taken arbitrarily, we claim that $\tilde{q}_{(p,z)}^0 + \tilde{q}_{(p,z)}^1 = 1$.

For the case of $p = p_{1/2}^*$, we have $P_{t_2}^\epsilon \neq p_{1/2}^*$ some time later due to the continuity of solutions. Thus, the above analysis is still valid so that $\tilde{q}_{(p,z)}^0 + \tilde{q}_{(p,z)}^1 = 1$ for all $p \in [0, 1]$.

By the definitions of $\tau_{0,\delta}^\epsilon$ and $\tau_{1,\delta}^\epsilon$, $\tilde{q}_{(p,z)}^0 + \tilde{q}_{(p,z)}^1 = 1$ is equivalent to (2.16). This proves the second assertion of Lemma 2.

References

Arnold L (2001) Hasselmann's program revisited: the analysis of stochasticity in deterministic climate models. In: Imkeller P, von Storch JS (eds) Stochastic climate models, Birkhäuser, Basel, pp 141–157. https://doi.org/10.1007/978-3-0348-8287-3_5

Arnold L, Kliemann W (1983) Qualitative theory of stochastic systems. In: Bharucha-reid AT (ed) Probabilistic Analysis and Related Topics, Academic Press, pp 1–79, https://doi.org/10.1016/B978-0-12-095603-6.50006-3

Benaïm M (2018) Stochastic persistence. arXiv e-prints arXiv:1806.08450

Benaïm M, Schreiber SJ (2009) Persistence of structured populations in random environments. Theor Popul Biol 76(1):19–34. https://doi.org/10.1016/j.tpb.2009.03.007

Benaïm M, Schreiber SJ (2019) Persistence and extinction for stochastic ecological models with internal and external variables. J Math Biol 79(1):393–431. https://doi.org/10.1007/s00285-019-01361-4

Cantrell RS, Cosner C, Lam KY (2017) Resident-invader dynamics in infinite dimensional systems. J Differ Eq 263(8):4565–4616. https://doi.org/10.1016/j.jde.2017.05.029

Caraballo T, Han XY (2016) Applied nonautonomous and random dynamical systems: applied dynamical systems. Springer International Publishing AG. https://doi.org/10.1007/978-3-319-49247-6

Chesson PL (1982) The stabilizing effect of a random environment. J Math Biol 15(1):1–36. https://doi.org/10.1007/BF00275786

Chesson PL (1986) Environmental variation and the coexistence of species. In: Diamond J, Case T (eds) Community Ecology. Harper and Row, New York, pp 240–256

Chesson PL, Ellner S (1989) Invasibility and stochastic boundedness in monotonic competition models. J Math Biol 27(2):117–138. https://doi.org/10.1007/BF00276099

Dercole F (2003) Remarks on branching-extinction evolutionary cycles. J Math Biol 47(6):569–580. https://doi.org/10.1007/s00285-003-0236-4

Dercole F (2016) The ecology of asexual pairwise interactions: the generalized law of mass action. Theor Ecol 9(3):299–321. https://doi.org/10.1007/s12080-015-0287-3

Dercole F, Geritz SAH (2016) Unfolding the resident-invader dynamics of similar strategies. J Theor Biol 394:231–254. https://doi.org/10.1016/j.jtbi.2015.11.032

Dercole F, Rinaldi S (2002) Evolution of cannibalistic traits: scenarios derived from adaptive dynamics. Theor Popul Biol 62(4):365–374. https://doi.org/10.1016/S0040-5809(02)00008-4

Dercole F, Rinaldi S (2008) Analysis of evolutionary processes: the adaptive dynamics approach and its applications. Princeton University Press, Princeton

Dercole F, Ferriere R, Rinaldi S (2002) Ecological bistability and evolutionary reversals under asymmetrical competition. Evolution 56(6):1081–1090. https://doi.org/10.1111/j.0014-3820.2002.tb01422.x

Dieckmann U, Law R (1996) The dynamical theory of coevolution: a derivation from stochastic ecological processes. J Math Biol 34(5):579–612. https://doi.org/10.1007/BF02409751

Diekmann O, Gyllenberg M, Huang H, Kirkilionis M, Metz JAJ, Thieme HR (2001) On the formulation and analysis of general deterministic structured population models ii. Nonlinear theory. J Math Biol 43(2):157–189. https://doi.org/10.1007/s002850170002

Doebeli M (1998) Invasion of rare mutants does not imply their evolutionary success: a counterexample from metapopulation theory. J Evol Biol 11(3):389–401. https://doi.org/10.1046/j.1420-9101.1998.11030389.x

d'Onofrio A (2013) Bounded noises in physics, biology, and engineering. Birkhäuser, Basel. https://doi.org/10.1007/978-1-4614-7385-5

Durinx M, (Hans) Metz JAJ, Meszéna G, (2008) Adaptive dynamics for physiologically structured population models. J Math Biol 56(5):673–742. https://doi.org/10.1007/s00285-007-0134-2

Ferriere R, Gatto M (1995) Lyapunov exponents and the mathematics of invasion in oscillatory or chaotic populations. Theor Popul Biol 48(2):126–171. https://doi.org/10.1006/tpbi.1995.1024

Ferris C, Best A (2019) The effect of temporal fluctuations on the evolution of host tolerance to parasitism. Theor Popul Biol. https://doi.org/10.1016/j.tpb.2019.07.015

Freidlin MI, Wentzell AD (2012) Random perturbations of dynamical systems, 3rd edn. Springer, Berlin. https://doi.org/10.1007/978-3-642-25847-3

Geritz SAH (2005) Resident-invader dynamics and the coexistence of similar strategies. J Math Biol 50(1):67–82. https://doi.org/10.1007/s00285-004-0280-8

Geritz SAH, Metz JAJ, Kisdi É, Meszéna G (1997) Dynamics of adaptation and evolutionary branching. Phys Rev Lett 78:2024–2027. https://doi.org/10.1103/PhysRevLett.78.2024

Geritz SAH, Meszéna G, Metz JAJ (1998) Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. Evol Ecol 12(1):35–57. https://doi.org/10.1023/A:1006554906681

Geritz SAH, van der Meijden E, Metz JAJ (1999) Evolutionary dynamics of seed size and seedling competitive ability. Theor Popul Biol 55(3):324–343. https://doi.org/10.1006/tpbi.1998.1409

Geritz SAH, Gyllenberg M, Jacobs FJA, Parvinen K (2002) Invasion dynamics and attractor inheritance. J Math Biol 44(6):548–560. https://doi.org/10.1007/s002850100136

Geritz SAH, Kisdi É, Yan P (2007) Evolutionary branching and long-term coexistence of cycling predators: Critical function analysis. Theor Popul Biol 71(4):424–435. https://doi.org/10.1016/j.tpb.2007.03.006

Khas'minskii RZ (1966) On stochastic processes defined by differential equations with a small parameter. Theory Probab Appl 11(2):211–228. https://doi.org/10.1137/1111018

Kifer Y (2001) Averaging and climate models. In: Imkeller P, von Storch JS (eds) Stochastic climate models. Birkhäuser, Basel, pp 171–188. https://doi.org/10.1007/978-3-0348-8287-3_7

Kisdi É (2015) Construction of multiple trade-offs to obtain arbitrary singularities of adaptive dynamics. J Math Biol 70(5):1093–1117. https://doi.org/10.1007/s00285-014-0788-5

Kisdi É, Geritz SAH (2016) Adaptive dynamics of saturated polymorphisms. J Math Biol 72(4):1039–1079. https://doi.org/10.1007/s00285-015-0948-2

Kisdi É, Liu S (2006) Evolution of handling time can destroy the coexistence of cycling predators. J Evol Biol 19(1):49–58. https://doi.org/10.1111/j.1420-9101.2005.00993.x

Kisdi É, Jacobs FJA, Geritz SAH (2001) Red Queen evolution by cycles of evolutionary branching and extinction. Selection 2(1–2):161–176. https://doi.org/10.1556/Select.2.2001.1-2.12

Kliemann W (1983) Qualitative theory of stochastic dynamical systems–applications to life sciences. Bull Math Biol 45(4):483–506. https://doi.org/10.1007/BF02459584

Kussell E, Leibler S (2005) Phenotypic diversity, population growth, and information in fluctuating environments. Science 309(5743):2075–2078. https://doi.org/10.1126/science.1114383

Lehtinen SO, Geritz SAH (2019) Cyclic prey evolution with cannibalistic predators. J Theor Biol 479:1–13. https://doi.org/10.1016/j.jtbi.2019.06.025

Li XY, Mao XR (2009) Population dynamical behavior of non-autonomous Lotka-Volterra competitive system with random perturbation. Discret Contin Dyn Syst 24:523–545. https://doi.org/10.3934/dcds.2009.24.523

Lion S (2018a) Class structure, demography, and selection: reproductive-value weighting in nonequilibrium, polymorphic populations. Am Nat 191(5):620–637. https://doi.org/10.1086/696976

Lion S (2018b) Theoretical approaches in evolutionary ecology: environmental feedback as a unifying perspective. Am Nat 191(1):21–44. https://doi.org/10.1086/694865

Liu SJ, Krstic M (2012) Stochastic averaging and stochastic extremum seeking. Springer, London. https://doi.org/10.1007/978-1-4471-4087-0

Matsuda H (1985) Evolutionarily stable strategies for predator switching. J Theor Biol 115(3):351–366. https://doi.org/10.1016/S0022-5193(85)80197-1

Melbinger A, Vergassola M (2015) The impact of environmental fluctuations on evolutionary fitness functions. Sci Rep 5:15211. https://doi.org/10.1038/srep15211

Meszéna G, Metz JAJ (1999) Species diversity and population regulation: the importance of environmental feedback dimensionality. IIASA Interim Report IR-99-045, International Institute for Applied Systems Analysis, Laxenburg, Austria. http://pure.iiasa.ac.at/id/eprint/5897/

Meszéna G, Gyllenberg M, Jacobs FJ, Metz JAJ (2005) Link between population dynamics and dynamics of Darwinian evolution. Phys Rev Lett 95:078105. https://doi.org/10.1103/PhysRevLett.95.078105

Metz JAJ, Nisbet RM, Geritz SAH (1992) How should we define 'fitness' for general ecological scenarios? Trends Ecol Evol 7(6):198–202. https://doi.org/10.1016/0169-5347(92)90073-K

Metz JAJ, Geritz SAH, Meszéna G, Jacobs FJA, Van H, Joost S (1996) Adaptive dynamics: a geometrical study of the consequences of nearly faithful reproduction. In: van Strien SJ, Venduyn Lunel SM (eds) Sotchastic and spatial structures of dynamical systems. North Holland, Amsterdam, pp 183–231

Metz JAJ, Mylius SD, Diekmann O (2008) When does evolution optimize? Evol Ecol Res 10:629–654

Miyahara Y (1972) Ultimate boundedness of the systems governed by stochastic differential equations. Nagoya Math J 47:111–144. https://doi.org/10.1017/S0027763000014951

Miyahara Y (1973) Invariant measures of ultimately bounded stochastic processes. Nagoya Math J 49:149–153. https://doi.org/10.1017/S0027763000015361

Mylius SD, Diekmann O (2001) The resident strikes back: invader-induced switching of resident attractor. J Theor Biol 211(4):297–311. https://doi.org/10.1006/jtbi.2001.2349

Nisbet RM, Gurney W (1982) Modellin fluctuating populations. Wiley, New York

Oba T, Kigami J (2018) Why does invasion imply substitution? Beyond the paradigm of invasion fitness. J Math Biol 77(5):1493–1532. https://doi.org/10.1007/s00285-018-1261-7

Parvinen K (1999) Evolution of migration in a metapopulation. Bull Math Biol 61(3):531–550. https://doi.org/10.1006/bulm.1999.0100

Priklopil T (2012) On invasion boundaries and the unprotected coexistence of two strategies. J Math Biol 64(7):1137–1156. https://doi.org/10.1007/s00285-011-0448-y

Priklopil T, Lehmann L (2019) Invasion implies substitution in ecological communities with class-structured populations. bioRxiv https://doi.org/10.1101/773580

Ripa J, Dieckmann U (2013) Mutant invasions and adaptive dynamics in vriable environments. Evolution 67(5):1279–1290. https://doi.org/10.1111/evo.12046

Roth G, Schreiber SJ (2014) Persistence in fluctuating environments for interacting structured populations. J Math Biol 69(5):1267–1317. https://doi.org/10.1007/s00285-013-0739-6

Sæther BE, Engen S (2015) The concept of fitness in fluctuating environments. Trends Ecol Evol 30(5):273–281. https://doi.org/10.1016/j.tree.2015.03.007

Schreiber SJ (2012a) Persistence for stochastic difference equations: a mini-review. J Differ Eq Appl 18(8):1381–1403. https://doi.org/10.1080/10236198.2011.628662

Schreiber SJ (2012b) The evolution of patch selection in stochastic environments. Am Nat 180(1):17–34. https://doi.org/10.1086/665655

Schreiber SJ, Benaïm M, Atchadé KAS (2011) Persistence in fluctuating environments. J Math Biol 62(5):655–683. https://doi.org/10.1007/s00285-010-0349-5

van Tienderen P, de Jong G (1986) Sex ratio under the haystack model: polymorphism may occur. J Theor Biol 122(1):69–81. https://doi.org/10.1016/S0022-5193(86)80225-9

Tuljapurkar S (1990) Population dynamics in variable environments. Springer, Berlin. https://doi.org/10.1007/978-3-642-51652-8

Vilar JMG, Rubi JM (2018) Determinants of population responses to environmental fluctuations. Sci Rep 8(1):887. https://doi.org/10.1038/s41598-017-18976-6

Wakano JY, Iwasa Y (2013) Evolutionary branching in a finite population: deterministic branching vs. stochastic branching. Genetics 193(1):229–241. https://doi.org/10.1534/genetics.112.144980