Demographic Stochasticity in Evolutionary Biology

by

Yen Ting Lin

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Physics)
in The University of Michigan
2013

Doctoral Committee:

Professor Charles R. Doering, Chair
Post-Doc Assistant Professor Arash Fahim
Assistant Professor Emanuel Gull
Professor Mark E. J. Newman
Professor John C. Schotland

5/1, 2013 at Navy Pier, Chicago.

To Chia-Hui and Amber

ACKNOWLEDGEMENTS

I met Prof. Charles Doering, who prefers to be called "Charlie", in my very first class—Complex System 541—here at the UM. Immediately I was attractive to his teaching charisma—until then I had not realized the meaning of the famous saying "teaching is an art". Charlie really masters the "art". After 541 I took his Asymptotic Analysis offered by the Department of Mathematics and learned a great deal about asymptotic analysis, which turned out to be one of the main tools for my later research.

If I can only thank one person in this Acknowledgement, Charlie is the one. Without him, I would probably never finish my Ph. D. study. He kindly accepted me transferring from Prof. Leonard Sander's group, and has been a wonderful mentor to me since then. He reshaped my "thinking" so that I can perform a more mathematical research, pushed me to take graduate courses in mathematics to broaden my view, encouraged me to attend to conferences (with financial support!) and supported me—financially and mentally—to work on subjects that I really enjoy. Charlie was the one who taught me everything about research—from performing theoretical analysis, comparing numerical data to theoretical conjectures, organizing literature reviews, drafting the dissertation, to refining the articles (he also carefully polished almost every sentences of my lame English writing). This dissertation would never exist without his guide. In every step I learned from him the rigor of scientific research, and he certainly set up a high-standard paradigm for me to follow in the future. Most important of all, as a well-respected Professor with a happy family, Charlie showed me

that taking care of my family never contradicts the excellence in scientific research.

Words cannot express my gratitude to Charlie. He has made this journey a wonderful experience in my life. I am very fortunate to have him as my mentor, and I sincerely hope to become the scientist he tried to shape.

I would also like to thank to Prof. Gull, who joined the Department of Physics last year. He taught me a lot about computational physics, which had never been organized so well in the curriculum of the Department before he joined. He also made substantial suggestions to this dissertation; with his advice the dissertation turned out to be much more complete. In addition, I would like to thank him for the wonderful referral to his previously-affiliated Max Planck Institute for the Physics of Complex Systems. With his and certainly with Charlie's recommendations, I am very blessed to have the opportunity to work at such a world-class research institute in the near future.

Starting from the beginning of this project, Dr. Hyejin Kim and I have been working closely together as a group. I would like to thank Hyejin for a wonderful experience of interdisciplinary research. Particularly, I enjoy our discussions, in which I gained a lot of knowledge in mathematics. Hyejin was always so patient to explain rigorous mathematics to me. Aside from the research, in the past few years our families have become very close friends, especially her 2-year-old son Minjoon and our 1-year-old daughter Amber. I certainly hope this friendship will last in the future.

I appreciate Dr. Arash Fahim for giving me many excellent suggestions in the final revision of the dissertation. I also enjoy very much his Math 526, in which he showed me a comprehensive picture about general stochastic processes.

Thanks very much to Prof. Mark Newman and Prof. John Schotland for serving as my committee members with a very short notice.

Many of my friends helped me in the process of earning the Ph. D. Here I thank my friends Paul Bierdz, Tim Saucer, and Ran Lu for helping me improve the writing

in my job applications and the dissertation. My groupmates in the Sweetland writing group, especially Ben Yee, also helped me to write a lot. I would also like to thank the Department of Physics for offering teaching positions several times to support me when the funding is short.

Finally, special thanks to my dear wife Chia-Hui for being with me all the time and raising our lovely daughter Amber almost independently. Both Chia-Hui and I thank all of our family members who emotionally support us in Taiwan.

At this time I would like to share the joy and the honor of earning this Ph. D. with all the people who ever helped—without any of them this would never be possible.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

LIST OF APPENDICES

Appendix

ABSTRACT

Demographic Stochasticity in Evolutionary Biology

by

Yen Ting Lin

Chair: C. R. Doering

Demographic stochasticity, the random fluctuations arising from the intrinsic discreteness of populations and the uncertainty of individual birth and death events, is an essential feature of population dynamics. Nevertheless theoretical investigations often neglect this naturally occurring noise due to the mathematical complexity of stochastic models. This dissertation reports the results of analytical and computational investigations of models of competitive population dynamics, specifically the competition between species in homogeneous or heterogeneous environments with different phenotypes of longevity or dispersal, fully accounting for demographic stochasticity. A novel asymptotic approximation is introduced and applied to derive remarkably simple analytical forms for key statistical quantities describing the populations' dynamical evolution. These formulas characterize the selection processes that determine which (if either) competitor has an evolutionary advantage. The theory is verified by conventional asymptotic analysis and large-scale numerical simulations.

After introducing demographic stochasticity into the deterministic models and motivating our mathematical approach to the analysis, we discover that the fluctuations can (1) break dynamical degeneracies, (2) support polymorphism that does not

exist in deterministic models, (3) reverse the direction of the weak selection and cause shifts in selection regimes, and (4) allow for the emergence of evolutionarily stable dispersal rates. Both dynamical mechanisms and time scales of the fluctuation-induced phenomena are identified within the theoretical approach. The analysis highlights the fundamental physical effect of the fluctuations and provides an intuitive interpretation of the complex dynamics. An interaction between stochasticity and nonlinearity is the foundation of noise-driven dynamical selection.

CHAPTER I

Introduction

1.1 General introduction

For centuries, theorists have been trying to develop mathematical models that describe the dynamics of populations. The effort began in the 18th century with T. R. Malthus' philosophical argument, "population, when unchecked, increases in a geometric ratio." Later in the 19th century P. F. Verhulst successfully captured more features of population growth including effects of intraspecies competition. In the early 20th century, A. J. Lotka and V. Volterra generalized the model to include interspecies interactions. Soon after Lotka–Volterra model was developed, a special model—the "competitive Lotka–Volterra model"—was proposed. Such model describes the dynamics of multiple species with both intraspecies and interspecies competition, and it soon became a modeling framework of competitive population dynamics.

The original models only considered population dynamics in well-mixed pools, i.e., they neglected spatial distributions of the populations. Needless to say, a more realistic demographic model of most ecological systems should account for the effect of spatial variations on interactions. There are two ways proposed to introduce spatially-dependent population dynamics. The first approach, patchy-like models, connect a number of "patches" each of which is a well-mixed pool. The population dynamics on

each patch are assumed to be described by a single well-mixed-pool model (Malthusian, Verhulst, Lotka–Volterra or Competitive Lotka-Volterra model) and exchange (or transport) terms are added to account for the changes of the local population due to dispersal from patch to patch. This approach results in sets of coupled ordinary differential equations that describe the population dynamics in a geometry composed of patches. The second approach, reaction–diffusion models (and in some more sophisticated cases with intelligent species, reaction–advection–diffusion models), consider continuous spatial domains. The population density is then described by a continuous function of both space and time. The local birth-and-death processes are characterized by "reaction" terms, and changes of populations due to dispersal are characterized by the "diffusion" terms and the "advection" terms if active transport effects are present as well. The evolution of the populations in reaction–diffusion models is therefore described by a set of partial differential equations.

The inclusion of the spatial dependence raises an interesting problem, often referred to as the "dispersal problem": is there a "best way" for a species to disperse in a given space, given the intrinsic birth and death dynamics everywhere within the system?

The inquiry to this problem originated from a novel idea proposed by W. D. Hamilton in 1967, the concept of evolutionary stability. The theory of evolutionary stability states that a currently existing species should have phenotypic traits which resist the invasion of any species with other phenotypic traits (assuming the selection is not neutral). The theory should apply on an evolutionary time scale, on which mutations must have occurred in successive reproductions, and the most fit phenotype(s) should be stronger competitors so that the dominant one(s) eventually exclude any other phenotypic traits.

At an abstract level, this dissertation will investigate the dispersal problem in the following framework. We place two "almost identical" species in a patchy-like

environment. The species have identical birth and death process, and compete for the same resource. Therefore the demographic dynamics in each location can be modeled by competitive Lotka—Volterra dynamics. The only difference between these two species is their propensities to move in the space: one of the species moves "faster" and the other moves "slower". We study the dynamics with particular emphasis on evolutionary time scales. In other words, we try to predict the populations of the species after many generations. Furthermore, we search for an optimal,"unbeatable" propensity to move such that the population of the species with such a mobility does not decrease when competing with any other varieties with different mobilities. We will refer such an unbeatable propensity as the evolutionarily stable propensity to move. If such evolutionarily stable propensity to move exists, we can also explore how it depends upon the structure and diversity of the environment.

As will be shown, the answer to the dispersal problem may be somewhat counter-intuitive: when the dispersal is passive, i.e., when the per capita dispersal rates are constants, both patchy-like models and reaction–diffusion models always select the slowest dispersers in heterogeneous environments. These robust observations suggest that the evolutionarily stable dispersal rate for passive dispersers is zero. Clearly these mathematical models fail to explain the existence of the species which utilize passive dispersal, for example, plants with airborne seeds.

On the other hand, agent-based models were developed to simulate the individual birth, death, and relocation events with the inclusion of demographic stochasticity. Surprisingly, in almost all such models, the inclusion of demographic stochasticity favors the fast dispersers. Owing to the complexity of the stochastic models, most of the studies relied on observations from numerical simulations and intuitive reasoning. The relation between demographic stochasticity and the selection of the faster dispersers remains unclear.

These observations pose interesting questions in theoretical dynamical systems

research: what is the mechanism by which faster dispersers are selected when demographic stochasticity is included? Can such mechanism provide with a driving force to establish a finite evolutionarily stable dispersal rate? How does the environment of the system affect such an evolutionarily stable dispersal rate?

The aim of this dissertation is to answer these questions analytically with the support of numerical simulations.

1.2 Literature review

Dispersion is a vital process in biology and ecology and has drawn the attention of many theorists and mathematicians. Various models have been proposed to explore the nature and effect of various dispersal processes. One of the essential questions is whether an evolutionarily stable dispersal rate (or strategy) exists for a species living in a heterogeneous environment. A dispersal rate (or strategy) is defined to be evolutionarily stable if species with such a dispersal rate (or strategy) can withstand invasion by species with other dispersal rates or (strategies), eventually excluding them in an isolated environment[1]. The existence of an evolutionarily stable dispersal rate (or strategy) is an important issue in evolutionary biology, since using the rationale "survival of the fittest", the species with the evolutionarily stable rate (or strategy) will prevail over the course of time.

Population dynamics in spatially heterogeneous habitats has been studied for a long time. Gadgil [13] first showed the complex nature of dispersal population dynamics on a set of globally connected patches. Hamilton and May [16] developed a discrete-generation and discrete-state model to demonstrate the existence of a nonzero evolutionarily stable dispersal rate. Comins *et al.* generalized Hamilton and May's model soon after it was proposed, and showed that the evolutionarily stable dispersal

[1]Hamilton first proposed the concept of a certain trait being "evolutionarily stable" to discuss the stability of natural sex ratios [15], and soon applied the same philosophy to a biological dispersal problem. [16]

rate converges to zero as the population scale goes to infinity [4]. Hastings proposed a continuous-time, continuous-state, and continuous-space model—a reaction–diffusion model with a passive diffusion mechanism—specifically to explore competitive dynamics between two species with different dispersal rates [19]. "Passive diffusion" refers to the case where the per capita migration rate of a species is a constant. By local stability analysis, Hastings proved that in any heterogeneous environment an established population distribution with a slower dispersal rate (the slow "residents") will drive an infinitesimal population of intruders with a faster dispersal rate (the fast "mutants") to eventual extinction. Hastings' conclusion hinted that the only evolutionarily stable dispersal rate should be zero, and the conclusion was coherent with the large population limit in Comins *et al.* [4]. Dockery *et al.* further analyzed Hastings' model and proved that the stay-at-home strategy is a globally stable fixation in pairwise competition with arbitrary initial populations [8].

In a separate line of thinking, Holt adopted Gadgil's patchy model [13] and discovered that connecting two separate patches increases the complexity of both competitive and predator-prey dynamics [20]. When competitive dispersers are subject to unconditional dispersal (i.e., passive diffusion), Holt showed the species with lower dispersal rate has a higher fitness based upon the idea "ideal free distribution" proposed by Fretwell and Lucas [12]. McPeek and Holt later found that the simplified two-patch dynamics could be adopted to explore population dynamics in various environmental settings. One of their numerical discoveries is, the trait with slowest dispersal always wins the competition with other different traits in heterogeneous environments, assuming all species are passive dispersers [21]. The evidence suggested the evolutionary stability of zero dispersal rate is robust, in the sense that such a trait always wins even when facing multiple types of intruders. Similar features were also reported by Cohen and Levin [3].

The idea of evolutionary stable dispersal *strategy* intrigued game theorists and

inspired several other investigations. We refer the interested reader to the review article by Cressman [6], which discusses a distinct approach to study the problem. Although game theoretic approaches seem remote from previous analyses, they all have a common hinge on the concept "ideal free distribution" [12]. For example, Cantrell *et al.* applied the concept to a model with a set of irreducibly connected patches, similar to Gadgil's model. He established theorems about the evolutionary stability of a certain dispersal strategies [2]. The theorems are coherent with the conclusion of McPeek's and Holt's numerical study on the 2-patch model [21]. Recent works continue to apply the idea of ideal free distributions in searching evolutionary stable *strategy* in dispersal problems.

Except for Hamilton and May's model, the above-mentioned models are mostly deterministic and ignore demographic stochasticity. It is well known that demographic stochasticity plays important roles in population dynamics. For example, Doering *et al.* identified spontaneous extinctions as large-deviations phenomena [9, 10]. Holt and McPeek first attempted to adopt Ricker model in the chaotic regime to mimic fluctuating populations on two connected patches. They reported that the chaotic fluctuation favors dispersal. Travis *et al.* generalized the McPeek–Holt model [29] to an individual-based description and also discovered that the inclusion of stochasticity favors the more frequent mover in competitive dynamics [37]. In the context of adaptive dynamical systems, Metz and Gyllenberg proposed a model that is conceptually identical to Hamilton and May's and defined a unique measure of the fitness, namely the "metapopulation reproduction rate" [30]. Following Metz and Gyllenberg's framework, both Cadet *et al.* [1] and Parvinen [33] numerically demonstrated that stochasticity enhances dispersal. More recently, Kessler and Sander [25] proposed an individual-based model and numerically discovered a regime shift between the dynamics of fast and slow dispersers. With a heuristic argument, they deduced a scaling law for the transition manifold in the parameter space. Waddell *et al.* [41]

showed that, on globally connected patches, in the limit of an infinite fast species and infinite slow species, it is possible to derive a set of closed moment equations that accurately describe the population dynamics. It should be remarked that the work of Waddell *et al.* took a dynamical approach, instead of the previously adopted local stability analysis [16, 4, 19, 30, 1, 33]; the latter approach generally does not describe global dynamics or provide conditions of coexistence or polymorphism.

The cumulative research suggests the generic features of competitive dynamics with passive dispersal: In the infinite population limit the competitive dynamics favors slow, or even no, dispersal. On the other hand, numerical studies suggest that demographic stochasticity, which is inevitable when the population is finite, favors the fast dispersers. The underlying mechanism remains unclear and merits further investigation. The purpose of this dissertation is to provide a novel dynamical approach to analytically study competitive population dynamics with demographic stochasticity. The goal of the dissertation is to study stochastic competitive population dynamics of two species that only differ in their propensity to move, in heterogeneous environments. For simplicity, we only consider passive dispersal, i.e., the case where the per capita dispersal rate is constant. Specifically, we are interested in the following questions:

1. Given the temporally fixed environment, does the dynamics tend to select one of the species?

2. When the answer of 1 is positive, which species—the fast moving or the slow moving one—has the competitive advantage? What is the mechanism responsible for the selection?

3. When the answer of 1 is negative, it suggests the dynamics is neutral and these two species can coexist. What are the conditions so that coexistence exists? Is coexistence stable or unstable? How are the populations of the

species distributed?

4. Is there an "optimal dispersal rate", or an evolutionarily stable dispersal rate, for a species to adopt? (We define the dispersal rate to be evolutionarily stable if a species with this dispersal rate always dominates species with other dispersal rates.)

5. If there is an optimal dispersal rate, how does it functionally depend on the parameters in the models?

By investigating the global dynamics, our analyses confirms the profound influences that demographic stochasticity has in the competitive dynamics, which in turn exhibits complicated transitions between different evolutionary outcomes. Our predictions and conclusions naturally extend the existing ones from local stability analyses.

1.3 Outline of the dissertation

Owing to the complexity of the problem, it is generally difficult to answer the questions above by directly applying conventional analyses. In order to develop novel techniques to resolve these difficulties, we performed multi-stage model reductions. The intuition behind the top-down model reductions are

1. Complexity potentially arises from the geometry of the domains. To focus our attention to the aspect of global stochastic dynamics, we developed two specific patchy models. The first model, which contains a domain of two patches, characterize a conceptually minimal model of the dispersal problems. The second model, which contains a countably infinite number of globally-connected patches, corresponds to the "mean-field" limit of the dispersal problem. With these specifications, we avoid complexity which is due to the geometry of the domain, for example, spatial pattern formation [5, 17, 18] or front propagation [26].

2. To further reduce the complexity of the model, we can consider the homogeneous spatial environments. That is, models where the patches are all identical. In this limit, we discover the deterministic limit of the models are degenerate. That is, the system has an infinite number of stable solutions.

3. To reduce the complexity of spatially-distributed models, we may further develop a seemingly unrelated model to the dispersal problem, competitive population dynamics of a fast and a slow living species on a single patch. This model is constructed with the intention of preserving the deterministically degenerate feature of the dispersal problems in homogeneous environments.

Having said that, the dissertation is organized in a bottom-up manner. A schematic diagram is shown in Fig. 1.1. In Chapter 2, we first discuss the most reduced model regarding to the competition of a fast and a slow living species. An essential and novel theoretical approach to resolve the analytical difficulty in the model is developed and verified by rigorous asymptotic analysis and direct numerical simulations. In Chapter 3, the developed technique is generalized and applied to the dispersal problems in homogeneous environments. We will see that when the population is large but finite, the stochastic models exhibit a selection for the fast-moving species on an identifiable time scale. In Chapter 4, we digress and perform regular asymptotic analysis to identify the time scale of the deterministic limits of the dispersal problems in heterogeneous environments. We conclude that in the deterministic limit, competitive dynamics favors the slow dispersers on another identifiable time scale. In Chapter 5, we investigate the stochastic dispersal problems in heterogeneous environments. As it turns out, the models combine the complexities in Chapter 3 and Chapter 4. We further generalized the developed technique to compute the boundaries of regime shifts between domination by faster or slower dispersing species. From Chapter 2 to Chapter 5, direct numerical simulations are presented to support the analyses. In Chapter 6, we discuss the bifurcation of the models in Chapter 5 and predict the

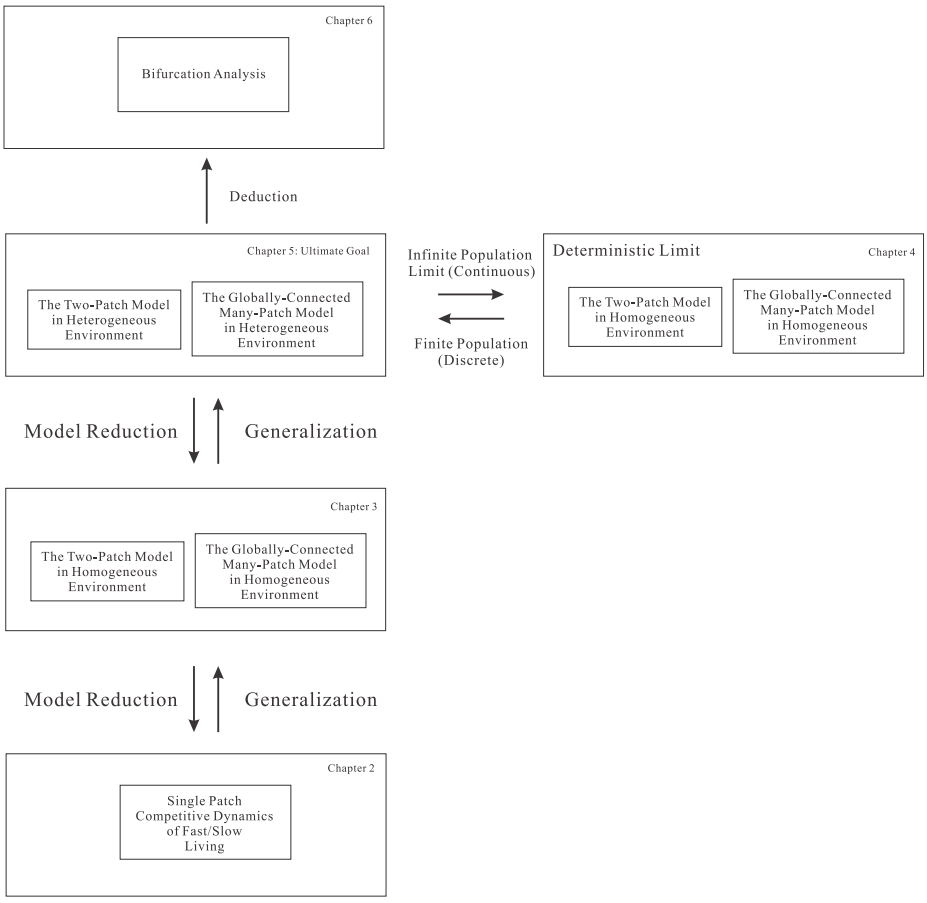

Figure 1.1: Outline of dissertation.

existence of an evolutionarily stable dispersal rate. Finally, in Chapter 7 we conclude the dissertation and propose future work on this class of the problems. The analyses of several remotely related problems are collected into the Appendices.

CHAPTER II

Features of Fast Living: On the Weak Selection for Longevity in Degenerate Birth-Death Processes

In this Chapter we develop and analyze a minimal model of degenerate competitive dynamics. The Chapter is organized as follows. In the next section 1 we describe the model in detail, discuss the degenerate deterministic dynamics, and define the non-degenerate stochastic evolution[1]. The following section 2 contains the physically-motivated analysis of the large carrying capacity (K) behavior and a derivation of the asymptotic forms of the drift and diffusion along the deterministically degenerate coexistence line. The subsequent section 3 contains the results of direct numerical simulations of the full birth-death process verifying our asymptotic theory. Section 4 performs a conventional asymptotic analysis on the problem to support the theory we develop in section 2. In the concluding section 5 we briefly conclude and discuss the results.

[1]We refer the "degeneracy" to the property that a system has an infinite number of stable solutions. As will be shown, in the deterministic model in section 2.1.1 is degenerate, and the individual-level asymmetry in the corresponding stochastic model brakes the degeneracy (section 2.1.2).

2.1 The models

2.1.1 Deterministic dynamics

Consider populations $X(t)$ and $Y(t)$ evolving according to

$$\dot{X} = \gamma_X X \left(1 - \frac{X+Y}{K}\right), \tag{2.1a}$$

$$\dot{Y} = \gamma_Y Y \left(1 - \frac{X+Y}{K}\right). \tag{2.1b}$$

This is a deterministic rate equation model of two species that compete equally for the available resources and differ only in the time scales of their evolution (birth and death rates) defined by the low-density growth rates γ_X and γ_Y. The two species' common carrying capacity K indicates the total number of individuals in the non-empty steady state. Such deterministic continuum descriptions are presumably applicable when X and Y are $\mathcal{O}(K)$ and $K \gg 1$.

It is convenient to introduce the scaled population variables $x = X/K$ and $y = Y/K$ and rescale the time variable by one growth rate (γ_Y) to write the system

$$\dot{x} = \gamma x(1 - x - y), \tag{2.2a}$$

$$\dot{y} = y(1 - x - y) \tag{2.2b}$$

where the ratio of time scales is $\gamma = \gamma_X/\gamma_Y$.

The dynamics of this system are elementary: the trajectories in the x-y phase plane solve the first-order equation

$$\gamma \frac{dy}{dx} = \frac{y}{x} \tag{2.3}$$

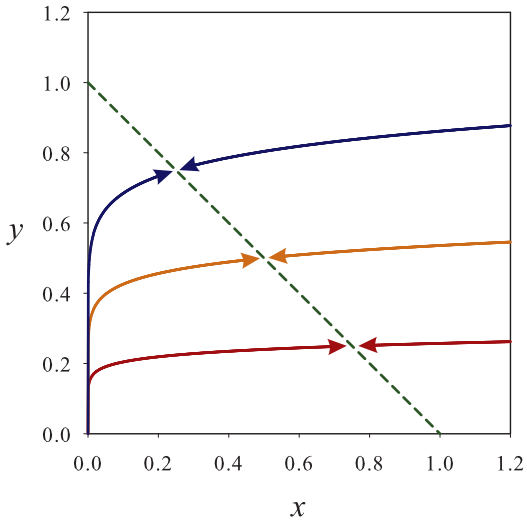

Figure 2.1: Deterministic trajectories and the (dashed) line of fixed points for Eq.(2.2) with $\gamma = 10$.

so at each instant of time during the evolution

$$\frac{x(t)}{x(t_0)} = \left(\frac{y(t)}{y(t_0)}\right)^{\gamma} \tag{2.4}$$

as illustrated in Fig. 2.1. Starting from any initial point in the first quadrant solutions converge within an $\mathcal{O}(\log K)$ time to a $1/K$-neighborhood of the coexistence line $x + y = 1$, each point of which is a marginally stable fixed point. Hence the eventual division of the population into fast and slow individuals is completely determined by the initial conditions, and once determined it remains fixed evermore. Demographic stochasticity introduces two more essential time scales into the system: the intermediate time until one or the other species goes extinct, which we will see is $\mathcal{O}(K)$, and the longer exponential-in-K time until the surviving species disappears [9, 10].

Independent process	Corresponding (per capita) rate
Birth of X	β_X
Birth of Y	β_Y
Death of X	$\delta_X[1+(n+m)/\tilde{K}]$
Death of Y	$\delta_Y[1+(n+m)/\tilde{K}]$

Table 2.1: The stochastic processes and the corresponding rates when the random populations are $X_t = n$ and $Y_t = m$.

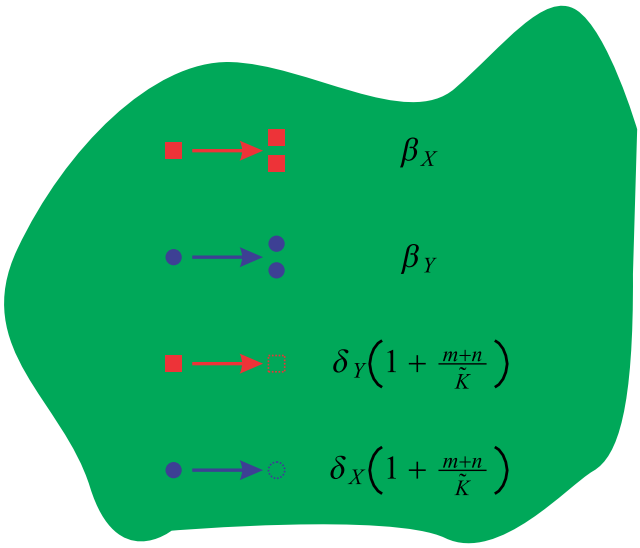

Figure 2.2: Schematic diagram of the stochastic processes and the corresponding per capita rate. The random populations are $X_t = n$ and $Y_t = m$ at this time.

2.1.2 Stochastic dynamics

The next more primitive level of description involves specifying stochastic evolution of the integer-valued random processes X_t and Y_t. We consider a Markov model, the competitive birth-and-death processes, as listed in Table 2.1. A schematic diagram is shown in Fig. 2.2. In addition, define $\tilde{K} = K/(\rho-1)$ with $\beta_X/\delta_X = \beta_Y/\delta_Y = \rho > 1$. This means that $p_{n,m}(t) = \mathbb{P}\{X_t = n \text{ and } Y_t = m\}$ evolves according to the master equation

$$
\begin{aligned}
(2.5) \qquad \frac{d}{dt}p_{n,m}(t) = &- \left(\beta_X + \delta_X[1 + (n+m)/\tilde{K}]\right) n\, p_{n,m} \\
&- \left(\beta_Y + \delta_Y[1 + (n+m)/\tilde{K}]\right) m\, p_{n,m} \\
&+ \beta_X(n-1)p_{n-1,m} + \delta_X[1 + (n+1+m)/\tilde{K}](n+1)p_{n+1,m} \\
&+ \beta_Y(m-1)p_{n,m-1} + \delta_Y[1 + (n+m+1)/\tilde{K}](m+1)p_{n,m+1}.
\end{aligned}
$$

The low-density growth rates appearing in the deterministic differential equations are $\gamma_X = \beta_X - \delta_X$ and $\gamma_Y = \beta_Y - \delta_Y$, and the ratio of evolution time scales is $\gamma = \delta_X/\delta_Y = \beta_X/\beta_Y$.

Now consider the carrying capacity to be large but finite, $K \gg 1$. For $\mathcal{O}(K)$ initial data $X(0)$ and $Y(0)$, fluctuations in the time-scaled "continuum" variables $x_t = K^{-1}X_{t/\gamma_Y}$ and $y_t = K^{-1}Y_{t/\gamma_Y}$ are relatively small and their evolution closely follows Eq.(2.2) for increasingly long times t as $K \to \infty$ [27]. For large but finite carrying capacities the discrete state space process' continuum variables are well-approximated by the Markov diffusion processes solving the Itô stochastic differential

equations

$$dx_t = \gamma\, x_t\,(1 - x_t - y_t)\,dt + \epsilon \sqrt{\gamma\, x_t \left(\frac{\rho+1}{\rho-1} + x_t + y_t\right)}\, dW_t^x \tag{2.6}$$

$$dy_t = y_t\,(1 - x_t - y_t)\,dt + \epsilon \sqrt{y_t \left(\frac{\rho+1}{\rho-1} + x_t + y_t\right)}\, dW_t^y \tag{2.7}$$

where W_t^x and W_t^y are independent Wiener processes and the "small" noise amplitude is $\epsilon = 1/\sqrt{K}$ [28].

When K is large and ϵ is small, a trajectory of the diffusion process (x_t, y_t) starting from (x_0, y_0) follows the deterministic dynamics[2] (2.2) to a neighborhood of the coexistence line in $\mathcal{O}(\log K)$ time and then performs $\mathcal{O}(1/\sqrt{K})$ fluctuations about the coexistence line. We are particularly interested in the reduced dynamics of the location of the joint process on the coexistence line:

$$z_t = x_t - y_t \in [-1, 1]. \tag{2.8}$$

Considering the initial value of z_t to be the position on the coexistence line where the deterministic dynamics lands starting from (x_0, y_0), the subsequent unavoidable and irreversible absorption of z_t at the right or left boundary (i.e., at ± 1) corresponds to extinction of one species or the other (i.e., y_t or x_t). Then there are two key questions to address:

1. What are the probabilities of absorption at ± 1 starting from $z \in (-1, 1)$?

2. What is the mean time to absorption at either ± 1 starting from $z \in (-1, 1)$?

(Note regarding notation: in the following the diffusion process will be distinguished by the time subscript, e.g., z_t, while undecorated quantities such as z indicate variables.)

[2]The effect due to the fluctuations can be ignored since the strength ϵ is small with respect to the strength of the deterministic flow.

We approximate the one-dimensional process z_t as a Markov diffusion process in its own right and seek drift $v(z)$ and diffusion $D(z)$ functions depending parametrically on $\epsilon = 1/\sqrt{K}$ so that its statistics are faithfully approximated by solutions of the Itô stochastic differential equation

$$dz_t = v(z_t)\,dt + \sqrt{2D(z_t)}\;dW_t \tag{2.9}$$

where W_t is a Wiener process.

The drift and diffusion for the reduced process z_t determine the desired statistical features of the competitive exclusion dynamics. Indeed, let $\tau(z) = \inf\{t : |z_t| = 1 \,|\, z_0 = z\}$ denote the random extinction time of one or the other species starting from position $z \in (-1, 1)$ on the coexistence line. Then the probability that x_t reaches 0 before y_t starting from z, i.e., the probability of domination of the Y-species over the X-species, is

$$u(z) \;\equiv \mathbb{P}\{z_{\tau(z)} = -1 \,|\, z_0 = z\}. \tag{2.10}$$

This probability satisfies the boundary value problem

$$0 = D(z)\frac{d^2u}{dz^2} + v(z)\frac{du}{dz}, \tag{2.11}$$

$$1 = u(-1), \tag{2.12}$$

$$0 = u(+1). \tag{2.13}$$

with explicit solution

$$u(z) \;=\; \frac{\int_z^1 e^{\Phi(z')}dz'}{\int_{-1}^1 e^{\Phi(z'')}dz''} \tag{2.14}$$

where

$$\Phi(z) = -\int_0^z \frac{v(\zeta)}{D(\zeta)} d\zeta. \tag{2.15}$$

The mean time to extinction of one or the other species,

$$m(z) \equiv \mathbb{E}\{\tau(z)\}, \tag{2.16}$$

satisfies the boundary value problem

$$-1 = D(z)\frac{d^2m}{dz^2} + v(z)\frac{dm}{dz}, \tag{2.17}$$

$$0 = m(\pm 1) \tag{2.18}$$

with solution

$$m(z) \;=\; G(z)H(1)/G(1) - H(z) \tag{2.19}$$

where

$$G(z) = \int_{-1}^{z} e^{\Phi(z')}\, dz' \tag{2.20}$$

$$H(z) = \int_{-1}^{z} e^{\Phi(z')} \left(\int_0^{z'} \frac{e^{-\Phi(z'')}}{D(z'')}\, dz'' \right) dz'. \tag{2.21}$$

The task now is to determine v and D.

2.2 Drift and diffusion along the coexistence line

2.2.1 A physically motivated asymptotic analysis

In the large K (small ϵ) regime, drift and diffusion along the coexistence line $x + y = 1$ are the result of interaction between demographic fluctuations and the nonlinear dynamics occuring near the coexistence line. In order to quantitatively estimate v and D, we will average the displacements and mean-square displacements of $z = x - y$ following birth and death events in the populations under the admittedly crude approximation that deviations are due to the stochastic terms in the stochastic differential equations (2.6) and (2.7) while the returns are dominated by the deterministic flow. As will be seen, this physically motivated approach produces the same result as a more traditional asymptotic singular perturbation analysis [31, 32].

Suppose at some instant of time the system is at position $z_0 = 2x_0 - 1$, i.e., at (x_0, y_0) with $y_0 = x_0 - 1$. We presume that in a small time interval dt the system is "kicked" by the demographic fluctuations to position $(x', y') = (x_0 + \phi a, y_0 + \eta b)$ where the independent random variables ϕ and η each take values ± 1 with probability $1/2$. To be consistent with the noise terms in the stochastic differential equations (2.6) and (2.7), the amplitudes should be

$$a \sim \epsilon \sqrt{\gamma x_0 \frac{2\rho}{\rho - 1} dt}, \tag{2.22}$$

$$b \sim \epsilon \sqrt{y_0 \frac{2\rho}{\rho - 1} dt}. \tag{2.23}$$

The system then quickly flows according to the deterministic dynamics along the lines $x^\gamma y = \text{constant}$ to position $(x_0 - \xi, y_0 + \xi)$ on the coexistence line. The net displacement along the coexistence line during this event is -2ξ so the drift and diffusion on the coexistence line are $v = \langle dz_t \rangle / dt = -2\langle \xi \rangle / dt$ and $2D = \langle dz_t^2 \rangle / dt = 4\langle \xi^2 \rangle / dt$ where $\langle \cdot \rangle$ indicates an average over the fluctuations. These events are illustrated in

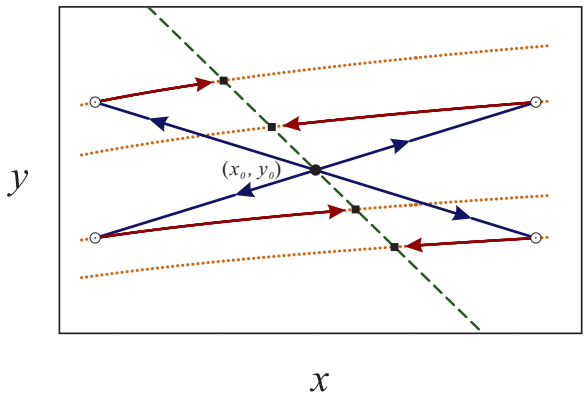

Figure 2.3: Mechanism for demographic fluctuation-induced drift and diffusion along deterministic coexistence line. From (x_0, y_0)—solid dot—the system fluctuates to (x', y')—indicated by the open circles—and subsequently relaxes back to corresponding $(x_0 - \xi, y_0 + \xi)$ point—indicated by the filled squares—on the coexistence line along a deterministic trajectory.

Fig. 2.3.

Make the ansatz that the elementary displacement along the coexistence line has an asymptotic expansion

$$\xi \sim \epsilon\,\xi_1 + \epsilon^2\,\xi_2 + \dots \text{ as } \epsilon \to 0. \tag{2.24}$$

Writing $(x_0 - \xi)(y_0 + \xi)^{-\gamma} = (x_0 + a\phi)(y_0 + b\eta)^{-\gamma}$ and expanding in powers of ϵ we quickly find

$$-\epsilon\xi_1\left(\frac{1}{x_0} + \frac{\gamma}{y_0}\right) = \frac{a\phi}{x_0} - \frac{\gamma b\eta}{y_0} \tag{2.25}$$

and

$$-\epsilon^2\xi_2\left(\frac{1}{x_0} + \frac{\gamma}{y_0}\right) + \epsilon^2\xi_1^2\left[\frac{\gamma(\gamma+1)}{2y_0^2} + \frac{\gamma}{x_0 y_0}\right] = -\frac{\gamma ab\phi\eta}{x_0 y_0} + \frac{\gamma(\gamma+1)b^2\eta^2}{2y_0^2}. \tag{2.26}$$

Then using $\langle\phi\rangle = 0 = \langle\eta\rangle$, $\langle\phi\eta\rangle = 0$, and $\langle\phi^2\rangle = 1 = \langle\eta^2\rangle$, the mean displacement $\langle\xi\rangle$

and mean-square displacement $\langle \xi^2 \rangle$ are determined to order ϵ^2:

$$\langle \xi \rangle \sim \epsilon \langle \xi_1 \rangle + \epsilon^2 \langle \xi_2 \rangle = \frac{-\frac{\gamma(\gamma+1)b^2}{2y_0^2} + \epsilon^2 \langle \xi_1^2 \rangle \left[\frac{\gamma(\gamma+1)}{2y_0^2} + \frac{\gamma}{x_0 y_0} \right]}{\frac{1}{x_0} + \frac{\gamma}{y_0}} \tag{2.27}$$

and

$$\langle \xi^2 \rangle \sim \epsilon^2 \langle \xi_1^2 \rangle = \frac{\frac{a^2}{x_0^2} + \frac{\gamma^2 b^2}{y_0^2}}{(\frac{1}{x_0} + \frac{\gamma}{y_0})^2}. \tag{2.28}$$

Now we write

$$a^2 = C\epsilon^2 \gamma x_0 \frac{2\rho}{\rho - 1} dt \tag{2.29}$$

$$b^2 = C\epsilon^2 y_0 \frac{2\rho}{\rho - 1} dt \tag{2.30}$$

where the $\mathcal{O}(1)$ proportionality constant $C > 0$ is a pure number that should not depends on ϵ, γ, or ρ.

It should be clear that, at least exactly on the line of the fixed points, the constant C is equal to 1: the processes x_t and y_t are purely diffusive in Eqs.(2.6) and (2.7), from which we can read off the *exact* strengths and the directions. It is our hypothesis that the constant C is asymptotically equal to 1 in the $\mathcal{O}(\epsilon)$-neighborhood of coexistence line. To support this hypothesis, we will perform an alternative analysis of a particular simplified version of the problem in section 2.2.1.

Recalling $x_0 = (1+z)/2$ and $y_0 = (1-z)/2$ and inserting (2.30) into (2.28) and then (2.27), we deduce that the drift and diffusion along the deterministic coexistence line are

$$v(z) = -2\frac{\langle \xi \rangle}{dt} \sim C\epsilon^2 \frac{2\rho\gamma(1-\gamma)}{\rho - 1} \times \frac{1 - z^2}{[(1-z) + \gamma(1+z)]^2} \tag{2.31}$$

and

$$D(z) \;=\; 2\frac{\langle \xi^2 \rangle}{dt} \;\sim\; C\epsilon^2 \frac{2\rho\gamma}{\rho - 1} \times \frac{1 - z^2}{(1 - z) + \gamma(1 + z)}. \tag{2.32}$$

The theory thus produces the conjecture that the drift and diffusion are both $\mathcal{O}(\epsilon^2) = \mathcal{O}(1/K)$ as $\epsilon = 1/\sqrt{K} \to 0$, and therefore that the time scale of the drift and diffusion along the coexistence line, i.e., the typical time it takes to select between the species is $\mathcal{O}(K)$ as $K \to \infty$. This is much longer than the $\mathcal{O}(\log K)$ time required for the deterministic flow to drive the system to the coexistence line but much less than the $\mathcal{O}(e^{cK})$ time to ultimate extinction.

The precise value of the mean time to extinction of one species or the other is inversely proportional to yet-to-be-determined number C but, interestingly, the probability of domination of one species over the other does not depend on it. Indeed, $u(z) = \text{Prob}\{z_{\tau(z)} = -1 | z_0 = z\}$ in Eq.(2.14) depends only on $\Phi(z) = -\int_0^z v(\zeta)/D(\zeta)\, d\zeta$ and according to the theory developed here,

$$\Phi(z) = \ln\left[1 + \left(\frac{\gamma - 1}{\gamma + 1}\right) z\right] \tag{2.33}$$

independent of C—and independent of ρ and K, too. Hence without further analysis we predict the probability of domination of the Y-species over the X-species in the large population limit with the simple expression

$$u(z) \;=\; \frac{1 - z}{2}\left[1 + \left(\frac{\gamma - 1}{\gamma + 1}\right)\frac{1 + z}{2}\right]. \tag{2.34}$$

This implies that even in the $K \to \infty$ "deterministic" limit, demographic fluctuations influence the selection of one of the species. Although it may not be immediately evident from Eq.(2.34), as will be seen in the next section the slower, longer-lived species is favored.

2.2.2 Truly degenerate case: $\gamma = 1$

In order to determine C we examine the truly degenerate situation $\gamma = 1$ where the two species are identical in every way except labeling. In this case the drift (2.31) vanishes so (2.32) implies that the mean time to extinction of one or the other species starting from $z_0 = z$, $m(z) = \mathbb{E}\{\tau(z)\}$ satisfies

$$C\epsilon^2 \frac{\rho}{\rho - 1}(1 - z^2)\frac{d^2 m}{dz^2} \;=\; -1 \tag{2.35}$$

with boundary conditions $m(\pm 1) = 0$. The strategy is to derive the differential equation for the leading approximation of the mean first passage time of the two-dimensional process (x_t, y_t) to the axes by conventional asymptotic methods and, by comparing it to (2.35), read off the value of C.

The stochastic differential equations (2.6) and (2.7) with $\gamma = 1$ imply that the mean time $T(x, y)$ for the two-dimensional process to hit either axis starting from $x_0 = x > 0$ and $y_0 = y > 0$ satisfies the boundary value problem

$$-1 = LT, \tag{2.36}$$

$$0 = T(x, 0), \tag{2.37}$$

$$0 = T(0, y) \tag{2.38}$$

where the backward Kolmogorov operator is $L \equiv \epsilon^2 L_0 + L_1$ with

$$L_0 \;=\; \frac{1}{2}\left(\frac{\rho + 1}{\rho - 1} + x + y\right)\left[x\frac{\partial^2}{\partial x^2} + y\frac{\partial^2}{\partial y^2}\right] \tag{2.39}$$

and

$$L_1 \;=\; (1 - x - y)\left[x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right]. \tag{2.40}$$

Inserting the asymptotic expansion

$$T(x,y) \sim \frac{1}{\epsilon^2} T_0 + T_1 + \epsilon^2 T_2 + \ldots, \tag{2.41}$$

where the leading term $\epsilon^{-2} T_0$ corresponds to m defined by (2.35), into (2.36) we find, order by order,

$$0 = L_1 \, T_0 \tag{2.42}$$

$$-1 = L_0 \, T_0 + L_1 \, T_1 \tag{2.43}$$

$$0 = L_0 \, T_n + L_1 \, T_{n+1} \text{ for } n \geq 1 \tag{2.44}$$

where each T_n satisfies homogeneous Dirichlet conditions on the x- and y-axes.

In order to make progress we transform to polar coordinates $r = \sqrt{x^2 + y^2}$ and $\theta = \arctan(y/x) \in (0, \pi/2)$. Then the advection operator simplifies to

$$L_1 \;=\; [1 - r(\cos\theta + \sin\theta)] \, r \frac{\partial}{\partial r} \tag{2.45}$$

and the leading equation (2.42) implies $0 = \partial_r T_0$ so T_0 is a function of θ alone. This is because when $\gamma = 1$ the deterministic trajectories are rays, radial lines from the initial condition to the stable coexistence line. Then the next leading order equation (2.43) in polar coordinates is

$$\begin{aligned} -1 = \frac{1}{2} \left[\frac{\rho + 1}{\rho - 1} + r(\cos\theta + \sin\theta) \right] \frac{\cos\theta \, \sin\theta}{r} \times \\ \left[2(\cos\theta - \sin\theta) \frac{d}{d\theta} + (\cos\theta + \sin\theta) \frac{d^2}{d\theta^2} \right] T_0 + \\ (1 - r(\cos\theta + \sin\theta)) \, r \, \frac{\partial T_1}{\partial r}. \end{aligned} \tag{2.46}$$

This equation is valid throughout the first quadrant of the phase plane so we may consider it restricted to the coexistence line $r(\cos\theta + \sin\theta) = 1$ where the T_1-term

disappears. This established that T_0 satisfies

$$-1 = \left(\frac{\rho}{\rho-1}\right)\cos\theta\,\sin\theta\,(\cos\theta+\sin\theta)\times \left[2(\cos\theta-\sin\theta)\frac{d}{d\theta}+(\cos\theta+\sin\theta)\frac{d^2}{d\theta^2}\right]T_0. \tag{2.47}$$

Change the independent variable from θ back to $z = x - y$, which is

$$z \;=\; \frac{\cos\theta-\sin\theta}{\cos\theta+\sin\theta} \tag{2.48}$$

when $r(\cos\theta+\sin\theta)=1$. Noting that

$$\frac{d}{d\theta} = -(1+z^2)\frac{d}{dz} \tag{2.49}$$

and

$$\cos\theta\sin\theta = \frac{1}{2}\left(\frac{1-z^2}{1+z^2}\right), \tag{2.50}$$

a little algebra reveals that T_0 satisfies

$$-1 \;=\; \left(\frac{\rho}{\rho-1}\right)(1-z^2)\,\frac{d^2T_0}{dz^2}. \tag{2.51}$$

Comparing this with equation (2.35) for $m = \epsilon^{-2}T_0$ we conclude that

$$C = 1. \tag{2.52}$$

2.2.3 Summary

In this section, we have developed a theory based on the physical insight of the dynamics. The theory implies that, asymptotically as the carrying capacity $K =$

$1/\epsilon^2 \to \infty$, demographic fluctuations induce the stochastic dynamics for $z_t = x_t - y_t$ of the form

$$dz_t \;=\; v(z_t)dt \;+\; \sqrt{2D(z_t)}\; dW_t \tag{2.53}$$

on the coexistence line $x + y = 1$ where the drift and diffusion are

$$v(z) \;=\; \frac{2\rho\,\gamma\,(1-\gamma)}{K(\rho-1)} \times \frac{1-z^2}{[(1-z)+\gamma(1+z)]^2} \tag{2.54}$$

and

$$D(z) \;=\; \frac{2\rho\,\gamma}{K(\rho-1)} \times \frac{1-z^2}{(1-z)+\gamma(1+z)}. \tag{2.55}$$

The remaining parameters in the theory are $\rho = \beta_X/\delta_X = \beta_Y/\delta_Y > 1$ and $0 < \gamma = \beta_X/\beta_Y = \delta_X/\delta_Y < \infty$, and that time in (2.53) is measured in units of $1/\gamma_Y = (\beta_Y - \delta_Y)^{-1}$.

2.3 Numerical simulations and asymptotic verification

We now turn to the numerical evaluation of the theoretical predictions. Continuous time Markov chain simulations [35] are carried out by performing exact simulations of the continuous time, discrete state space Markov process defined by the master equation (2.5). That is, the waiting time in a given state (n, m) is exponentially distributed with rate equalling the sum of the coefficients of $p_{n,m}$ in Eq.(2.5) followed by a transition to $(n \pm 1, m)$ or $(n, m \pm 1)$ with probability proportional to the corresponding coefficient of $p_{n\pm1,m}$ or $p_{n,m\pm1}$. The carrying capacity K and species' life-cycle/longevity ratio γ are varied while the low-density birth-to-death ratio is fixed at $\rho = 2$.

First we consider the probability of domination of the Y-species over the X-

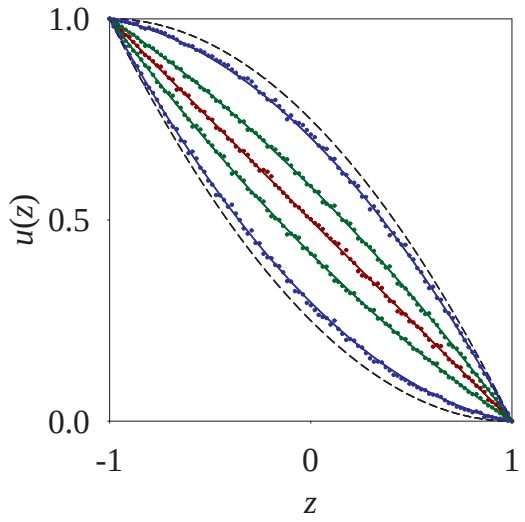

Figure 2.4: Probability of domination of the Y-species over the X-species starting from position z on the coexistence line, i.e., $u(z)$ from (2.56), for life-cycle ratios $\gamma = .1, .5, 1, 2, 10$ (solid lines bottom to top) as a function of starting position z on the coexistence line. The discrete data are from 10^4 independent simulations with $K = 1000$, and the dashed lines are the $\lim_{\gamma\to 0}$ and $\lim_{\gamma\to\infty}$ forms for $u(z)$.

species $u(z) = \mathbb{P}\{z_{\tau(z)} = -1 | z_0 = z\}$ in the large population limit given by (2.34) and reproduced here for reference:

$$u(z) = \frac{1-z}{2}\left[1 + \left(\frac{\gamma-1}{\gamma+1}\right)\frac{1+z}{2}\right]. \tag{2.56}$$

Figure 2.4 is a plot the theoretical predictions and the simulation results for $K = 1000$. In the totally degenerate case $\gamma = 1$, the probability that the Y-species outlives the X-species is simply proportional to the initial fraction of the Y-species in the total population ($u = (1-z)/2 = y$ when $\gamma = 1$) and the simulations in this situation simply serve to indicate the level of statistical noise associated with 10^4 samples at carrying capacity $K = 1000$. It is evident that the asymptotic theory is in excellent agreement with the data.

Figure 2.4 also illustrates how demographic fluctuations break the degeneracy in the deterministic dynamics, endowing the longer-lived species with a competitive

advantage. When $\gamma > 1$ the X-species reproduces and dies faster, and at every value of $z \in (-1, 1)$, the probability of the slower Y-species dominating is increased over that in the $\gamma = 1$ case. Likewise, when $\gamma < 1$ the X-species reproduces and dies slower than the Y-species, and at every value of the initial populations the probability of the Y-species dominating is decreased over that in the $\gamma = 1$ case. In both cases the slower-to-reproduce/longer-to-live species has an enhanced probability of winning the competition battle.

It is also interesting to notice that the leading approximation of the competitive advantage is limited as the longevity ratio varies between the extremes $\gamma = 0$ and $\gamma = \infty$. In a head-to-head competition starting from a large population comprised of exactly 50% X-species and 50% Y-species, the probability of Y outliving X is $u(0) = \frac{3\gamma+1}{4\gamma+4}$ so the probability of Y winning is never less than 25% or more than 75%. We stress that this conclusion applies in the $K \to \infty$ limit preceding the $\gamma \to 0$ or $\gamma \to \infty$ limit; simulations (not presented here) show clearly that the competitive advantage enjoyed by the longer-lived species may be significantly greater at finite K.

The leading large-K approximation for the mean time to extinction of one species or the other starting from z on the coexistence line follows from equations (2.19) and (2.21) given (2.54) and (2.55). The result is

$$
\begin{aligned}
(2.57) \qquad m(z) \; = \frac{K(\rho-1)(\gamma-1)}{2\rho\gamma} \Bigg[& 2\frac{\gamma+1}{\gamma-1}\log 2 - \frac{\gamma-1}{\gamma+1}\frac{(1-z^2)}{2} \; + \\
& + \; (1+z)\log(1+z)\left(\frac{1-z}{2} - \frac{\gamma+1}{\gamma-1}\right) \\
& - \; (1-z)\log(1-z)\left(\frac{1+z}{2} + \frac{\gamma+1}{\gamma-1}\right)\Bigg].
\end{aligned}
$$

Figure 2.5 shows data for the mean first passage time to $X_t = 0$ or $Y_t = 0$ from simulations of the discrete two-dimensional (X_t, Y_t) process along with the predictions of the asymptotic theory from equation (2.57). The parameters used in Figure 2.5

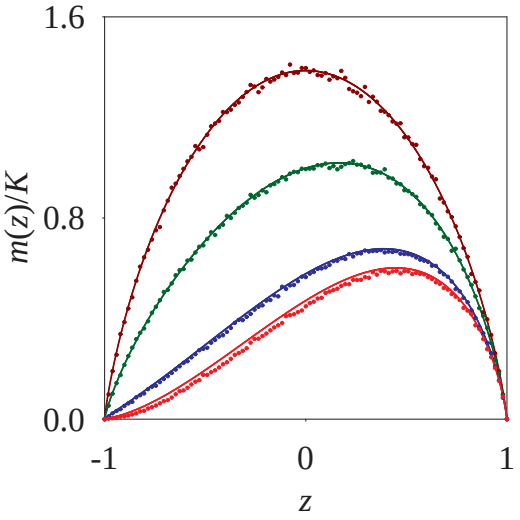

Figure 2.5: Mean extinction time $m(z) = \mathbb{E}\{\tau(z)\}$ of one species or the other starting from position z on the coexistence curve. The solid lines are the theoretical predictions of (2.57) and the discrete data are from 10^4 independent simulations at $K = 1000$. The other parameters are $\rho = 2$ and $\gamma = 1,\ 2,\ 10,\ 50$ (top to bottom). Note that the $\mathcal{O}(1)$ vertical axis is $m(z)/K$, i.e., the mean time in units of the total carrying capacity K.

are $K = 1000$, $\rho = 2$, and $\gamma = 1$ (top curve), 2, 10, and 50 (bottom), and the exit times for 10^4 independent simulations were averaged at each value of the parameters and for each initial starting position z on the coexistence line. The agreement, especially for $\gamma = 1$ and 2, is excellent. The theory systematically overestimates the simulation results at higher values of the longevity ratio, a sign that the approach to the asymptotic limit as $K \to \infty$ is not uniform in γ.

2.4 Asymptotic analysis for arbitrary γ

In this section we show how to extract the leading order equation for the mean exit time from the first quadrant in the general case $\gamma \neq 1$. The boundary value

problem is

$$-1 = LT, \tag{2.58}$$

$$0 = T(0, y), \tag{2.59}$$

$$0 = T(x, 0) \tag{2.60}$$

where the backward Kolmogorov operator $L \equiv \epsilon^2 L_0 + L_1$ contains with

$$L_0 = \frac{1}{2}\left(\frac{\rho+1}{\rho-1} + x + y\right)\left[\gamma x \partial_x^2 + y\partial_y^2\right] \tag{2.61}$$

and

$$L_1 = (1 - x - y)\left(\gamma x \partial_x + y\partial_y\right). \tag{2.62}$$

Inserting the asymptotic ansatz $T \sim \epsilon^{-2}T_0 + T_1 + \epsilon^2 T_2 + \ldots$ implies, order by order, that the T_n satisfy

$$0 = L_1 T_0, \tag{2.63}$$

$$-1 = L_0 T_0 + L_1 T_1, \tag{2.64}$$

$$0 = L_0 T_n + L_1 T_{n+1} \text{ for } n \geq 1. \tag{2.65}$$

Transform the problem into coordinates (Ψ, Φ) defined by

$$\Psi \equiv x + y, \tag{2.66}$$

$$\Phi \equiv \frac{x}{y^\gamma}. \tag{2.67}$$

This transformation is a one-to-one mapping of the open first quadrant into itself, and the degenerate coexistence manifold is the set $\{\Psi = 1, 0 \leq \Phi \leq \infty\}$. The important

point is that the coordinate Φ is constant on deterministic trajectories defined by the drift. In these new coordinates the operators are

$$L_0 = \frac{1}{2}\left(\frac{\rho+1}{\rho-1}+\Psi\right)\left[\gamma\frac{x}{y^{2\gamma}}\partial_\Phi^2 + 2\gamma\frac{x}{y^\gamma}\partial_\Psi\partial_\Phi + \gamma x\partial_\Psi^2 + \gamma\left(\gamma+1\right)\frac{x}{y^{\gamma+1}}\partial_\Phi + \gamma^2\frac{x^2}{y^{2\gamma+1}}\partial_\Phi^2 - 2\gamma\frac{x}{y^\gamma}\partial_\Phi\partial_\Psi + \partial_\Psi^2\right], \tag{2.68}$$

$$L_1 = \left(1-\Psi\right)\left(\gamma x + y\right)\partial_\Psi \tag{2.69}$$

where we refer to the inverse mapping from (Ψ,Φ) back to (x,y) by

$$x \equiv x(\Psi,\Phi), \tag{2.70}$$

$$y \equiv y(\Psi,\Phi) \tag{2.71}$$

remarking that, at least for the leading asymptotic approximation of the mean first passage time, the explicit form of these functions are irrelevant.

The leading equation $0 = L_1T_0$ implies $0 = \partial_\Psi T_0$ so that T_0 is a function of Φ only. Then the next (inhomogeneous) equation $-1 = L_0T_0 + L_1T_1$ is

$$-1 = \frac{1}{2}\left(\frac{\rho+1}{\rho-1}+\Psi\right)\left[\gamma\left(\gamma+1\right)\frac{x}{y^{\gamma+1}}\partial_\Phi + \frac{\gamma x}{y^{2\gamma}}\left(1+\gamma\frac{x}{y}\right)\partial_\Phi^2\right]T_0(\Phi) + L_1T_1. \tag{2.72}$$

This is valid throughout the first quadrant of the (Ψ,Φ) plane, and in particular on the coexistence line $\Psi = 1$ where L_1 vanishes implying

$$-1 = \frac{\rho}{\rho-1}\left[\gamma\left(\gamma+1\right)\frac{x}{y^{\gamma+1}}\partial_\Phi + \frac{\gamma x}{y^{2\gamma}}\left(1+\gamma\frac{x}{y}\right)\partial_\Phi^2\right]T_0(\Phi). \tag{2.73}$$

Now transform to the variable $z = x - y$, noting that

$$\Phi = \frac{x}{y^\gamma} = 2^{\gamma-1}\frac{1+z}{(1-z)^\gamma} \tag{2.74}$$

is a monotone function of z on the coexistence line, and finally we deduced

$$-1 = \frac{2\rho\gamma}{\rho-1}\left[\frac{1-z^2}{(1-z)+\gamma(1+z)}\partial_z^2 + (1-\gamma)\frac{1-z^2}{\left[(1-z)+\gamma(1+z)\right]^2}\partial_z\right]T_0(z) \tag{2.75}$$

from which the drift (2.54) and diffusion (2.55) on the coexistence line may be read off.

2.5 Summary and discussion

Demographic fluctuations break the degeneracy displayed by the deterministic rate equation description of the dynamics of two competing species differing only in the time scales of their life cycles. The theoretical analysis presented here, along with its confirmation via direct numerical simulations, shows that the longer-lived-slower-to-reproduce species enjoys a slight competitive advantage over the shorter-lived-but-faster-reproducing species. This in itself may not be surprising given the asymmetry of the stochastic dynamics when birth-death noise is incorporated into the model, but what is remarkable is that the effect persists all the way to the continuum limit. That is, the competitive disadvantage of fast living remains an $\mathcal{O}(1)$ effect in the infinite carrying capacity $K \to \infty$ limit where the deterministic dynamics is in fact valid, albeit over sufficiently bounded time intervals. The simple resolution of this apparent dilemma is that the time for the distinction between the species to be realized, i.e., the time required for the demographic fluctuations to substantially affect the population balance, diverges $\sim K$ as $K \to \infty$.

The singular behavior of the species selection is illustrated in Figure 2.6 where

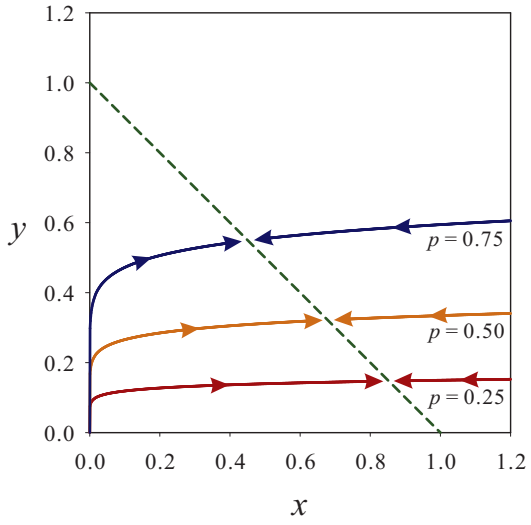

Figure 2.6: Isoprobability curves for survival of the slow species (in this case the Y-species) for $\gamma = 10$ in the limit $K \to \infty$. Compare with Fig. 2.3.

we plot isoprobability curves of dominance of the longer-lived species during times $\mathcal{O}(K) \leq t \leq \mathcal{O}(e^{cK})$ in the deterministic $K \to \infty$ limit. The degenerate dynamics described by differential equations faithfully carries the two-dimensional continuum system from its initial position in the phase plane to the coexistence line in an $\mathcal{O}(\log K)$ where the stochastic dynamics takes charge and determines the victor of the competition who subsequently survives until the ultimate extinction. If time is measured in units proportional to K, the continuum limit is deterministic only during a vanishingly small transient after which it is a Markov diffusion process restricted to the coexistence line, eventually being absorbed and remaining ever thereafter at $z = \pm 1$.

It is our original motivation to construct the model as a minimal model (stochastic) degenerate competitive dynamics, as illustrated in section 1.3. Interestingly, this sort of degenerate dynamics has been considered by other researchers. Over two decades ago Katzenberger [24] studied the behavior of solutions of stochastic differential equations with strong drift driving the system onto a submanifold fixed points of the deterministic dynamics in a formal setting. More recently Parsons and Quince

[31] and Parsons, Quince, and Plotkin [32] considered the system analyzed here, using conventional asymptotic methods similar to those employed in Section 2.3 to fix the constant C (and more generally in section 2.4), to evaluate both the probability and mean time to "fixation". Most recently Durrett and Popovic [11] studied the stochastic dynamics of a different degenerate model where the coexistence curve is not a simple straight line segment while the deterministic trajectories are.

The theoretical method introduced here consists of using the elementary fluctuations away from the degenerate manifold and the subsequent deterministic relaxation back to evaluate the effective drift and diffusion in a reduced description. This constitutes a novel approach to the quantitative analysis with two advantages. First is that it intuitively incorporates the physical processes that produce the drift and diffusion on the coexistence curve. Such insight contributes substantially to our understanding of these dynamics and their quantitative description. Second is that it produces accurate asymptotic predictions relatively quickly. Indeed, as developed in detail in section 2.4, conventional perturbation theory analysis of the $\gamma \neq 1$ situation is significantly more involved than the relatively straightforward $\gamma = 1$ calculations presented in Section 2.3. This is because in order to implement the projection onto the coexistence curve, the two-dimensional dynamics must be formulated in coordinates incorporating the deterministic trajectories.

CHAPTER III

Demographic Stochasticity and Evolution of Dispersal in Homogeneous Environments

In this Chapter, we generalized the established physically motivated asymptotic analysis in Chapter II to investigate competitive population dynamics in homogeneous environments with demographic fluctuations. We investigate two specific models: the two-patch model, presented in section 3.1, and the many-patch model, presented in section 3.2. In section 3.3 we discuss and summarize the analyses in this Chapter.

3.1 The two-patch model

3.1.1 The model

We begin with a simple model of competing species in a spatially extended yet homogeneous environment as illustrated in schematic Figure 3.1. The model consists of two identical patches, each of which is a well-mixed pool with carrying capacity K. Two species X and Y reside in the patches and compete locally for limited resources. In this continuous-time Markov model, each individual randomly reproduces, dies, or moves to the other patch at precisely defined rates. To explore the relation between survival probability and mobility under identical environment conditions it is assumed that each individual of both species has the same birth rate and, in the same

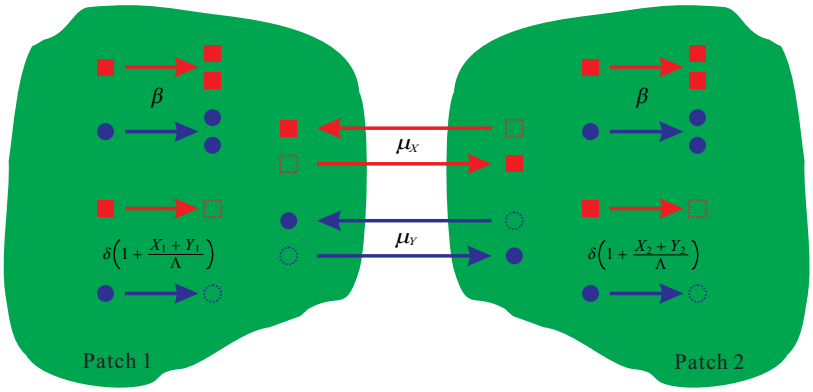

Figure 3.1: Dynamics of the interacting species distributed on two patches. The X and Y populations compete locally, and individuals randomly move from one patch to the other at rates μ_X and μ_Y.

Independent process	Corresponding (per capita) rate
Birth of X on patch i	β
Birth of Y on patch i	β
Death of X on patch i	$\delta[1+(X_i+Y_i)/\Lambda]$
Death of Y on patch i	$\delta[1+(X_i+Y_i)/\Lambda]$
Dispersal of X from patch i to patch j, $i \neq j$	μ_X
Dispersal of Y from patch i to patch j, $i \neq j$	μ_Y

Table 3.1: The stochastic processes and the corresponding rates.

competitive environment, the same death rate. However, the hopping rates of the two species are not restricted to be identical and they will be treated as two independent parameters hereafter. From here on we refer this model as the (homogeneous) 2–patch model.

More precisely, let $(X_1(s), X_2(s), Y_1(s), Y_2(s))$ be the non-negative integer-valued populations of the X and Y species on patch 1 or 2, respectively, at (dimensional) time s. Let i be the patch index, i.e., $i \in \{1, 2\}$. Per capita birth and death rates of both species in patch i are, respectively, β and $\delta\,[1 + (X_i(s) + Y_i(s))/\Lambda]$. Here Λ is a population scale, which will be shown to be proportional to the carrying capacity of the patch. We will always consider parameter values where the low-density birth-to-

death rate ratio $\rho := \beta/\delta > 1$. The term $X_i + Y_i/\Lambda$ in the death rates characterize the feature of competition to limited resources $K := (\rho - 1)\,\Lambda$. The hopping (dispersal) rates of species X and Y are μ_X and μ_Y respectively. Although we refer these rates as "faster" or "slower" they do not represent movement speed, but rather the propensity for individuals of either species to migrate to a new location. We consider both $\mu_X > 0$ and $\mu_Y > 0$ although empirically we shall see that $\max\{\mu_X, \mu_Y\} > 0$ is sufficient to draw the conclusions. Table 3.1 summarizes the random processes.

Let the probability of the state (a, b, c, d) at time s be

$$p_{a,b,c,d}(s) = \mathbb{P}\left[\{X_1(s) = a\} \cap \{Y_1(s) = b\} \cap \{X_2(s) = c\} \cap \{Y_2(s) = d\}\right], \tag{3.1}$$

and its evolution be given by the master equation

$$\begin{aligned}
\frac{d}{ds}p_{a,b,c,d} = & -\left(\beta + \delta\left[1 + (a+b)/\Lambda\right] + \mu_X\right) a p_{a,b,c,d} \\
& -\left(\beta + \delta\left[1 + (a+b)/\Lambda\right] + \mu_Y\right) b p_{a,b,c,d} \\
& -\left(\beta + \delta\left[1 + (c+d)/\Lambda\right] + \mu_X\right) c p_{a,b,c,d} \\
& -\left(\beta + \delta\left[1 + (c+d)/\Lambda\right] + \mu_Y\right) d p_{a,b,c,d} \\
& + \beta\,(a-1)\,p_{a-1,b,c,d} + \delta\left[1 + (a+b+1)/\Lambda\right](a+1)\,p_{a+1,b,c,d} \\
& + \beta\,(b-1)\,p_{a,b-1,c,d} + \delta\left[1 + (a+b+1)/\Lambda\right](b+1)\,p_{a,b+1,c,d} \\
& + \beta\,(c-1)\,p_{a,b,c-1,d} + \delta\left[1 + (c+d+1)/\Lambda\right](c+1)\,p_{a,b,c+1,d} \\
& + \beta\,(d-1)\,p_{a,b,c,d-1} + \delta\left[1 + (c+d+1)/\Lambda\right](d+1)\,p_{a,b,c,d+1} \\
& + \mu_X\,(a+1)\,p_{a+1,b,c-1,d} + \mu_X\,(c+1)\,p_{a-1,b,c+1,d} \\
& + \mu_Y\,(b+1)\,p_{a,b+1,c,d-1} + \mu_Y\,(d+1)\,p_{a,b-1,c,d+1}.
\end{aligned} \tag{3.2}$$

Denote the carrying capacity $K := (\rho - 1)\Lambda$ with $\rho := \beta/\delta > 1$. As Λ and thus $K \to \infty$, fluctuations in the time-scaled continuum variables $x_i(t) = X_i\left(t/(\beta - \delta)\right)/K$ and $y_i(t) = Y_i\left(t/(\beta - \delta)\right)/K$ are relatively small. For large but finite K the dynamics are

well-described by a diffusion process with probability density $f(x_1, x_2, y_1, y_2, t)$ governed by the Kolmogorov forward (a.k.a. Fokker–Planck) equation [28, 14] obtained by Taylor expanding (3.2) in powers of $1/K$,

$$
\begin{aligned}
\frac{\partial f}{\partial t} = & -\frac{\partial}{\partial x_1}\left\{\left[x_1\left(1-x_1-y_1\right)+\mu_x\left(x_2-x_1\right)\right] f\right\} \\
& -\frac{\partial}{\partial x_2}\left\{\left[x_2\left(1-x_2-y_2\right)+\mu_x\left(x_1-x_2\right)\right] f\right\} \\
& -\frac{\partial}{\partial y_1}\left\{\left[y_1\left(1-x_1-y_1\right)+\mu_y\left(y_2-y_1\right)\right] f\right\} \\
& -\frac{\partial}{\partial y_2}\left\{\left[y_2\left(1-x_2-y_2\right)+\mu_y\left(y_1-y_2\right)\right] f\right\} \\
& +\frac{1}{2K}\frac{\partial^2}{\partial x_1^2}\left\{\left[x_1\left(\frac{\rho+1}{\rho-1}+x_1+y_1\right)+\mu_x\left(x_1+x_2\right)\right] f\right\} \\
& +\frac{1}{2K}\frac{\partial^2}{\partial x_2^2}\left\{\left[x_2\left(\frac{\rho+1}{\rho-1}+x_2+y_2\right)+\mu_x\left(x_1+x_2\right)\right] f\right\} \\
& +\frac{1}{2K}\frac{\partial^2}{\partial y_1^2}\left\{\left[y_1\left(\frac{\rho+1}{\rho-1}+x_1+y_1\right)+\mu_y\left(y_1+y_2\right)\right] f\right\} \\
& +\frac{1}{2K}\frac{\partial^2}{\partial y_2^2}\left\{\left[y_2\left(\frac{\rho+1}{\rho-1}+x_2+y_2\right)+\mu_y\left(y_1+y_2\right)\right] f\right\} \\
& -\frac{1}{K}\frac{\partial^2}{\partial x_1\,\partial x_2}\left[\mu_x\left(x_1+x_2\right) f\right]-\frac{1}{K}\frac{\partial^2}{\partial y_1\,\partial y_2}\left[\mu_x\left(y_1+y_2\right) f\right]
\end{aligned}
\tag{3.3}
$$

where the dimensionless time $t = (\beta - \delta)s$ with (now) dimensionless hopping rates μ_x and μ_y suitably scaled by $(\beta - \delta)$.

(Note regarding notations: throughout rest of the dissertation the parameters with capital decoration, for example μ_X, stand for the *unscaled* parameters, and the parameters with lower-case decoration, for example μ_x, stand for the *scaled* ones. In addition, the variable s is always the dimensional time, and the variable t is reserved to be the scaled and dimensionless time.)

This is the evolution equation for the probability transition density of the solution

of the coupled Itô stochastic differential equations

$$dx_1 = [x_1(1-x_1-y_1)+\mu_x(x_2-x_1)]\,dt + \epsilon\sqrt{x_1\left(\frac{\rho+1}{\rho-1}+x_1+y_1\right)}\,dW_1 + \epsilon\sqrt{\mu_x(x_1+x_2)}\,dW_2, \tag{3.4a}$$

$$dx_2 = [x_2(1-x_2-y_2)+\mu_x(x_1-x_2)]\,dt + \epsilon\sqrt{x_2\left(\frac{\rho+1}{\rho-1}+x_2+y_2\right)}\,dW_3 - \epsilon\sqrt{\mu_x(x_1+x_2)}\,dW_2, \tag{3.4b}$$

$$dy_1 = [y_1(1-x_1-y_1)+\mu_y(y_2-y_1)]\,dt + \epsilon\sqrt{y_1\left(\frac{\rho+1}{\rho-1}+x_1+y_1\right)}\,dW_4 + \epsilon\sqrt{\mu_y(y_1+y_2)}\,dW_5 \tag{3.4c}$$

$$dy_2 = [y_2(1-x_2-y_2)+\mu_y(y_1-y_2)]\,dt + \epsilon\sqrt{y_2\left(\frac{\rho+1}{\rho-1}+x_2+y_2\right)}\,dW_6 - \epsilon\sqrt{\mu_y(y_1+y_2)}\,dW_5, \tag{3.4d}$$

where $\epsilon \equiv 1/\sqrt{K}$ and the $W_i(t)$ are independent Wiener processes. (Pardon the conventional abuse of notation here: in (3.3) the x_i and y_i are independent variables while in (3.4) they random processes. Context inevitably resolves any possible confusion.) In the infinite carrying capacity limit the continuum variables $x_i(t) = X_i(t/(\beta-\delta))/K$ and $y_i = Y_i(t/(\beta-\delta))/K$ evolve according to the classical deterministic rate (ordinary differential) equations [27]

$$\dot{x_1} = x_1(1-x_1-y_1)+\mu_x(x_2-x_1), \tag{3.5a}$$

$$\dot{x_2} = x_2(1-x_2-y_2)+\mu_x(x_1-x_2), \tag{3.5b}$$

$$\dot{y_1} = y_1(1-x_1-y_1)+\mu_y(y_2-y_1), \tag{3.5c}$$

$$\dot{y_2} = y_2(1-x_2-y_2)+\mu_y(y_1-y_2). \tag{3.5d}$$

As shown in the following theorem, solutions of Eqs.(3.5) are in equilibrium if and

only if the state is on the line (x_0, x_0, y_0, y_0) with $x_0 \in [0,1]$ and $x_0 + y_0 = 1$ in the four-dimensional phase space. On the interior of this line segment, i.e., for $x_0 \in (0,1)$, the two species coexist in the deterministic limit. We will refer the open segment as the coexistence line. In the following subsection we deduce that for large but finite K the 2–patch model exhibits a weak selection for the fast disperser on an $\mathcal{O}(K)$ time scale as a result of the fluctuations from individual-level processes.

Theorem III.1. *[Fixed points of the deterministic dynamics of the 2-patch model.] The state (x_1, x_2, y_1, y_2) is a coexistence fixed point of the deterministic 2-patch model if and only if $(x_1, x_2, y_1, y_2) = (x_0, x_0, y_0, y_0)$ with $x_0, y_0 \in (0,1)$ and $x_0 + y_0 = 1$.*

Proof. It is trivial to check that states of the form (x_0, x_0, y_0, y_0) with $x_0 + y_0 = 1$ are stationary. We prove that every stationary state (x_1, x_2, y_1, y_2) has $x_1 = x_2$ and $y_1 = y_2$, which then also necessarily satisfy $x_i + y_i = 1$, by contradiction. Assume (x_1, x_2, y_1, y_2) satisfy

$$0 = x_1 (1 - x_1 - y_1) + \mu_x (x_2 - x_1), \tag{3.6a}$$

$$0 = y_1 (1 - x_1 - y_1) + \mu_y (y_2 - y_1), \tag{3.6b}$$

$$0 = x_2 (1 - x_2 - y_2) + \mu_x (x_1 - x_2), \tag{3.6c}$$

$$0 = y_2 (1 - x_2 - y_2) + \mu_y (y_1 - y_2), \tag{3.6d}$$

and $x_1 > x_2 > 0$. Then (3.6a) and (3.6c) imply $x_2 + y_2 > 1 > x_1 + y_1$ so that $y_2 - y_1 > x_1 - x_2 > 0$. But then (3.6b) and (3.6d) require, in contradiction, that $x_1 + y_1 > 1$ and $1 > x_2 + y_2$. Thus $x_1 \leq x_2$ and, by symmetry, $x_2 \leq x_1$ so that $x_1 = x_2$ and $y_1 = y_2$. □

3.1.2 Asymptotic analysis

In this section we generalize the physically motivated asymptotic analysis developed in section 2.2.1 to the Itô stochastic differential equations (3.4). There are two

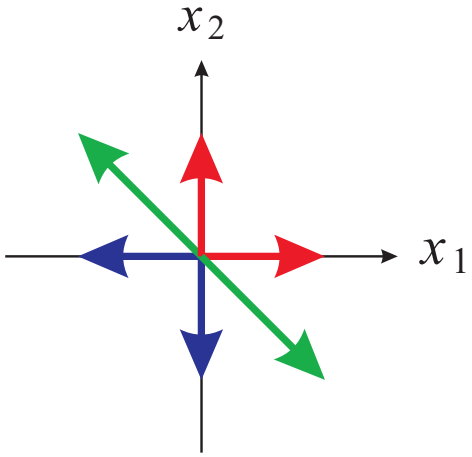

Figure 3.2: Directions of independent events of species X. Red and blue arrows (along the axes) represent birth and death events on specific patch respectively. Green arrows (diagonal) shows the direction of the hopping events in the phase space. Species Y has similar diagram but the diagonal arrows may have different strength.

complications in the 2-patch model as compared to the model in Chapter II. First, the approach developed in section 2.2.1 utilizes an analytically closed form for the deterministic trajectories in order to connect the noise-perturbed state and its final destination on the coexistence line. In the 2-patch model competition couples the variables $\{x_i, y_i\}$ locally while the hopping process couples populations on different patches. Thus the entire 4-dimensional states are dynamically and nonlinearly dependent and, owing to this complexity, an exact expression for the deterministic trajectories is not known. To unravel this difficulty, we combine a regular perturbation analysis with the proposed "intuitive" asymptotic approach. Second, the diffusive terms contain cross derivatives $\partial^2/\partial x_1 \partial x_2$ and $\partial^2/\partial y_1 \partial y_2$ indicating that the fluctuations along the coordinates (x_1, x_2) and (y_1, y_2) are coupled. The fact can also be seen in the coupled noises that are proportional to dW_2 and dW_5 in the Itô stochastic differential equations (3.4).

As illustrated in Fig. 3.2, independent noise processes at the individual level kick the system in 12 distinct directions in the phase space. These directions can be paired

Balanced Processes	Direction	Magnitude of the Fluctuation
Birth/Death of X on patch 1	$(\pm 1, 0, 0, 0)$	$\triangle_1 \equiv \left[\frac{2\rho}{\rho-1}\frac{x_0}{K}\,dt\right]^{1/2}$
Birth/Death of X on patch 2	$(0, \pm 1, 0, 0)$	$\triangle_2 \equiv \left[\frac{2\rho}{\rho-1}\frac{x_0}{K}\,dt\right]^{1/2}$
Birth/Death of Y on patch 1	$(0, 0, \pm 1, 0)$	$\square_1 \equiv \left[\frac{2\rho}{\rho-1}\frac{y_0}{K}\,dt\right]^{1/2}$
Birth/Death of Y on patch 2	$(0, 0, 0, \pm 1)$	$\square_2 \equiv \left[\frac{2\rho}{\rho-1}\frac{y_0}{K}\,dt\right]^{1/2}$
Hopping of X	$(\pm 1, \mp 1, 0, 0)$	$\blacktriangle \equiv \left[2\mu_x\frac{x_0}{K}\,dt\right]^{1/2}$
Hopping of Y	$(0, 0, \pm 1, \mp 1)$	$\blacksquare \equiv \left[2\mu_y\frac{y_0}{K}\,dt\right]^{1/2}$

Table 3.2: Fluctuation strengths and shorthand notations of the homogeneous 2–patch model.

into 6 balanced groups on the coexistence line: four from demographic birth and death processes and two from the hopping between patches. As a result, the fluctuations can effectively perturb a coexistent state into 2^6 possible directions. As $K \to \infty$ the effective strength of the fluctuations due to random birth and death events in a time interval dt were conjectured and verified by rigorous asymptotic analysis in section 2.4. The effective strength of the fluctuations due to random hopping events are obtained by evaluating their relative strengths to the birth and death fluctuations from Eq.(3.3). In summary, the effective strength of each fluctuation, as well as the shorthand notations are listed in the Table 3.2.

Next we present the outline of asymptotic calculation as $K \to \infty$. A heuristic diagram is provided in Fig. 3.3 and the detail derivation can be found in the following section 3.1.3. Suppose at some instance the system is at a coexistent state (x_0, x_0, y_0, y_0) with $x_0 + y_0 = 1$. Define a coordinate $z := x - y \equiv 2x - 1$ on the coexistence line such that z uniquely specifies a coexistence state. In a time interval

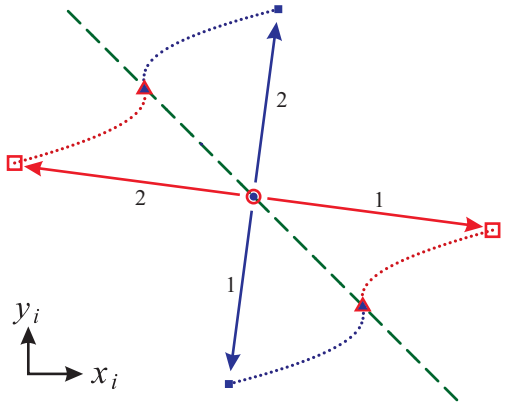

Figure 3.3: Heuristic diagram of physical asymptotic analysis. 4–dimensional states in this 2–dimensional diagram are represented by 2 points: open red (x_1, y_1) and closed blue (x_2, y_2). Dashed green line represents the coexistence line. Start from coexistent state (circles), two fluctuations (represented by arrows and noted by 1 and 2) kick the state out of the coexistence line (squares), then the states flow back to the coexistence line along deterministic trajectory (dotted curve) to the final destinations (triangles).

dt, one of the fluctuations drives the system out of the equilibrium to a new state

$$\text{(3.7)} \qquad \begin{pmatrix} x_0 \\ x_0 \\ y_0 \\ y_0 \end{pmatrix} \xrightarrow{\text{Impact of fluctuation}} \begin{pmatrix} x_0 + \phi_1 \triangle_1 + \phi_5 \blacktriangle \\ x_0 + \phi_2 \triangle_2 - \phi_5 \blacktriangle \\ y_0 + \phi_3 \square_1 + \phi_6 \blacksquare \\ y_0 + \phi_4 \square_2 - \phi_6 \blacksquare \end{pmatrix}$$

where each independent variable $\phi_j, j \in \{1, 2, 3, 4, 5, 6\}$ takes values ± 1 with probability $1/2$. We will use $\vec{\phi}$ to denote a specific combination of the ϕ_j's. Then this perturbed state is treated as the initial condition of the deterministic rate equations (3.5) and we seek for the final destination in the $1/K$-neighborhood of coexistence line after relaxation. To achieve this goal, we note that each perturbation carries a

small parameter $\epsilon \equiv 1/\sqrt{K}$ that suggests the perturbative ansatz

$$x_i(t) \equiv x_i^{(0)} + \epsilon\, x_i^{(1)} + \epsilon^2\, x_i^{(2)} + \mathcal{O}\left(\epsilon^3\right), \tag{3.8a}$$

$$y_i(t) \equiv y_i^{(0)} + \epsilon\, y_i^{(1)} + \epsilon^2\, y_i^{(2)} + \mathcal{O}\left(\epsilon^3\right), \tag{3.8b}$$

where $x_i^{(j)}$ and $y_i^{(j)}$ are ϵ-independent at every order j. The system then returns to the coexistence line along the deterministic flow

$$\begin{pmatrix} x_0 + \phi_1 \triangle_1 + \phi_5 \blacktriangle \\ x_0 + \phi_2 \triangle_2 - \phi_5 \blacktriangle \\ y_0 + \phi_3 \square_1 + \phi_6 \blacksquare \\ y_0 + \phi_4 \square_2 - \phi_6 \blacksquare \end{pmatrix} \xrightarrow{\text{Relaxation of rate equations}} \begin{pmatrix} x_{\vec{\phi}} \\ x_{\vec{\phi}} \\ y_{\vec{\phi}} \\ y_{\vec{\phi}} \end{pmatrix}. \tag{3.9}$$

In the above, $x_{\vec{\phi}}$ and $y_{\vec{\phi}}$ represent the x- and y-coordinate of the relaxed state previously perturbed by $\vec{\phi}$. In the z-coordinate, the displacement of the final state $z_{\vec{\phi}}$ from the original state z_0 is computed as $\Delta z_{\vec{\phi}} \equiv 2\left(x_{\vec{\phi}} - x_0\right)$. Consequently, in the large K limit, the dynamics in the entire 4–dimensional space is approximated by an effective one dimensional Markov process on the coexistence line, $z(t)$, defined by the Itô stochastic differential equation

$$dz = v(z)\, dt + \sqrt{2D(z)}\, dW \tag{3.10}$$

with drift $v(z_0) = \langle \Delta z_{\vec{\phi}} \rangle / dt$ and diffusion $D(z_0) = \langle \Delta {z_{\vec{\phi}}}^2 / 2 \rangle / dt$ (see section 2.2.1) where $\langle A_{\vec{\phi}} \rangle$ denotes the expectation value of random variable $A_{\vec{\phi}}$.

Somewhat surprisingly, after unraveling the straightforward but nontrivial calculations described in Appendix B, we deduce the remarkably simple forms of the drift

and diffusion:

$$v(z) = \frac{1-z^2}{K}\left[C_0 + \frac{C_1}{(\mu_x - \mu_y)\, z - C_2}\right], \tag{3.11}$$

$$D(z) = \frac{\rho}{2(\rho-1)}\frac{1-z^2}{K}, \tag{3.12}$$

where the coefficients C_0, C_1, and C_2 are functions of the parameters only:

$$C_0 = -\frac{\mu_x - \mu_y}{2}\left[\frac{1}{1+2(\mu_x+\mu_y)}\right], \tag{3.13a}$$

$$C_1 = -\frac{\mu_x - \mu_y}{2}\left[\frac{\rho}{\rho-1} + \frac{4\mu_x\,\mu_y}{1+2(\mu_x+\mu_y)}\right], \tag{3.13b}$$

$$C_2 = 4\mu_x\,\mu_y + \mu_x + \mu_y. \tag{3.13c}$$

When $\mu_X \neq \mu_Y$, the $\mathcal{O}(1/K)$ drift induced by demographic fluctuations breaks the degeneracy of the deterministic dynamics (3.5) and there is a preference for the fast disperser on an $\mathcal{O}(K)$ time scale. That is, if $\mu_x > \mu_y$, the drift $v(z)$ is strictly positive for $z \in (-1, 1)$ (ref: Chapter 6). On the other hand the diffusion $D(z)$ does not depend on μ_x and μ_y. Therefore, there is no evolutionary stable dispersal rate in the 2–patch model. That is, a faster-disperser will always have an advantage over a slower disperser. Given these drift and diffusion functions we can also derive integral forms of both the probability that one species outlives the other and the mean "fixation" time as function of initial position by the formulae provided in section 2.1.2.

3.1.3 Detailed computation of the physically motivated asymptotic analysis

In this section, we present details of the computations in the physically motivated asymptotic analysis.

We begin with applying regular perturbation theory to compute the time displacement of an equilibrium point (x_0, x_0, y_0, y_0) on the coexistence manifold $(x_0 + y_0 = 1)$

due to small initial displacements off the coexistence manifold in the deterministic 2-patch model. Following the intuition described in section 3.1.4 we presume an asymptotic expansion of the solutions to (3.5) of the form

$$x_i = x_i^{(0)} + \epsilon\, x_i^{(1)} + \epsilon^2\, x_i^{(2)} + \mathcal{O}\left(\epsilon^3\right) \tag{3.14a}$$

$$y_i = y_i^{(0)} + \epsilon\, y_i^{(1)} + \epsilon^2\, y_i^{(2)} + \mathcal{O}\left(\epsilon^3\right) \tag{3.14b}$$

where $\epsilon := 1/\sqrt{K}$ is the magnitude of stochastic kicks from birth, death, and hopping events. That is, we consider initial conditions within an $\mathcal{O}(\epsilon)$ displacement from the coexistence line:

$$x_i\,(0) = x_0 + \epsilon\, x_i^{(1)}(0) \tag{3.15a}$$

$$y_i\,(0) = y_0 + \epsilon\, y_i^{(1)}(0). \tag{3.15b}$$

with $x_i^{(n)}(0) = 0 = y_i^{(n)}(0)$ for all $n \geq 2$.

It is convenient to transform variables to total population n_x, n_y on both the patches, and the population difference w_x, w_y between the patches

$$n_x(t) = x_1(t) + x_2(t) = 2x_0 + \epsilon\, n_x^{(1)} + \epsilon^2\, n_x^{(2)} + \ldots \tag{3.16a}$$

$$n_y(t) = y_1(t) + y_2(t) = 2y_0 + \epsilon\, n_y^{(1)} + \epsilon^2\, n_y^{(2)} + \ldots \tag{3.16b}$$

$$w_x(t) = x_1(t) - x_2(t) = 0 + \epsilon\, w_x^{(1)} + \epsilon^2\, v_x^{(2)} + \ldots \tag{3.16c}$$

$$w_y(t) = y_1(t) - y_2(t) = 0 + \epsilon\, w_y^{(1)} + \epsilon^2\, v_y^{(2)} + \ldots \tag{3.16d}$$

with initial conditions of the form

$$n_x(0) = 2x_0 + \epsilon\, n_x^{(1)}(0) \tag{3.17a}$$

$$n_y(0) = 2y_0 + \epsilon\, n_y^{(1)}(0) \tag{3.17b}$$

$$w_x(0) = \ 0 \ + \epsilon\, w_x^{(1)}(0) \tag{3.17c}$$

$$w_y(0) = \ 0 \ + \epsilon\, w_y^{(1)}(0) \tag{3.17d}$$

and vanishing $n_x^{(n)}(0)$, $n_y^{(n)}(0)$, $w_x^{(n)}(0)$, $w_y^{(n)}(0)$ for $n \geq 2$. At $\mathcal{O}(\epsilon)$ the dynamical equations (3.5) become the linear homogeneous systems

$$\dot{n_x}^{(1)} = -x_0 \left(n_x^{(1)} + n_y^{(1)}\right) \tag{3.18a}$$

$$\dot{n_y}^{(1)} = -y_0 \left(n_x^{(1)} + n_y^{(1)}\right) \tag{3.18b}$$

and

$$\dot{w_x}^{(1)} = -x_0 \left(w_x^{(1)} + w_y^{(1)}\right) - 2\mu_x w_x^{(1)} \tag{3.19a}$$

$$\dot{w_y}^{(1)} = -y_0 \left(w_x^{(1)} + w_y^{(1)}\right) - 2\mu_y w_y^{(1)}. \tag{3.19b}$$

The solutions for (3.18) are

$$n_x^{(1)}(t) = y_0\, n_x^{(1)}(0) - x_0\, n_y^{(1)}(0) + x_0 \left(n_x^{(1)}(0) + n_y^{(1)}(0)\right) e^{-t} \tag{3.20a}$$

$$n_y^{(1)}(t) = x_0\, n_y^{(1)}(0) - y_0\, n_x^{(1)}(0) + y_0 \left(n_x^{(1)}(0) + n_y^{(1)}(0)\right) e^{-t} \tag{3.20b}$$

and the solutions for (3.19) are

$$w_x^{(1)}(t) = \frac{1}{2\kappa}\left\{\left[(\kappa-\eta)\,w_x^{(1)}(0) - x_0\,w_y^{(1)}(0)\right]e^{\lambda_+ t} + \left[(\kappa+\eta)\,w_x^{(1)}(0) + x_0\,w_y^{(1)}(0)\right]e^{\lambda_- t}\right\} \tag{3.21a}$$

$$w_y^{(1)}(t) = \frac{1}{2\kappa}\left\{\left[(\kappa+\eta)\,w_y^{(1)}(0) - y_0\,w_x^{(1)}(0)\right]e^{\lambda_+ t} + \left[(\kappa-\eta)\,w_y^{(1)}(0) + y_0\,w_x^{(1)}(0)\right]e^{\lambda_- t}\right\} \tag{3.21b}$$

where

$$\kappa = \frac{1}{2}\sqrt{1 + 4\left(\mu_x - \mu_y\right)^2 + 4\left(\mu_x - \mu_y\right)\left(x_0 - y_0\right)}\,, \tag{3.22}$$

$$\eta = \mu_x - \mu_y + \frac{1}{2}\left(x_0 - y_0\right), \tag{3.23}$$

and the (strictly negative) eigenvalues in the exponents are

$$\lambda_\pm = -\frac{1}{2}\left(1 + 2\mu_x + 2\mu_y\right) \pm \kappa. \tag{3.24}$$

Thus

$$\begin{aligned}\lim_{t\to\infty} x_1^{(1)}(t) &= \lim_{t\to\infty}\frac{1}{2}\left(n_x^{(1)}(t) + w_x^{(1)}(t)\right)\\ &= y_0\left(x_1^{(1)}(0) + x_2^{(1)}(0)\right) - x_0\left(y_1^{(1)}(0) + y_2^{(1)}(0)\right)\end{aligned} \tag{3.25}$$

and so on for the $x_2^{(1)}(t)$, $y_1^{(1)}(t)$, and $y_2^{(1)}(t)$. Because the initial perturbations $\epsilon x_i^{(1)}(0)$ and $\epsilon y_i^{(1)}(0)$ are symmetrically distributed, to first order in ϵ, *on average* they relax back to the starting point (x_0, x_0, y_0, y_0) on the equilibrium manifold. This means that the leading order terms do not contribute to the drift on the equilibrium manifold at $\mathcal{O}(\epsilon)$. But because the average of their squares do not vanish, they do constitute the leading approximation to the diffusion on the equilibrium manifold at $\mathcal{O}(\epsilon^2)$ as shown later. In order to calculate the leading $\mathcal{O}(\epsilon^2)$ contribution to the drift, we must

compute $\lim_{t\to\infty} x_i^{(2)}$ and $\lim_{t\to\infty} y_i^{(2)}$.

At second order the differential equations are

$$\dot{n_x}^{(2)} = -x_0\left(n_x^{(2)} + n_y^{(2)}\right) - \frac{1}{2}\left({n_x^{(1)}}^2 + {w_x^{(1)}}^2 + n_x^{(1)}n_y^{(1)} + w_x^{(1)}w_y^{(1)}\right) \tag{3.26a}$$

$$\dot{n_y}^{(2)} = -y_0\left(n_x^{(2)} + n_y^{(2)}\right) - \frac{1}{2}\left({n_y^{(1)}}^2 + {w_y^{(1)}}^2 + n_x^{(1)}n_y^{(1)} + w_x^{(1)}w_y^{(1)}\right) \tag{3.26b}$$

and

$$\dot{w_x}^{(2)} = -x_0\left(w_x^{(2)} + w_y^{(2)}\right) - 2\mu_x w_x^{(2)} - \frac{1}{2}\left(2n_x^{(1)}w_x^{(1)} + n_y^{(1)}w_x^{(1)} + n_x^{(1)}w_y^{(1)}\right) \tag{3.27a}$$

$$\dot{w_y}^{(2)} = -y_0\left(w_x^{(2)} + w_y^{(2)}\right) - 2\mu_x w_y^{(2)} - \frac{1}{2}\left(2n_y^{(1)}w_y^{(1)} + n_x^{(1)}w_y^{(1)} + n_y^{(1)}w_x^{(1)}\right). \tag{3.27b}$$

The matrix of coefficients of the linear terms on the right hand side of equations (3.27) immediately above is invertible, and the last inhomogeneous terms on the right hand side vanish as $t \to \infty$, so $\lim_{t\to\infty} w_x^{(2)}(t) = \lim_{t\to\infty} w_y^{(2)}(t) = 0$. Thus the ultimate displacements on the equilibrium manifold are determined, to $\mathcal{O}(\epsilon^2)$, by $\lim_{t\to\infty} x_i^{(2)} = \lim_{t\to\infty} n_x^{(2)}/2$ and $\lim_{t\to\infty} y_i^{(2)} = \lim_{t\to\infty} n_y^{(2)}/2$.

We solve equations (3.26) as follows. First add them to obtain a closed linear inhomogeneous differential equation for total population of both species $n^{(2)}(t) \equiv n_x^{(2)}(t) + n_y^{(2)}(t)$

$$\dot{n}^{(2)} = -n^{(2)} - \frac{1}{2}\left[\left(n_x^{(1)} + n_y^{(1)}\right)^2 + \left(w_x^{(1)} + w_y^{(1)}\right)^2\right], \tag{3.28}$$

and equivalently,

$$\frac{d}{dt}\left[e^t\, n^{(2)}(t)\right] = -\frac{e^t}{2}\left[\left(n_x^{(1)} + n_y^{(1)}\right)^2 + \left(w_x^{(1)} + w_y^{(1)}\right)^2\right]. \tag{3.29}$$

Since $n^{(2)}(0) = 0$, the solution of Eq.(3.29) is

$$n^{(2)}(t) = -\frac{e^{-t}}{2} \int_0^t e^{t'} \left[\left(n_x^{(1)}(t') + n_y^{(1)}(t') \right)^2 + \left(w_x^{(1)}(t') + w_y^{(1)}(t') \right)^2 \right] dt'. \tag{3.30}$$

Next, the equation of motion of

$$\dot{n}_x^{(2)} = -x_0\, n^{(2)} - x_1^{(1)} \left(x_1^{(1)} + y_1^{(1)} \right) - x_2^{(1)} \left(x_2^{(1)} + y_2^{(1)} \right). \tag{3.31}$$

With the initial condition $n_x^{(2)}(0) = 0$ and Eq.(3.30), the evolution of the total population of species X is

$$\begin{aligned} n_x^{(2)}(t) = x_0 \int_0^t \int_0^{t'} & \left[\left(n_x^{(1)}(t'') + n_y^{(1)}(t'') \right)^2 \right. \\ & \left. + \left(w_x^{(1)}(t'') + w_y^{(1)}(t'') \right)^2 \right] e^{t''} dt'' e^{-t'} dt' \\ - \frac{1}{2} \int_0^t & \left[{n_x^{(1)}}^2(t') + {w_x^{(1)}}^2(t') \right. \\ & \left. + n_x^{(1)}(t') n_y^{(1)}(t') + w_x^{(1)}(t') w_y^{(1)}(t') \right] dt' \end{aligned} \tag{3.32}$$

To proceed analysis, we prove the following technical lemma.

Lemma III.2. *If* $\lambda < 0$,

$$\int_0^\infty e^{-t} \int_0^t e^{t'} e^{\lambda t'}\, dt'\, dt = -\frac{1}{\lambda}. \tag{3.33}$$

Proof. Suppose $\lambda \neq -1$,

$$\int_0^\infty e^{-t} \int_0^t e^{t'} e^{\lambda t'} dt'\, dt = \int_0^\infty e^{-t} \int_0^t e^{(\lambda+1)t'} dt'\, dt \tag{3.34}$$
$$= \frac{1}{\lambda+1} \int_0^\infty \left[e^{\lambda t} - e^{-t}\right] dt$$
$$= \frac{1}{\lambda+1} \left[\frac{e^{\lambda t}}{\lambda} + e^{-t}\right]_0^\infty$$
$$= \frac{-1}{\lambda+1} \left[\frac{1}{\lambda} + 1\right] = -\frac{1}{\lambda}.$$

If $\lambda = -1$,

$$\int_0^\infty e^{-t} \int_0^t e^{t'} e^{-t'} dt'\, dt = \int_0^\infty e^{-t} t dt = -\int_0^\infty t de^{-t} \tag{3.35}$$
$$= \left[-te^{-t}\right]_0^\infty + \int_0^\infty e^{-t} dt = 1 = -\frac{1}{\lambda}.$$

□

We are now in the position to apply the physically motivated asymptotic analysis. To alleviate the lengthy expressions in the computations, it is convenient to define the following variables:

$$w_{x+}^{(1)} \equiv \frac{1}{2\kappa} \left[(\kappa - \eta)\, w_x^{(1)}(0) - x_0 w_y^{(1)}(0)\right], \tag{3.36a}$$
$$w_{x-}^{(1)} \equiv \frac{1}{2\kappa} \left[(\eta + \kappa)\, w_x^{(1)}(0) + x_0 w_y^{(1)}(0)\right], \tag{3.36b}$$
$$w_{y+}^{(1)} \equiv \frac{1}{2\kappa} \left[(\kappa + \eta)\, w_y^{(1)}(0) - y_0 w_x^{(1)}(0)\right], \tag{3.36c}$$
$$w_{y-}^{(1)} \equiv \frac{1}{2\kappa} \left[(\kappa - \eta)\, w_y^{(1)}(0) + y_0 w_x^{(1)}(0)\right], \tag{3.36d}$$

so the $\mathcal{O}(\epsilon)$ solutions of the population differences (3.21) are

$$w_x^{(1)}(t) = w_{x+}^{(1)} e^{\lambda_+ t} + w_{x-}^{(1)} e^{\lambda_- t}, \tag{3.37a}$$

$$w_y^{(1)}(t) = w_{y+}^{(1)} e^{\lambda_+ t} + w_{y-}^{(1)} e^{\lambda_- t}. \tag{3.37b}$$

By reading off the initial conditions from Eq.(3.7), it is clear that the initial perturbed states in the coordinates (n, w) are

$$n_x^{(1)}(0) = \phi_1 \Delta_1 + \phi_2 \Delta_2, \tag{3.38a}$$

$$n_y^{(1)}(0) = \phi_3 \square_1 + \phi_4 \square_2, \tag{3.38b}$$

$$w_x^{(1)}(0) = \phi_1 \Delta_1 - \phi_2 \Delta_2 + 2\phi_5 \blacktriangle, \tag{3.38c}$$

$$w_y^{(1)}(0) = \phi_3 \square_1 - \phi_4 \square_2 + 2\phi_6 \blacksquare. \tag{3.38d}$$

Note regarding the notations that the "initial conditions" $n_i^{(1)}(0)$ and $w_i^{(1)}(0)$ are inherently random variables. Their values depend on the stochastic perturbation vector $\vec{\phi}$.

It is straightforward to evaluate the effective diffusion (to $\mathcal{O}(\epsilon^2)$) on the coexistence line,

$$\begin{aligned} D(z_0) \equiv \frac{\left\langle \Delta z_{\vec{\phi}}^2 \right\rangle}{2dt} &= \lim_{t\to\infty} \frac{1}{2dt} \left\langle \Delta \left(x_{\vec{\phi}} - y_{\vec{\phi}} \right)^2 \right\rangle \\ &= \frac{1}{2} \left\langle \left[y_0\, n_x^{(1)}(0) - x_0\, n_y^{(1)}(0) \right]^2 \right\rangle. \end{aligned} \tag{3.39}$$

and by exploiting the balance condition $\langle \phi_i \rangle = 0$, $\langle \phi_i^2 \rangle = 1$, and the independency of the fundamental fluctuations $\langle \phi_i \phi_j \rangle = \delta_{ij}$, we arrive at a simple form of the effective diffusion coefficient

$$D(z) = \frac{\rho}{2(\rho - 1)} \frac{1 - z^2}{K}. \tag{3.40}$$

The computation of the drift coefficient is more complicated. We start by plugging the $\mathcal{O}(\epsilon^1)$ solutions into the inhomogeneous terms in Eq.(3.32). The expressions involve the total population (of both species) and the population differences (of both species) among patches

$$n_x^{(1)}(t) + n_y^{(1)}(t) = \left[n_x^{(1)}(0) + n_y^{(1)}(0)\right] e^{-t} \tag{3.41a}$$
$$= \left[\phi_1\Delta_1 + \phi_2\Delta_2 + \phi_3\Box_1 + \phi_4\Box_2\right] e^{-t},$$

$$w_x^{(1)}(t) + w_y^{(1)}(t) = \left[w_{x+}^{(1)} + w_{y+}^{(1)}\right] e^{\lambda_+ t} + \left[w_{x-}^{(1)} + w_{y-}^{(1)}\right] e^{\lambda_- t}, \tag{3.41b}$$

so that

$$\left[n_x^{(1)}(t) + n_y^{(1)}(t)\right]^2 = n_0^2\, e^{-2t}, \tag{3.42a}$$

$$\left[w_x^{(1)}(t) + w_y^{(1)}(t)\right]^2 = w_{++}e^{2\lambda_+ t} + 2w_{+-}e^{(\lambda_+ + \lambda_-)t} + w_{--}e^{2\lambda_- t}, \tag{3.42b}$$

with the defined random variables

$$n_0 \equiv \left[\phi_1\Delta_1 + \phi_2\Delta_2 + \phi_3\Box_1 + \phi_4\Box_2\right], \tag{3.43}$$

$$w_{++} \equiv \left[w_{x+}^{(1)} + w_{y+}^{(1)}\right]^2, \tag{3.44}$$

$$w_{+-} \equiv \left[w_{x+}^{(1)} + w_{y+}^{(1)}\right] \times \left[w_{x-}^{(1)} + w_{y-}^{(1)}\right], \tag{3.45}$$

$$w_{--} \equiv \left[w_{x-}^{(1)} + w_{y-}^{(1)}\right]^2. \tag{3.46}$$

Plug the inhomogeneous terms (3.1.3) into the $\mathcal{O}(\epsilon^2)$ solution (3.32), one immediately

has

$$n_x^{(2)}(t) = \frac{x_0}{2}\int_0^t e^{-t'}\int_0^{t'}\Big(n_0^2 e^{-2t'} + w_{++}e^{2\lambda_+ t''} + \\ + 2w_{+-}e^{(\lambda_+ + \lambda_-)t''}\, w_{--}e^{2\lambda_- t''}e^{t''}\Big)\, dt''\, dt' \\ - \frac{1}{2}\int_0^t \Big(n_{s1}e^{-t'} + n_{s2}e^{-2t'} + w_{s++}e^{2\,\lambda_+ t'} \\ + w_{s+-}e^{(\lambda_+ + \lambda_-)t'} + w_{s--}e^{2\lambda_- t'}\Big)\, dt' \tag{3.47}$$

where the source terms in the second integration are defined as

$$n_{s1} \equiv n_0 \times \left[y_0\, n_x^{(1)}(0) - x_0\, n_y^{(1)}(0)\right], \tag{3.48}$$

$$n_{s2} \equiv x_0\, n_0^2, \tag{3.49}$$

$$w_{s++} \equiv w_{x+}^{(1)} \times \left[w_{x+}^{(1)} + w_{y+}^{(1)}\right], \tag{3.50}$$

$$w_{s+-} \equiv w_{x+}^{(1)} \times \left[w_{x-}^{(1)} + w_{y-}^{(1)}\right] + w_{x-}^{(1)} \times \left[w_{x+}^{(1)} + w_{y+}^{(1)}\right], \tag{3.51}$$

$$w_{s--} \equiv w_{x-}^{(1)} \times \left[w_{x-}^{(1)} + w_{y-}^{(1)}\right]. \tag{3.52}$$

By applying Lemma III.2 to Eq.(3.47), one obtain an intermediate expression of the drift

$$n_x^{(2)}(t \to \infty) = \frac{1}{2}\left[-n_{s1} + \frac{x_0\, w_{++} - w_{s++}}{-2\lambda_+} \right. \\ \left. + \frac{2\, x_0\, w_{+-} - w_{s+-}}{-\lambda_+ - \lambda_-} + \frac{x_0\, w_{--} - w_{s--}}{-2\,\lambda_-}\right]. \tag{3.53}$$

Note that the $\lim_{t\to\infty} n_x^{(2)}(t)$ in above expression (3.53) is again a random variable whose value depends on the stochastic perturbative vector $\vec{\phi}$. It is elementary to

evaluate $\langle \lim_{t\to\infty} n_x^{(2)}(t)\rangle$ by inserting (3.38) into (3.53) to obtain

$$
\begin{aligned}
(3.54) \qquad v_c\left(z_0\right) = \frac{1-z_0^2}{K\kappa^2}\Bigg\{ & \frac{\left(\mu_x-\mu_y\right)}{4}\frac{z_0^2-2\eta\, z_0-1}{1+2\mu_x+2\mu_y} \\
& + \frac{2\left(1+2\mu_x+2\mu_y\right)}{4\kappa^2-\left(1+2\mu_x+2\mu_y\right)^2}\left(\mu_x-\mu_y\right)\left(z_0^2+4\kappa^2+4\eta^2-4\eta z_0-1\right) \\
& + \frac{\kappa^2}{4\kappa^2-\left(1+2\mu_x+2\mu_y\right)^2}\left(\frac{\rho}{\rho-1}+\mu_x+\mu_y\right)\left(z_0-2\eta\right)\Bigg\}.
\end{aligned}
$$

Further simplifications can be made by plugging (3.79) with $z_0 = x_0 - y_0$, and finally we obtain Eq.(3.11).

3.1.4 Simulations and numerical computations

Exact continuous time Markov chain (CTMC) simulations [35] of the homogeneous 2–patch model were carried out. The fixed low density birth-death ratio is $\rho = 2$ and the hopping rates μ_X and μ_Y were varied along with the carrying capacity $K\,(=\Lambda$ when $\rho = 2)$. For each of the parameter sets, 30 uniformly distributed points on the coexistence line were sampled as initial populations, and 10^4 realizations were performed for each initial condition. The simulations ran until one of the species had total population 0 on both patches, and the winning probability $\pi(z)$ of species X was computed, where z denotes the initial difference of scaled population. Because the measurement of winning probability is essentially a Bernoulli trial, the sample error can be computed by the sample mean, i.e. $\pi(z)$. The error of the mean of the trials was bounded by 0.5×10^{-2} so for neatness we omit error bars in the figures. The mean extinction time of either species, $\tau(z)$, was also measured.

Theoretical winning probabilities and mean extinction times were computed numerically by using (3.11) and (3.12) in the general formulae provided in Chapter II, section 2.1.2. We observe that the simulation results converge to the theoretical predictions as K increases. With $K = 125$ the asymptotic result differs from the

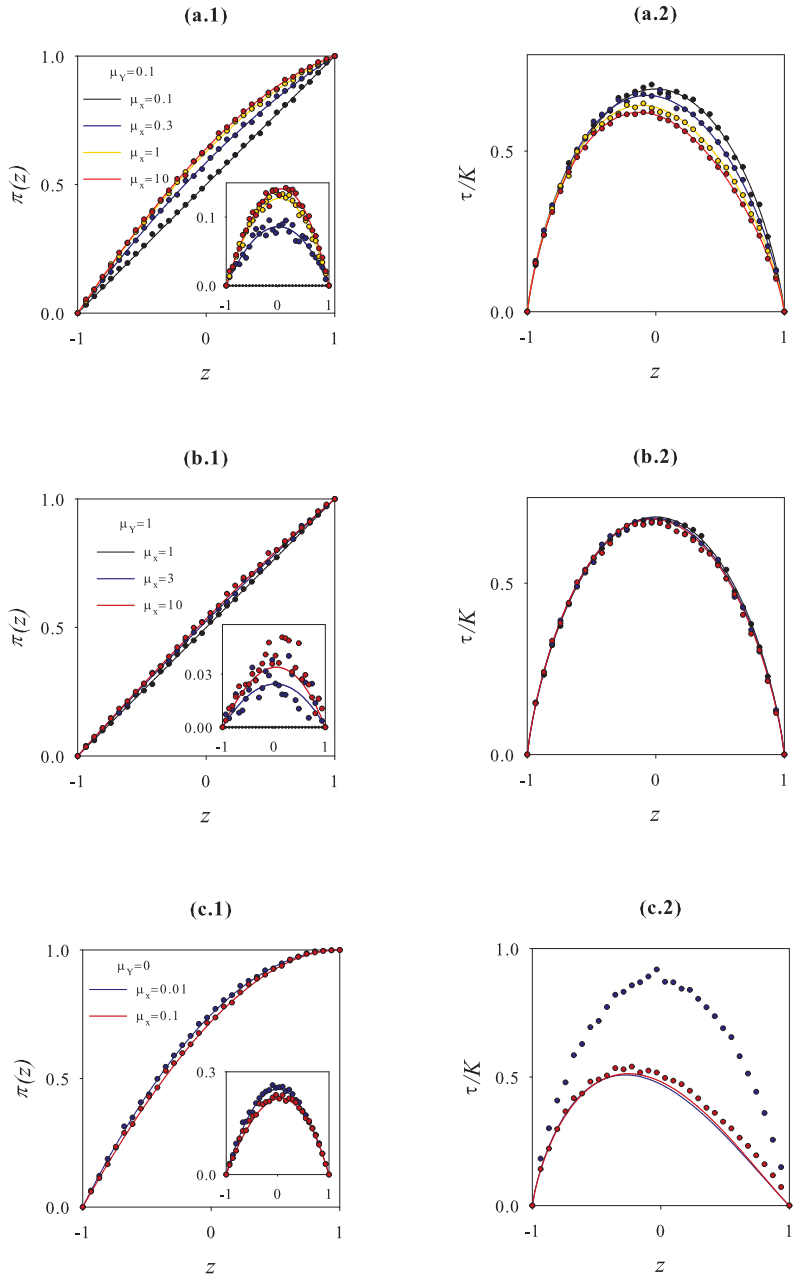

Figure 3.4: Comparisons of simulations (discrete dots) and theoretical predictions (solid lines). Left column: winning probability π of species X as function of initial state z; the inset of the left column shows the gained winning probability of species X from the microscopic symmetric system $\mu_X = \mu_Y$. Right column: scaled mean extinction time τ/K of any of the species as function of initial state z.

simulations by less than 5% in general (not presented). However, it is noted that the convergence is not uniform in (μ_X, μ_Y).

We present results of characteristic sets of parameters in Fig. 3.4. The carrying capacity $K = 200$ in these sets of parameters and $\mu_Y \leq \mu_X$. When the hopping rate of the slow species is small ($\mu_Y \approx 0.1$) there exist significant gains of the winning probability of the fast species X for various initial condition z. However, as soon as $\mu_X \approx 1$, asymptotic analysis confirmed by numerical evidence suggests the advantage of the more mobile species X saturates. On the other hand, if the slow species Y increases the hopping rate to about unity, the advantage of the fast species X drops down to be less than 5%.

The condition $\mu_X, \mu_Y > 0$ is relaxed to $\max(\mu_X, \mu_Y) > 0$ in the last set in Fig. 3.4. The winning probability of the theoretical analysis fits the simulations remarkably well, but the analysis does not give quantitatively correct predictions for mean extinction time. The reason for this is that both $\mu_X > 0$ and $\mu_Y > 0$ are necessary for asymptotic convergence within the time scale $\mathcal{O}(K)$. When this condition is relaxed, convergence breaks near the boundary of the coexistence line at $z = \pm 1$. When computing the winning probability, due to the effective diffusion in this region, the states near $z = \pm 1$ are still absorbed to the boundary $z = \pm 1$ without accumulating significant errors. But when computing the mean exit time, the error in convergence time builds up to invalidate the asymptotic analysis; in reality it takes longer time for either of the species to go extinct.

Because the asymptotic approach faithfully reproduces many feature of the simulations we can confidently fix the initial condition $z = 0$ and explore the winning probability of X in the parameter space (μ_X, μ_Y) theoretically. The result is presented in Fig. 3.5. In such head-to-head competition, analysis suggests that the maximum winning probability of the fast species is at most 75% among all possible (μ_X, μ_Y). It is noteworthy that this sort of saturation (of the advantage) phenomenon was ob-

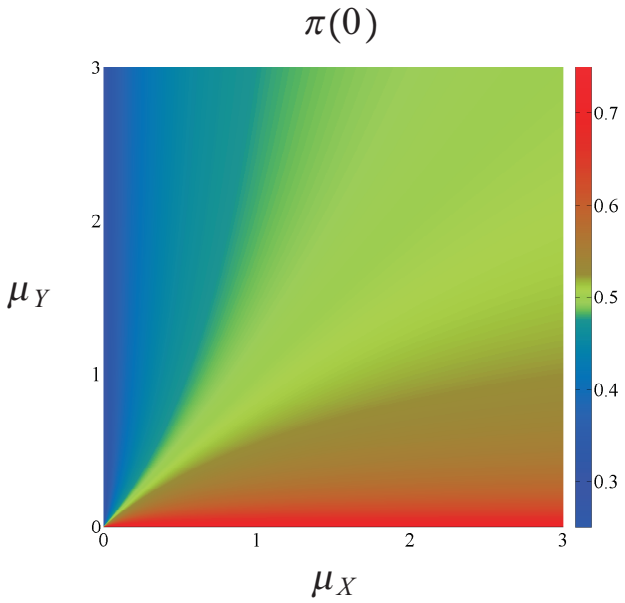

Figure 3.5: Landscape of the winning probability of the species X in a head–to–head competition.

served and analyzed in Chapter II. The analytical form of the winning probability $\pi(0)$ involves an incomplete Γ function and is beyond our interest.

In a short conclusion, the 2-patch model exhibits a weak preference for the higher-mobility species on an $\mathcal{O}(K)$ time scale. The fast species is more likely to win in head-to-head competition, but when the hopping rate of the slow species increases to about unity, the winning probability of the fast species is not significantly greater than 50%. There exists no finite evolutionarily stable rate for dispersion.

3.2 The many-patch model

3.2.1 The model

In this section we construct an extended model consisting of a countably infinite number of identical patches. The motivation is to study how the number of patches changes the behavior of the dynamics, and how demographic fluctuations affect the

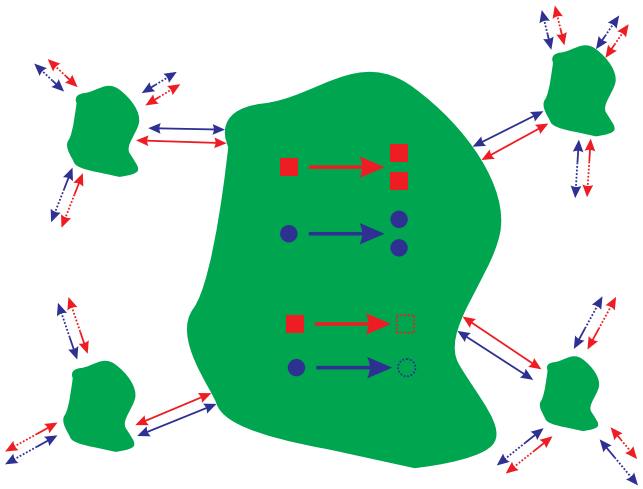

Figure 3.6: Dynamics of the interacting species distributed on many patches. The X and Y populations compete locally, as in the 2-patch model, while individuals randomly move from any patch to any other at rates μ_X and μ_Y.

competition outcome in such a system.

As in the 2-patch model we consider two species moving among patches and competing locally for limited resources. The universe is homogeneous in the sense that the carrying capacities of the patches are identical. Each species has identical per capita rates of birth and death under identical environmental conditions, but the per capita hopping rates of the species are not in general identical. When an individual relocates, it moves to any other patch with equal probability. We choose this global hopping dynamics to avoid local effects and exploit the extra level of averaging to study the interplay of mobility and birth-death fluctuations.

It is natural to adopt a theoretical approach similar to that used for the homogeneous 2-patch model: write down the master equation based on the individual processes, derive the Kolmogorov forward equation (Fokker-Planck equation) within the large carrying capacity expansion, and perform the "intuitive" asymptotic analysis. When the number of patches is countably infinite, fluctuations perturb the system into uncountably infinite many directions which potentially presents analytical chal-

lenges.

To address these challenges we note several facts:

- Because the patches are identical, when there are identical initial configurations on each patch the distribution of the random populations on each patch will always be identical.

- Correlations of the dynamics on different patches arise from the hopping events when the number of patches is finite, but decorrelate in the infinite-patch limit.

- The hopping events can be viewed as birth and death events with ensemble-averaged rates in the infinite-patch limit.

Moreover, as the number of patches increases the fluctuations between the population on any one patch and the total population on all other patches should be approximately independent.

Consider a simple example to illustrate how these properties help us to build an effective model. Let $X_k(s)$ be the random population of X species at time s in k patch. When $X_k(s) = a_k \in \mathbb{N}$ $(\forall k \in \mathbb{N})$ the total hopping rate of species X out of site 1 is $\mu_X \cdot a_1$. On the other hand, the total hopping rate of species into patch 1 is $\mu_X \cdot \lim_{N\to\infty} \sum_{k\neq 1}^{N} a_k/(N-1)$, which converges to $\mu_X \cdot \langle X_1(s)\rangle$ since the $X_k(s)$ are identically distributed. The point is that we can construct a model with only one patch where the populations evolve at each instant of time according not only to the local populations at that instant, but also in accord with the expectation values of the populations.

With this in mind we consider the following effective homogeneous many-patch model (referred to simply as the many-patch model later): there is *only one patch* in the space and X_s and Y_s represent the discrete nonnegative integer valued random populations of species X and Y at time s. When $X_s = a$ and $Y_s = b$, the following

transitions characterize the birth, death, and hopping events as a whole:

$$a \to a+1, \quad \text{with rate } \beta\, a + \mu_X \cdot \langle X_s \rangle \tag{3.55a}$$

$$a \to a-1, \quad \text{with rate } \delta\, a \left(1 + \frac{a+b}{\Lambda}\right) + \mu_X\, a \tag{3.55b}$$

$$b \to b+1, \quad \text{with rate } \beta\, b + \mu_Y \cdot \langle Y_s \rangle \tag{3.55c}$$

$$b \to b-1, \quad \text{with rate } \delta\, b \left(1 + \frac{a+b}{\Lambda}\right) + \mu_Y\, b, \tag{3.55d}$$

with low-density per capita birth rate β, death rate δ and hopping rates μ_X and μ_Y, and $\langle X_s \rangle$ and $\langle Y_s \rangle$ the expectation values of the random variables at time s. The probability $p_{a,b}(s) := \mathbb{P}(\{X_s = a\} \cap \{Y_s = b\})$ evolves with the "nonlinear" master equation

$$\begin{aligned}
\frac{d}{ds} p_{a,b} = &- \left\{ a \left[\beta - \delta \left(1 + \frac{a+b}{\Lambda}\right) \right] + \mu_X (\langle X_s \rangle - a) \right\} p_{a,b} \\
&- \left\{ b \left[\beta - \delta \left(1 + \frac{a+b}{\Lambda}\right) \right] + \mu_Y (\langle Y_s \rangle - b) \right\} p_{a,b} \\
&+ \left[\delta\,(a+1) \left(1 + \frac{a+b+1}{\Lambda}\right) + \mu_X\,(a+1) \right] p_{a+1,b} \\
&+ \left[\delta\,(b+1) \left(1 + \frac{a+b+1}{\Lambda}\right) + \mu_Y\,(b+1) \right] p_{a,b+1} \\
&+ \left[\beta\,(a-1) + \mu_X\ \langle X_s \rangle \right] p_{a-1,b} \\
&+ \left[\beta\,(b-1) + \mu_Y\ \langle Y_s \rangle \right] p_{a,b-1}
\end{aligned} \tag{3.56}$$

and for enough large but finite carrying capacities (3.56) is approximated by the

nonlinear Kolmogorov forward (Fokker-Plank) equation

$$
\begin{aligned}
\frac{\partial f(x,y,t)}{\partial t} =& -\frac{\partial}{\partial x}\left\{\left[x(1-x-y)+\mu_x\left(\langle x\rangle - x\right)\right] f\right\} \\
& -\frac{\partial}{\partial y}\left\{\left[y(1-x-y)+\mu_y\left(\langle y\rangle - y\right)\right] f\right\} \\
& +\frac{1}{2K}\frac{\partial^2}{\partial x^2}\left\{\left[x\left(\frac{\rho+1}{\rho-1}+x+y\right)+\mu_x(\langle x\rangle + x)\right] f\right\} \\
& +\frac{1}{2K}\frac{\partial^2}{\partial y^2}\left\{\left[y\left(\frac{\rho+1}{\rho-1}+x+y\right)+\mu_y(\langle y\rangle + y)\right] f\right\}
\end{aligned} \tag{3.57}
$$

referring to time–scaled continuum variables $x_t = X_s/K$ and $y_t = Y_s/K$ and dimensionless time $t = (\beta - \delta)s$. The parameters are defined identically to their counterparts in the 2-patch model: $K = (\rho - 1)\Lambda$ with $\rho = \beta/\delta > 1$, $\mu_x = \mu_X/(\beta - \delta)$ and $\mu_y = \mu_Y/(\beta - \delta)$. The time-dependent site-averages are $\langle x\rangle = \iint x\, f(x,y,t)\, dydx$ and $\langle y\rangle = \iint y\, f(x,y,t)\, dx\, dy$. We will now demonstrate that this many-patch model also has a weak selection of the fast disperser in a time scale $\mathcal{O}(K)$.

3.2.2 Asymptotic analysis

We generally follow the strategy used for the 2-patch model to study this many–patch model but because the distribution f evolves according to the expectation values of random variables, the interpretation of each step must be carefully modified. In this section we outline the analysis and the physical interpretations. Details of the calculations are presented in the following section 3.2.3.

First, the rate equations for the effective many-patch model, the analogs of (3.5) for the 2-patch model, are

$$\dot{x} = x(1-x-y) + \mu_x\left(\langle x\rangle - x\right), \tag{3.58a}$$

$$\dot{y} = y(1-x-y) + \mu_y\left(\langle y\rangle - y\right). \tag{3.58b}$$

It is tempting to refer this system of ordinary differential equations as "deterministic"

but it is important to realize that x and y are variables at one site (within one sample path) of the random processes (x_t, y_t) whose dynamics are described by (3.57) (albeit without the second order derivatives). The evolution of this particular sample path actually depends upon other sites (realizations) through the coupling with $\langle x\rangle$ and $\langle y\rangle$. Once the initial distributions of x_0 and y_0 are specified, i.e., once $f(x, y, 0)$ is selected, the distribution is deterministically evolved by (3.57). We will call (3.58) the "mean field rate equations".

Furthermore, the dynamics of x and y in (3.58) consists of two parts: the first part is the competitive dynamics from birth–death processes and the second part is the effect of hopping events. Without hopping events, the competitive dynamics at each site (in every realization) is independent of the others and the variables relax to a $1/K$-neighborhood of coexistence line on an $\mathcal{O}(\log K)$ time scale. The effect from hopping events is to push x_t and y_t toward an "equilibrium point" $\langle x_t\rangle$ and $\langle y_t\rangle$ among all equilibria on the $\mathcal{O}(\log K)$ time scale. In the following theorem, parallel to Theorem III.1, proves that the rate equations drive every realization onto the coexistence line with identical x and y.

Theorem III.3. *[Equilibrium of the deterministic many-patch model.] The population of each species are the same on all sites when the many-patch model is in equilibrium. That is, $x = \langle x\rangle$ and $y = \langle y\rangle$ at each site (equivalently, $f_{eq}(x, y) = \delta(x - \langle x\rangle)\,\delta(y - \langle y\rangle)$ in the distributional setting). Moreover, $\langle x\rangle + \langle y\rangle = 1$ unless $\langle x\rangle = 0 = \langle y\rangle$.*

Proof. The steady states at each site satisfy $0 \le x \le 1$ and $0 \le y \le 1$ and

$$0 = x\,(1 - x - y) + \mu_x\,(\langle x\rangle - x) \tag{3.59}$$

$$0 = y\,(1 - x - y) + \mu_y\,(\langle y\rangle - y)\,. \tag{3.60}$$

Suppose there is a non-trivial equilibrium with $\langle x\rangle + \langle y\rangle > 0$ and, without loss of

generality, $\langle x \rangle > 0$. Then (3.59) implies $x > 0$ everywhere.

Now suppose that $x \neq \langle x \rangle$ on a finite fraction of sites so that there is a finite fraction of sites with $1 \geq x > \langle x \rangle > 0$. On those sites (3.59) guarantees that $x + y < 1$.

If $y = 0$ on those sites then (3.60) requires that $\langle y \rangle = 0$ so $y = 0$ *everywhere* in which case x is the unique positive solution to $0 = x^2 + (\mu_x - 1)x - \mu_x \langle x \rangle$ at each site, contradicting the assumption $x \neq \langle x \rangle$ on a finite fraction of sites.

If $y \neq 0$ on a finite fraction of the finite fraction of sites where $1 \geq x > \langle x \rangle > 0$, then $\langle y \rangle > 0$. Because $x + y < 1$ on those sites, (3.60) implies that $y > \langle y \rangle$ on those sites. Thus on those sites $x + y > \langle x \rangle + \langle y \rangle$ and since $x + y < 1$ there, we deduce that the averages satisfy $\langle x \rangle + \langle y \rangle < 1$. Therefore at *every* site (recalling that $\langle x \rangle > 0$ guarantees that $x > 0$ everywhere),

$$0 = 1 - x - y + \mu_x \left(\frac{\langle x \rangle}{x} - 1 \right) \tag{3.61}$$

and, averaging over all sites,

$$0 = 1 - \langle x \rangle - \langle y \rangle + \mu_x \left(\langle x \rangle \langle \frac{1}{x} \rangle - 1 \right). \tag{3.62}$$

But the Cauchy-Schwarz inequality guarantees that $\langle x \rangle \langle \frac{1}{x} \rangle \geq 1$ so (3.62) implies $\langle x \rangle + \langle y \rangle > 1$ contradicting the deduction above that $\langle x \rangle + \langle y \rangle < 1$.

Hence we conclude that $x = \langle x \rangle$ everywhere, and by symmetry, that $y = \langle y \rangle$ everywhere. It immediately follows from (3.59) and (3.60) that $x + y = 1$ everywhere and $\langle x \rangle + \langle y \rangle = 1$. □

When the system is far away from a fixed point the dynamics is mainly governed by the rate equations. Once the system approaches an equilibrium, as for the 2-patch model, the second order derivative terms of (3.57) become comparable to the drift and the local asymptotic approach is adopted to analyze the nonlinear fluctuation-driven

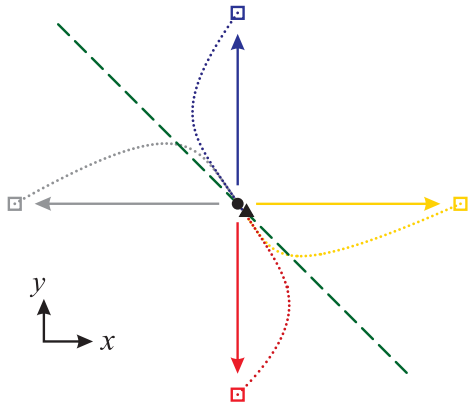

Figure 3.7: Heuristic diagram of of the physical asymptotic analysis of the homogeneous many patch model. Dashed green line represents the coexistence line. The effective 1–patch system consists of 4 points, and start from a coexistent state (closed circle). Fluctuations (denoted by arrows) perturb each point to each of the 4 characteristic directions to the "kicked–out" states (open squares). The rate equations evolves these 4 points back to the coexistence line (denoted by dashed curves) to the final state (closed triangle).

Balanced Processes	Direction	Magnitude of the Fluctuation
Birth/Death of X	$(\pm 1, 0)$	$\triangle_X \equiv \left[\left(\frac{2\rho}{\rho-1} + 2\mu_x\right) \frac{x_0}{K}\, dt\right]^{1/2}$
Birth/Death of Y	$(0, \pm 1)$	$\triangle_Y \equiv \left[\left(\frac{2\rho}{\rho-1} + 2\mu_y\right) \frac{y_0}{K}\, dt\right]^{1/2}$

Table 3.3: Fluctuation strengths and shorthand notations of the homogeneous many patch model.

dynamics.

To begin local asymptotic analysis we notice that individual-based processes (3.55) suggest that fundamental fluctuations are along $(x, y) = (\pm 1, 0)$ and $(0, \pm 1)$. At each equilibrium, these paired directions balance, and the strength of the fluctuations can be determined. Table 3.3 summarizes the directions and the strength of the demographic noises. As a first order approximation to examine the effect of fluctuations we approximate the many-patch model as if there are only 4 realizations. Starting at an equilibrium, each realization has identical x and y on the coexistence line, as inferred

by Theorem III.3. When fluctuations act on the system, these variables are kicked in 4 different directions with precisely defined strengths. Then the whole system flows back to the coexistence line by rate equations (3.58) where the arithmetic average position of the four points is obviously substituted for the position of the mean field $(\langle x_t \rangle, \langle y_t \rangle)$. As a consequence, the process from perturbed states back to equilibrium is described by 4×2 nonlinear ordinary differential equations which can be solved with similar regular perturbation analysis in the 2-patch model. We remark that the many-patch model does not have any effective diffusion on the coexistence line; the system has self-attractions in the phase space, described by terms $-(x - \langle x \rangle)$ and $-(y - \langle y \rangle)$ so the final states must be again an equilibrium. A heuristic diagram is shown in Fig. 3.7. In the end we derive drift for $z = \langle x \rangle - \langle y \rangle$, $\bar{v}(z)$, similar to that for the 2-patch model:

$$\bar{v}(z) = \frac{1 - z^2}{K} \left[\bar{C}_0 + \frac{\bar{C}_1}{(\mu_x - \mu_y) z - \bar{C}_2} \right], \tag{3.63}$$

where the constants $\bar{C}_k$ with $k \in \{0, 1, 2\}$ are functions of parameters only:

$$\bar{C}_0 \equiv -\frac{\mu_x - \mu_y}{2} \left[\frac{1}{1 + \mu_x + \mu_y} \right], \tag{3.64a}$$

$$\bar{C}_1 \equiv -(\mu_x - \mu_y) \left[\frac{\rho}{\rho - 1} + \frac{\mu_x \mu_y}{1 + \mu_x + \mu_y} \right], \tag{3.64b}$$

$$\bar{C}_2 \equiv 2\mu_x \mu_y + \mu_x + \mu_y. \tag{3.64c}$$

On the coexistence line, an effective evolution can be formulated as a nonlinear ordinary differential equation:

$$\frac{d}{dt} z(t) = \bar{v}(z), \tag{3.65}$$

In the infinite patch limit the evolution of the system is then conceptually deter-

ministic: starting at any $z = z_0$, at any later time t the site-averaged state $z(t)$ is predictable with probability 1—but we have included the effect of fluctuations. Solutions of (3.65) will be compared to the (exact) numerical simulations of actual many-patch system in section 3.2.4.

3.2.3 Detailed computation of the physically motivated asymptotic analysis

As depicted in the previous section, we approximate the system with only $N \equiv 4$ realizations. Each realization will be released from a specific perturbed state (due to stochastic " kick"). As a consequence, the original field equations (3.58) are approximated by

$$\dot{x}_i = x_i\left(1 - x_i - y_i\right) + \mu_x\left(\frac{1}{N}\sum_{k=1}^{N} x_k - x_i\right), \tag{3.66a}$$

$$\dot{y}_i = y_i\left(1 - x_i - y_i\right) + \mu_y\left(\frac{1}{N}\sum_{k=1}^{N} y_k - y_i\right), \tag{3.66b}$$

with $i \in \{1 \ldots N\}$ and the initial conditions

$$\begin{aligned} &x_1(0) = \Delta_X,\ x_2(0) = -\Delta_X,\ x_3(0) = -\Delta_X,\ x_4(0) = -\Delta_X, \\ &y_1(0) = \Delta_Y,\ y_2(0) = \Delta_Y,\ y_3(0) = -\Delta_Y,\ y_4(0) = -\Delta_Y. \end{aligned} \tag{3.67}$$

The strengths of the noises Δ_X and Δ_Y can be found in Table 3.3. In addition, define the "ensemble average" of a variable A to be

$$\langle A \rangle \equiv \frac{1}{N}\sum_{i=1}^{N} A_i. \tag{3.68}$$

The final goal is to evaluate the deviation of $\langle z(t) \rangle \equiv \langle x(t) - y(t) \rangle$ as $t \to \infty$. We

begin by plugging in the ansatz

$$x_i \equiv x_0 + \epsilon^1 x_i^{(1)} + \epsilon^2 x_i^{(2)} + \mathcal{O}\left(\epsilon^3\right), \tag{3.69a}$$

$$y_i \equiv y_0 + \epsilon^1 y_i^{(1)} + \epsilon^2 y_i^{(2)} + \mathcal{O}\left(\epsilon^3\right), \tag{3.69b}$$

into Eq.(3.66). The linearized equations, i.e. to the order $\mathcal{O}(\epsilon)$, are

$$\dot{x}_i^{(1)} = -x_0\left(x_i^{(1)} + y_i^{(1)}\right) + \mu_x\left(\left\langle x^{(1)}\right\rangle - x_i^{(1)}\right), \tag{3.70a}$$

$$\dot{y}_i^{(1)} = -y_0\left(x_i^{(1)} + y_i^{(1)}\right) + \mu_y\left(\left\langle y^{(1)}\right\rangle - y_i^{(1)}\right), \tag{3.70b}$$

which can be further simplified by noticing at $\mathcal{O}(\epsilon)$, the "total population" $n^{(1)}(t) \equiv \sum_{i=1}^{N}[x_i^{(1)} + y_i^{(1)}]$ satisfies

$$\dot{n}^{(1)}(t) = -n^{(1)}(t). \tag{3.71}$$

Therefore, the solution of the total population exponentially saturates to zero $n^{(1)}(t) = n^{(1)}(0)\exp(-t)$. Since the stochastic perturbation is symmetric (ref: Table 3.3), $n^{(1)}(0) = 0 = n^{(1)}(t)$. In turn, consider $n_x^{(1)} \equiv \sum_{i=1}^{N} x_i^{(1)}$, the "total population of species X", which satisfies

$$\dot{n}_x^{(1)}(t) = -x_0(n^{(1)}(t) = 0. \tag{3.72}$$

The solution of $n_x^{(1)}(t)$ is conserved. With symmetrical initial conditions, we deduce that $n_x^{(1)}(t) = n_x^{(1)}(0) = 0$, which implies $\langle x^{(1)}\rangle = 0$, and the equations of motion at $\mathcal{O}(\epsilon)$ is

$$\dot{x}_i^{(1)} = -x_0\left(x_i^{(1)} + y_i^{(1)}\right) - \mu_x\, x_i^{(1)}, \tag{3.73a}$$

$$\dot{y}_i^{(1)} = -y_0\left(x_i^{(1)} + y_i^{(1)}\right) - \mu_y\, y_i^{(1)}, \tag{3.73b}$$

which have similar functional forms to the linearized equations of population difference of the 2-patch model, i.e. Eq.(3.19). The solutions can be obtained immediately from (3.21) by the following transformations for $i \in 1 \dots N$:

$$w_x \to x_i, \tag{3.74}$$

$$w_y \to y_i, \tag{3.75}$$

$$\mu_x \to \frac{\mu_x}{2}, \tag{3.76}$$

$$\mu_y \to \frac{\mu_y}{2}. \tag{3.77}$$

For the reference of the reader, the solutions of (3.73) are

$$\begin{aligned} x_i^{(1)}(t) = \frac{1}{2\kappa} \Big\{ & \Big[(\kappa - \eta)\, x_i^{(1)}(0) - x_0\, y_i^{(1)}(0) \Big] e^{\lambda_+ t} \\ & + \Big[(\kappa + \eta)\, x_i^{(1)}(0) + x_0\, y_i^{(1)}(0) \Big] e^{\lambda_- t} \Big\} \end{aligned} \tag{3.78a}$$

$$\begin{aligned} y_i^{(1)}(t) = \frac{1}{2\kappa} \Big\{ & \Big[(\kappa + \eta)\, y_i^{(1)}(0) - y_0\, x_i^{(1)}(0) \Big] e^{\lambda_+ t} \\ & + \Big[(\kappa - \eta)\, y_i^{(1)}(0) + y_0\, x_i^{(1)}(0) \Big] e^{\lambda_- t} \Big\} \end{aligned} \tag{3.78b}$$

where

$$\kappa = \frac{1}{2} \sqrt{1 + (\mu_x - \mu_y)^2 + 2\,(\mu_x - \mu_y)\,(x_0 - y_0)}\ , \tag{3.79}$$

$$\eta = \frac{1}{2}\,(\mu_x - \mu_y + x_0 - y_0)\,, \tag{3.80}$$

and the (strictly negative) eigenvalues in the exponents are

$$\lambda_\pm = -\frac{1}{2}\,(1 + \mu_x + \mu_y) \pm \kappa. \tag{3.81}$$

Note that to this order, $\lim_{t\to\infty} x_i^{(1)} = \lim_{t\to\infty} y_i^{(1)} = 0$ and therefore there exists no effective diffusion to $\mathcal{O}(\epsilon^2)$.

For simplicity, we define the following constants (up to the initial conditions)

$$x_i^+ \equiv \frac{1}{2\kappa}\left[(\kappa-\eta)\; x_i^{(1)}(0) - x_0\, y_i^{(1)}(0)\right], \tag{3.82a}$$

$$x_i^- \equiv \frac{1}{2\kappa}\left[(\kappa+\eta)\; x_i^{(1)}(0) + x_0\, y_i^{(1)}(0)\right], \tag{3.82b}$$

$$y_i^+ \equiv \frac{1}{2\kappa}\left[(\kappa+\eta)\; y_i^{(1)}(0) - y_0\, x_i^{(1)}(0)\right], \tag{3.82c}$$

$$y_i^- \equiv \frac{1}{2\kappa}\left[(\kappa-\eta)\; y_i^{(1)}(0) + y_0\, x_i^{(1)}(0)\right], \tag{3.82d}$$

$$n_i^+ \equiv x_i^+ + y_i^+, \tag{3.82e}$$

$$n_i^- \equiv x_i^- + y_i^-, \tag{3.82f}$$

so that

$$x_i^{(1)}(t) \equiv x_i^+\, e^{\lambda_+ t} + x_i^-\, e^{\lambda_- t}, \tag{3.83a}$$

$$y_i^{(1)}(t) \equiv y_i^+\, e^{\lambda_+ t} + y_i^-\, e^{\lambda_- t}, \tag{3.83b}$$

$$x_i^{(1)}(t) + y_i^{(1)}(t) \equiv n_i^+\, e^{\lambda_+ t} + n_i^-\, e^{\lambda_- t}. \tag{3.83c}$$

We move on to $\mathcal{O}(\epsilon^2)$. The equations of motions are

$$\dot{x}_i^{(2)} = -x_0\left(x_i^{(2)} + y_i^{(2)}\right) - x_i^{(1)}\left(x_i^{(1)} + y_i^{(1)}\right) + \mu_x\left(\left\langle x^{(2)}\right\rangle - x_i^{(2)}\right), \tag{3.84a}$$

$$\dot{y}_i^{(2)} = -y_0\left(x_i^{(2)} + y_i^{(2)}\right) - y_i^{(1)}\left(x_i^{(1)} + y_i^{(1)}\right) + \mu_y\left(\left\langle x^{(2)}\right\rangle - y_i^{(2)}\right). \tag{3.84b}$$

Define the total populations

$$n_x^{(2)} \equiv \sum_{i=1}^{N} x_i^{(2)}, \tag{3.85a}$$

$$n_y^{(2)} \equiv \sum_{i=1}^{N} y_i^{(2)}, \tag{3.85b}$$

$$n^{(2)} \equiv n_x^{(2)} + n_y^{(2)}. \tag{3.85c}$$

The evolution of the total population of X and Y follows

$$n^{(2)} = -n^{(2)} - \sum_{i=1}^{N} \left(x_i^{(1)} + y_i^{(1)} \right)^2 = -n^{(2)} - N \left\langle x^{(1)} + y^{(1)} \right\rangle , \tag{3.86}$$

which has solution

$$n^{(2)}(t) = -Ne^{-t} \int_0^t e^{t'} \left\langle \left[x^{(1)}(t') + y^{(1)}(t') \right]^2 \right\rangle dt' \tag{3.87}$$

because $n^{(2)}(0) = 0$. Finally, the equation of motion of $n_x^{(2)}(t)$ is

$$\dot{n}_x^{(2)} = -x_0\, n^{(2)} - \sum_{i=1}^{N} x_i^{(1)} \left(x_i^{(1)} + y_i^{(1)} \right), \tag{3.88}$$

and the solutions can be obtained by integration:

$$\begin{aligned} n_x^{(2)}(t) =& x_0 N \int_0^t \int_0^{t'} \left\langle \left[x^{(1)}(t'') + y^{(1)}(t'') \right]^2 \right\rangle e^{t''} dt'' e^{-t'} dt' \\ & - N \int_0^t \left\langle x^{(1)}(t') \left[x^{(1)}(t') + y^{(1)}(t') \right] \right\rangle dt' \end{aligned} \tag{3.89}$$

or equivalently

$$\begin{aligned} \left\langle x^{(2)}(t) \right\rangle =& x_0 \int_0^t \int_0^{t'} \left\langle \left[x^{(1)}(t'') + y^{(1)}(t'') \right]^2 \right\rangle e^{t''} dt'' e^{-t'} dt' \\ & - \int_0^t \left\langle x^{(1)}(t') \left[x^{(1)}(t') + y^{(1)}(t') \right] \right\rangle dt'. \end{aligned} \tag{3.90}$$

Now we plug (3.83) into the final $\mathcal{O}(\epsilon^2)$ solution (3.89) and apply Lemma III.2 to

obtain

$$\lim_{t\to\infty}\left\langle {n_x}^{(2)}\left(t\right)\right\rangle = \sum_{i=1}^{N}\left\{x_0\left[\frac{\left(n_i^+\right)^2}{-2\lambda_+}+2\,\frac{n_i^+n_i^-}{-\lambda_+-\lambda_-}+\frac{\left(n_i^-\right)^2}{-2\lambda_-}\right]\right. \\ \left. -\left[\frac{x_i^+\,n_i^+}{-2\lambda_+}+\frac{x_i^+n_i^-+x_i^-n_i^+}{-\lambda_+-\lambda_-}+\frac{x_i^-\,n_i^-}{-2\lambda_-}\right]\right\}, \tag{3.91}$$

or equivalently

$$\lim_{t\to\infty}\left\langle x^{(2)}\left(t\right)\right\rangle = \left\langle x_0\left[\frac{\left(n^+\right)^2}{-2\lambda_+}+2\,\frac{n^+n^-}{-\lambda_+-\lambda_-}+\frac{\left(n^-\right)^2}{-2\lambda_-}\right]\right. \\ \left. -\left[\frac{x^+\,n^+}{-2\lambda_+}+\frac{x^+n^-+x^-n^+}{-\lambda_+-\lambda_-}+\frac{x^-\,n^-}{-2\lambda_-}\right]\right\rangle. \tag{3.92}$$

From (3.87) we know that $\lim_{t\to\infty}\langle x^{(2)}(t)+y^{(2)}(t)\rangle = 0$, hence the effective coordinate $\langle z^{(2)}\rangle \equiv \langle x^{(2)}-y^{(2)}\rangle = \langle 2x^{(2)}\rangle$. Finally, inserting the initial conditions (3.67) and taking the average in (3.92), we deduce test

$$\bar{v}_c\left(z_0\right) = \frac{1-z_0^2}{K\kappa^2}\left\{\frac{\left(\mu_x-\mu_y\right)}{4}\frac{z_0^2-2\eta\,z_0-1}{1+\mu_x+\mu_y}\right. \\ +\frac{2\left(1+\mu_x+\mu_y\right)}{4\kappa^2-\left(1+\mu_x+\mu_y\right)^2}\left(\mu_x-\mu_y\right)\left(z_0^2+4\kappa^2+4\eta^2-4\eta z_0-1\right) \\ \left.+\frac{2\kappa^2}{4\kappa^2-\left(1+\mu_x+\mu_y\right)^2}\left(\frac{\rho}{\rho-1}+\frac{\mu_x}{2}+\frac{\mu_y}{2}\right)\left(z_0-2\eta\right)\right\}. \tag{3.93}$$

Further simplifications can be made by plugging (3.79) with $z_0 = x_0 - y_0$, and finally we obtain Eq.(3.63).

3.2.4 Simulations and numerical computations

We performed exact continuous time Markov chain simulations of many-patch competitive systems. The simulation are almost identical to those for the 2-patch problem except that the number of patches N is increased. The birth, death, and hopping rates are exactly the same as in 2-patch model with an additional rule that

when a hopping event occurs the individual lands in any patch with the same probability (excluding the one it is currently on). With different parameter sets, the average populations of each species as function of time were recorded and compared with the corresponding asymptotic predictions, i.e., the numerical integration of (3.65).

Selected parameter sets are presented in Fig. 3.8. In particular, $K = 200$ in these simulations. As shown in Fig. 3.8(a), the simulations confirm that the processes of average populations converge to limit processes as the number of patches N increases. The inset of Fig. 3.8(a) verifies that the total population stays on the coexistence line almost all the time. In the simulations we fixed the number of realization to be $10^4/N$, and it is clear in Fig. 3.8(a) that $N = 1000$ patches produces less noisy data than the other N's with the same "total number of samples" $10^4 = N \times (10^4/N)$. This suggests the many-patch system has an intrinsic averaging effect, which is again a feature of such globally coupled systems.

When $\mu_X, \mu_Y > 0$, we observe that the qualitative behavior of systems with different (μ_X, μ_Y) are similar. One particular set, $\mu_X = 1$ and $\mu_Y = 0.1$, with different initial conditions are presented in Fig. 3.8(b-c) along with the corresponding theoretical predictions. It is clear that the asymptotic analysis produces excellent quantitative predictions. Even when K as small as 100, the quantitative predictions have less than 10% error over the course of time for various parameter sets we have tested.

When the condition $\mu_X, \mu_Y > 0$ is relaxed to $\max(\mu_X, \mu_Y) > 0$, as shown in Fig. 3.8(d), we observe a divergence of the theoretical prediction due to the previously mentioned break-down (non-uniformity) of the asymptotic analysis near the boundaries $z = \pm 1$. Nevertheless, the analysis still provides quantitative predictions accurately until $y(t)$ gets close to 0 (i.e., $z(t) \to -1$) in the large K limit.

To summarize, as long as $\mu_X > \mu_Y$ there is a drift along the coexistence line that persistently favors the X species. In this effective many-patch model it takes infinite amount of time for any species to become extinct due to the fact that there

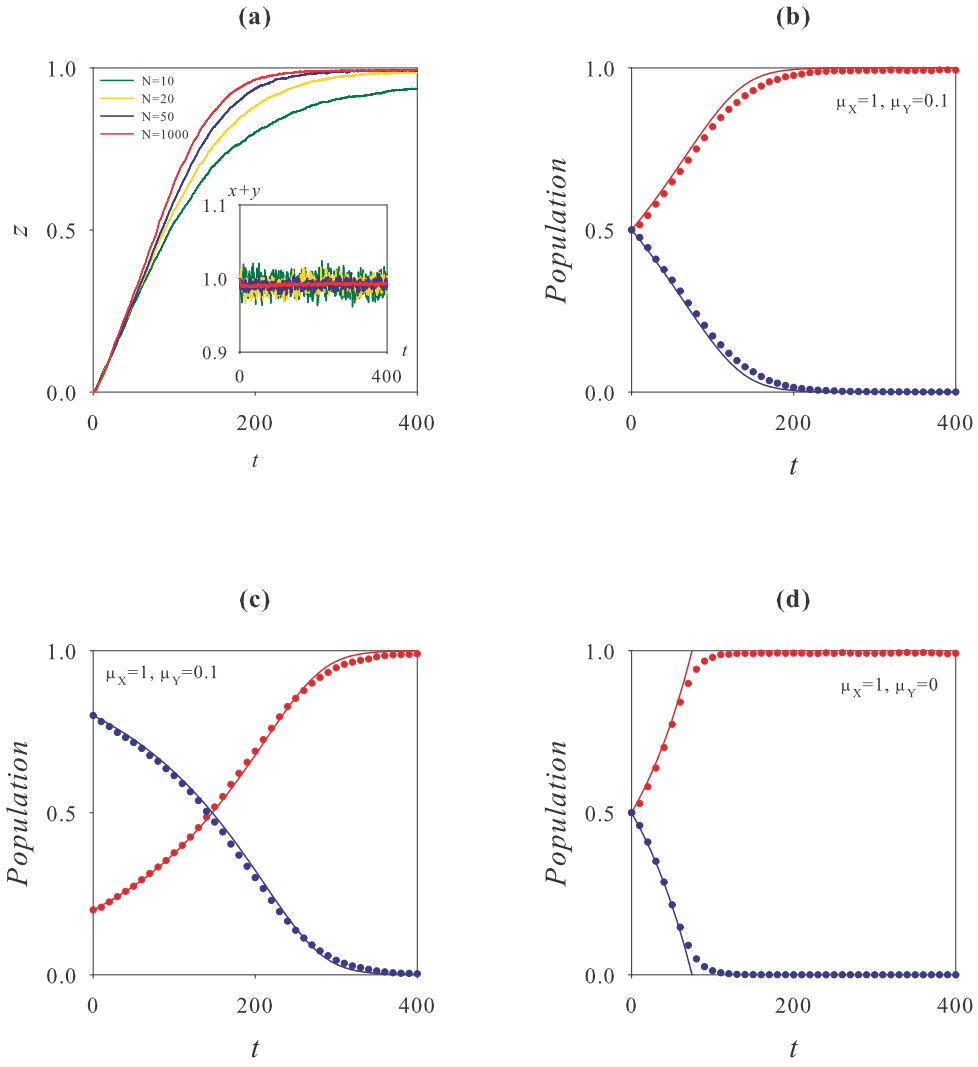

Figure 3.8: (a) Continuous time Markov chain (CTMC) simulation results: position on the coexistence line $z = x - y$ as function of time t for different numbers of patches N. These simulations started with the same number of X and Y individuals (i.e., $z(0) = 0$), $\mu_X = 1$, and $\mu_Y = 0.1$. Inset: total population $x + y$ as function of time. (b) Comparison of CTMC simulations with $N = 1000$ (discrete dots) and the asymptotic prediction (solid curves). Red: population of fast species and blue: population of slow species. $(\mu_X, \mu_Y) = (1, 0.1)$ and the initial condition $z\,(0) = 0$. (c) Similar to (b) with different initial condition $z(0) = 0.6$. (d) The asymptotic theory breaks down when $\mu_Y = 0$ but remains quantitatively predictive before the theoretical Y population vanishes.

are always colonies surviving somewhere among the infinite set of patches available to repopulate locally depleted sites. For a finite but large number of patches N, a species is effectively extinct as soon as its average population drops down to $1/N$, i.e., as soon at there is an expectation of less than one individual remaining anywhere. For finite N, due to demographic fluctuations the slow dispersers will go extinct in a finite order $\mathcal{O}(K)$ time in the large K limit with probability close to 1. This means that there is no finite evolutionarily stable rate for dispersion. It is always better to move more often. There is no evolutionarily stable rate in the infinite-N many-patch model either; the faster a species hops around the greater its evolutionary advantage is.

3.3 Discussion and conclusion

In this Chapter, we have constructed, analyzed, and simulated two multi-patch discrete population competitive dynamics models. Both are deterministically degenerate in the continuum, i.e., infinite carrying capacity K, limit in the sense that the rate equations possess an infinite number of stable coexistence states. Asymmetry of the disperse rates coupled with individual-level fluctuations breaks the degeneracy. In both models, an evolutionary advantage for the faster disperser emerges on an $\mathcal{O}(K)$ time scale. The preference originates in the interactions between demographic fluctuations and the nonlinearity of the deterministic dynamics.

In the 2-patch model, the probability of the fast species winning is enhanced by at most 25% throughout the phase space so the slow species still has positive probability to win the competition with positive initial population. On the other hand, in the globally coupled many-patch model the fast dispersers always out-compete the slower dispersers.

We emphasize that in both these spatially homogeneous systems, demographic fluctuations enhance the survival probability of the *faster* dispersers. This is notable

because for deterministic dynamics in heterogeneous environments Hastings [19] and Dockery *et al.* [8] have shown that the *slower* dispersers typically have the evolutionary advantage. The level of demographic fluctuations, depending also on the degree of environmental variations, determines whether faster or slower dispersion is favored. We will generalize the models in this Chapter to include spatial heterogeneity in Chapter V.

CHAPTER IV

Nonlinear Dynamics of Heterogeneous Patchy Models

In this Chapter we take a short digression to investigate *deterministic* and *nonlinear* competitive population dynamics models. When the spatial resource is heterogeneously distributed, Hastings [19], Holt [20], and Dockery *et al.* [8] showed that the species with low mobility (with passive diffusion) is vulnerable to the invasion of species with high mobility in both patchy-like (in Holt [20]) and continuous-space (in Hastings [19] and Dockery *et al.* [8]) models. Nevertheless, their analyses did not provide dynamical insights, i.e., the time scale, the strength, and the mechanism of the selection. The motivation of this Chapter is to investigate the deterministic, nonlinear, and patch-like models with dynamical approaches. With a novel asymptotic expansion with respect to the normalized environmental variance σ^2, we deduce that the slow species has advantage in a quantitatively identified time scale $\mathcal{O}(\sigma^{-2})$. The deterministic time scale $\mathcal{O}(\sigma^{-2})$ along with the time scale $\mathcal{O}(K)$ identified in Chapter III predict the strengths of two competing effects (slow or fast species having the advantage), and they will be matched in the following Chapter V when the most general stochastic population dynamics with competition is considered.

This Chapter is organized into three parts. In section 4.1, we perform both analytical and numerical analyses on the deterministic two-patch model. In section 4.2,

we perform parallel analysis on the spatially extended many patch model. In section 4.3 the conclusions and a discussion are presented.

4.1 The deterministic two-patch model

The space consists of two patches, labeled by 1 and 2, with distinct carrying capacities. Each patch is a well-mixed pool. Two species, X and Y, compete for limited resource on the patches. In order to explore only the effect of spatial dispersions on evolutionary advantages, we assume X and Y have identical demographic dynamics. With proper scaling of time and population scales ref: Chapter II, the model becomes:

$$\dot{x}_1 = x_1\left[1-(1+\sigma)(x_1+y_1)\right]+\mu_x(x_2-x_1), \tag{4.1a}$$

$$\dot{y}_1 = y_1\left[1-(1+\sigma)(x_1+y_1)\right]+\mu_y(y_2-y_1), \tag{4.1b}$$

$$\dot{x}_2 = x_2\left[1-(1-\sigma)(x_2+y_2)\right]+\mu_x(x_1-x_2), \tag{4.1c}$$

$$\dot{y}_2 = y_2\left[1-(1-\sigma)(x_2+y_2)\right]+\mu_y(y_1-y_2). \tag{4.1d}$$

$x_i, y_i \geq 0$ for $i \in 1,2$ are the continuous population variables of species X and Y respectively on patch i. Parameters $\mu_x, \mu_y > 0$ are respectively the species' (scaled) symmetric hopping rates between patches. The scale of the populations is normalized by the harmonic mean of the carrying capacities, so patch 1 and 2 have carrying capacity $(1-\sigma)^{-1}$ and $(1+\sigma)^{-1}$ respectively. Notably, the parameter $\sigma > 0$ characterizes the standard deviation of the inhomogeneous resource distribution (scaled by the harmonic mean of the environmental distribution).

The complexity of the dynamics comes from the following facts: (1) Two species interact via competitions, i.e. terms like $(1+x_i+y_i)/(1\pm\sigma)$. (2) The populations among the patches are coupled via dispersion, i.e. terms like $\mu_x(x_i-x_j)$ and $\mu_y(y_i-y_j)$. As a result, the whole 4-dimentional phase space is coupled in a nonlinear manner. To

our knowledge, there exists no closed-form solution of (4.1). Therefore, we perform perturbation analysis to solve the problem with the expansion of small environmental variation, assuming $\sigma \ll 1$. Take the ansatz that the solution takes the form with constants x_0 and y_0:

$$x_1(t) = x_0 + \sigma^1 x_1^{(1)}(t) + \sigma^2 x_1^{(2)}(t) + \mathcal{O}(\sigma^3), \tag{4.2a}$$

$$x_2(t) = x_0 + \sigma^1 x_2^{(1)}(t) + \sigma^2 x_2^{(2)}(t) + \mathcal{O}(\sigma^3), \tag{4.2b}$$

$$y_1(t) = y_0 + \sigma^1 y_1^{(1)}(t) + \sigma^2 y_1^{(2)}(t) + \mathcal{O}(\sigma^3), \tag{4.2c}$$

$$y_2(t) = y_0 + \sigma^1 y_2^{(1)}(t) + \sigma^2 y_2^{(2)}(t) + \mathcal{O}(\sigma^3), \tag{4.2d}$$

It is convenient to define total populations $n_x^{(k)}$ and $n_y^{(k)}$ among patches in various orders $k = 1, 2, 3 \ldots$

$$n_x^{(k)} := x_1^{(k)} + x_2^{(k)}, \tag{4.3a}$$

$$n_y^{(k)} := y_1^{(k)} + y_2^{(k)}, \tag{4.3b}$$

and the difference of the populations $w_x^{(k)}$ and $w_y^{(k)}$ among patches

$$w_x^{(k)} := x_1^{(k)} - x_2^{(k)}, \tag{4.4a}$$

$$w_y^{(k)} := y_1^{(k)} - y_2^{(k)}, \tag{4.4b}$$

We plug the ansatz (4.2) into the equations of motion (4.1), and then perform the regular asymptotic analysis by the orders of σ. The lowest order $\mathcal{O}(\sigma^0)$ does not involve σ, and the dynamics is identical to the degenerate dynamics of the 2-patch model in homogeneous space (section 3.1.3):

$$x_0 + y_0 = 1. \tag{4.5}$$

At the next order $\mathcal{O}(\sigma^1)$ the dynamics reads

$$\dot{x}_1^{(1)} = -x_0\left(1 + x_1^{(1)} + y_1^{(1)}\right) - \mu_x\left(x_1^{(1)} - x_2^{(1)}\right), \tag{4.6a}$$

$$\dot{y}_1^{(1)} = -y_0\left(1 + x_1^{(1)} + y_1^{(1)}\right) - \mu_y\left(y_1^{(1)} - y_2^{(1)}\right), \tag{4.6b}$$

$$\dot{x}_2^{(1)} = -x_0\left(-1 + x_2^{(1)} + y_2^{(1)}\right) - \mu_x\left(x_2^{(1)} - x_1^{(1)}\right), \tag{4.6c}$$

$$\dot{y}_2^{(1)} = -y_0\left(-1 + x_2^{(1)} + y_2^{(1)}\right) - \mu_y\left(y_2^{(1)} - y_1^{(1)}\right). \tag{4.6d}$$

After the transformation to the variables (n, w), the dynamics is decoupled:

$$\dot{n}_x^{(1)} = -x_0\left(n_x^{(1)} + n_y^{(1)}\right), \tag{4.7a}$$

$$\dot{n}_y^{(1)} = -y_0\left(n_x^{(1)} + n_y^{(1)}\right), \tag{4.7b}$$

and

$$\dot{w}_x^{(1)} = -x_0\left(2 + w_x^{(1)} + w_y^{(1)}\right) - 2\mu_x w_x^{(1)}, \tag{4.8a}$$

$$\dot{w}_y^{(1)} = -y_0\left(2 + w_x^{(1)} + w_y^{(1)}\right) - 2\mu_x w_y^{(1)}. \tag{4.8b}$$

It is clear that (4.7) is identical to its counterparts (3.18) in the homogeneous 2-patch model. In addition, it is elementary to show that the fixed point of (4.8) is

$$w_x^* := -\frac{2x_0\mu_y}{y_0\mu_x + x_0\mu_y + 2\mu_x\mu_y}, \tag{4.9a}$$

$$w_y^* := -\frac{2y_0\mu_x}{x_0\mu_x + y_0\mu_y + 2\mu_x\mu_y}. \tag{4.9b}$$

Next, we linearize $w_x^{(1)}$ and $w_y^{(1)}$ with respect to the fix point (4.9),

$$\tilde{w}_x^{(1)}(t) \equiv w_x^{(1)}(t) - w_x^*, \tag{4.10a}$$

$$\tilde{w}_y^{(1)}(t) \equiv w_y^{(1)}(t) - w_y^*, \tag{4.10b}$$

and deduce the equations of motion of the linearized fields, $\tilde{w}_x^{(1)}$ and $\tilde{w}_y^{(1)}$, to be

$$\dot{\tilde{w}}_x^{(1)} = -x_0 \left(\tilde{w}_x^{(1)} + \tilde{w}_y^{(1)}\right) - 2\mu_x \tilde{w}_x^{(1)} \tag{4.11a}$$

$$\dot{\tilde{w}}_y^{(1)} = -y_0 \left(\tilde{w}_x^{(1)} + \tilde{w}_y^{(1)}\right) - 2\mu_y \tilde{w}_y^{(1)}. \tag{4.11b}$$

which is identical to their counterpart (3.19) in the homogeneous 2-patch model.

We will be interested in the case[1] when the initial conditions are of order $\mathcal{O}(\sigma^1)$, therefore the solutions depend on the initial conditions. Let the initial conditions to be

$$n_x^{(1)}(0) = x_1^{(1)}(0) + x_2^{(1)}(0), \tag{4.12a}$$

$$n_y^{(1)}(0) = y_1^{(1)}(0) + y_2^{(1)}(0), \tag{4.12b}$$

$$w_x^{(1)}(0) = x_1^{(1)}(0) - x_2^{(1)}(0), \tag{4.12c}$$

$$w_y^{(1)}(0) = y_1^{(1)}(0) - y_2^{(1)}(0). \tag{4.12d}$$

The analytical results are quoted from section 3.1.3: the solutions of the total populations are

$$n_x^{(1)}(t) = y_0\, n_x^{(1)}(0) - x_0\, n_y^{(1)}(0) + x_0 \left(n_x^{(1)}(0) + n_y^{(1)}(0)\right) e^{-t} \tag{4.13a}$$

$$n_y^{(1)}(t) = x_0\, n_y^{(1)}(0) - y_0\, n_x^{(1)}(0) + y_0 \left(n_x^{(1)}(0) + n_y^{(1)}(0)\right) e^{-t} \tag{4.13b}$$

and the differences of the populations are

$$w_x^{(1)}(t) = w_x^* + \tilde{w}_x = \frac{-2x_0\mu_y}{y_0\mu_x + x_0\mu_y + 2\mu_x\mu_y} + \tilde{w}_x(t; w_{x0}, w_{y0}), \tag{4.14a}$$

$$w_y^{(1)}(t) = w_y^* + \tilde{w}_y = \frac{-2y_0\mu_x}{y_0\mu_x + x_0\mu_y + 2\mu_x\mu_y} + \tilde{w}_y(t; w_{x0}, w_{y0}), \tag{4.14b}$$

[1]We only consider the initial conditions with order $\mathcal{O}(\sigma^1)$ because in the stochastic models, the demographic fluctuations is of the order $\mathcal{O}(1/\sqrt{K})$ and we are interested in the cased when these two effects are comparable; see Chapter V.

where the transient terms are

$$\tilde{w}_x^{(1)}(t) = \frac{1}{2\kappa}\left\{\left[(\kappa-\eta)\,\tilde{w}_x^{(1)}(0) - x_0\,\tilde{w}_y^{(1)}(0)\right]e^{\lambda_+ t} + \left[(\kappa+\eta)\,\tilde{w}_x^{(1)}(0) + x_0\,\tilde{w}_y^{(1)}(0)\right]e^{\lambda_- t}\right\} \tag{4.15a}$$

$$\tilde{w}_y^{(1)}(t) = \frac{1}{2\kappa}\left\{\left[(\kappa+\eta)\,\tilde{w}_y^{(1)}(0) - y_0\,\tilde{w}_x^{(1)}(0)\right]e^{\lambda_+ t} + \left[(\kappa-\eta)\,\tilde{w}_y^{(1)}(0) + y_0\,\tilde{w}_x^{(1)}(0)\right]e^{\lambda_- t}\right\}. \tag{4.15b}$$

For the convenience of the readers we reproduce the parameters κ and η:

$$\kappa = \frac{1}{2}\sqrt{1 + 4\left(\mu_x - \mu_y\right)^2 + 4\left(\mu_x - \mu_y\right)\left(x_0 - y_0\right)},$$
$$\eta = \mu_x - \mu_y + \frac{1}{2}\left(x_0 - y_0\right),$$

and the eigenvalues $\lambda_\pm$:

$$\lambda_\pm = -\frac{1}{2}\left(1 + 2\mu_x + 2\mu_y\right) \pm \kappa.$$

At the order $\mathcal{O}(\sigma^2)$, the equations of motions are

$$\dot{x}_1^{(2)} = x_0\left(-x_1^{(2)} - y_1^{(2)}\right) + \mu_x\left(x_2^{(2)} - x_1^{(2)}\right) + x_1^{(1)}\left(-x_1^{(1)} - y_1^{(1)}\right) - x_0\left(x_1^{(1)} + y_1^{(1)}\right) - x_1^{(1)}, \tag{4.16a}$$

$$\dot{x}_2^{(2)} = x_0\left(-x_2^{(2)} - y_2^{(2)}\right) + \mu_x\left(x_1^{(2)} - x_2^{(2)}\right) + x_2^{(1)}\left(-x_2^{(1)} - y_2^{(1)}\right) + x_0\left(x_2^{(1)} + y_2^{(1)}\right) + x_2^{(1)}, \tag{4.16b}$$

$$\dot{y}_1^{(2)} = y_0\left(-x_1^{(2)} - y_1^{(2)}\right) + \mu_y\left(y_2^{(2)} - y_1^{(2)}\right) + y_1^{(1)}\left(-x_1^{(1)} - y_1^{(1)}\right) - y_0\left(x_1^{(1)} + y_1^{(1)}\right) - y_1^{(1)}, \tag{4.16c}$$

$$\dot{y}_2^{(2)} = y_0\left(-x_2^{(2)} - y_2^{(2)}\right) + \mu_y\left(y_1^{(2)} - y_2^{(2)}\right) + y_2^{(1)}\left(-x_2^{(1)} - y_2^{(1)}\right) + y_0\left(x_2^{(1)} + y_2^{(1)}\right) + y_2^{(1)}. \tag{4.16d}$$

A parallel computation to section 3.1.3 shows that the dynamics of the total population of X, i.e., $n_x \equiv x_1 + x_2$ is described by

$$\begin{aligned} \dot{n}_x^{(2)} = & -x_0 n^{(2)} - x_1^{(1)}\left(x_1^{(1)} + y_1^{(1)}\right) - x_{2'}^{(1)}\left(x_2^{(1)} + y_2^{(1)}\right) \\ & - x_0\left(w_x^{(1)} + w_y^{(1)}\right) - w_x^{(1)}, \end{aligned} \tag{4.17}$$

where the total population of X and Y, i.e., $n^{(2)} \equiv x_1 + x_2 + y_1 + y_2$ satisfies

$$\dot{n}^{(2)} = -n^{(2)} - \frac{1}{2}\left[\left(n_x^{(1)} + n_y^{(1)}\right)^2 + \left(w_x^{(1)} + w_y^{(1)}\right)^2\right] - 2\left(w_x^{(1)} + w_y^{(1)}\right). \tag{4.18}$$

Since the initial conditions for the $\mathcal{O}(\sigma^2)$-variables are zero, the solution of (4.18) is

$$\begin{aligned} n^{(2)}(t) = & -\frac{e^{-t}}{2}\int_0^t e^{t'}\left[\left(n_x^{(1)}(t') + n_y^{(1)}(t')\right)^2 + \left(w_x^{(1)}(t') + w_y^{(1)}(t')\right)^2\right] dt' \\ & - 2e^{-t}\int_0^t e^{t'}\left(w_x^{(1)} + w_y^{(1)}\right) dt', \end{aligned} \tag{4.19}$$

and the solution of (4.17) is

$$
\begin{aligned}
(4.20)\qquad n_x^{(2)}(t) = {} & + n_x^{(2)}(0) + x_0 \left[n^{(2)}(0)\right] t \\
& - 2x_0 \int_0^t \int_0^{t'} \left(w_x^{(1)}(t'') + w_y^{(1)}(t'')\right) e^{t''} dt'' e^{-t'} dt' \\
& - \int_0^t \left[x_0 \left(w_x^{(1)}(t') + w_y^{(1)}(t')\right) + w_x^{(1)}(t')\right] dt' \\
& x_0 \int_0^t \int_0^{t'} \Big[\left(n_x^{(1)}(t'') + n_y^{(1)}(t'')\right)^2 \\
& \qquad + \left(w_x^{(1)}(t'') + w_y^{(1)}(t'')\right)^2\Big] e^{t''} dt'' e^{-t'} dt' \\
& - \frac{1}{2} \int_0^t \Big[{n_x^{(1)}}^2(t') + {w_x^{(1)}}^2(t') \\
& \qquad + n_x^{(1)}(t') n_y^{(1)}(t') + w_x^{(1)}(t') w_y^{(1)}(t')\Big] dt'
\end{aligned}
$$

Note that the finite separation of the populations between patches, i.e. first terms in (4.14), contributes an effective drift after the transient parts converge to zero in a time scale $\mathcal{O}(\log \sigma^{-1})$:

$$
(4.21)\qquad \lim_{t \to t_0 \gg 1} n_x^{(2)}(t) = -\frac{4\mu_x \mu_y (\mu_x - \mu_y) x_0 y_0}{(y_0 \mu_x + x_0 \mu_y + 2\mu_x \mu_y)^2} t_0.
$$

We then identify an effective drift $v_{\text{eff}}(x_0, y_0)$ to be the rate of change of the total populations of species X in this order

$$
(4.22)\qquad v_{\text{eff}}(x_0, y_0) = -\frac{4\mu_x \mu_y (\mu_x - \mu_y) x_0 y_0}{(y_0 \mu_x + x_0 \mu_y + 2\mu_x \mu_y)^2}.
$$

Define an effective coordinate, the difference of the average population per patch $z = (x_1 + x_2 - y_1 - y_2)/2$ to order $\mathcal{O}(\sigma^2)$. Then the effective dynamics to the order

$\mathcal{O}(\sigma^1)$ can then be formulated as

$$\frac{dz(t)}{dt} = -\frac{4\mu_x\mu_y(\mu_x-\mu_y)(1-z^2)}{\left[\mu_x+\mu_y-(\mu_x-\mu_y)z+4\mu_x\mu_y\right]^2}\sigma^2. \tag{4.23}$$

This analysis provides us with a detailed picture of the system. In the phase space, starting from any initial conditions far away from the center manifold L: $(x_1, x_2, y_1, y_2) = (x_0 + \sigma w_x^*/2, x_0 - \sigma w_x^*/2, y_0 + \sigma w_y^*/2, y_0 - \sigma w_y^*/2)$ with $x_0, y_0 \in (0, 1)$ and $x_0 + y_0 = 1$, the state converges to the σ^2-neighborhood of L in a $\log(\sigma^{-1})$ time scale. States on L are "metastable" in the sense that they are stable in a time scale of $\mathcal{O}(\sigma^{-1})$. The dispersion rates, μ_x and μ_y, serve a sort of the chemical potential for the system. When the hopping rates μ_x and μ_y are zero, there is no coupling among patches and the total populations on each patch ($x_1 + y_1$ and $x_2 + y_2$) converge to the carrying capacity of the patch. On the other hand, as $\mu_x, \mu_y \to \infty$, the system mixes the population more and more efficiently, and the populations (of X and Y) on each patch are identical ($x_1 = x_2$ and $y_1 = y_2$) in the limit. We aim to study the nontrivial case $0 < \mu_x, \mu_y < \infty$, where the metastable population distributions fail to match to the resource distribution. As a consequence, at the order $\mathcal{O}(\sigma^2)$ there exists a slow drift (4.22) near the center manifold L. Since the strength of the effective drift is of $\mathcal{O}(\sigma^2)$, the time scale of the dynamics along L is of order $\mathcal{O}(\sigma^{-2})$.

Direct numerical verifications, performed by integrating the equations of motion (4.1), are presented in Fig.(4.1).

4.2 The deterministic many patch model

In this section we generalize the 2-patch model in section 4.1 to a globally connected model with countably infinite patches. The space in the model consists of countably infinite patches. Denote $i \in \mathbb{N}$ to be the patch index. Each patch has constant carrying capacity $(1 + \sigma_i)^{-1}$, where σ_i's are independent and identically dis-

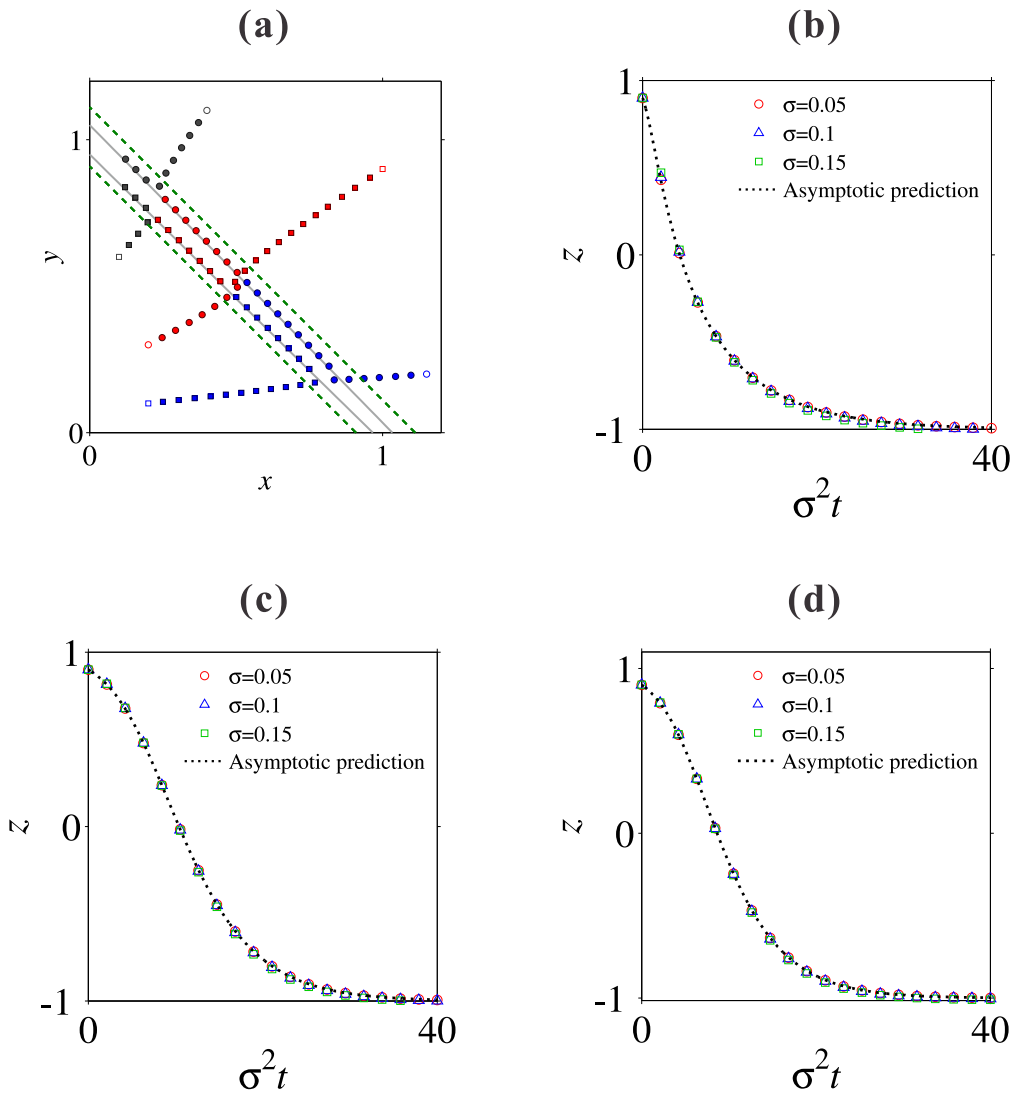

Figure 4.1: (a) Three trajectories of the deterministic 2-patch model with different initial conditions are labeled by grey, red and blue markers. Circle and square markers represent populations on patch 1 and patch 2 respectively. Open markers represent the initial configurations. The dashed green lines are $x + y = (1 \pm \sigma)^{-1}$, which are the solutions when the patches are isolated ($\mu_x = \mu_y = 0$). Pale grey lines represent the center manifold derived from the asymptotic analysis. $\mu_x = 1$, $\mu_y = 0.5, \sigma = 0.1$. (b-d) Direct numerical simulations with $\sigma = 0.05$, 0.1, and 0.15 (discrete markers) and the prediction of the asymptotic analysis (dotted line). (b) $\mu_x = 1$, $\mu_y = 0.1$. (c) $\mu_x = 5$, $\mu_y = 1$. (d) $\mu_x = 10$, $\mu_y = 1$.

tributed (i.i.d.) random variables which characterize the spatial heterogeneity. We assume all the moments of σ_i to exist. Similar to the 2-Patch model, the populations are normalized by the harmonic mean of the carrying capacities, so the first moments of σ_i vanishes:

$$\langle \sigma_i \rangle \equiv \mathbb{E}[\sigma_i] := \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N} \sigma_i = 0. \tag{4.24}$$

Hence, the second moment is the variance of the random variable σ_i:

$$\mathrm{Var}(\sigma_i) \equiv \left\langle \sigma_i^2 \right\rangle \equiv \mathbb{E}\left[\sigma_i^2\right] := \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N} \sigma_i^2. \tag{4.25}$$

We also assume the magnitude of σ_i's are small compared to identity: $|\sigma_i| \ll 1$, although the results are not restricted by such limitation[2]. In the following analysis, we adopt the notation that for any observable A_i on patch $i \in \mathbb{N}$, the first moment is defined as

$$\langle A \rangle := \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N} A_i. \tag{4.26}$$

Species X and Y live on the patches and have identical demographic dynamics. The model is globally connected in the sense that respectively to its species, each individual hops with rates μ_x or $\mu_y > 0$ to each patch with equal probability. Therefore, the dynamics of the system is

$$\dot{x}_i = x_i \left[1 - (1+\sigma_i)(x_i + y_i)\right] + \mu_x(\langle x \rangle - x_i), \tag{4.27a}$$

$$\dot{y}_i = y_i \left[1 - (1+\sigma_i)(x_i + y_i)\right] + \mu_y(\langle y \rangle - y_i), \tag{4.27b}$$

$$i \in \mathbb{N}. \tag{4.27c}$$

[2]The claim is verified by numerical observations in Fig. 4.2.

The model represents a spatial case of Gadgil's globally connected model [13]. In the following paragraph, we apply the intuition from section 4.1—first solve for the metastable distribution at the order $\mathcal{O}(\sigma)$, and then plug it into the $\mathcal{O}(\sigma^2)$ to obtain the effective drift along the center manifold.

Parallel to the analysis in section 4.1, we start with plugging in the ansatz

$$x_i = x_0 + x_i^{(1)} + x_i^{(2)} + \mathcal{O}(\sigma_i^3), \tag{4.28a}$$

$$y_i = y_0 + y_i^{(1)} + y_i^{(2)} + \mathcal{O}(\sigma_i^3), \tag{4.28b}$$

into Eqs.(4.27). Note that in the expansion, distinct from the previous analysis, we assume $x_i^{(j)}$ and $y_i^{(j)}$ are of order $\mathcal{O}(\sigma_i^j)$.

At the order $\mathcal{O}(\sigma_i^0)$, the constraint is

$$x_0 + y_0 = 1. \tag{4.29}$$

We are interested in the nontrivial domain $0 < x_0, y_0 < 1$.

At the order $\mathcal{O}(\sigma_i^1)$, the dynamics are

$$\dot{x}_i^{(1)} = x_0 \left(-\sigma_i - x_i^{(1)} - y_i^{(1)}\right) + \mu_x \left(\left\langle x^{(1)} \right\rangle - x_i^{(1)}\right), \tag{4.30a}$$

$$\dot{y}_i^{(1)} = y_0 \left(-\sigma_i - x_i^{(1)} - y_i^{(1)}\right) + \mu_y \left(\left\langle y^{(1)} \right\rangle - y_i^{(1)}\right). \tag{4.30b}$$

We now solve for the stable distribution of 4.30 by applying standard technique to solve the implicit mean-field equations. Let $\dot{x}_i = \dot{y}_i = 0$, and express x_i and y_i in terms of $\left\langle x^{(1)} \right\rangle$ and $\left\langle y^{(1)} \right\rangle$:

$$x_i^{(1)} = \frac{-x_0\,\mu_y\,\sigma_i + (\mu_x\,\mu_y + \mu_x\,y_0)\,\left\langle x^{(1)} \right\rangle - x_0\,\mu_y\,\left\langle y^{(1)} \right\rangle}{\mu_x\,\mu_y + \mu_x\,y_0 + \mu_y\,x_0}, \tag{4.31a}$$

$$y_i^{(1)} = \frac{-y_0\,\mu_x\,\sigma_i + (\mu_x\,\mu_y + \mu_y\,x_0)\,\left\langle y^{(1)} \right\rangle - y_0\,\mu_x\,\left\langle x^{(1)} \right\rangle}{\mu_x\,\mu_y + \mu_x\,y_0 + \mu_y\,x_0}. \tag{4.31b}$$

Take the average $\langle\cdot\rangle$ over the patch index i, and notice $\langle\sigma_i\rangle = 0$ by construction, we arrive at

$$\left\langle x^{(1)}\right\rangle = \frac{\left(\mu_x\,\mu_y + \mu_x\,y_0\right)\,\left\langle x^{(1)}\right\rangle - x_0\,\mu_y\,\left\langle y^{(1)}\right\rangle}{\mu_x\,\mu_y + \mu_x\,y_0 + \mu_y\,x_0}, \tag{4.32a}$$

$$\left\langle y^{(1)}\right\rangle = \frac{\left(\mu_x\,\mu_y + \mu_y\,x_0\right)\,\left\langle y^{(1)}\right\rangle - y_0\,\mu_x\,\left\langle x^{(1)}\right\rangle}{\mu_x\,\mu_y + \mu_x\,y_0 + \mu_y\,x_0}. \tag{4.32b}$$

and then we deduce

$$\mu_y\,x_0\left[\left\langle x^{(1)}\right\rangle + \left\langle y^{(1)}\right\rangle\right] \equiv 0, \tag{4.33a}$$

$$\mu_x\,y_0\left[\left\langle x^{(1)}\right\rangle + \left\langle y^{(1)}\right\rangle\right] \equiv 0. \tag{4.33b}$$

Since we assume μ_y, μ_x, x_0 and $y_0 \neq 0$ in general, we have

$$\left\langle x^{(1)}\right\rangle + \left\langle y^{(1)}\right\rangle = 0 \tag{4.34}$$

and the $\mathcal{O}(\sigma)$ solutions

$$x_i^{(1)} = -\sigma_i \frac{x_0\,\mu_y}{\mu_x\,\mu_y + \mu_x\,y_0 + \mu_y\,x_0} + \left\langle x^{(1)}\right\rangle, \tag{4.35a}$$

$$y_i^{(1)} = -\sigma_i \frac{y_0\,\mu_x}{\mu_x\,\mu_y + \mu_x\,y_0 + \mu_y\,x_0} + \left\langle y^{(1)}\right\rangle. \tag{4.35b}$$

In addition, note that the dynamics preserve $\left\langle x^{(1)}\right\rangle - \left\langle y^{(1)}\right\rangle$:

$$\begin{aligned}\frac{d}{dt}\left\langle x^{(1)} - y^{(1)}\right\rangle &= \left\langle x_0\left(-\sigma_i - x^{(1)} - y^{(1)}\right) - y_0\left(-\sigma_i - x^{(1)} - y^{(1)}\right)\right\rangle \\ &= \left\langle x_0\left(-x^{(1)} - y^{(1)}\right) - y_0\left(-x^{(1)} - y^{(1)}\right)\right\rangle = 0.\end{aligned} \tag{4.36}$$

Therefore, at the order $\mathcal{O}(\sigma)$, we have conserved $\left\langle x^{(1)}\right\rangle - \left\langle y^{(1)}\right\rangle$ for any time $t \geq 0$. The physical motivated asymptotic analysis (see Chapter V) will require $\left\langle x^{(1)}(0)\right\rangle - \left\langle y^{(1)}(0)\right\rangle = 0$, which in turn indicates $\left\langle x^{(1)}\right\rangle = 0 = \left\langle y^{(1)}\right\rangle$, and the metastable

distribution

$$x_i^{(1)} = -\sigma_i \frac{x_0\,\mu_y}{\mu_x\,\mu_y + \mu_x\,y_0 + \mu_y\,x_0}, \tag{4.37a}$$

$$y_i^{(1)} = -\sigma_i \frac{y_0\,\mu_x}{\mu_x\,\mu_y + \mu_x\,y_0 + \mu_y\,x_0}. \tag{4.37b}$$

At the second order $\mathcal{O}(\sigma_i^2)$, the equations of motions are

$$\dot{x}_i^{(2)} = x_0\left(-n_i^{(2)} - \sigma_i\,n_i^{(1)}\right) + x_i^{(1)}\left(-\sigma_i - n_i^{(1)}\right) + \mu_x\left(\left\langle x_i^{(2)}\right\rangle - x_i^{(2)}\right), \tag{4.38a}$$

$$\dot{y}_i^{(2)} = y_0\left(-n_i^{(2)} - \sigma_i\,n_i^{(1)}\right) + y_i^{(1)}\left(-\sigma_i - n_i^{(1)}\right) + \mu_y\left(\left\langle y_i^{(2)}\right\rangle - y_i^{(2)}\right). \tag{4.38b}$$

and the equation of motion of the average total population $\left\langle n_i^{(2)}\right\rangle \equiv \left\langle x_i^{(2)} + y_i^{(2)}\right\rangle$ is

$$\frac{d}{dt}\left\langle n^{(2)}\right\rangle = -\left\langle n^{(2)}\right\rangle - 2\left\langle \sigma n^{(1)}\right\rangle - \left\langle \left(n^{(1)}\right)^2\right\rangle, \tag{4.39}$$

With the initial condition $\left\langle n^{(2)}\,(0)\right\rangle = 0$, the solution can be obtained

$$\left\langle n^{(2)}\right\rangle \equiv -e^{-t}\int_0^t e^{t'}\left(2\left\langle \sigma n^{(1)}\right\rangle + \left\langle \left(n^{(1)}\right)^2\right\rangle\right)dt'. \tag{4.40}$$

Ignoring the transient parts, $\left\langle n^{(2)}\,(t)\right\rangle$ is then identified as

$$\lim_{t\to t_0\gg 1}\left\langle n^{(2)}\,(t)\right\rangle \to -\lim_{t\to t_0\gg 1} 2\left\langle \sigma n^{(1)}\,(t)\right\rangle + \left\langle \left(n^{(1)}\,(t)\right)^2\right\rangle. \tag{4.41}$$

Therefore, for a sufficiently long time (so long as $\mathcal{O}\,(t) > \mathcal{O}\log\left(\sigma^{-1}\right)$), the dynamics

of the average population difference $\left\langle z^{(2)}\right\rangle \equiv \left\langle x^{(2)} - y^{(2)}\right\rangle$ converges to

$$\begin{aligned}\frac{d}{dt}\left\langle z^{(2)}\right\rangle = &+\left\langle y_0\left(n^{(2)}+\sigma n^{(1)}\right) - y^{(1)}\left(\sigma + n^{(1)}\right)\right\rangle \\ &-\left\langle x_0\left(n^{(2)}+\sigma n^{(1)}\right) + x^{(1)}\left(\sigma + n^{(1)}\right)\right\rangle \\ &\to -2\left\langle\left(\sigma + n^{(1)}\right)\left(y_0\, x^{(1)} - \, x_0\, y^{(1)}\right)\right\rangle\end{aligned} \tag{4.42}$$

To compute these averages, we exploit the facts

$$\sigma_i + n_i^{(1)} = \sigma_i \frac{2\mu_x\,\mu_y}{2\mu_x\,\mu_y + \mu_x\, y_0 + \mu_y\, x_0}, \tag{4.43a}$$

$$y_0\, x_i^{(1)} = -\sigma_i \frac{y_0\, x_0\, \mu_y}{2\mu_x\,\mu_y + \mu_x\, y_0 + \mu_y\, x_0}, \tag{4.43b}$$

$$x_0\, y_i^{(1)} = -\sigma_i \frac{x_0\, y_0\, \mu_x}{2\mu_x\,\mu_y + \mu_x\, y_0 + \mu_y\, x_0}. \tag{4.43c}$$

Finally we deduce $\langle z\rangle$ has nontrivial dynamics at the order $\mathcal{O}\left(\sigma^2\right)$:

$$\lim_{t\to t_0\gg 1}\left\langle z^{(2)}(t)\right\rangle = -2\frac{x_0 y_0 \mu_x \mu_y(\mu_x-\mu_y)}{(y_0\mu_x + x_0\mu_y + \mu_x\mu_y)^2}\left\langle\sigma^2\right\rangle t_0, \tag{4.44}$$

hence the effective evolution of $\langle z\rangle$ is identified as

$$\frac{d\langle z\rangle}{dt} = -\frac{2\mu_x\mu_y(\mu_x-\mu_y)(1-\langle z\rangle^2)}{\left[\mu_x+\mu_y-(\mu_x-\mu_y)\langle z\rangle + 2\mu_x\mu_y\right]^2}\left\langle\sigma^2\right\rangle. \tag{4.45}$$

Eqs.(4.23) and (4.45) take similar form. The difference comes from different hopping mechanism: in the globally connected model, the influx of specific site is proportional to *the mean* of the population, while in the 2-patch model it is proportional to the populations *on the other patch*. If the space consists of only 2 patches with carrying capacity $1 \pm \sigma$ and the hopping rates are doubled, Eq.(4.45) reduces to Eq.(4.23).

Direct numerical verifications are performed by integrating (4.27) are presented in Fig. 4.2.

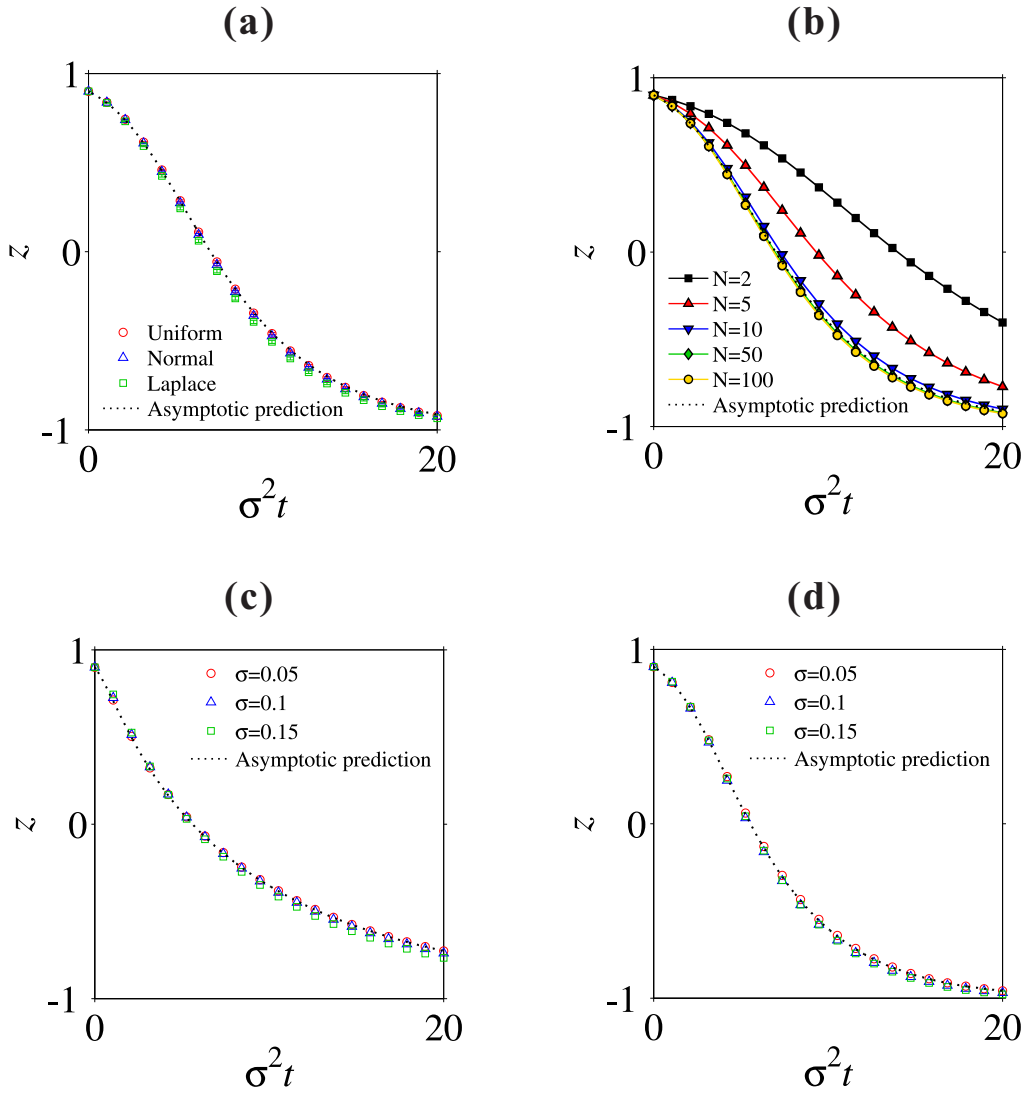

Figure 4.2: (a) Numerical simulation of the system (discrete markers) and the prediction of the asymptotic analysis (dotted line). The simulated system has 100 patches. $\{\sigma_i\}_{i=1}^{100}$ are i.i.d. r.v. with bounded uniform, truncated normal and truncated Laplace (double-exponential) distributions (we truncate the tails of the distributions to avoid negative carrying capacities.) 10 samples are measured to compute the mean and the sample error. $\mu_x = 5, \mu_y = 1$. (b) The convergence to the infinite-patch model as the number of patches $N \to \infty$. $\{\sigma_i\}_{i=1}^{N}$ are normal distributed, $\mu_x = 5$, $\mu_y = 1$. (c-d) Numerical simulations of the systems with $\sigma = 0.05, 0.1, 0.15$ (discrete markers) and the prediction of asymptotic analysis (dotted line). $N = 100$. $\sigma_{i\,i=1}^{N}$ are normal distributed. (c) $\mu_x = 1, \mu_y = 0.1$ and (d) $\mu_x = 5, \mu_y = 1$.

4.3 Discussion and conclusion

We come to the conclusion that the 2-patch and the many patch models shares similar features in deterministic description. In the 2-patch model, the parameter σ serves as a natural measure of heterogeneity. In the many patch model, we define the corresponding inhomogeneity σ to be the standard deviation of the environmental distribution $(\mathbb{E}[\sigma_i^2])^{1/2}$. Started with arbitrary initial conditions, in a time scale $\mathcal{O}(\log \sigma^{-1})$ the systems equilibrate to a metastable distribution. The metastable distribution remains stable in a time scale $\mathcal{O}(\sigma^{-1})$, and there exists a slow drift in a longer time scale $\mathcal{O}(\sigma^{-2})$, predicted by (4.22) or (4.44). One important feature of the drifts in (4.22) and Eq.(4.44) are always negative if $\mu_x > \mu_y$ (the claim will be proved in Chapter VI). This result, along with the lowest order constraint $x_0 + y_0 = 1$, shows the deterministic dynamics always favors the slow species in the competition as long as the spatial inhomogeneity σ is not equal to 0. In addition, the analysis shows such evolutionary advantage prevails in a time scale of order $\mathcal{O}(\sigma^{-2})$. Our quantitative analysis on the models confirms qualitative theorems Hasting [19] and Dockery *et al.* [8] proved for similar PDE models with distinct settings.

CHAPTER V

Demographic Stochasticity and Evolution of Dispersal in Heterogeneous Environments

Two limits of the competitive population dynamics models have been investigated thoroughly—in Chapter III we considered the dynamics *with demographic stochasticity* in *homogeneous environments*, and in Chapter IV we investigated the dynamics *without demographic stochasticity* in *heterogeneous environments*. The time scales of the weak selection of the faster and slower mobility species are as respectively $\mathcal{O}(K)$ and $\mathcal{O}(\sigma^{-2})$ respectively. In the first case, demographic stochasticity favors the *fast* dispersers, and in the second one, nonlinear dynamics favors the *slow* dispersers.

It is a natural conjecture that the population dynamics *with demographic stochasticity* in *heterogeneous environments* ought to be the outcome of the competition between the effects—due to the stochasticity and due to the nonlinearity. Indeed, numerical studies by Kessler and Sander [25], and Waddel *et al.* [41] revealed a clear regime shifts in their models.

In this Chapter, we generalize the individual-based models developed in Chapter III to investigate the competitive population dynamics with demographic stochasticity in heterogeneous environments. As will be seen, the boundaries of regime shifts can be analytically predicted. This Chapter is organized as follows. In sections 5.1 and 5.2, we present analyses of the (heterogeneous) 2-patch model and the (heterogeneous)

Independent process	Corresponding (per capita) rate
Birth of X on patch i	β
Birth of Y on patch i	β
Death of X on patch i	$\delta[1-(1-(-1)^i\sigma)(X_i+Y_i)/\Lambda]$
Death of Y on patch i	$\delta[1-(1-(-1)^i\sigma)(X_i+Y_i)/\Lambda]$
Dispersal of X from patch i to patch j, $i \neq j$	μ_X
Dispersal of Y from patch i to patch j, $i \neq j$	μ_Y

Table 5.1: The stochastic processes and the corresponding rates.

many patch model, respectively. The analyses are parallel to those in Chapter III. In section 5.3 we verify the predictions from sections 5.1 and 5.2 via exact numerical simulations. In the final section 5.4 we summarize and conclude the features of the dynamics. Further discussion will be presented in Chapter VII.

5.1 Stochastic two-patch model

5.1.1 The model

In the stochastic 2-patch model, the universe consists of two well-mixed patches. We denote the population of species X and Y on patch $i \in \{1,2\}$ at time s by non-negative integer-valued random variables $X_i(s)$ and $Y_i(s)$. Both species have identical demographic dynamics. On patch i, every individual waits exponentially distributed random times with rate β to reproduce one offspring. Similarly, individuals decease at rate $\delta\{1+[1-(-1)^i\sigma][X_i(s)+Y_i(s)]/\Lambda\}$. The parameters β, δ, and Λ are positive; in addition, the environment is assumed to be able to sustain large but finite populations, hence $\rho := \beta/\delta > 1$. As for dispersions, individuals of species X (or Y) hop to the other patch with rate μ_X (or μ_Y). The detailed processes are listed in Table 5.1, and Fig. 5.1 schematically demonstrates the processes.

Denote the probability of the system at state $(X_1(s) = a, X_2(s) = b, Y_1(s) = c, Y_2(s) = d)$ at time s by $p_{a,b,c,d}(s)$, then the evolution of $p_{a,b,c,d}(s)$ can be described

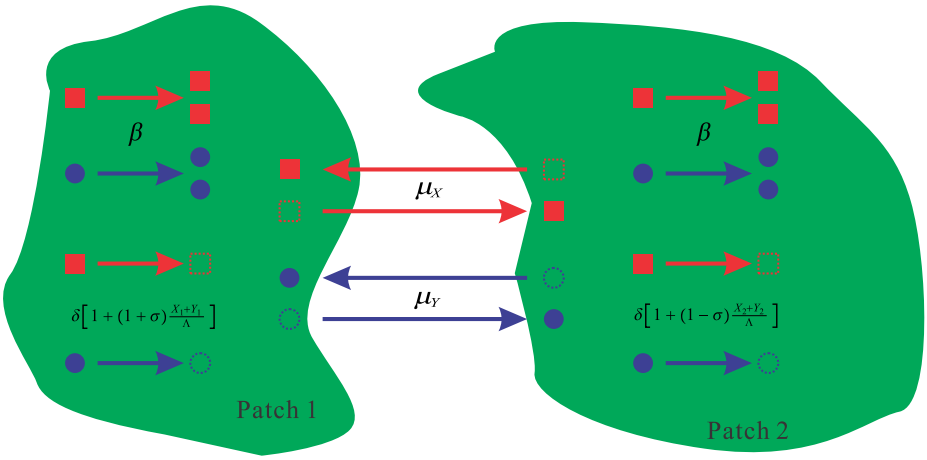

Figure 5.1: Dynamics of the interacting species distributed on two patches. The X and Y populations compete locally, and individuals randomly move from one patch to the other at rates μ_X and μ_Y.

by the following master equation:

$$
\begin{aligned}
(5.1)\quad \frac{d}{ds}p_{a,b,c,d} = &-\{\beta+\delta[1+(1+\sigma)(a+b)/\Lambda]+\mu_X\}ap_{a,b,c,d}\\
&-\{\beta+\delta[1+(1+\sigma)(a+b)/\Lambda]+\mu_Y\}bp_{a,b,c,d}\\
&-\{\beta+\delta[1+(1-\sigma)(c+d)/\Lambda]+\mu_X\}cp_{a,b,c,d}\\
&-\{\beta+\delta[1+(1-\sigma)(c+d)/\Lambda]+\mu_Y\}dp_{a,b,c,d}\\
&+\beta(a-1)p_{a-1,b,c,d}+\delta\left[1+(1+\sigma)(a+b+1)/\Lambda\right](a+1)p_{a+1,b,c,d}\\
&+\beta(b-1)p_{a,b-1,c,d}+\delta\left[1+(1+\sigma)(a+b+1)/\Lambda\right](b+1)p_{a,b+1,c,d}\\
&+\beta(c-1)p_{a,b,c-1,d}+\delta\left[1+(1-\sigma)(c+d+1)/\Lambda\right](c+1)p_{a,b,c+1,d}\\
&+\beta(d-1)p_{a,b,c,d-1}+\delta\left[1+(1-\sigma)(c+d+1)/\Lambda\right](d+1)p_{a,b,c,d+1}\\
&+\mu_X(a+1)p_{a+1,b,c-1,d}+\mu_X(c+1)p_{a-1,b,c+1,d}\\
&+\mu_Y(b+1)p_{a,b+1,c,d-1}+\mu_Y(d+1)p_{a,b-1,c,d+1}.
\end{aligned}
$$

After scaling the population by $K := (\rho-1)\Lambda$, time by $t := s(\beta-\delta)$, and assuming that the patches sustains large but finite population scale $\mathcal{O}(K)$, the solution of above

Eq.(5.1) can be well-approximated by the following Fokker–Planck equation

$$
\begin{aligned}
(5.2)\qquad \frac{\partial f}{\partial t} =& -\frac{\partial}{\partial x_1}\{[x_1(1-(1+\sigma)(x_1+y_1))-\mu_x(x_1-x_2)]f\} \\
&-\frac{\partial}{\partial x_2}\{[x_2(1-(1-\sigma)(x_2+y_2))-\mu_x(x_2-x_1)]f\} \\
&-\frac{\partial}{\partial y_1}\{[y_1(1-(1+\sigma)(x_1+y_1))-\mu_y(y_1-y_2)]f\} \\
&-\frac{\partial}{\partial y_2}\{[y_2(1-(1-\sigma)(x_2-y_2))-\mu_y(y_2-y_1)]f\} \\
&+\frac{1}{2K}\frac{\partial^2}{\partial x_1^2}\left\{\left[x_1(\frac{\rho+1}{\rho-1}+(1+\sigma)(x_1+y_1))+\mu_x(x_1+x_2)\right]f\right\} \\
&+\frac{1}{2K}\frac{\partial^2}{\partial x_2^2}\left\{\left[x_2(\frac{\rho+1}{\rho-1}+(1-\sigma)(x_2+y_2))+\mu_x(x_1+x_2)\right]f\right\} \\
&+\frac{1}{2K}\frac{\partial^2}{\partial y_1^2}\left\{\left[y_1(\frac{\rho+1}{\rho-1}+(1+\sigma)(x_1+y_1))+\mu_y(y_1+y_2)\right]f\right\} \\
&+\frac{1}{2K}\frac{\partial^2}{\partial y_2^2}\left\{\left[y_2(\frac{\rho+1}{\rho-1}+(1-\sigma)(x_2+y_2))+\mu_y(y_1+y_2)\right]f\right\} \\
&-\frac{1}{K}\frac{\partial^2}{\partial x_1\,\partial x_2}\left[\mu_x(x_1+x_2)f\right]-\frac{1}{K}\frac{\partial^2}{\partial y_1\,\partial y_2}\left[\mu_x(y_1+y_2)f\right]
\end{aligned}
$$

with scaled population x_i, y_i, scaled probability density f, scaled hopping rates μ_x and μ_y. When taking $K \to \infty$ limit, Eq.(5.2) reduced to the deterministic description in Chapter IV.

In addition, from Chapters III and IV, we know that the time scales of the effective drifts in the center manifold are $\mathcal{O}(K)$ and $\mathcal{O}(\sigma^{-2})$. If one of the effect has much shorter time scale than the other one, the effect with shorter time scale will prevail and dominate the other one. We are interested in the boundary of regime shifts, that is, when the time scales of the effect due to stochasticity and the effect of heterogeneity are comparable. To investigate such parameter region, we impose a constraint on the parameters that

$$
(5.3)\qquad \mathcal{O}(\sigma^{-2}) \approx \mathcal{O}(K) = \mathcal{O}((\rho-1)\Lambda) = \mathcal{O}(\epsilon^2).
$$

Balanced Processes	Direction	Magnitude of the Fluctuation
Birth/Death of X on patch 1	$(\pm 1, 0, 0, 0)$	$\triangle_1' \equiv \left[\frac{2\rho}{\rho-1}\frac{1}{K}(x_0 + \frac{w_x^*\sigma}{2})\,dt\right]^{1/2}$
Birth/Death of X on patch 2	$(0, \pm 1, 0, 0)$	$\triangle_2' \equiv \left[\frac{2\rho}{\rho-1}\frac{1}{K}(x_0 - \frac{w_x^*\sigma}{2})\,dt\right]^{1/2}$
Birth/Death of Y on patch 1	$(0, 0, \pm 1, 0)$	$\square_1' \equiv \left[\frac{2\rho}{\rho-1}\frac{1}{K}(y_0 + \frac{w_y^*\sigma}{2})\,dt\right]^{1/2}$
Birth/Death of Y on patch 2	$(0, 0, 0, \pm 1)$	$\square_2' \equiv \left[\frac{2\rho}{\rho-1}\frac{1}{K}(y_0 - \frac{w_y^*\sigma}{2})\,dt\right]^{1/2}$
Hopping of X	$(\pm 1, \mp 1, 0, 0)$	$\blacktriangle \equiv \left[2\mu_x\frac{x_0}{K}\,dt\right]^{1/2}$
Hopping of Y	$(0, 0, \pm 1, \mp 1)$	$\blacksquare \equiv \left[2\mu_y\frac{y_0}{K}\,dt\right]^{1/2}$

Table 5.2: Fluctuation strengths and shorthand notations of the homogeneous 2–patch model. w_x^* and w_y^* are defined in Eq.(4.9).

5.1.2 Physically motivated asymptotic analysis

Spatial inhomogeneity changes the structure of the deterministic flow of the 2-patch model as compared to the homogeneous 2-patch model in Chapter III. The stable line of fixed points exhibited in homogeneous 2-patch model becomes a center manifold when the spatial variance is not equal to zero (verified in Chapter IV). As a consequence, in the "physical" asymptotic analysis, we need to modify (compared to Chapter III) (1) the initial conditions after the fluctuations "kick" the state out of the center manifold and (2) the convergent trajectories back to the center manifold.

We first investigate the positions after the "kicks" of the fluctuations and argue that the strengths of the stochastic "kicks" near the center manifold are asymptotically identical to their counterparts in homogeneous model. Assuming the system starts from the $\mathcal{O}(\sigma)$-metastable distribution L

$$L \equiv \left\{(x_1, x_2, y_1, y_2) = (x_0 + \frac{\sigma w_x^*}{2}, x_0 - \frac{\sigma w_x^*}{2}, y_0 + \frac{\sigma w_y^*}{2}, y_0 - \frac{\sigma w_y^*}{2})\right\} \tag{5.4}$$

where w_x and w_y are defined in Eq.(4.9). Table 5.3 lists the directions and the strength of the stochastic perturbations. Because the order $\mathcal{O}(\epsilon)$ is matched to $\mathcal{O}(\sigma)$, to adopt

Table 5.3 instead of Table 3.2 results in a correction which is of higher order $\mathcal{O}(\epsilon^2) = \mathcal{O}(\sigma^2)$. On the other hand, in the asymptotic calculation in Chapter III, we learned that the $\mathcal{O}(\epsilon)$ "kicks" already result in the higher order $\mathcal{O}(\epsilon^2)$ drift and diffusion. Therefore, it is sufficient to approximate Table 5.3 by Table 3.2, considering the asymptotic functionals of effective drift and diffusion to $\mathcal{O}(\epsilon^2)$; the initial conditions, which are parallel to (3.38):

$$n_x^{(1)}(0) = \phi_1\Delta_1 + \phi_2\Delta_2 + \mathcal{O}(\epsilon^2), \tag{5.5a}$$

$$n_y^{(1)}(0) = \phi_3\square_1 + \phi_4\square_2 + \mathcal{O}(\epsilon^2), \tag{5.5b}$$

$$w_x^{(1)}(0) = w_x^*\sigma + \phi_1\Delta_1 - \phi_2\Delta_2 + 2\phi_5\blacktriangle + \mathcal{O}(\epsilon^2), \tag{5.5c}$$

$$w_y^{(1)}(0) = w_y^*\sigma + \phi_3\square_1 - \phi_4\square_2 + 2\phi_6\blacksquare + \mathcal{O}(\epsilon^2). \tag{5.5d}$$

where $(\Delta_1, \Delta_2, \square_1, \square_2, \blacktriangle, \blacksquare)$ are defined in Table 3.2.

Next we construct the effective diffusion and the effective drift in the center manifold from the analytical results in Chapters III and IV. Asymptotically the trajectories of the deterministic flow (4.1) near the center manifold had been analyzed in section 4.1. With the initial conditions (5.5), it is elementary to show that the effective diffusion is identical to (3.40). The intuition behind the result is, the effective diffusion is due to the fluctuations of the total populations n_x and n_y, and n_x and n_y in the heterogeneous model and homogeneous model share the same dynamics and initial conditions (that is, Eqs. (3.18) and (4.7), (3.38), and (5.5)). With the same initial conditions, the outcomes to the order $\mathcal{O}(\epsilon)$ must be identical—which in turn produce identical $\mathcal{O}(\epsilon^2)$ effective diffusion. As for the effective drift, we first denote the state

by the 4×1 vector $\vec{\psi}(t)$ by

$$\vec{\psi}(t) \equiv \begin{pmatrix} n_x^{(1)}(t) \\ n_y^{(1)}(t) \\ w_x^{(1)}(t) \\ w_y^{(1)}(t) \end{pmatrix}. \tag{5.6}$$

Then the expression of Eq.(3.32) can be rewritten in the matrix form

$$\begin{aligned} n_x^{(2)} \sim & \int_0^t \int_0^{t'} \vec{\psi}^{\,T}(t'') \cdot \mathbf{A}(\mathbf{t}', \mathbf{t}'') \cdot \vec{\psi}(t'')\, dt''\, dt' + \int_0^t \vec{\psi}^{\,T}(t') \cdot \mathbf{B} \cdot \vec{\psi}(t')\, dt' \\ & + \int_0^t \int_0^{t'} \vec{C}^{\,T}(t', t'') \cdot \vec{\psi}(t'')\, dt''\, dt' + \int_0^t \vec{D}^{\,T} \cdot \vec{\psi}(t')\, dt'. \end{aligned} \tag{5.7}$$

where the 4×4 matrix $\mathbf{A}$ and the 2×1 vector $\vec{C}$ involve the temporal propagators $e^{t''-t'}$, and the 4×4 matrix $\mathbf{B}$ and the 2×1 vector $\vec{D}$ are temporally constant. In addition, the $\mathcal{O}(\epsilon)$ solutions were derived in Chapter IV, i.e. Eqs. (4.13), (4.14), and (4.15). It has the form

$$\vec{\psi}(t) = \begin{pmatrix} n_x^{(1)}(t) \\ n_y^{(1)}(t) \\ w_x^{(1)}(t) \\ w_y^{(1)}(t) \end{pmatrix} = \begin{pmatrix} n_x^{(1)}(t) \\ n_y^{(1)}(t) \\ \tilde{w}_x^{(1)}(t) \\ \tilde{w}_y^{(1)}(t) \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ w_x^* \\ w_y^* \end{pmatrix}. \tag{5.8}$$

Define the meta-stable "fixed point" of ψ

$$\psi^* \equiv \begin{pmatrix} 0 \\ 0 \\ w_x^* \\ w_y^* \end{pmatrix} \tag{5.9}$$

and the "transient part" of ψ

$$\tilde{\psi} \equiv \begin{pmatrix} n_x^{(1)}(t) \\ n_y^{(1)}(t) \\ \tilde{w}_x^{(1)}(t) \\ \tilde{w}_y^{(1)}(t), \end{pmatrix} \tag{5.10}$$

then Eq.(5.7) becomes

$$\begin{aligned} n_x^{(2)} \sim & \int_0^t \int_0^{t'} (\psi^* + \tilde{\psi}(t''))^T \cdot \mathbf{A}(\mathbf{t}', \mathbf{t}'') \cdot (\psi^* + \tilde{\psi}(t'')) \, dt'' \, dt' \\ & + \int_0^t (\psi^* + \tilde{\psi}(t'))^T \cdot \mathbf{B} \cdot (\psi^* + \tilde{\psi}(t')) \, dt' \\ & + \int_0^t \int_0^{t'} \vec{C}^{\,T}(t', t'') \cdot \left(\psi^* + \tilde{\psi}(t'')\right) \, dt'' \, dt' \\ & + \int_0^t \vec{D}^{\,T} \cdot \left(\psi^* + \tilde{\psi}(t')\right) \, dt'. \end{aligned} \tag{5.11}$$

In the physically motivated asymptotic analysis, we need to perform an "ensemble average" $\langle \cdot \rangle$ to compute effective drift and diffusion. The average takes a finite number of the characteristic directions and calculates their means. Because of the number of the directions are finite (in this model, $2^6 = 64$), the average $\langle \cdot \rangle$ and the integrations are order-exchangeable. In addition, we notice that the nontrivial contributions comes from the *quadratic* terms of $\tilde{\psi}$; because the initial conditions are symmetrical (see

Table 5.3 for reference), the averages involving odd order of $\tilde{\psi}$ must be zero:

$$0 = \langle \tilde{\psi} \rangle, \tag{5.12a}$$

$$0 = \langle \psi^* \cdot \mathbf{A} \cdot \tilde{\psi} \rangle, \tag{5.12b}$$

$$0 = \langle \psi^* \cdot \mathbf{B} \cdot \tilde{\psi} \rangle, \tag{5.12c}$$

$$0 = \langle \vec{C}^{\,T} \cdot \tilde{\psi} \rangle, \text{and} \tag{5.12d}$$

$$0 = \langle \vec{D}^{\,T} \cdot \tilde{\psi} \rangle \tag{5.12e}$$

With the identities, after the ensemble average, (5.11) becomes

$$\begin{aligned} \left\langle n_x^{(2)} \right\rangle = & \int_0^t \int_0^{t'} \psi^{*\,T}\, \mathbf{A}(\mathbf{t}', \mathbf{t}'')\, \psi^*\, dt''\, dt' + \int_0^t \psi^{*\,T}\, \mathbf{B}\, \psi^*\, dt' \\ & + \int_0^t \int_0^{t'} \vec{C}^{\,T}(t', t'') \cdot \psi^*\, dt''\, dt' + \int_0^t \vec{D}^{\,T} \cdot \psi^*(t')\, dt' \\ & + \left\langle \int_0^t \int_0^{t'} \tilde{\psi}^{\,T}\, \mathbf{A}(\mathbf{t}', \mathbf{t}'')\, \tilde{\psi}\, dt''\, dt' \right\rangle + \left\langle \int_0^t \tilde{\psi}^{\,T}\, \mathbf{B}\, \tilde{\psi}\, dt' \right\rangle. \end{aligned} \tag{5.13}$$

After a sufficiently long time ($\gtrsim \mathcal{O}(\log K)$), the terms

$$\begin{aligned} & \int_0^t \int_0^{t'} \psi^{*\,T}\, \mathbf{A}(\mathbf{t}', \mathbf{t}'')\, \psi^*\, dt''\, dt' + \int_0^t \psi^{*\,T}\, \mathbf{B}\, \psi^*\, dt' \\ & + \int_0^t \int_0^{t'} \vec{C}^{\,T}(t', t'') \cdot \psi^*\, dt''\, dt' + \int_0^t \vec{D}^{\,T} \cdot \psi^*(t')\, dt' \end{aligned}$$

are identified to be the drift due to the nonlinearity, i.e. Eq.(4.21), and the terms

$$\left\langle \int_0^t \int_0^{t'} \tilde{\psi}^{\,T}\, \mathbf{A}(\mathbf{t}', \mathbf{t}'')\, \tilde{\psi}\, dt''\, dt' \right\rangle + \left\langle \int_0^t \tilde{\psi}^{\,T}\, \mathbf{B}\, \tilde{\psi}\, dt' \right\rangle$$

are identified to be the drift due to the stochasticity, i.e. Eq.(3.53), since the transient

terms in the heterogeneous model have identical dynamics to the corresponding ones in the homogeneous model (compare Eqs.(3.19), (4.11), (3.38), and (5.5)).

The physically motivated asymptotic analysis therefore yields the following conclusions. For any given initial conditions, within $\mathcal{O}(\log K)$ time, the system flows along the deterministic trajectories to the $1/K-$neighborhood of the one dimensional center manifold L: $(x_1, x_2, y_1, y_2) = (x_0 + \sigma\omega_x^*/2, x_0 - \sigma\omega_x^*/2, y_0 + \sigma\omega_y^*/2, y_0 - \sigma\omega_y^*/2)$ with $x_0, y_0 \in (0,1)$ and $x_0 + y_0 = 1$. After convergence, the state exhibits a slow motion with time scale $O(K)$, due to the interaction of stochasticity and nonlinearity near the center manifold. The effective drift v and diffusion D near the center manifold are determined by $K, \rho, \mu_x, \mu_y, \sigma$, and an effective coordinate z in the center manifold:

$$v(z) = \frac{1-z^2}{K}\left\{C_0 + \frac{C_1}{(\mu_x-\mu_y)z - C_2} + \frac{C_3}{\left[(\mu_x-\mu_y)z - C_2\right]^2}\right\}, \tag{5.14}$$

$$D(z) = \frac{\rho}{2(\rho-1)}\frac{1-z^2}{K}, \tag{5.15}$$

where the effective coordinate z is defined as

$$z \equiv \frac{(x_1+x_2)-(y_1+y_2)}{2}, \tag{5.16}$$

and the parameters are

$$C_0 = -\frac{1}{2}\frac{\mu_x-\mu_y}{1+2(\mu_x+\mu_y)}, \tag{5.17a}$$

$$C_1 = -\frac{(\mu_x-\mu_y)}{2}\left[\frac{\rho}{\rho-1} + \frac{4\mu_x\mu_y}{1+2(\mu_x+\mu_y)}\right], \tag{5.17b}$$

$$C_2 = 4\mu_x\mu_y + \mu_x + \mu_y, \tag{5.17c}$$

$$C_3 = -4\mu_x\mu_y(\mu_x-\mu_y)K\sigma^2. \tag{5.17d}$$

Incidentally, the analysis shows the 1-dimensional effective diffusion in the center

manifold, Eq.(5.15), is identical to the effective diffusion (3.12) in the homogeneous 2-patch model, and the drift, Eq.(5.14) turns out to be the linear superposition of the effective drift due to the heterogeneity (4.22) and the effective drift due to the stochasticity (3.11). We emphasize that the physically motivated analysis, which only assumes $\mathcal{O}(\epsilon^2) = \mathcal{O}(\sigma^2)$ from matching two dynamical time scales, *does not* assume the separation of the two drifts a priori. In fact, the approach analyzes the nonlinearity and stochasticity simultaneously, and predicts the separation of the drifts mathematically. The fundamental reason for the separation is that both drifts come from the nonlinearity to the same order, $\mathcal{O}(\epsilon^2) = \mathcal{O}(\sigma^2)$.

The drift due to the heterogeneity always favors the slow dispersers, and the drift due to the stochasticity always favors the fast dispersers in the model (the claim will be proved in Chapter VI). In Chapter VI, we will discuss the bifurcations of the combined flow, which in turn determine which species has evolutionary advantage in the competition.

With the predicted drift (5.14) and diffusion (5.15), standard analysis formulated in section 2.1.2 is carried out to compute the extinction probability of the species, and the mean extinction time of any species in the reduced center manifold. In section 5.3 we present numerical evidence to support the analysis in this section.

5.2 Stochastic many patch model

5.2.1 The model

The stochastic many patch model consists of countably infinite number of patches, on which individuals of species X and Y lives. The carrying capacity of patch $i \in \mathbb{N}$ is $K/(1+\sigma_i)$ with i.i.d. random variable σ_i's. The distribution of $\{\sigma_i\}_{i=1}^{\infty}$ are assumed to be bounded and $\mathbb{E}[\sigma_i] = 0$. The heterogeneity of the environment is measured by

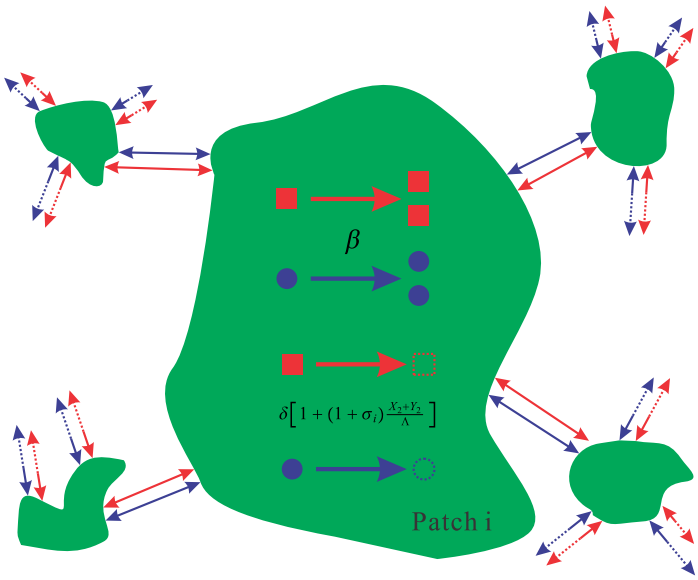

Figure 5.2: Dynamics of the interacting species distributed on many patches. The X and Y populations compete locally, and individuals randomly move from one patch to the other at rates μ_X and μ_Y.

$\sigma := (\mathbb{E}[\sigma_i^2])^{1/2}$ and we also assume $|\sigma_i| \ll 1 \; \forall i \in \mathbb{N}$ for analytical convenience[1]. Each individual has identical demographic processes as in the stochastic 2-patch model (section 5.1.1). The low density birth rate β, low density death rate δ, and birth-to-death rate ratio ρ are defined in parallel to the ones in the stochastic 2-patch model. As for dispersion, each individual of species X (or Y) waits an exponentially distributed random time with rate $\mu_X > 0$ (or $\mu_Y > 0$), and then moves to another patch chosen with equal probability.

In addition, we match the time scale of the effective drifts due to the stochasticity and effective drift due to the heterogeneity of the environment, so $\mathcal{O}(K) = \mathcal{O}(\sigma^{-2})$, where $K \equiv 1/((\rho - 1)\Lambda)$ is the harmonic mean carrying capacity of the model. The detailed processes are listed in Table 5.3, and Fig. 5.2 schematically demonstrates the processes.

[1]The analytical conclusion is not limited by restriction, however, as verified by numerical simulations in section 5.3.

Independent process	Corresponding (per capita) rate
Birth of X on patch i	β
Birth of Y on patch i	β
Death of X on patch i	$\delta[1-(1-(-1)^i\sigma_i)(X_i+Y_i)/\Lambda]$
Death of Y on patch i	$\delta[1-(1-(-1)^i\sigma_i)(X_i+Y_i)/\Lambda]$
Dispersal of X from patch i to patch j, $i \neq j$	μ_X
Dispersal of Y from patch i to patch j, $i \neq j$	μ_Y

Table 5.3: The stochastic processes and the corresponding rates in the many patch model.

5.2.2 Physically motivated asymptotic analysis

After scaling the population, time and hopping rates, in the infinite K limit the dynamics converges to the deterministic dynamics discussed in Chapter IV, from which we obtained the center manifold in the large K limit:

$$\begin{aligned} x_i &= \langle x\rangle \left[1-\sigma_i \frac{\mu_y}{\langle y\rangle\mu_x + \langle x\rangle\mu_y + \mu_x\mu_y}\right] + \mathcal{O}(\sigma_i^2), \\ y_i &= \langle y\rangle \left[1-\sigma_i \frac{\mu_x}{\langle y\rangle\mu_x + \langle x\rangle\mu_y + \mu_x\mu_y}\right] + \mathcal{O}(\sigma_i^2). \end{aligned} \tag{5.18}$$

with the mean populations per patch $\langle x\rangle$ and $\langle y\rangle$ and

$$\langle x\rangle, \langle y\rangle \in (0,1), \tag{5.19}$$

$$\langle x\rangle + \langle y\rangle = 1. \tag{5.20}$$

The constraint (5.20) suggests the difference of mean populations per patch $\langle z\rangle \equiv \langle x\rangle - \langle y\rangle$ is an effective coordinate in the one-dimensional center manifold. Similar to the analysis of the stochastic 2-patch models in section 5.1.2, near the center manifold, the strengths of the demographic fluctuations in homogeneous and heterogeneous models are identical to $\mathcal{O}(1/K) = \mathcal{O}(\epsilon^2) = \mathcal{O}(\sigma^2)$, so we only need to modify the computations regarding the convergence to the center manifold in homogeneous many

patch model. The abstract constructions developed in section 5.1.2 can be adopted to carry out the parallel computations of the many patch model, and it yields the same conclusion—the effective drift in the center manifold is a linear superposition of the drift due to stochasticity and the drift due to nonlinearity. As a consequence, the reduced dynamics of the state in the center manifold is

$$\frac{d\langle z\rangle}{dt} = \bar{v}(\langle z\rangle), \tag{5.21}$$

where the effective velocity field $\bar{v}$ in the center manifold is

$$\bar{v}(\langle z\rangle) = \frac{1-\langle z\rangle^2}{K}\left\{\bar{C}_0 + \frac{\bar{C}_1}{(\mu_x-\mu_y)\langle z\rangle - \bar{C}_2} + \frac{\bar{C}_3}{[(\mu_x-\mu_y)\langle z\rangle - \bar{C}_2]^2}\right\}, \tag{5.22}$$

with

$$\bar{C}_0 = -\frac{1}{2}\frac{\mu_x-\mu_y}{1+\mu_x+\mu_y}, \tag{5.23a}$$

$$\bar{C}_1 = -(\mu_x-\mu_y)\left(\frac{\rho}{\rho-1} + \frac{\mu_x\mu_y}{1+\mu_x+\mu_y}\right), \tag{5.23b}$$

$$\bar{C}_2 = 2\mu_x\mu_y + \mu_x + \mu_y, \tag{5.23c}$$

$$\bar{C}_3 = -2\mu_x\mu_y(\mu_x-\mu_y)K\sigma^2. \tag{5.23d}$$

As remarked in section 5.1.2, the separation is predicted by the analysis rather than being assumed *a priori*.

The difference of the spatial average populations, $\langle z\rangle$, evolves according to (5.21), which has incorporated the effect of the deterministic nonlinear dynamics—the term that proportional to $K\sigma^2$ in (5.22)—and the effect of the stochastic dynamics. Furthermore, the evolution of the metastable distributions (5.18) is predicted by $\langle z\rangle$ and the constraint (5.20). Similar to the homogeneous many patch model in Chapter III, the stochasticity—terms that proportional to $\bar{C}_0$ and $\bar{C}_1$ in (5.22)—has a non-trivial

contribution to the overall deterministic evolution of the state, i.e. Eq.(5.21).

Eq.(5.21) resembles an one dimensional deterministic dynamical system for $\langle z \rangle \in (-1, 1)$. As will be shown in Chapter VI, the effective drift due to stochasticity always favors the fast dispersers, and the effective drift due to the deterministic nonlinear dynamics always favors the slow dispersers. Interestingly, only four independent parameters, ρ, μ_x, μ_y and $K\sigma^2$ determine the solution as $t \to \infty$. The expression involving product of the population scale K and spatial environmental variance σ^2 confirms the heuristic argument Kessler and Sander presented [25].

5.3 Simulations and numerical computations

Continuous time Markov chain simulation are constructed to simulate the sample path of the processes in this Chapter. In the simulation, each individual belongs to either species X or Y with corresponding parameters. Let i to be the patch index, X_i and Y_i to be the populations of species X and Y respectively on patch i. Every agent waits for an exponentially distributed random time, and then proceed one of the following possible processes with appropriate distribution [35]: (1) reproduces one new agent of the same kind with rate β, (2) demises with rate $\delta[1+(1+\sigma_i)(X_i+Y_i)/\Lambda]$, and (3) hops to another patch with rate μ_X or μ_Y. Many sample paths of the system with the same initial conditions are generated to measure the ensemble averages of the observables.

5.3.1 Stochastic two-patch model

Recall that $\sigma_1 \equiv \sigma$ and $\sigma_2 \equiv -\sigma$ with heterogeneity $\sigma > 0$ in the 2-patch model. Parameters $\Lambda \in \mathbb{N}$, $\mu_X > 0$, $\mu_Y > 0$, $\rho > 1$ and $\sigma > 0$ are varied to examine the prediction in section 5.1. Several sample paths are shown in Fig. 5.3. A typical character observed in the 2-patch model is that the populations fluctuates dramatically—there does not seem to be a significant trend that a certain species out-competes the other.

In some sample paths, the fast moving species X extinct before the end of the simulations, and in others Y extincts first.

Figs. 5.4 and 5.5 present the evolution of the system configurations in the two-dimensional projected phase space, (X, Y), for $\Lambda = 500$ and $\Lambda = 200$ respectively. One hundred sample paths were generated to demonstrate the phase flow. One of the observation of these preliminary simulations are, the time scale of the extinction events does not scale exponentially with respect to the population scale K. This was our principle motivation to develop the "physical" asymptotic approach—to explain a phenomenon that is neither a large-deviations phenomenon, nor a spontaneous extinction phenomenon.

To be more precise, we measure the winning probability of species X and the mean extinction time (of any of the species) by measuring the observables in 10^4 identical sample paths. The initial populations are set to be

$$X_1(0) = X_2(0) = \frac{1 + z_0}{2}(\rho - 1)\Lambda, \tag{5.24}$$

$$Y_1(0) = Y_2(0) = \frac{1 - z_0}{2}(\rho - 1)\Lambda. \tag{5.25}$$

where the initial effective coordinate z_0 in the center manifold is varied. In most of the parameter sets we examined, the analysis in section 5.1 agrees with the continuous time Markov chain simulations[2]— see for example, selective data of $(\beta, \delta) = (2, 1)$ in Fig. 5.6.

The asymptotic analysis faithfully predicts the winning probability and mean extinction time. We can furthermore explore the landscape of the weak selection in parameter space in a head-to-head competition by numerically computing the winning probability of species X. Figs. 5.7 and 5.8 demonstrate the landscape of X- or

[2]The only exceptions is again when μ_x and $\mu_Y \ll 1$, as pointed out in Chapter IV. With such parameters the patches are almost decoupled. The physical asymptotic analysis breaks down because the separation of time scale no longer holds.

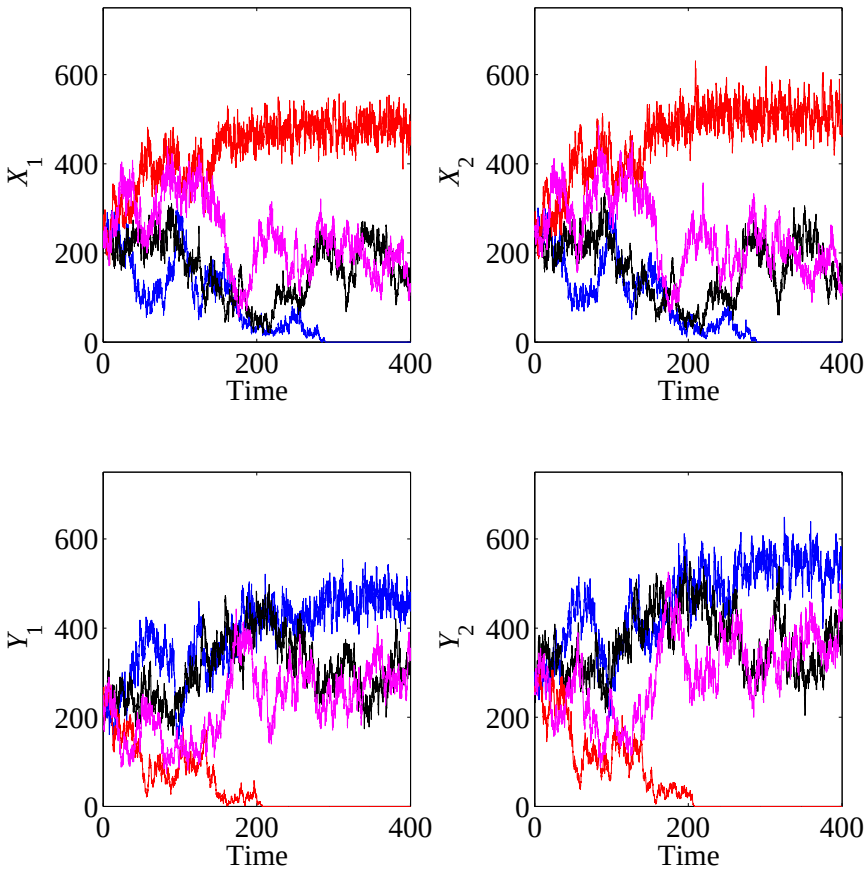

Figure 5.3: Sample paths of the stochastic (heterogeneous) 2-patch model. The parameters are: $\beta = 2$, $\delta = 1$, $\Lambda = 500$, $\sigma = 0.09$, $\mu_X = 1$, $\mu_Y = 0.1$.

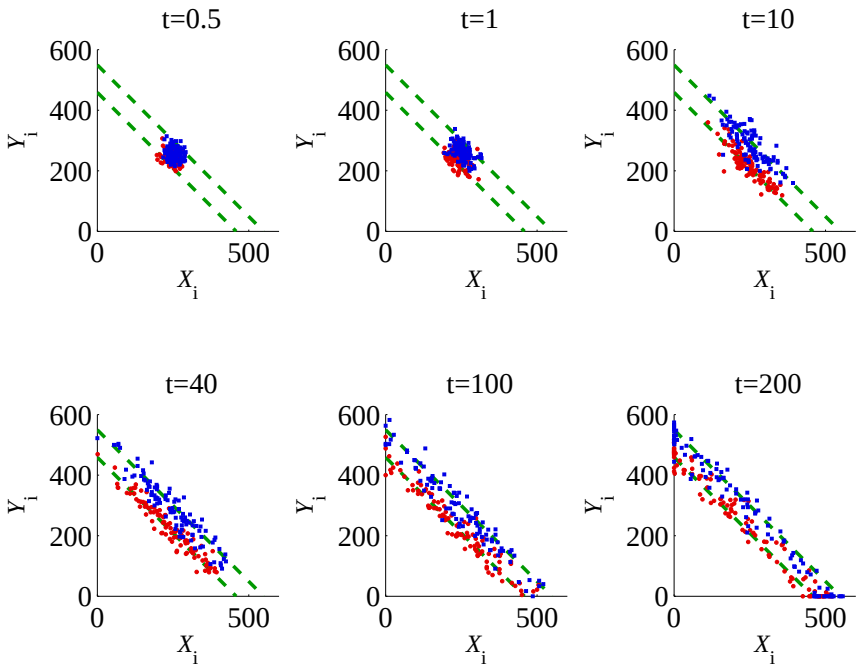

Figure 5.4: The evolution of the stochastic 2-patch model. The ensemble has 100 sample paths, which are identically initiated with $X_i(0) = Y_i(0) = 0.5\Lambda(= 0.5K)$. The parameters are: $\beta = 2$, $\delta = 1$, $\Lambda = 500\,(= K)$, $\sigma = 0.09$, $\mu_X = 1$, $\mu_Y = 0.1$. Red and blue markers are populations on patch 1 and 2 respectively. Dashed green lines denotes $X + Y$ equals to the carrying capacity of both patches.

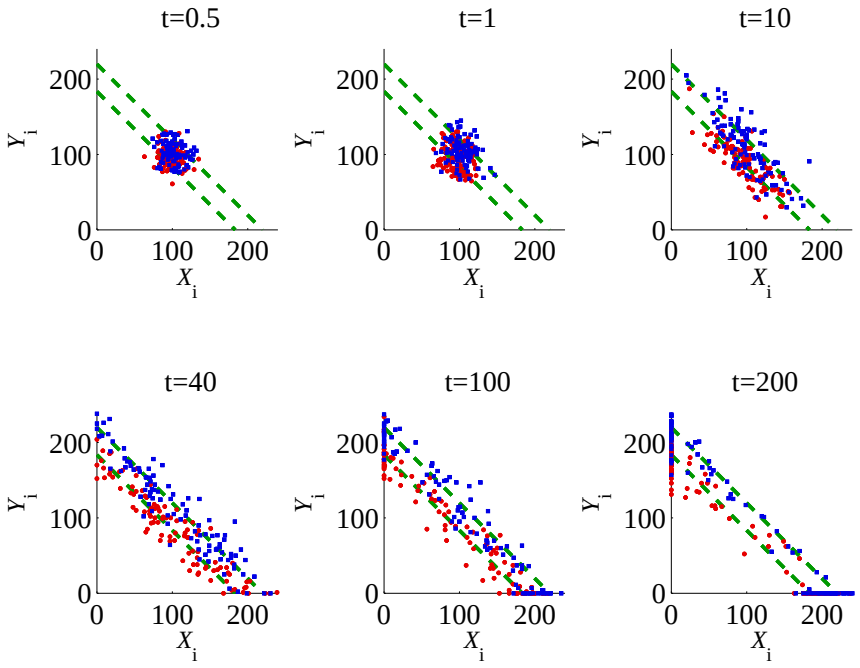

Figure 5.5: Reproduction of Fig. 5.4. All the representations and parameters are identical except for $\Lambda\,(=K)$, the population scale, is lowered to 200.

Y-dominating regimes in the parameter space. In Fig. 5.7, (μ_X, μ_Y) are fixed, and we plot the winning probability as function of the population scale Λ and the heterogeneity σ. In the log–log plot, the transition boundary of X- and Y-dominating regimes is revealed to be $\Lambda\sigma^2 = const$. It is now clear that the fundamental mechanism of the transition is the competition between effective drift due to stochasticity and effective drift due to nonlinear dynamics in inhomogeneous environment, i.e. Eq.(5.14). In Fig. 5.8, we fix (ρ, Λ, σ) and plot the winning probability of species X in a head-to-head competition (z_0=0) . The landscape in (μ_X, μ_Y) qualitatively changes as the value of $\Lambda\sigma^2$ changes and has the following characteristics:

- When $\Lambda\sigma^2$ is very small, the faster disperser has greater chance to survive—Fig. 5.8.(a).
- When $\Lambda\sigma^2$ is very large, the slower disperser has greater chance to survive—Fig. 5.8.(b).

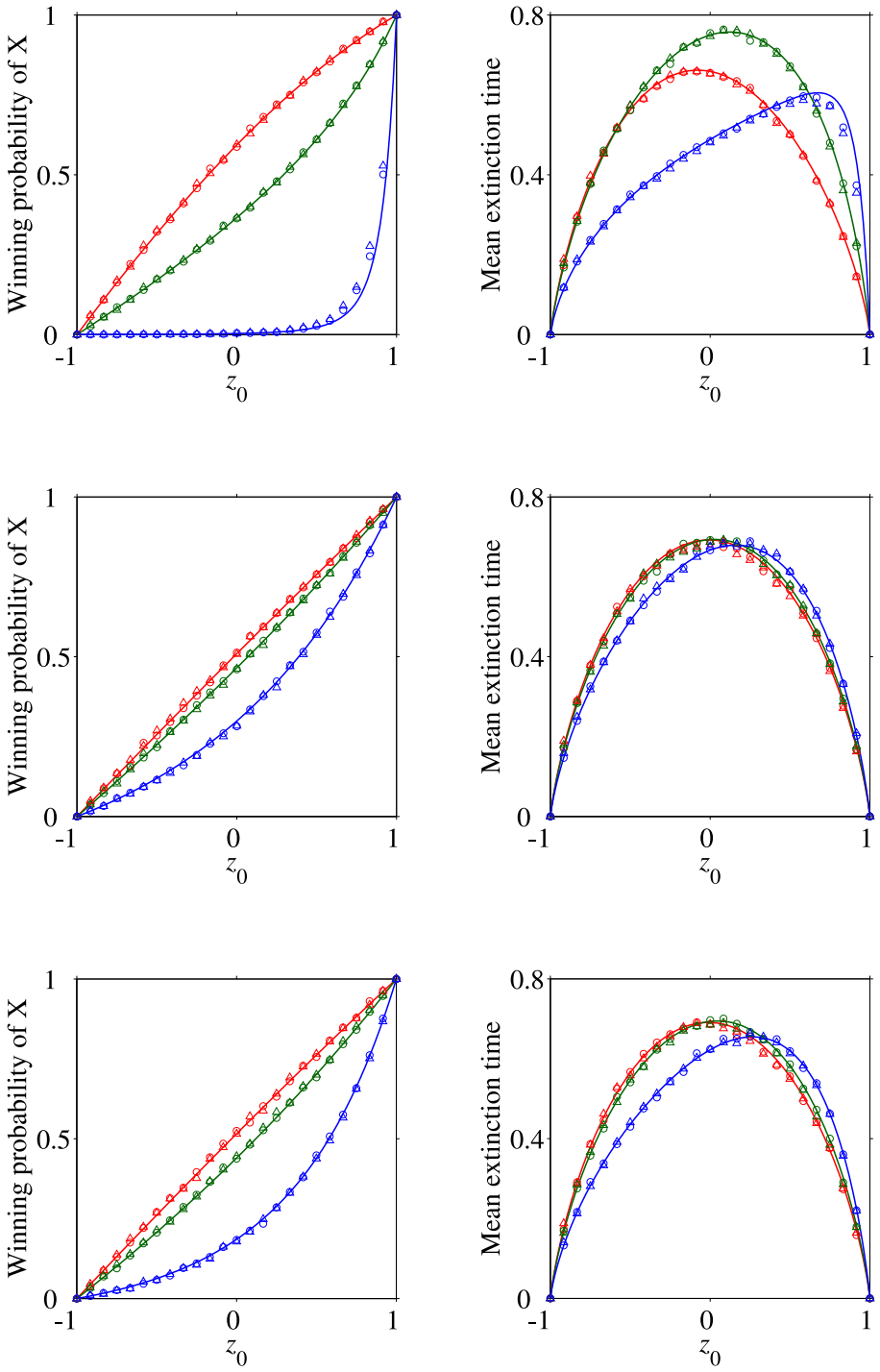

Figure 5.6: Comparisons of exact continuous time Markov chain simulations (discrete data) and the predictions of asymptotic analysis, color coded by red ($\Lambda\sigma^2 = 0.05$), green ($\Lambda\sigma^2 = 4.05$) and blue ($\Lambda\sigma^2 = 20$). $\rho = 2$. The circle markers are for $\Lambda = 500$ and the triangle markers are for $\Lambda = 245$. Left: winning probability of species X vs. initial condition z_0. Right: Mean extinction time (normalized by Λ) vs. initial condition z_0. Top pair: $\mu_X = 2$, $\mu_Y = 0.2$. Middle pair: $\mu_X = 4$, $\mu_Y = 2$. Bottom pair: $\mu_X = 10$, $\mu_Y = 2$.

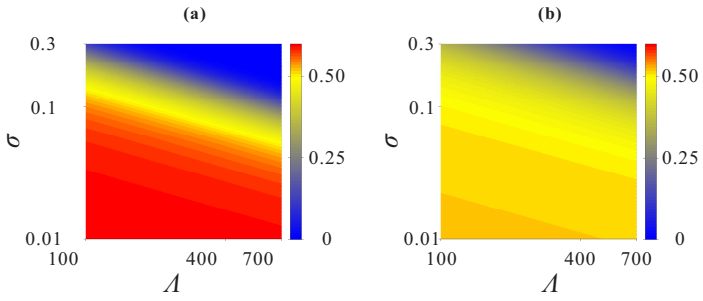

Figure 5.7: The numerical calculated landscape of winning probability of X. (Λ, σ) is plotted in log-log scale. (a) $\mu_X = 2$, $\mu_Y = 0.2$ and (b) $\mu_X = 10$, $\mu_Y = 2$. $\rho = 2$.

- In Fig. 5.8.(b-c), there exists a very shallow saddle. The saddle indicates the **evolutionarily stable rate for dispersion**—for any species, choosing the rate at the saddle guarantees *advantage in probability* when competing to *any* another species with any other rate of dispersal.

Unfortunately the winning probability cannot be explicitly expressed. Nevertheless, since the weak selection is driven by the effective drift in the center manifold, the location of the saddle can be predicted by analyzing the bifurcations of the effective drift, as will be shown in Chapter VI.

5.3.2 Stochastic many-patch model

We generalize the simulations to have a large number of patches to verify the analysis in section 5.2. In the following text we denote the number of patches to be $N \gg 1$. The dynamically fixed spatial distribution is determined by generating i.i.d. random variables $\{\sigma_i\}_{i=1}^N$ drawn from specific distributions. The standard deviation of the sample sequence $\{\sigma_i\}_{i=1}^N$ is computed and will be denoted to be σ. We adopt different distributions to generate σ_i, including (1) bounded uniform distribution, (2) truncated normal distribution, and (3) truncated Laplace (bounded double-exponential) distribution.

In contrast to the 2-patch model, an apparent characteristic observed from the

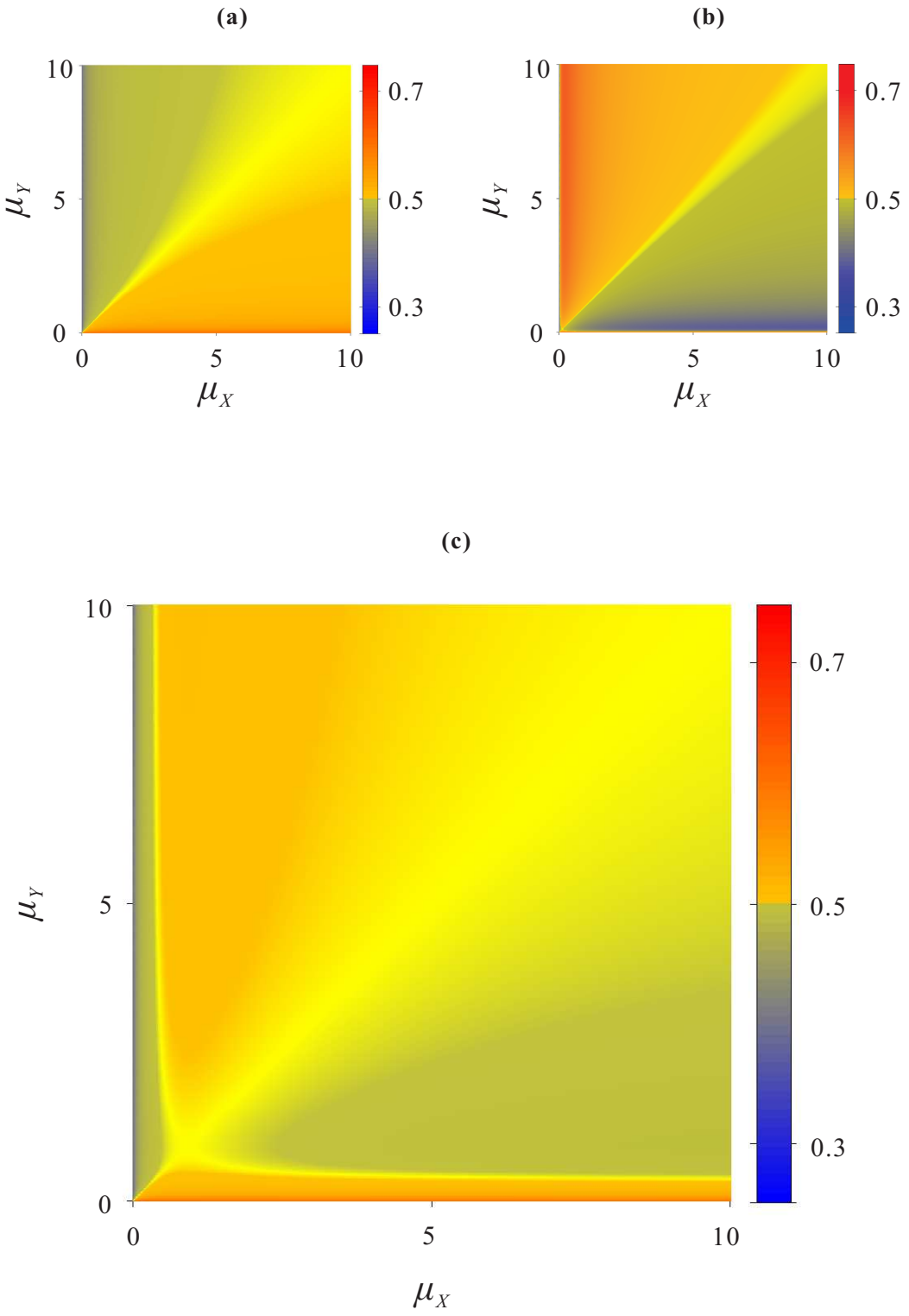

Figure 5.8: The numerical calculated landscape of winning probability of X in the space (μ_X, μ_Y) with (a) $\sigma = 0.01$, (b) $\sigma = 0.10$, and (c) $\sigma = 0.05$. $\rho = 2$ and $\Lambda = 500$.

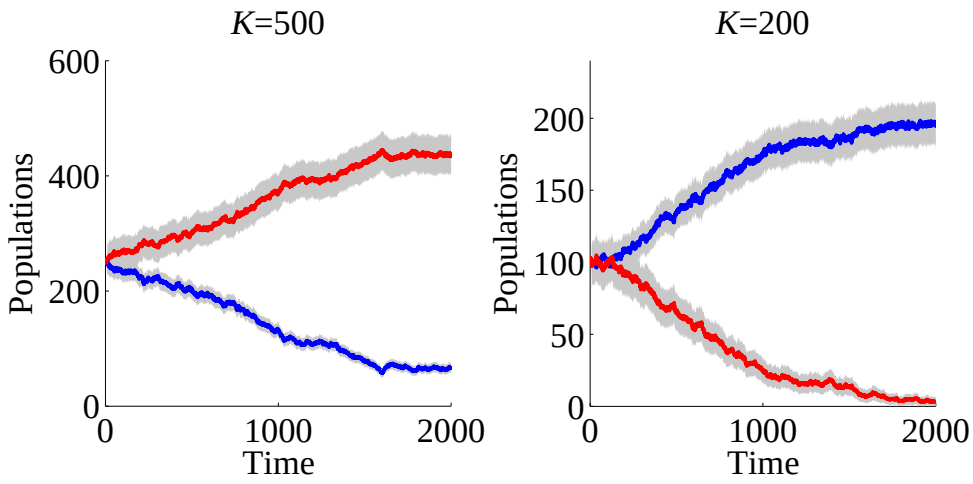

Figure 5.9: One sample path of the stochastic (heterogeneous) many patch model. The parameters are: $\beta = 2$, $\delta = 1$, $\sigma = 0.09$, $\mu_X = 1$, $\mu_Y = 0.1$, $N = 500$. The blue and red curves denote respectively the mean populations of the slow species Y and the fast species X among the patches. Grey band represents the mean populations (among the patches) plus / minus one standard deviation of the distribution.

simulations is that the *stochastic dynamics* in an environment with globally connected patches seem more *deterministic*. For example, Fig. 5.9 shows *one* sample path of the simulation with uniformly distributed $\{\sigma_i\}_{i=1}^{N=500}$, for systems with $K = \Lambda = 200$ and 500. In the case of $K = 200$ the fast species gains the population and in the case of $K = 500$ it loses. In Figs. 5.10 and 5.11 the configurations of dynamical system are projected onto a two-dimensional plane to show the evolution of a *single* sample path.

In the simulation we discover that the final results are indeed not sensitive to the choice of distributions, which confirms the prediction of our asymptotic analysis: as long as the distributions have the same variance, the final dynamics are indistinguishable.

With large N, the mean-field effect decreases the probability of any species goes

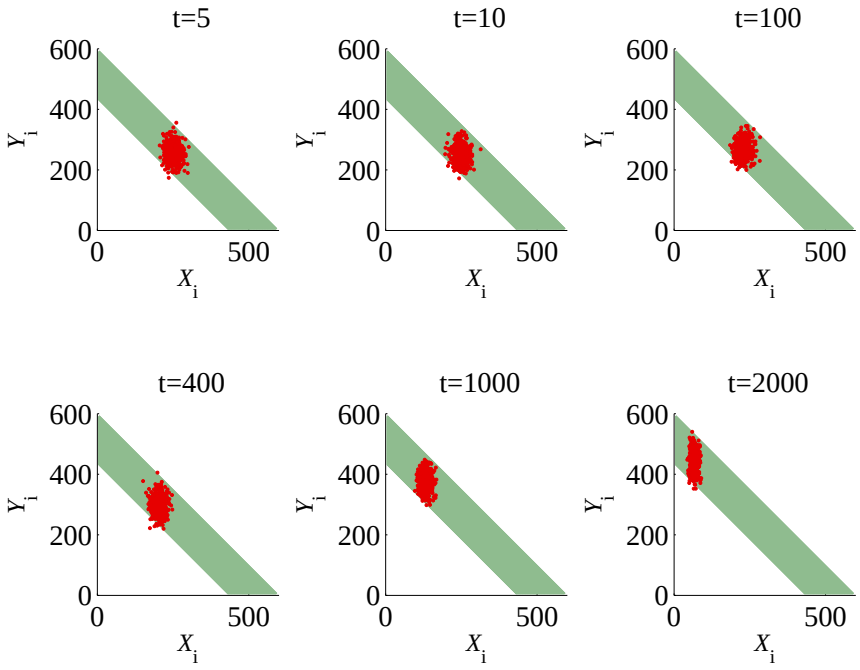

Figure 5.10: The evolution of a sample path of the stochastic many model. The system starts with $X_i(0) = Y_i(0) = 0.5\Lambda(= 0.5K)$ for $i \in \{1 \dots N\}$. The parameters are: $N = 500$, $\beta = 2$, $\delta = 1$, $\Lambda = 500\,(= K)$, $\sigma = 0.09$, $\mu_X = 1$, $\mu_Y = 0.1$. The green band is the environmental distribution (where the uniform σ_i's are generated). Each red dot denotes the populations (X_i, Y_i) on patch i.

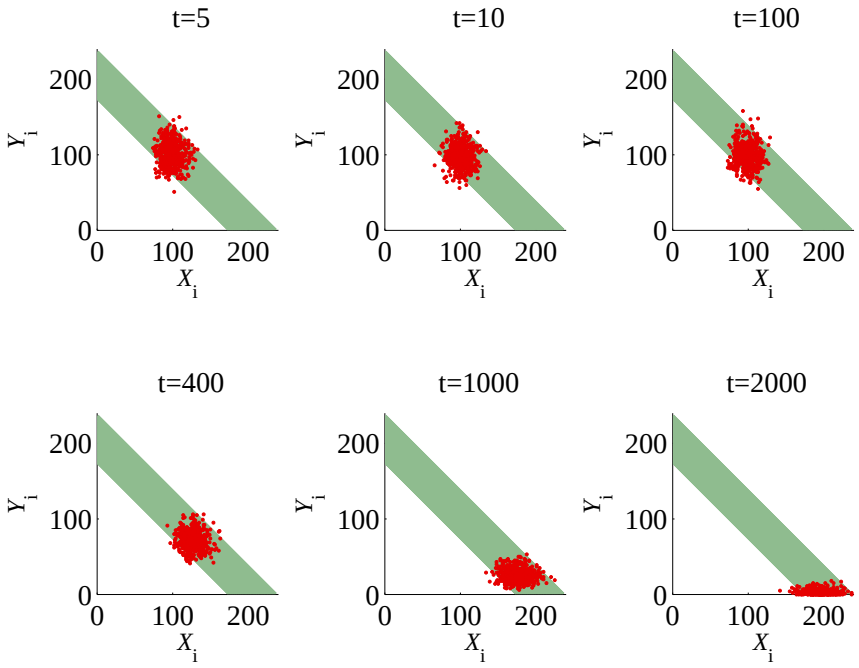

Figure 5.11: Reproduction of Fig. 5.10. All the representations and parameters are identical except for $\Lambda\,(=K)$, the population scale, is lowered to 200.

to extinct in finite time. Hence, we adopt the average populations among the patches

$$\langle X(t)\rangle := \frac{1}{N}\sum_{i=1}^{N} X_i(t),$$
$$\langle Y(t)\rangle := \frac{1}{N}\sum_{i=1}^{N} Y_i(t),$$

to be the observables. Fig. 5.12 shows the comparison between results from large-scale continuous time Markov chain simulations of a head-to-head competition, together with the analytical prediction from numerically integrating Eq.(5.21).

5.4 Summary

In this Chapter, we generalize the physically motivated asymptotic analysis to analyze the most general model of competitive dynamics with demographic fluctua-

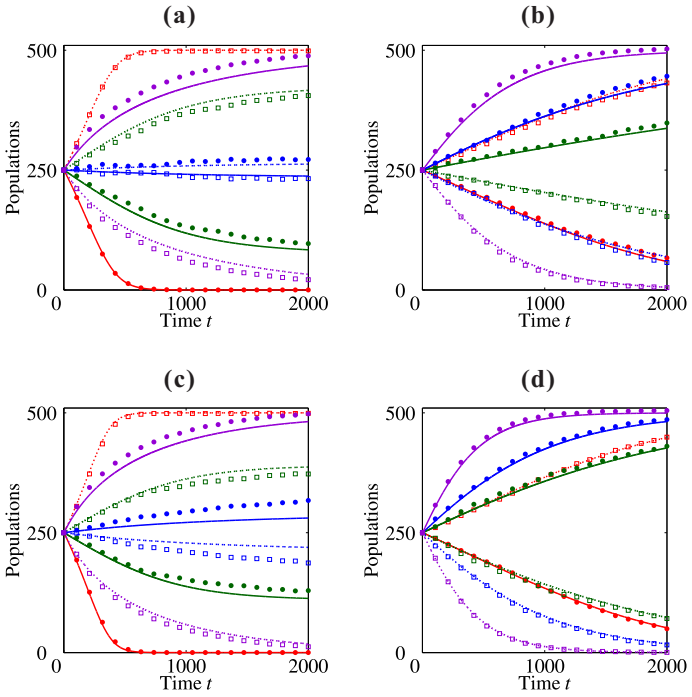

Figure 5.12: σ's are color coded by red (0.01), green (0.09), blue (0.11) and purple (0.15). Open squares and filled circles are respectively the average population per patch of the fast species X and the slow species Y from exact continuous time Markov chain simulations. Dotted and solid lines are respectively the populations of species X and Y derived from the asymptotic prediction. $N = 1000$, $\rho = 2$, $\Lambda = 500$, $\{\sigma_i\}_{i=1}^{N}$ are generated by bounded normal distributions, and 8 sample paths were generated to compute the sample mean. (a) $\mu_x = 2$, $\mu_y = 0.2$. (b) $\mu_x = 5$, $\mu_y = 0.2$. (c) $\mu_x = 2$, $\mu_y = 1$. (d) $\mu_x = 10$, $\mu_y = 2$.

tion in heterogeneous environments. In the two patch model, the fluctuations of the dynamics are large enough for the system to exhibit probabilistic behavior—it is not possible for any of the species to win the competition as $t \to \infty$ with probability 1. On the other hand, the mean-field effect of the many patch model drives the system to behave more "deterministically", even though the effect due to demographic stochasticity is taking into account.

In both the two-patch model and the many patch model, the dynamics show a weak selection to a certain species (unless the two effects balances—see Chapter VI). In general, when the *environmental distribution* is more *heterogeneous*, or when the *population scale* is *larger* (that is, when the population density is more "continuous" and the effect due to *demographic stochasticity* is *smaller*), the *slower* dispersers enjoys the advantage in the competition. On the contrary, when the *environmental distribution* is more *uniform*, or when the *population scale* is *smaller* (that is, the population density is more "discrete" and the *effect of stochasticity* is *larger*), the *faster* dispersers enjoys the advantage in the competition.

The time scale of the selection is proportional to the size of the population per patch—and equivalently in our assumption of time-scale matching, the inverse of the environmental variance. In a large but finite population size, the time scale is long, but exponentially smaller than the time scale of spontaneous extinction (which is a large-deviations phenomenon). The argument shows that the probability that one species exclude the other before a spontaneous extinction event occurs converges to 1 as the population scale increases in the 2-patch model. We will show in Chapters VI and VII that in the many patch model it is possible to reach to a coexistent configuration in a suitable parameter set.

In both models, the critical parameters controlling the transition is identified as $\Lambda\sigma^2 \propto K\sigma^2$. The advantage of our analysis, comparing to the heuristic argument provided by Kessler and Sander [25], is that we have identified the underlying mech-

anism of the dynamics—it is due to the competition between two effects in the center manifold: the drift caused by nonlinearity in heterogeneous models and the drift caused by the demographic stochasticity (with its interaction with the nonlinear flow near the center manifold). As a consequence, in addition to the scaling relation of the regime shift boundary, we are able to predict the winning probability in the two-patch model, and the global dynamics of the many patch model.

CHAPTER VI

Bifurcation Analysis

One of the objectives of this dissertation is to understand how advantages or disadvantages emerge and assert themselves to species with different dispersal rates on an evolutionary time scale. We consider both stochastic and nonlinear effects given the low-density birth-to-death ratio ρ, the population scale $(\rho - 1)\Lambda$, and the environmental variance σ^2. More precisely, in the many patch model, we are interested in the following questions:

- Given ρ, Λ, σ, μ_x, and μ_y, what is (are) the possible solution(s) of Eq.(5.22) as $t \to \infty$?
- As $t \to \infty$, given ρ, Λ and σ^2, what is the "landscape" of selection in the space (μ_x, μ_y) as $t \to \infty$?
- Does the pairwise-competition models exhibit an evolutionary stable dispersal rate? If it does, how does the evolutionarily stable dispersal rate depend upon the parameters?

For the 2-patch model, we can ask similar questions in a probabilistic setting:

- Given ρ, Λ, σ, μ_x, and μ_y, which species has greater probability of winning a competition?

- Given ρ, Λ and σ^2, what is the landscape of the winning probability in the space (μ_x, μ_y)?

- Does the model exhibit an evolutionarily stable dispersal rate (in probability)? If it does, how does the evolutionarily stable dispersal rate depend upon the parameters?

Since we have identified that the effective drifts in the center manifolds are the driving force of the weak selection, the key objective of this Chapter is to investigate the structure of effective drifts and how they parametrically depend upon parameters of the system. In other words, we will perform **bifurcation analysis** of the effective drifts. With knowledge of the effective drifts we can deduce which species has the advantage for any given set of parameters. This analysis is standard, straightforward, but technical; the Chapter also serves for the purpose of documenting details of the computations.

The Chapter is organized as follows. In section 6.1 we identify the problem and show that the qualitative features of the effective drift in the center manifold are determined by four independent conditions. Section 6.2 identifies the sets satisfying each condition in the four-dimensional parameter space, followed by section 6.3 where the properties of the sets are presented. In section 6.4 we combine the results from previous sections in this Chapter describe the bifurcations of the selection landscape in the parameter space and show the existence of an evolutionarily stable rate of dispersion.

6.1 Analysis of the effective drifts in the center manifold

For the reader's reference, we first reproduce the effective drifts in the center manifold of the heterogeneous 2-patch model, Eq.(5.14),

$$v(z) = \frac{1-z^2}{K}\left\{C_0 + \frac{C_1}{(\mu_x-\mu_y)z - C_2} + \frac{C_3}{[(\mu_x-\mu_y)z - C_2]^2}\right\}, \tag{6.1a}$$

and the effective drifts in the center manifold of the heterogeneous many patch model, Eq.(5.22),

$$\bar{v}(\langle z\rangle) = \frac{1-\langle z\rangle^2}{k}\left\{\bar{C}_0 + \frac{\bar{C}_1}{(\mu_x-\mu_y)\langle z\rangle - \bar{C}_2} + \frac{\bar{C}_3}{[(\mu_x-\mu_y)\langle z\rangle - \bar{C}_2]^2}\right\}. \tag{6.1b}$$

The parameters are

$$C_0 = -\frac{1}{2}\frac{\mu_x-\mu_y}{1+2(\mu_x+\mu_y)}, \tag{6.2a}$$

$$C_1 = -\frac{(\mu_x-\mu_y)}{2}\left[\frac{\rho}{\rho-1} + \frac{4\mu_x\mu_y}{1+2(\mu_x+\mu_y)}\right], \tag{6.2b}$$

$$C_2 = 4\mu_x\mu_y + \mu_x + \mu_y, \tag{6.2c}$$

$$C_3 = -4\mu_x\mu_y(\mu_x-\mu_y)K\sigma^2, \tag{6.2d}$$

and

$$\bar{C}_0 = -\frac{1}{2}\frac{\mu_x-\mu_y}{1+\mu_x+\mu_y}, \tag{6.2e}$$

$$\bar{C}_1 = -(\mu_x-\mu_y)\left(\frac{\rho}{\rho-1} + \frac{\mu_x\mu_y}{1+\mu_x+\mu_y}\right), \tag{6.2f}$$

$$\bar{C}_2 = 2\mu_x\mu_y + \mu_x + \mu_y, \tag{6.2g}$$

$$\bar{C}_3 = -2\mu_x\mu_y(\mu_x-\mu_y)K\sigma^2. \tag{6.2h}$$

We now show that it is sufficient to analyze the bifurcation of (6.1b). The effective

drifts of these models, (6.1a) and (6.1b), share similar functional form. In fact, (6.1b) are transformed to (6.1a) by the rescaling of variables

$$\mu_x \to 2\mu_x,\ \mu_y \to 2\mu_y, \text{ and } K \to 2K. \tag{6.3}$$

So we only present the analysis of the effective drift in the many patch model. All the derived conclusions will apply to the 2-patch model after the parameter transformation (6.3).

The reduced dynamics of the many patch model, Eq.(5.21), is an ordinary differential equation. The fate of the system depends on the structure of the one-dimensional velocity field, (6.1b), and the initial condition. For the one-dimensional dynamics, there can be only stable or unstable fixed points (it is impossible to have oscillations in an one-dimensional dynamics [36]). Moreover, it is sufficient to analyze the *sign* of the velocity field (6.1b) if we are interested in the fate of the dynamical system as $t \to \infty$ [36]. Finally, it is clear that the sign of (6.1b) depends only four free parameters: μ_x, μ_y, ρ, and $K \times \sigma$.

Now we can set up the domain of the variable and the parameters. The variable, the effective coordinate $\langle z \rangle$, is in the open interval $(-1, 1)$ because we are not interested in the cases when $\langle z \rangle = \pm 1$ which are the certain fixations of one species or the other. For the parameters, the model imposes the following constraint (see Chapter V):

- $\rho \in (1, \infty)$,
- μ_x and $\mu_y \in (0, \infty)$,
- $K \in [1, \infty)$,
- $\sigma \in [0, 1)$.

Therefore, the parameter space of our interest is

$$(\mu_x, \mu_y, \rho, \xi) \in (0, \infty) \times (0, \infty) \times (1, \infty) \times [0, \infty), \tag{6.4}$$

with

$$\xi \equiv K\sigma^2. \tag{6.5}$$

In addition, we only need to analyze the domain when μ_x is strictly greater than μ_y, since the case $\mu_x < \mu_y$ can be derived by symmetry (that is, the exchange $x \to y$ and $y \to x$.) In conclusion, the parameter space Ω of our interest is

$$(\mu_x, \mu_y, \rho, \xi) \in \Omega \equiv \{(0, \infty) \times (0, \infty) \times (1, \infty) \times [0, \infty) \text{ and } \mu_x > \mu_y\} \tag{6.6}$$

To simplify the expression of the calculations, we will adopt the set-theoretic notation: $\{\mathcal{A} \text{ is true}\} \in \Omega$ is defined to be

$$\{\mathcal{A} \text{ is true}\} \equiv \{(\mu_x, \mu_y, \rho, \xi) \in \Omega \,|\, \mathcal{A} \text{ is true}\}. \tag{6.7}$$

We now demonstrate that the task is equivalent to determining the sign of a quadratic polynomial. Define

$$\varpi\left(q; \mu_x, \mu_y, \rho, \xi\right) \equiv -\left[\frac{1}{1+\mu_x+\mu_y}\frac{q^2}{2} + \left(\frac{\rho}{\rho-1} + \frac{\mu_x\,\mu_y}{1+\mu_x+\mu_y}\right) q + 2\,\mu_x\,\mu_y\,\xi\right] \tag{6.8}$$

with

$$q\left(\langle z \rangle\right) \equiv \left(\mu_x - \mu_y\right) \langle z \rangle - \bar{C}_2. \tag{6.9}$$

It is straightforward to show

$$\{\varpi(q) > 0\} = \{\bar{v}(\langle z \rangle) > 0\}, \tag{6.10a}$$

$$\{\varpi(q) = 0\} = \{\bar{v}(\langle z \rangle) = 0\}, \tag{6.10b}$$

$$\{\varpi(q) < 0\} = \{\bar{v}(\langle z \rangle) < 0\}, \tag{6.10c}$$

since in Ω,

$$\frac{1 - \langle z \rangle^2}{K} \frac{\mu_x - \mu_y}{\left[(\mu_x - \mu_y)\langle z \rangle - \bar{C}_2\right]^2} > 0. \tag{6.11}$$

In addition, $(\mu_x - \mu_y)\langle z \rangle - \bar{C}_2$ in the denominator of (6.11) is never 0 for $\langle z \rangle \in (0, 1)$. We will analyze ϖ instead of $\bar{v}$ because ϖ is a quadratic polynomial in q, and it is elementary to analyze the sign of a quadratic function.

The domain of q is obtained directly from (6.9):

$$q \in (-2\mu_x(\mu_y + 1), -2\mu_y(\mu_x + 1)) \equiv (q_{\text{left}}, q_{\text{right}}). \tag{6.12}$$

where q_{left} and q_{right} are respectively the lower and upper bound of the interval. We remind the reader the $q_{\text{left}} < q_{\text{right}}$ only if $\mu_x > \mu_y$. In addition, ϖ is concave in q because the leading order has a negative coefficient. The strategy to analyze the concave function ϖ is therefore as follows. We first enumerate the possible scenarios of the sign change(s) of the concave ϖ in the domain $(q_{\text{left}}, q_{\text{right}})$:

S1 ϖ is *always* positive.

S2 ϖ changes its sign once in $q \in (q_{\text{left}}, q_{\text{right}})$; *in addition*, $\varpi(q_{\text{left}}) < 0$ and $\varpi(q_{\text{right}}) > 0$.

S3 ϖ changes its sign once in $q \in (q_{\text{left}}, q_{\text{right}})$; *in addition*, $\varpi(q_{\text{left}}) > 0$ and $\varpi(q_{\text{right}}) < 0$.

S4 ϖ changes its sign twice in $q \in (q_{\text{left}}, q_{\text{right}})$.

S5 ϖ is *always* negative.

Then, these scenarios (S1-S5) respectively indicate the following physical interpretation

S1 $\bar{v}(\langle z \rangle)$ has one unstable fixed point at $\langle z \rangle = -1$ and one stable fixed point at $\langle z \rangle = 1$.

S2 $\bar{v}(\langle z \rangle)$ has two stable fixed point at $\langle z \rangle = \pm 1$ and one unstable fixed point in $\langle z \rangle \in (-1, 1)$.

S3 $\bar{v}(\langle z \rangle)$ has two unstable fixed point at $\langle z \rangle = \pm 1$ and one stable fixed point in $\langle z \rangle \in (-1, 1)$.

S4 $\bar{v}(\langle z \rangle)$ has four fixed points: stable $\langle z \rangle = -1$, unstable $\langle z \rangle = z_1$, stable $\langle z \rangle = z_2$, and unstable $z = +1$ such that $-1 \leq z_1 \leq z_2 \leq 1$.

S5 $\bar{v}(\langle z \rangle)$ has one stable fixed point at $\langle z \rangle = -1$ and one unstable fixed point at $\langle z \rangle = 1$.

Fig. 6.1 shows the effective drift in each scenario.

Bifurcations occurs between the transitions of the scenarios. For example, in Ω, crossing the "boarder" from scenario 4 to scenario 5 represents the transition where two roots z_1 and z_2 in scenario 4 approaches and annihilates each other—that is, a typical saddle–node bifurcation.

Lastly, we define the maxima of the polynomial ϖ to be $\varpi_{\max}$ and $\varpi(q_{\max}) = \varpi_{\max}$ for simplicity. Then we can finally formulate the criterion of the sets which correspond above scenarios respectively:

S1 $\{\varpi(q_{\text{left}}; \mu_x, \mu_y, \rho, \xi) > 0\} \cap \{\varpi(q_{\text{right}}; \mu_x, \mu_y, \rho, \xi) > 0\}$.

S2 $\{\varpi(q_{\text{left}}; \mu_x, \mu_y, \rho, \xi) < 0\} \cap \{\varpi(q_{\text{right}}; \mu_x, \mu_y, \rho, \xi) > 0\}$.

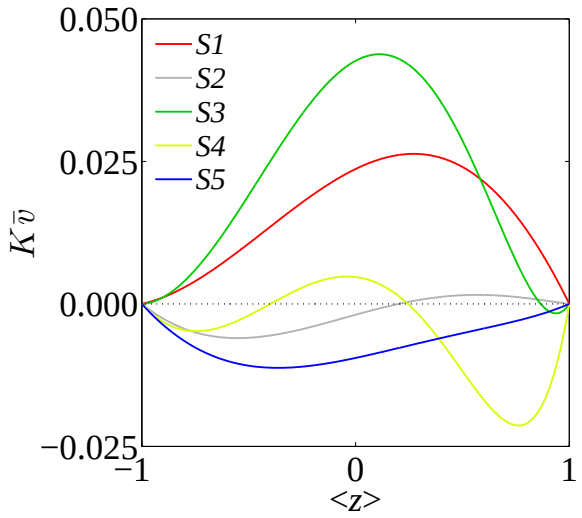

Figure 6.1: Reduced velocity fields of the system in different scenarios with $\rho = 2, \Lambda = 500$. $S1 : \mu_x = 10, \mu_y = 2, \sigma = 6.00\%$, $S2 : \mu_x = 10, \mu_y = 2, \sigma = 6.35\%$, $S3 : \mu_x = 40, \mu_y = 1, \sigma = 6.5\%$, $S4 : \mu_x = 40, \mu_y = 1, \sigma = 6.72\%$, and $S5 : \mu_x = 10, \mu_y = 2, \sigma = 6.45\%$. Dotted line is the reference $\bar{v} = 0$.

S3 $\{\varpi(q_{\text{left}}; \mu_x, \mu_y, \rho, \xi) > 0\} \cap \{\varpi(q_{\text{right}}; \mu_x, \mu_y, \rho, \xi) < 0\}$.

S4 $\{\varpi(q_{\text{left}}; \mu_x, \mu_y, \rho, \xi) < 0\} \cap \{\varpi(q_{\text{right}}; \mu_x, \mu_y, \rho, \xi) < 0\} \cap \{q_{\max} \in (q_{\text{left}}, q_{\text{right}})\} \cap \{\varpi_{\max} > 0\}$.

S5 The complement set of the above sets in Ω.

As a consequence, in section 5.2 we investigate how each pair of the sets

- $\{\varpi(q_{\text{right}}; \mu_x, \mu_y, \rho, \xi) \leq 0\}$ and $\{\varpi(q_{\text{right}}; \mu_x, \mu_y, \rho, \xi) \geq 0\}$,
- $\{\varpi(q_{\text{left}}; \mu_x, \mu_y, \rho, \xi) \leq 0\}$ and $\{\varpi(q_{\text{left}}; \mu_x, \mu_y, \rho, \xi) \geq 0\}$,
- $\{q_{\max} \in (q_{\text{left}}, q_{\text{right}})\}$ and $\{q_{\max} \notin (q_{\text{left}}, q_{\text{right}})\}$,
- $\{\varpi_{\max} \geq 0\}$ and $\{\varpi_{\max} \leq 0\}$

partition the entire parameter space Ω.

6.2 Identifying the sets in the parameter space

The fact that the parameter space Ω is four dimensional poses difficulty for visualization. Nevertheless, from the perspective of searching for the evolutionarily stable dispersal rate, it is natural to treat the parameter ρ as fixed, the parameter $\xi \equiv K\sigma^2$ as the control parameter, and to analyze how the sets partition the reduced two-dimensional μ-space, i.e. $(\mu_x, \mu_y) \in (0,\infty) \times (0,\infty)$.

Note regarding notation: In the following sections, we will use a short hand notation $\varpi(q)$ to denote $\varpi(q; \mu_x, \mu_y, \rho, \xi)$, and ϕ to denote the empty set.

6.2.1 $\{\varpi(q_{\text{right}}; \mu_x, \mu_y, \rho, \xi) \leq 0\}$ and $\{\varpi(q_{\text{right}}; \mu_x, \mu_y, \rho, \xi) \geq 0\}$

Let S_1 be the set $\{\varpi(q_{\text{right}}) \leq 0\} \cap \{(\mu_x, \mu_y) \in (0,\infty) \times (0,\infty)\}$. After some algebra, it can be shown that

$$S_1 = \left\{ 2(\mu_x+1)^2\mu_y^2 - 2\left[\frac{\rho}{\rho-1}(1+\mu_x+\mu_y) + \mu_x\,\mu_y\right](\mu_x+1)\,\mu_y \right. \\ \left. +2\mu_x\,\mu_y(1+\mu_x+\mu_y)\xi > 0 \right\} \tag{6.13}$$

and we notice in Ω, $\mu_y > 0$, therefore

$$S_1 = \left\{ 2(\mu_x+1)^2\mu_y - 2\left[\frac{\rho}{\rho-1}(1+\mu_x+\mu_y) + \mu_x\,\mu_y\right](\mu_x+1) \right. \\ \left. +2\mu_x\,(1+\mu_x+\mu_y)\xi > 0 \right\} \tag{6.14}$$

and further computations yield

$$S_1 = \{[1+\mu_x-(\rho-1)\mu_x\,\xi]\,\mu_y \leq [(\rho-1)\mu_x\,\xi - \rho(1+\mu_x)]\,(1+\mu_x)\}\,. \tag{6.15}$$

Depending on the value of ξ, the polynomial $1+\mu_x-(\rho-1)\mu_x\,\xi$ may change its sign and reverse the direction of the inequality in the end. The possibilities and the

corresponding implications are:

- $\xi \leq 1/(\rho - 1)$:

$$S_1 = \phi. \tag{6.16}$$

because $\mu_y > 0$.

- $1/(\rho - 1) < \xi \leq \rho/(\rho - 1)$:

$$S_1 = \left\{ \frac{1}{(\rho - 1)\xi - 1} < \mu_x \right\} \cap \left\{ \mu_y \geq \Gamma_1 \right\} \tag{6.17}$$

where Γ_1 is defined to be

$$\Gamma_1(\mu_x, \rho, \xi) \equiv \frac{(\rho - 1)\mu_x \xi - \rho(1 + \mu_x)}{1 + \mu_x - (\rho - 1)\mu_x \xi}(1 + \mu_x). \tag{6.18}$$

- $\rho/(\rho - 1) < \xi$:

$$\begin{aligned} S_1 = & \left(\left\{ \frac{1}{(\rho - 1)\xi - \rho} < \mu_x < \frac{\rho}{(\rho - 1)\xi - 1} \right\} \cap \left\{ \mu_y \geq \Gamma_1 \right\} \right) \\ & \cup \left\{ \frac{\rho}{(\rho - 1)\xi - \rho} \leq \mu_x \right\} \end{aligned} \tag{6.19}$$

A parallel computation shows the representation of $S_2 \equiv \{\varpi(q_{\text{right}}) \geq 0\} \cap \{(\mu_x, \mu_y) \in (0, \infty) \times (0, \infty)\}$ in each case.

- $\xi \leq 1/(\rho - 1)$:

$$S_2 = \{(\mu_x, \mu_y) \in (0,\infty) \times (0,\infty)\}\,. \tag{6.20}$$

- $1/(\rho-1) < \xi \leq \rho/(\rho-1)$:

$$S_2 = \left\{\mu_x \leq \frac{1}{(\rho-1)\xi - 1}\right\} \cup \left(\left\{\frac{1}{(\rho-1)\xi-1} < \mu_x\right\} \cap \left\{\mu_y \leq \Gamma_1\right\}\right). \tag{6.21}$$

- $\rho/(\rho-1) < \xi$:

$$S_2 = \left\{\mu_x \leq \frac{1}{(\rho-1)\xi - 1}\right\} \cup \left(\left\{\frac{1}{(\rho-1)\xi-1} < \mu_x < \frac{\rho}{(\rho-1)\xi-1}\right\} \cap \left\{\mu_y \leq \Gamma_1\right\}\right) \tag{6.22}$$

A more transparent way is presented in Fig. 6.2. When $\xi \leq 1/(\rho-1)$, $\varpi(q_{\text{right}})$ is positive in the entire μ–space. As soon as $\xi > 1/(\rho-1)$, the curve $\{\mu_y = \Gamma_1, \mu_y > 0\}$ separates the μ–space such that $\varpi(q_{\text{right}})$ becomes negative in the region above $\{\mu_y = \Gamma_1, \mu_x > 1/((\rho-1)\xi - 1)\}$ and remains positive in the region below. When $\xi > \rho/(\rho-1)$, $\{\mu_y = \Gamma_1, \mu_x > 1/((\rho-1)\xi-1)\}$ intersects with the μ_x–axis, hence $\varpi(q_{\text{right}})$ is always negative to the right of the intersection, $\mu_x = \rho/((\rho-1)\xi - \rho)$.

Finally, we point out that the set $\{\varpi(q_{\text{right}}) = 0\} = \{S_1 \cap S_2\}$ is empty when $\xi \leq 1/(\rho-1)$, and $\{\mu_y = \Gamma_1, \mu_x > 1/((\rho-1)\xi-1)\}$ when $\xi > 1/(\rho-1)$.

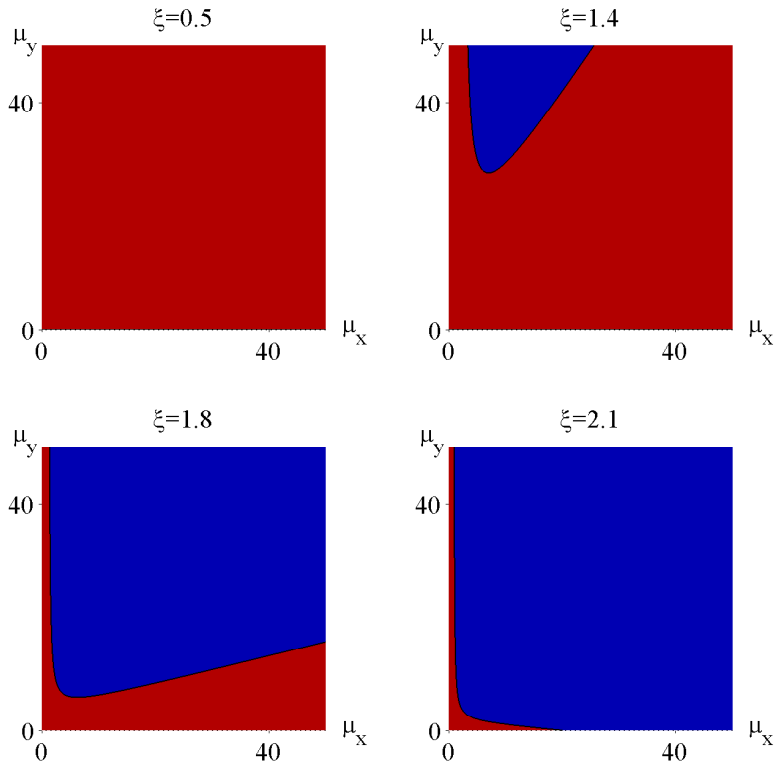

Figure 6.2: The landscape of the sign of $\varpi(q_{\text{right}})$ in the μ-space, with different values of the control parameter ξ. The sets $\{\varpi(q_{\text{right}}; \mu_x, \mu_y, \rho, \xi) > 0\}$ and $\{\varpi(q_{\text{right}}; \mu_x, \mu_y, \rho, \xi) < 0\}$ are plotted as red and blue respectively. $\rho = 2$, and the black boundary is $\{\mu_y = \Gamma_1(\mu_x, \rho, \xi), \mu_x > 1/((\rho-1)\xi - 1)\}$.

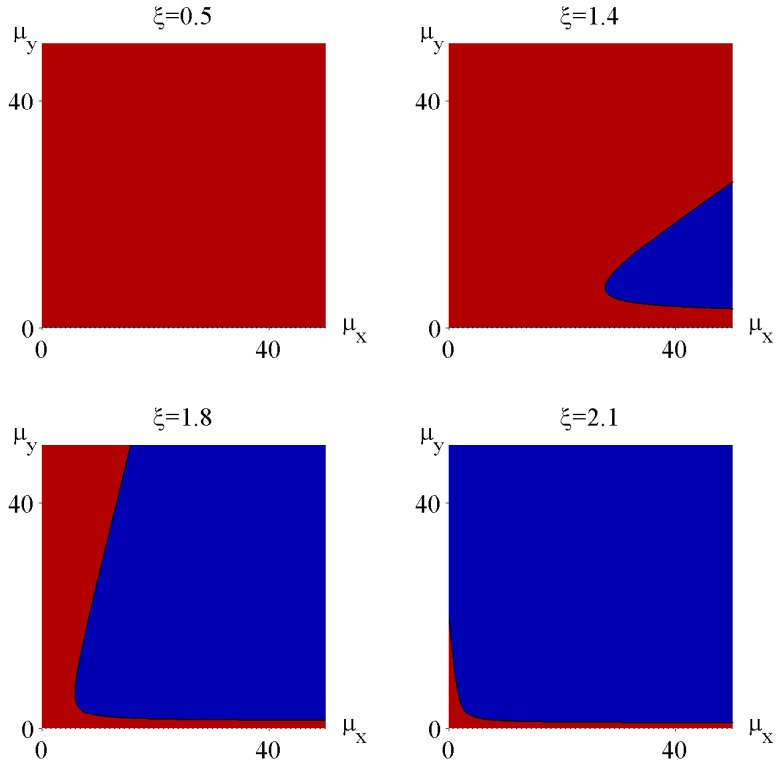

Figure 6.3: The landscape of the sign of $\varpi(q_{\text{left}})$ in the μ-space, with different values of the control parameter ξ. The sets $\{\varpi(q_{\text{left}}; \mu_x, \mu_y, \rho, \xi) > 0\}$ and $\{\varpi(q_{\text{left}}; \mu_x, \mu_y, \rho, \xi) < 0\}$ are plotted as red and blue respectively. $\rho = 2$, and the black boundary is $\{\mu_x = \Gamma_2(\mu_y, \rho, \xi), \mu_y > 1/((\rho-1)\xi - 1)\}$.

6.2.2 $\{\varpi(q_{\text{left}}; \mu_x, \mu_y, \rho, \xi) \geq 0\}$ and $\{\varpi(q_{\text{left}}; \mu_x, \mu_y, \rho, \xi) \leq 0\}$

It is not necessary to perform parallel computation for sets $\{\varpi(q_{\text{left}}; \mu_x, \mu_y, \rho, \xi) \geq 0\}$ and $\{\varpi(q_{\text{left}}; \mu_x, \mu_y, \rho, \xi) \leq 0\}$ once we realize the symmetry relation

$$\varpi(q_{\text{left}}; \mu_x, \mu_y, \rho, \xi) = \varpi(q_{\text{right}}; \mu_y, \mu_x, \rho, \xi); \tag{6.23}$$

In Fig. 6.3 we show the graphical representation of the sets. Define

$$\Gamma_2(\mu_y, \rho, \xi) \equiv \frac{(\rho-1)\mu_y\,\xi - \rho(1+\mu_y)}{1+\mu_y-(\rho-1)\mu_y\,\xi}(1+\mu_y). \tag{6.24}$$

When $\xi \leq 1/(\rho-1)$, $\varpi(q_{\text{left}})$ is positive in the entire μ–space. As soon as $\xi > 1/(\rho-1)$, the curve $\{\mu_x = \Gamma_2, \mu_x > 0\}$ separates the μ–space such that $\varpi(q_{\text{left}})$ become negative in the region above $\{\mu_x = \Gamma_2, \mu_y > 1/((\rho-1)\xi-1)\}$ and remains positive in the region below. When $\xi > \rho/(\rho-1)$, $\{\mu_x = \Gamma_2, \mu_y > 1/((\rho-1)\xi-1)\}$ intersects with the μ_y–axis, hence $\varpi(q_{\text{left}})$ is always negative above the intersection $\mu_y = \rho/((\rho-1)\xi-\rho)$.

The set $\{\varpi(q_{\text{left}}) = 0\}$ is empty when $\xi \leq 1/(\rho-1)$, and $\{\mu_x = \Gamma_2 = 0\}$ when $\xi > 1/(\rho-1)$.

6.2.3 $\{q_{\max} \in (q_{\text{left}}, q_{\text{right}})\}$ and $\{q_{\max} \notin (q_{\text{left}}, q_{\text{right}})\}$

We focus on the set $S_3 \equiv \{q_{\max} \in (q_{\text{left}}, q_{\text{right}})\}$ in this section (since $\{q_{\max} \notin (q_{\text{left}}, q_{\text{right}})\}$ is the complement set). First, it is elementary to show the maxima of the polynomial $\varpi(q)$ is at

$$q = q_{\max} = -\left[\frac{\rho}{\rho-1}(1+\mu_x+\mu_y)+\mu_x\,\mu_y\right], \tag{6.25}$$

therefore the set $\{q_{\max} \in (q_{\text{left}}, q_{\text{right}})\}$ is equivalent to

$$\begin{aligned}\{q_{\max} \in (q_{\text{left}}, q_{\text{right}})\} = &\left\{-q_{\text{left}} > \frac{\rho}{\rho-1}(1+\mu_x+\mu_y)+\mu_x\,\mu_y\right\} \\ &\cap \left\{\frac{\rho}{\rho-1}(1+\mu_x+\mu_y)+\mu_x\,\mu_y > -q_{\text{right}}\right\}\end{aligned} \tag{6.26}$$

For the reference of the reader, we reproduce the definition of q_{left} and q_{right} in (6.12):

$$\begin{aligned} q_{\text{left}} &= -2\mu_x(1+\mu_y) \\ q_{\text{right}} &= -2\mu_y(1+\mu_x) \end{aligned}$$

where $\mu_x > \mu_y$.

A little algebra establishes the following relations

$$\left\{-q_{\text{left}} > \frac{\rho(1+\mu_x+\mu_y)}{\rho-1} + \mu_x\,\mu_y\right\} = \left\{\left(\frac{\rho-2}{\rho-1}+\mu_y\right)\mu_x > \frac{\rho(1+\mu_y)}{\rho-1}\right\} \tag{6.27}$$

$$\left\{\frac{\rho(1+\mu_x+\mu_y)}{\rho-1} + \mu_x\,\mu_y > -q_{\text{right}}\right\} = \left\{\left(\frac{\rho-2}{\rho-1}+\mu_x\right)\mu_y < \frac{\rho(1+\mu_x)}{\rho-1}\right\} \tag{6.28}$$

Depending on the signs of $(\rho-2)/(\rho-1)+\mu_x$ and $(\rho-2)/(\rho-1)+\mu_y$, the directions of the inequalities may be reversed.

To proceed with the analysis we observe an important symmetry: if we swap μ_x and μ_y in the set in Eq.(6.27) *and* reverse the direction of the inequality, we obtain the set in Eq.(6.28). Therefore, it is sufficient to analyze the first set (6.27). We enumerate all possibilities in the following list:

- $\rho > 2$:

  In this case $(\rho-2)/(\rho-1) > 0$ and it is straightforward to obtain

  $$\left\{\left(\frac{\rho-2}{\rho-1}+\mu_y\right)\mu_x > \frac{\rho}{\rho-1}(1+\mu_y)\right\} = \left\{\mu_x > \Gamma_3\right\}, \tag{6.29}$$

  where

  $$\Gamma_3(\mu_y,\rho) \equiv \frac{\rho(1+\mu_y)}{\rho-2+(\rho-1)\mu_y}. \tag{6.30}$$

- $1 < \rho \leq 2$, or equivalently $(\rho-2)/(\rho-1) \leq 0$:

  - When $\mu_y > -(\rho-2)/(\rho-1)$,

    $$\left\{\left(\frac{\rho-2}{\rho-1}+\mu_y\right)\mu_x > \frac{\rho}{\rho-1}(1+\mu_y)\right\} = \left\{\mu_x > \Gamma_3\right\}, \tag{6.31}$$

– When $\mu_y \leq -(\rho-2)/(\rho-1)$,

$$\left\{\left(\frac{\rho-2}{\rho-1}+\mu_y\right)\mu_x > \frac{\rho}{\rho-1}(1+\mu_y)\right\} = \phi \tag{6.32}$$

because $\mu_x,\ \mu_y > 0$ in Ω.

Regardless of the value of ρ, the above relations can be organized into a compact form

$$\left\{\left(\frac{\rho-2}{\rho-1}+\mu_y\right)\mu_x > \frac{\rho}{\rho-1}(1+\mu_y)\right\} = \left\{\mu_x > \Gamma_3, \mu_y > \max\left\{0, \frac{2-\rho}{\rho-1}\right\}\right\}. \tag{6.33}$$

It is clear that $\{\mu_x = \Gamma_3, \mu_y > \max\{0, (2-\rho)/(\rho-1)\}\}$ is the boundary that divide the μ-space. To the right of $\mu_x = \Gamma_3$ is the set

$$\left\{-q_{\text{left}} > \frac{\rho(1+\mu_x+\mu_y)}{\rho-1} + \mu_x\,\mu_y\right\}$$

and to the left is the set

$$\left\{-q_{\text{left}} < \frac{\rho(1+\mu_x+\mu_y)}{\rho-1} + \mu_x\,\mu_y\right\}.$$

The symmetry argument prompts us to define the boundary

$$\left\{\mu_y = \Gamma_4, \mu_y > \max\{0, \frac{2-\rho}{\rho-1}\}\right\} \tag{6.34}$$

with

$$\Gamma_4(\mu_x, \rho) \equiv \frac{\rho(1+\mu_x)}{\rho-2+(\rho-1)\mu_x}. \tag{6.35}$$

Above $\{\mu_y = \Gamma_4, \mu_x > \max\{0, (2-\rho)/(\rho-1)\}\}$ is the set

$$\left\{\frac{\rho(1+\mu_x+\mu_y)}{\rho-1} + \mu_x\,\mu_y < -q_{\text{right}}\right\}$$

and below to the boundary is the set

$$\left\{ \frac{\rho(1+\mu_x+\mu_y)}{\rho-1} + \mu_x\,\mu_y > -q_{\text{right}} \right\}.$$

The analysis in this section focuses on $\mu_x > \mu_y$, therefore, the conclusion only applies to the half quadrant. The conclusion of the other half quadrant will follow naturally from the reflection $\mu_x \to \mu_y$ and $\mu_y \to \mu_x$, that is, the mirror image with respect to $\{\mu_x = \mu_y\}$.

We now prove a theorem asserting the set $\{q_{\max} \in (q_{\text{left}}, q_{\text{right}})\}$ is not empty.

Lemma VI.1. *Γ_4 and Γ_3 are monotonic decreasing functions of μ_x and μ_y, respectively.*

Proof.

$$\frac{d\,\Gamma_4}{d\,\mu_x} = \frac{-\rho}{(\rho-1)^2(\frac{\rho}{\rho-1}-\mu_x)^2} < 0. \tag{6.36}$$

Similarly,

$$\frac{d\,\Gamma_3}{d\,\mu_y} = \frac{-\rho}{(\rho-1)^2(\frac{\rho}{\rho-1}-\mu_y)^2} < 0. \tag{6.37}$$

□

Theorem VI.2. *$\{q_{\max} \in (q_{\text{left}}, q_{\text{right}})\}$ is not empty (with $\rho > 1$).*

Proof. We consider the curves $\mu_x = \Gamma_3$ and $\mu_y = \Gamma_4$. The intersections of the sets can be obtained by solving their roots, and the results are

$$\begin{aligned}\{\mu_x = \Gamma_3\} \cap \{\mu_y = \Gamma_4\} = &\left\{ \mu_x = \mu_y = \frac{1+\sqrt{1+\rho(\rho-1)}}{\rho-1} \right\} \\ &\cup \left\{ \mu_x = \mu_y = \frac{1-\sqrt{1+\rho(\rho-1)}}{\rho-1} \right\}.\end{aligned} \tag{6.38}$$

Nevertheless the second set is not in the domain of our interests, since when $\rho > 1$,

$$\frac{1 - \sqrt{1 - \rho(\rho - 1)}}{\rho - 1} < 0 \tag{6.39}$$

but μ_x and μ_y are strictly positive. Next, it is straightforward to show

$$\left\{\mu_x = \Gamma_3\right\} = \left\{\mu_y = \frac{(\rho - 2)\mu_x - \rho}{\rho - (\rho - 1)\mu_x}\right\}. \tag{6.40}$$

Comparing the asymptotic behavior as $\mu_x \to \infty$,

$$\lim_{\mu_x \to \infty} \frac{(\rho - 2)\mu_x - \rho}{\rho - (\rho - 1)\mu_x} = \frac{2 - \rho}{\rho - 1}, \tag{6.41}$$

$$\lim_{\mu_x \to \infty} \frac{\rho(1 + \mu_x)}{\rho - 2 + (\rho - 1)\mu_x} = \frac{\rho}{\rho - 1}, \tag{6.42}$$

and when $\rho > 1$,

$$\frac{2 - \rho}{\rho - 1} < \frac{\rho}{\rho - 1}. \tag{6.43}$$

This shows the curve $\mu_x = \Gamma_3$ is below the curve $\mu_y = \Gamma_4$ as $\mu_x \to \infty$. On the other hand, they have at most one intersection at

$$\mu_x = \mu_y = \frac{1 + \sqrt{1 + \rho(\rho - 1)}}{\rho - 1}, \tag{6.44}$$

we deduce that *for any* $\mu_x > \mu_y$, the curve $\mu_x = \Gamma_3$ is below the curve $\mu_y = \Gamma_4$. Finally it is clear that

$$\{q_{\max} \in (q_{\text{left}}, q_{\text{right}})\} = \left\{\mu_x > \Gamma_3, \mu_y > \max\left\{0, \frac{2 - \rho}{\rho - 1}\right\}\right\} \cap \{\mu_y < \Gamma_4\} \tag{6.45}$$

is not empty. □

It is again more transparent to present the conclusion graphically in Fig. 6.5. The

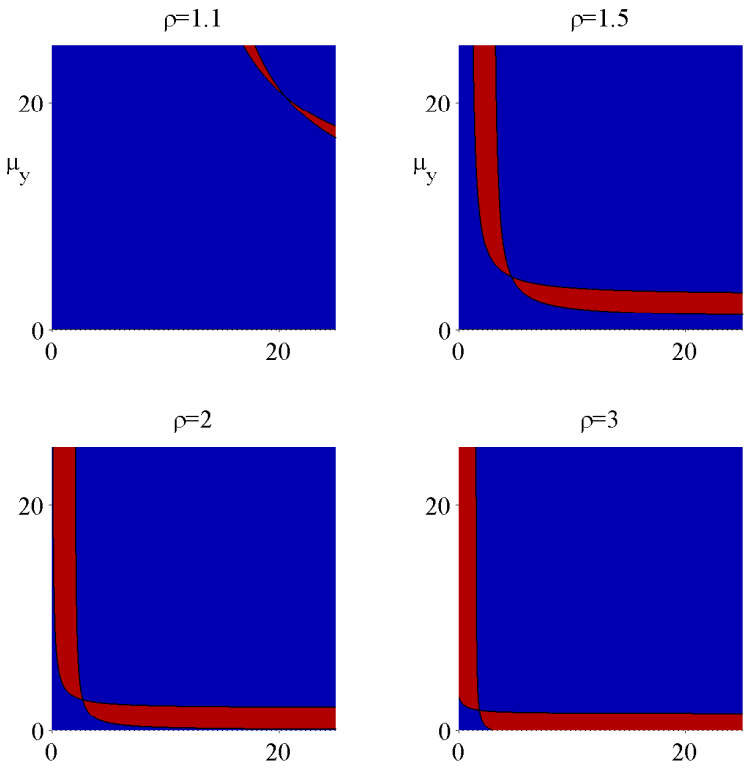

Figure 6.4: The landscape of the sets $\{q_{\max} \in (q_{\text{left}}, q_{\text{right}})\}$—plotted in blue—and $\{q_{\max} \notin [q_{\text{left}}, q_{\text{right}}]\}$—plotted in red—in the μ-space, with different values of ρ. The landscape does not involve in ξ. The black line represents the boundaries $\{\mu_x = \Gamma_3(\mu_y, \rho), \mu_y > \max\{0, (2-\rho)/(\rho-1)\}\}$ and $\{\mu_y = \Gamma_4(\mu_x, \rho), \mu_x > \max\{0, (2-\rho)/(\rho-1)\}\}$.

set $S_3 \equiv \{q_{\max} \in (q_{\text{left}}, q_{\text{right}})\}$ does not depend on the parameter ξ—instead, the set S_3 in the μ-space depends only on the value of ρ. It is clear that S_3 is the intersection of the open set in between the upper-right branches of the hyperbolas $\mu_x = \Gamma_3$ and $\mu_y = \Gamma_4$, and the open quadrant $\{\mu_x > 0, \mu_y > 0\}$.

6.2.4 $\{\varpi_{\max} \geq 0\}$ and $\{\varpi_{\max} \leq 0\}$

The analysis starts with two simple lemma.

Lemma VI.3. *Given real-valued $A, B > 0$, and $\zeta \leq 2$. Then*

$$(A + B)^2 - 2\zeta AB \geq 0.$$

Proof. By completing the square

$$A^2 + 2AB + B^2 - 2\zeta AB \geq A^2 + 2AB + B^2 - 4AB = (A - B)^2 \geq 0. \tag{6.46}$$

□

Lemma VI.4. *Given real-valued $A, B > 0$, and $\zeta > 2$. Then the set*

$$\left\{(A, B) \in \mathbb{R}^+ \times \mathbb{R}^+ | (A + B)^2 - 2\zeta AB > 0\right\}$$

is equal to

$$\left\{(A, B) \in \mathbb{R}^+ \times \mathbb{R}^+ | B - A\nu_+ > 0\right\} \cap \left\{(A, B) \in \mathbb{R}^+ \times \mathbb{R}^+ | B - A\nu_- < 0\right\}$$

with

$$\nu_\pm = \zeta - 1 \pm \sqrt{\zeta^2 - 2\zeta} > 0.$$

Proof.

$$\left\{(A+B)^2 - 2\zeta AB > 0\right\} = \left\{1 + \left(\frac{B}{A}\right)^2 + 2(1-\zeta)\frac{A}{B} > 0\right\} \tag{6.47}$$
$$= \left\{\left(\frac{B}{A} - \nu_+\right)\left(\frac{B}{A} - \nu_-\right) > 0\right\}$$
$$= \{B - A\nu_+ > 0\} \cap \{B - A\nu_- < 0\}$$

□

Next we identify the set $\{\varpi_{\max} > 0\}$. It is elementary to show

$$\varpi_{\max} = \varpi(q_{\max}) = \left[\frac{\rho}{\rho-1}(1+\mu_x+\mu_y) + \mu_x\,\mu_y\right]^2 - 4\mu_x\,\mu_y(1+\mu_x+\mu_y)\xi. \tag{6.48}$$

We identify that

$$A \equiv \frac{\rho}{\rho-1}(1+\mu_x+\mu_y) > 0, \tag{6.49a}$$

$$B \equiv \mu_x\,\mu_y > 0, \tag{6.49b}$$

$$\zeta \equiv 2\,\frac{\rho-1}{\rho}\xi > 0, \tag{6.49c}$$

$$\nu_\pm = 2\,\frac{\rho-1}{\rho}\xi - 1 \pm \sqrt{\left(2\,\frac{\rho-1}{\rho}\xi\right)^2 - 4\frac{\rho-1}{\rho}\xi} \tag{6.49d}$$

and then the lemmas yield

- When $\zeta \le 2$, i.e., when $\xi \le \rho/(\rho-1)$, $\varpi_{\max}$ is always non-negative. In addition, it is straightforward to show the equality only holds on the set

$$\left\{\xi = \frac{\rho}{\rho-1}\right\} \cap \left\{\mu_x > \frac{\rho}{\rho-1}\right\} \cap \left\{\mu_y = \frac{\rho(1+\mu_x)}{(\rho-1)\mu_x - \rho}\right\}. \tag{6.50}$$

- When $\zeta > 2$, i.e., when $\xi > \rho/(\rho-1)$,

$$\{\varpi_{\max} > 0\} = \{A > B\nu_+\} \cap \{A < B\nu_-\}\,. \tag{6.51}$$

After some computations, one can show the equalities

$$\left\{A > B\nu_+\right\} = \left\{\mu_x \le \frac{\rho}{\rho-1}\nu_-\right\} \cup \left(\left\{\mu_x > \frac{\rho}{\rho-1}\nu_-\right\} \cap \left\{\mu_y < \Gamma_5\right\}\right), \tag{6.52}$$

$$\left\{A < B\nu_-\right\} = \left(\left\{\mu_x > \frac{\rho}{\rho-1}\nu_+\right\} \cap \left\{\mu_y > \Gamma_6\right\}\right). \tag{6.53}$$

where Γ_5 and Γ_6 are define respectively

$$\Gamma_5(\mu_x, \rho, \xi) \equiv \frac{\rho\,(1+\mu_x)}{(\rho-1)\,\nu_+\,\mu_x - \rho}, \tag{6.54}$$

$$\Gamma_6(\mu_x, \rho, \xi) \equiv \frac{\rho\,(1+\mu_x)}{(\rho-1)\,\nu_-\,\mu_x - \rho}. \tag{6.55}$$

Furthermore, the boundaries $\{\mu_y = \Gamma_5\}$ and $\{\mu_y = \Gamma_6\}$ are monotonic decreasing function of μ_x:

$$\frac{d\,\Gamma_{5,6}}{d\,\mu_x} = -\rho^2 \frac{1}{\left[(\rho-1)\nu_{+,-}\,\mu_x - \rho\right]^2} < 0. \tag{6.56}$$

Fig. 6.5 shows the sets in the μ-space. When $\zeta \le 2$, i.e., when $\xi \le \rho/(\rho-1)$, $\varpi_{\max}$ is non-negative in the entire μ-space; it is zero only on the curve $\{\mu_x > \rho/(\rho-1), \mu_y = \rho(1+\mu_x)/((\rho-1)\mu_x - \rho)$ when $\zeta = 2$. As soon as ζ is greater than 2, the curve splits into two,

$$\left\{\mu_y > \frac{\rho}{\rho-1}\nu_-, \mu_y = \Gamma_5\right\} \text{ and } \left\{\mu_y > \frac{\rho}{\rho-1}\nu_+, \mu_y = \Gamma_6\right\} \tag{6.57a}$$

and $\varpi_{\max} < 0$ in the open gap in between the curves, and remains positive elsewhere. A final remark on these two curves is that they are symmetrical with respect to the

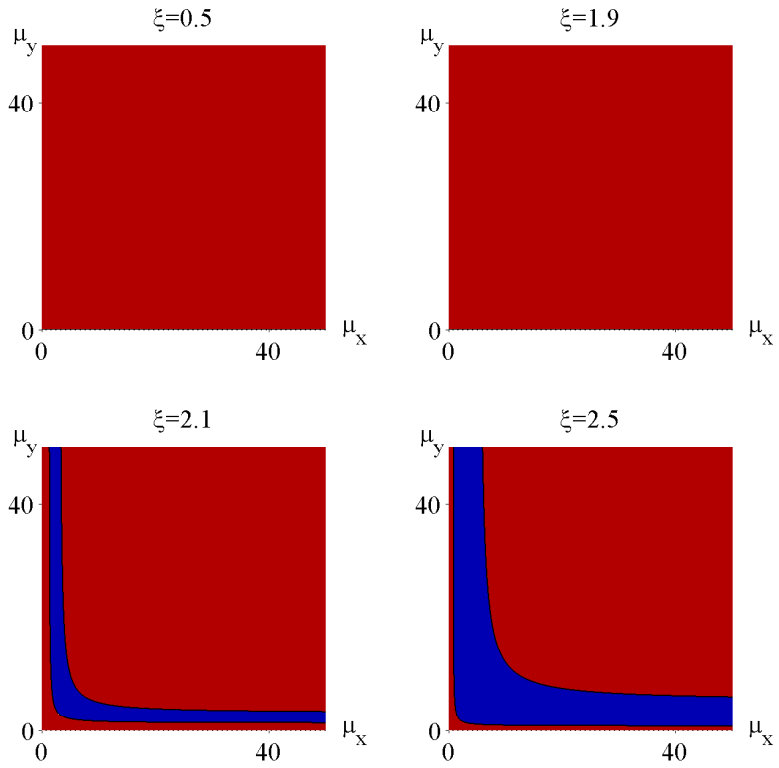

Figure 6.5: The landscape of the sign of $\varpi_{\max}$ in the μ-space, with different values of the control parameter ξ. $\varpi_{\max}$ is positive in the red region, and negative in the blue. $\rho = 2$, and the black line represents the boundaries $\{\mu_y > \rho\nu_-/(\rho-1), \mu_y = \Gamma_5\}$ and $\{\mu_y > \rho\nu_+/(\rho-1), \mu_y = \Gamma_6\}$.

Curve	Set	Definition of the boundary
$\mathcal{C}_1$	$\left\{\mu_x > \frac{\rho}{(\rho-1)\xi-1}, \mu_y > 0, \mu_y = \Gamma_1\right\}$	$\Gamma_1(\mu_x,\rho,\xi) \equiv \frac{(\rho-1)\mu_x\,\xi-\rho(1+\mu_x)}{1+\mu_x-(\rho-1)\mu_x\,\xi}(1+\mu_x)$
$\mathcal{C}_2$	$\left\{\mu_y > \frac{\rho}{(\rho-1)\xi-1}, \mu_x > 0, \mu_x = \Gamma_2\right\}$	$\Gamma_2(\mu_y,\rho,\xi) \equiv \frac{(\rho-1)\mu_y\,\xi-\rho(1+\mu_y)}{1+\mu_y-(\rho-1)\mu_y\,\xi}(1+\mu_y)$
$\mathcal{C}_3$	$\left\{\mu_y > \max\left\{0, \frac{2-\rho}{\rho-1}\right\}, \mu_x > 0, \mu_x = \Gamma_3\right\}$	$\Gamma_3(\mu_y,\rho) \equiv \frac{\rho(1+\mu_y)}{\rho-2+(\rho-1)\mu_y}$
$\mathcal{C}_4$	$\left\{\mu_x > \max\left\{0, \frac{2-\rho}{\rho-1}\right\}, \mu_y > 0, \mu_y = \Gamma_4\right\}$	$\Gamma_4(\mu_x,\rho) \equiv \frac{\rho(1+\mu_x)}{\rho-2+(\rho-1)\mu_x}$
$\mathcal{C}_5$	$\left\{\mu_x,\ \mu_y > \frac{\rho}{\rho-1}\nu_-, \mu_y = \Gamma_5\right\}$	$\Gamma_5(\mu_x,\rho,\xi) \equiv \frac{\rho\,(1+\mu_x)}{(\rho-1)\,\nu_+\,\mu_x-\rho}$
$\mathcal{C}_6$	$\left\{\mu_x,\ \mu_y > \frac{\rho}{\rho-1}\nu_+, \mu_y = \Gamma_6\right\}$	$\Gamma_6(\mu_x,\rho,\xi) \equiv \frac{\rho\,(1+\mu_x)}{(\rho-1)\,\nu_-\,\mu_x-\rho}$

Table 6.1: Boundaries predicted in section 6.2.

diagonal line $\{\mu_x = \mu_y\}$.

6.3 Properties of the sets in the parameter space

In this section, we provide technical lemmas and theorems of the relations between the boundaries of the sets in the parameter space. A list of important boundaries predicted in section 6.2 are provided in Table 6.1. We will adopt the code names defined in Table 6.1 in this section.

Lemma VI.5. *When $\zeta > 2$, i.e., $\xi > \rho/(\rho-1)$, $\mathcal{C}_6$ does not intersect with $\mathcal{C}_3$ and $\mathcal{C}_4$.*

Proof. Because $\mathcal{C}_4$ is the mirror image of $\mathcal{C}_3$ and $\mathcal{C}_6$ is symmetric with respect to $\{\mu_x = \mu_y\}$, it is sufficient to show that $\mathcal{C}_6$ has no intersection with $\mathcal{C}_3$ in the μ-space.

It is straightforward to solve the intersection of $\mathcal{C}_3$ and $\mathcal{C}_6$:

$$\mathcal{C}_6 \cap \mathcal{C}_3 = \left\{\mu_x = \frac{2}{\nu_- - 1}\right\} \cap \mathcal{C}_6. \tag{6.58}$$

Since $\nu_- < 1$, the intersection is not in the domain $\{\mu_x > 0\}$. □

Lemma VI.6. *$\mathcal{C}_6$ is always above (and right to) $\mathcal{C}_3$ and $\mathcal{C}_4$.*

Proof. Asymptotically, as $\mu_x \to \infty$,

$$\lim_{\mu_x \to \infty} \Gamma_4(\mu_x) = \frac{\rho}{\rho - 1} \text{ and } \lim_{\mu_x \to \infty} \Gamma_6(\mu_x) = \frac{\rho}{\rho - 1}\nu_+, \tag{6.59}$$

and since $\nu_+ > 1$, $\mathcal{C}_6$ is above $\mathcal{C}_4$ as $\mu_x \to \infty$. By lemma VI.5 $\mathcal{C}_6$ and $\mathcal{C}_4$ never intersects in the domain of interests; therefore, $\mathcal{C}_6$ is always above $\mathcal{C}_4$, and by symmetry, always above $\mathcal{C}_3$. □

Theorem VI.7. *$\mathcal{C}_6$ is irrelevant to bifurcations of the system.*

Proof. In the analysis, the conditions $\varpi_{\max} > 0$ and $q_{\max} \in (q_{\text{left}}, q_{\text{right}})$ always come together—the system could exhibit double roots if *both* the conditions are met. Since $\mathcal{C}_6$ is always above $\mathcal{C}_3$ and $\mathcal{C}_4$, the subsets of $\{\varpi_{\max} > 0, q_{\max} \in (q_{\text{left}}, q_{\text{right}})\}$ defined by these two boundaries are empty. Therefore $\mathcal{C}_6$ is not involved in determining the bifurcation. □

Theorem VI.8. *$\mathcal{C}_3$ and $\mathcal{C}_5$ have at most one intersection. In addition, $\mathcal{C}_3$ and $\mathcal{C}_5$ always intersect when $\rho \geq 2$, and when $\rho < 2$, $\mathcal{C}_3$ and $\mathcal{C}_5$ intersect iff*

$$\xi < \frac{\rho^2 - 2}{2(\rho - 2)(\rho - 1)}.$$

Proof. We prove this theorem by directly solving for the intersection,

$$\mathcal{C}_3 \cap \mathcal{C}_5 = \left\{\mu_x = \frac{2}{\nu_+ - 1}\right\} \cap \left\{\mu_y = \frac{\rho(\nu_+ + 1)}{\nu_+ (\rho - 2) + \rho}\right\} \tag{6.60}$$

which has only one solution. The solution exists when $\rho \geq 2$. If $\rho < 2$, the solution is in the μ-space only when the denominator in the expression of μ_y is greater than

0, and it is straightforward to show

$$\text{(6.61a)} \qquad \left\{\nu_+(\rho-2)+\rho>0\right\} = \left\{\nu_- > \frac{2-\rho}{\rho}\right\}$$
$$= \left\{1-2(\zeta-1)\left(\frac{2-\rho}{\rho}\right)+\left(\frac{2-\rho}{\rho}\right)^2>0\right\}$$
$$= \left\{\xi < \frac{\rho^2-2}{2(\rho-2)(\rho-1)}\right\}.$$

□

Corollary VI.9. *$\mathcal{C}_4$ and $\mathcal{C}_5$ have at most one intersection. In addition, $\mathcal{C}_4$ and $\mathcal{C}_5$ always intersect when $\rho \geq 2$, and when $\rho < 2$, $\mathcal{C}_3$ and $\mathcal{C}_5$ intersect iff*

$$\xi < \frac{\rho^2-2}{2(\rho-2)(\rho-1)}.$$

Proof. $\mathcal{C}_4$ is the mirror image of $\mathcal{C}_3$ with respect to $\{\mu_x = \mu_y\}$; on the other hand $\mathcal{C}_5$ is symmetric about $\{\mu_x = \mu_y\}$. The transformations $\mu_x \to \mu_y$ and $\mu_y \to \mu_x$ to Thm. VI.8 prove the corollary. □

Theorem VI.10. *$\mathcal{C}_1$ and $\mathcal{C}_5$ has at most one intersection, however, it is a double-root. The intersection is*

$$\left\{\mu_x = \frac{\rho(1-\nu_-)}{2(\rho-1)\xi-2}\right\} \cap \mathcal{C}_1,$$

The same statement applies to $\mathcal{C}_2$ and $\mathcal{C}_5$ with intersection

$$\left\{\mu_y = \frac{\rho(1-\nu_-)}{2(\rho-1)\xi-2}\right\} \cap \mathcal{C}_1,$$

Proof. Solving for the μ_x at the intersection yields

$$\text{(6.62)} \qquad \mathcal{C}_1 \cap \mathcal{C}_5 \subset \left\{0 = \left[(\rho-1)\xi-\rho\right]\nu_+\mu_x^2 + \rho(1-\nu_+)\mu_x + \rho\right\}.$$

The determinant of the quadratic equation turns out to be zero

$$\begin{aligned}(6.63)\qquad &\rho^2(1-\nu_+)^2 - 4\rho\left[(\rho-1)\xi-\rho\right]\nu_+ \\ &= \rho^2\left[\left(\zeta-2+\sqrt{\zeta^2-2\zeta}\right)^2 - (2\zeta-4)\left(\zeta-2+\sqrt{\zeta^2-2\zeta}\right)\right] = 0.\end{aligned}$$

Therefore it is a double-root, and the solution can be derived directly

$$(6.64)\qquad \mu_x = \frac{\rho(\nu_+-1)}{2\left[(\rho-1)\xi-\rho\right]\nu_+} = \frac{\rho(1-\nu_-)}{2(\rho-1)\xi-2}.$$

With symmetry argument the statement applies to $\mathcal{C}_2$ and $\mathcal{C}_5$. □

Lemma VI.11. *$\mathcal{C}_1$ intersects $\{\mu_x = \mu_y\}$ when $\xi > (\rho+1)/(2(\rho-1))$, and the intersection is at $\mu_x = \mu_y = \mu_*$ with*

$$\mu_* \equiv \frac{(2\rho+1)-(\rho-1)\xi+\sqrt{(\rho-1)^2\xi^2+2(2\rho-1)(\rho-1)\xi+1}}{4(\rho-1)\xi-2(\rho+1)}.$$

The statement applies to $\mathcal{C}_2$.

Proof. With the symmetry argument, it is sufficient to prove the claim for $\mathcal{C}_1$. It is elementary to show

$$\begin{aligned}(6.65)\quad &\mathcal{C}_1 \cap \{\mu_x = \mu_y\} = \{\mu_x = \mu_y\} \cap \\ &\qquad \left\{\left[(\rho+1)-2(\rho-1)\xi\right]\mu_x^2 + \left[2\rho+1-(\rho-1)\xi\right]\mu_x + \rho = 0\right\},\end{aligned}$$

and the (positive) solution is

$$(6.66)\qquad \mu_x = \mu_y = \frac{(2\rho+1)-(\rho-1)\xi+\sqrt{(\rho-1)^2\xi^2+2(2\rho-1)(\rho-1)\xi+1}}{4(\rho-1)\xi-2(\rho+1)}.$$

It is positive *iff* $\xi > (\rho+1)/(2(\rho-1))$. □

Lemma VI.12. *When $\xi > 1/(\rho-1)$, there is only one point on $\mathcal{C}_1$ at which the slope of the tangent line is -1. In addition, the point locates at*

$$\mu_x = \frac{(\rho-1)\xi + 1}{(\rho-1)\xi - 1}.$$

Proof. For simplicity we define $\theta \equiv (\rho-1)\xi$. Some calculations yield

$$\frac{d\,\Gamma_1}{d\,\mu_x} = \frac{(1-\theta)(\theta-\rho)\mu_x^2 + 2(\theta-\rho)\mu_x + (\theta-2\rho) + (1-\theta)\rho}{\left[1+(1-\theta)\mu_x\right]^2}, \tag{6.67}$$

and $d\Gamma_1/d\mu_x = -1$ can be simplified as

$$0 = (1-\theta)\mu_x^2 + 2\mu_x - (1+\theta). \tag{6.68}$$

The solutions are

$$\mu_x = \frac{-1 \pm \theta}{1-\theta}, \tag{6.69}$$

but because $\mu_x > 0$ and $\theta = (\rho-1)\xi > 1$, only the positive solution survives

$$\mu_x = \frac{-1-\theta}{1-\theta} = \frac{(\rho-1)\xi + 1}{(\rho-1)\xi - 1}. \tag{6.70}$$

□

Corollary VI.13. *When $\xi > 1/(\rho-1)$, there is only one point on $\mathcal{C}_2$ at which the slope of the tangent line is -1. In addition, the point locates at*

$$\mu_y = \frac{(\rho-1)\xi + 1}{(\rho-1)\xi - 1}.$$

Proof. Apply the symmetry argument ($\mathcal{C}_1$ and $\mathcal{C}_2$ with respect to $\mu_x = \mu_y$) to Thm. VI.12.

□

Lemma VI.14. *When $\xi > 1/(\rho - 1)$, $\mathcal{C}_1$ is convex (from above) and $\mathcal{C}_2$ is convex (from the right).*

Proof. With the symmetry argument, it is sufficient to prove $\mathcal{C}_1$ is convex from above. Rearranging the representation of Γ_1 yields

$$\Gamma_1(\mu_x, \rho) = \frac{\theta - \rho}{1 - \theta}\mu_x + \frac{-\theta^2 + (2\rho + 1)\theta - 2\rho}{(1 - \theta)^2} + \frac{(1 - \rho)\theta^2 - \theta + \rho}{(1 - \theta)^3}\frac{1}{\mu_x + \frac{1}{1-\theta}}, \tag{6.71}$$

therefore the sign of the second derivative of Γ_1 (with respect to μ_x) is the same to the sign of the prefactor

$$\frac{(1 - \rho)\theta^2 - \theta + \rho}{(1 - \theta)^3}.$$

The denominator is always less than 0 ($\theta > 1$ when $\xi > 1/(\rho-1)$), and the numerator is a quadratic polynomial of θ with negative leading coefficient. The maxima occurs at $\theta = 1/(2 - 2\rho) < 0$ and the value of the polynomial is equal to 0 at $\theta = 1$. Therefore the denominator is always negative for $\theta > 1$, or equivalently $\xi > 1/(\rho-1)$. In turn, the prefactor is positive so the second derivative of Γ_1 (with respect to μ_x) is positive. □

6.4 Bifurcation of the Landscape in Parameter Space and Evolutionarily Stable Dispersal Rate

In this section, we combine the results from section 6.3 to investigate change of the selection landscape in the μ-space.

Theorem VI.15. *When $\xi \leq 1/(\rho - 1)$, the* **fast species** *always wins.*

Proof. When $\xi < 1/(\rho-1)$, $\varpi(q_{\text{right}})$ and $\varpi(q_{\text{left}})$ are both positive (see sections 6.2.1 and 6.2.2). As a consequence, the concave function $\varpi(q) > 0$ for $q \in (q_{\text{left}}, q_{\text{right}})$, or equivalently $\bar{v}(\langle z \rangle) > 0$ for $\langle z \rangle \in (-1, 1)$ when $\mu_x > \mu_y$. The effective flow always drives the dynamical system to the X-dominating fixation $\langle z \rangle = 1$.

Symmetry shows when $\mu_y > \mu_x$, Y species always wins. □

Theorem VI.16. *When $1/(\rho-1) < \xi \leq \rho/(\rho-1)$, the system could exhibit single unstable fixed point in $\langle z \rangle \in (-1,1)$ with proper (μ_x, μ_y).*

Proof. When $1/(\rho-1) < \xi \leq \rho/(\rho-1)$, the μ-space is divided by $\mathcal{C}_1$ and $\mathcal{C}_2$. The region under $\mathcal{C}_1$ and to the right of $\mathcal{C}_2$ corresponds to $\varpi(q_{\text{left}}) < 0$ and $\varpi(q_{\text{right}}) > 0$ (when $\mu_x > \mu_y$). Therefore, the system must exhibit a unstable fixed point in the domain $\langle z \rangle \in (-1,1)$. □

Theorem VI.17. *In μ-space, the selection landscapes are qualitatively distinct when $1/(\rho-1) < \xi \leq (\rho+1)/(2\rho-2)$ and when $(\rho+1)/(2\rho-2) < \xi \leq \rho/(\rho-1)$.*

Proof. By Lemma VI.11, $\mathcal{C}_1$ and $\mathcal{C}_2$ intersects only when $\xi > (\rho+1)/(2\rho-2)$. □

Theorem VI.18. *When $\xi > (\rho+1)/(2\rho-2)$, the* **slow species** *could win the competition with sufficient large μ_x and μ_y.*

Proof. When $\xi > (\rho+1)/(2\rho-2)$, there exists a region which is above $\mathcal{C}_1$ and right to $\mathcal{C}_2$. In the region, $\varpi(q_{\text{left}}) < 1$ and $\varpi(q_{\text{right}}) < 1$; therefore, the Y-dominating fixation $\langle z \rangle = -1$ is stable. □

Theorem VI.19. *When $\xi > \rho/(\rho-1)$, there exists an* **evolutionarily stable dispersal rate**

$$\mu_{ES} = \mu_* \equiv \frac{(2\rho+1) - (\rho-1)\xi + \sqrt{(\rho-1)^2\xi^2 + 2(2\rho-1)(\rho-1)\xi + 1}}{4(\rho-1)\xi - 2(\rho+1)}.$$

Proof. As soon as $\xi > \rho/(\rho-1)$, $\mathcal{C}_1$ intersects with $\{\mu_y = 0\}$ and every boundary $\mathcal{C}_i$ is concave from above. The region above all the curves $\mathcal{C}_1, \ldots, \mathcal{C}_5$ is the region that the slow species always wins (when $\mu_x > \mu_y$). Similarly, the fast species always wins in the region below all the curves $\mathcal{C}_1, \ldots, \mathcal{C}_5$. As the consequence, the intersection of $\mathcal{C}_1$ and $\mathcal{C}_2$, i.e. $\{\mu_x = \mu_y = \mu_*\}$ is evolutionarily stable. □

Theorem VI.20. *As soon as $\xi > \rho/(\rho-1)$, the dynamics could exhibit single stable coexistent state and a pair of stable and unstable coexistent states (depending on the values of ξ, μ_x, and μ_y).*

Proof. Consider the asymptotic behavior of each curve:

$$\lim_{\mu_x\to\infty} \mathcal{C}_1 \to \left\{\mu_y = \frac{(\rho-1)\xi-\rho}{1-(\rho-1)\xi}\,\mu_x\right\}, \tag{6.72a}$$

$$\lim_{\mu_x\to\infty} \mathcal{C}_2 \to \left\{\mu_y = \frac{1}{(\rho-1)\xi-1}\right\}, \tag{6.72b}$$

$$\lim_{\mu_x\to\infty} \mathcal{C}_3 \to \left\{\mu_y = \frac{\rho-2}{\rho-1}\right\}, \tag{6.72c}$$

$$\lim_{\mu_x\to\infty} \mathcal{C}_4 \to \left\{\mu_y = \frac{\rho}{\rho-1}\right\}, \tag{6.72d}$$

$$\lim_{\mu_x\to\infty} \mathcal{C}_5 \to \left\{\mu_y = \frac{\rho}{\rho-1}\nu_-\right\}. \tag{6.72e}$$

It is clear that for a sufficiently large μ_x, $\mathcal{C}_1$ is below the other boundaries (because it has negative slope). We have proved that $\mathcal{C}_3$ is below $\mathcal{C}_4$ in Thm. VI.2. We now show that when ξ is close to $\rho/(\rho-1)$, for sufficiently large μ_x, $\mathcal{C}_4$ is above $\mathcal{C}_5$, which is above $\mathcal{C}_2$, and the bottom one (except for $\mathcal{C}_1$) is $\mathcal{C}_3$.

When $\xi \gtrsim \rho/(\rho-1)$, that is $\zeta = 2(\rho-1)\xi/\rho \gtrsim 2$, we define $\zeta \equiv 2 + d\zeta$ with $d\zeta \ll 1$. Then

$$\nu_- = 1 + d\zeta - \sqrt{d\zeta^2 + 2d\zeta} \approx 1 - \sqrt{d\zeta}, \tag{6.73}$$

and the asymptotes of the boundaries $\mathcal{C}_2$ and $\mathcal{C}_5$ are approximately

$$\lim_{\mu_x\to\infty} \mathcal{C}_2 \to \left\{\mu_y = \frac{2}{\rho(2+d\zeta)-2}\right\} \approx \left\{\mu_y = \frac{1}{\rho-1}\left(1 - \frac{\rho\,d\zeta}{2\rho-2}\right)\right\}, \tag{6.74a}$$

$$\lim_{\mu_x\to\infty} \mathcal{C}_5 \to \left\{\mu_y = \frac{\rho}{\rho-1}\left(1-\sqrt{d\zeta}\right)\right\}. \tag{6.74b}$$

Therefore $\mathcal{C}_2$ is below $\mathcal{C}_5$ because $\rho > 1$ and $d\zeta \ll 1$. In the region above $\mathcal{C}_2$, above $\mathcal{C}_3$,

and below $\mathcal{C}_5$ is the region where $\varpi(q_{\text{right}}) < 0$, $\varpi(q_{\text{left}}) < 0$, $q_{\max} \in (q_{\text{left}}, q_{\text{right}})$ and $\varpi_{\max} > 0$—which is the condition that $\bar{v}(\langle z \rangle) = 0$ has two roots in $\langle z \rangle \in (-1, 1)$. On the other hand, $\varpi(q_{\text{right}}) < 0$ and $\varpi(q_{\text{left}}) > 0$ in the region below $\mathcal{C}_2$ and above $\{\mu_x = 0\}$—which in turns means the system exhibits a single stable coexistent state. □

Theorem VI.21. *When*

$$\xi \geq \frac{2\rho - 1 + 2\sqrt{\rho^2 - \rho + 1}}{3(\rho - 1)}$$

the system does not exhibit unstable coexistence.

Proof. It is elementary to show when $\xi \gtrsim \rho/(\rho-1)$, $\mathcal{C}_1$ and $\mathcal{C}_2$ have three intersections. As ξ increases, the intersections approaches to the middle one ($\{\mu_x = \mu_y = \mu_*\}$) and eventually merges as one at a critical value of $\xi = \xi_c$. We solve for the critical value ξ_c.

At the critical point, the slope (of $\mathcal{C}_1$ and $\mathcal{C}_2$) at $\{\mu_x = \mu_y\}$ must be -1 by symmetry. By Lemma VI.12, we know the slope is -1 only at $\mu_x = \mu_\dagger$ with

$$\mu_\dagger \equiv \frac{(\rho - 1)\xi + 1}{(\rho - 1)\xi - 1}, \tag{6.75}$$

and by inserting it into the equation for μ_* (see proof of Thm.VI.11), we obtain

$$[(\rho + 1) - 2(\rho - 1)\xi_c]\,\mu_\dagger^2 + [2\rho + 1 - (\rho - 1)\xi_c]\,\mu_\dagger + \rho = 0 \tag{6.76}$$

at the critical point $\xi = \xi_c$. Directly solving this equation yields the (positive) solution

$$\xi_c = \frac{2\rho - 1 + 2\sqrt{\rho^2 - \rho + 1}}{3(\rho - 1)}. \tag{6.77}$$

□

Theorem VI.22. *When*

$$\frac{\rho}{\rho-1} < \xi < \frac{2\rho-1+2\sqrt{\rho^2-\rho+1}}{3(\rho-1)}$$

the dynamics could have a single unstable coexistent state (with proper μ_x and μ_y).

Proof. We show that the tangent line of $\mathcal{C}_1$ at μ_* has slope that is less than -1 when

$$\frac{\rho}{\rho-1} < \xi < \frac{2\rho-1+2\sqrt{\rho^2-\rho+1}}{3(\rho-1)}. \tag{6.78}$$

First, as long as $\xi > \rho/(\rho-1)$, both μ_* and $\mu_\dagger$ are continuous function of $(\rho-1)\xi$. From Thm. VI.20, we know that $\mu_* = \mu_\dagger$ iff $\xi = \xi_c$, so the sign of $\mu_* - \mu_\dagger$ must be the same when $\xi > \xi_c$ and when $\xi < \xi_c$. Let $\xi = \rho/(\rho-1) < \xi_c$, we find

$$\mu_* = \frac{\rho+1+\sqrt{5\rho^2-2\rho+1)}}{2(\rho-1)}, \tag{6.79a}$$

$$\mu_\dagger = \frac{\rho+1}{\rho-1}, \tag{6.79b}$$

therefore,

$$\operatorname{sgn}(\mu_*-\mu_\dagger) = \operatorname{sgn}(\sqrt{5\rho^2-2\rho+1)}-\rho-1). \tag{6.80}$$

Let $f \equiv \sqrt{5\rho^2-2\rho+1)}-\rho-1$ and since $\rho > 1$, $\rho^2 > \rho$,

$$\begin{aligned} f &= \sqrt{\rho^2+4\rho^2-2\rho+1)}-\rho-1 \\ &> \sqrt{\rho^2+4\rho-2\rho+1)}-\rho-1 \\ &> \sqrt{\rho^2+2\rho+1)}-\rho-1 = 0. \end{aligned} \tag{6.81}$$

This proves $\mu_* > \mu_\dagger$ on the set $\{\xi < \xi_c\}$. On the other hand, when $\xi > \xi_c$ we use

parallel argument to show as $\xi \to \infty$,

$$\mu_*(\xi \to \infty) = 0 < 1 = \mu_\dagger(\xi \to \infty), \tag{6.82}$$

which proves $\mu_* < \mu_\dagger$ on the set $\{\xi < \xi_c\}$.

Finally, from Lemma VI.14, we know $\mathcal{C}_1$ is convex. Since $\mathcal{C}_1$ is convex and $\mu_* > \mu_\dagger$, at $\{\mu_x = \mu_y = \mu_*\}$ the tangent slope must be greater than -1. By symmetry the slope of $\mathcal{C}_2$ at $\{\mu_x = \mu_y = \mu_*\}$ is less than -1. As a consequence we can find a region near $\{\mu_x = \mu_y = \mu_*\}$ that is above $\mathcal{C}_2$ but below $\mathcal{C}_1$, which corresponds to $\varpi(q_{\text{left}}) < 0$ and $\varpi(q_{\text{right}}) > 0$—or equivalently, the dynamics support a single unstable fixed points. □

Fig. 6.6 presents the bifurcations when $\rho = 2$. A brief description of the bifurcations is as follows:

- When $\xi = K\sigma^2 < 1/(\rho - 1)$, i.e. the population scale is low or the space is homogeneous, the fast species always wins the competition as $t \to \infty$, and the heterogeneous many patch model is qualitatively the same as the homogeneous many patch model discussed in Chapter III. See Fig. 6.6 (a).

- When $\xi > (\rho + 1)/(2\rho - 2)$ and provided with proper μ_x and μ_y, the slow dispersers could win the competition if initially their population is large enough (that is, the model has a single *unstable* fixed point.) See Fig. 6.6 (b).

- When $\xi > (\rho+1)/(2\rho-2)$ and with large enough dispersal rates of both species, the slow species could always enjoy advantage regardless of the initial population distribution. See Fig. 6.6 (c).

- As soon as $\xi > \rho/(\rho-1)$, a finite evolutionarily stable rate for dispersal μ_* (see Thm. VI.19 for functional dependence) emerges. In addition, the dynamics start

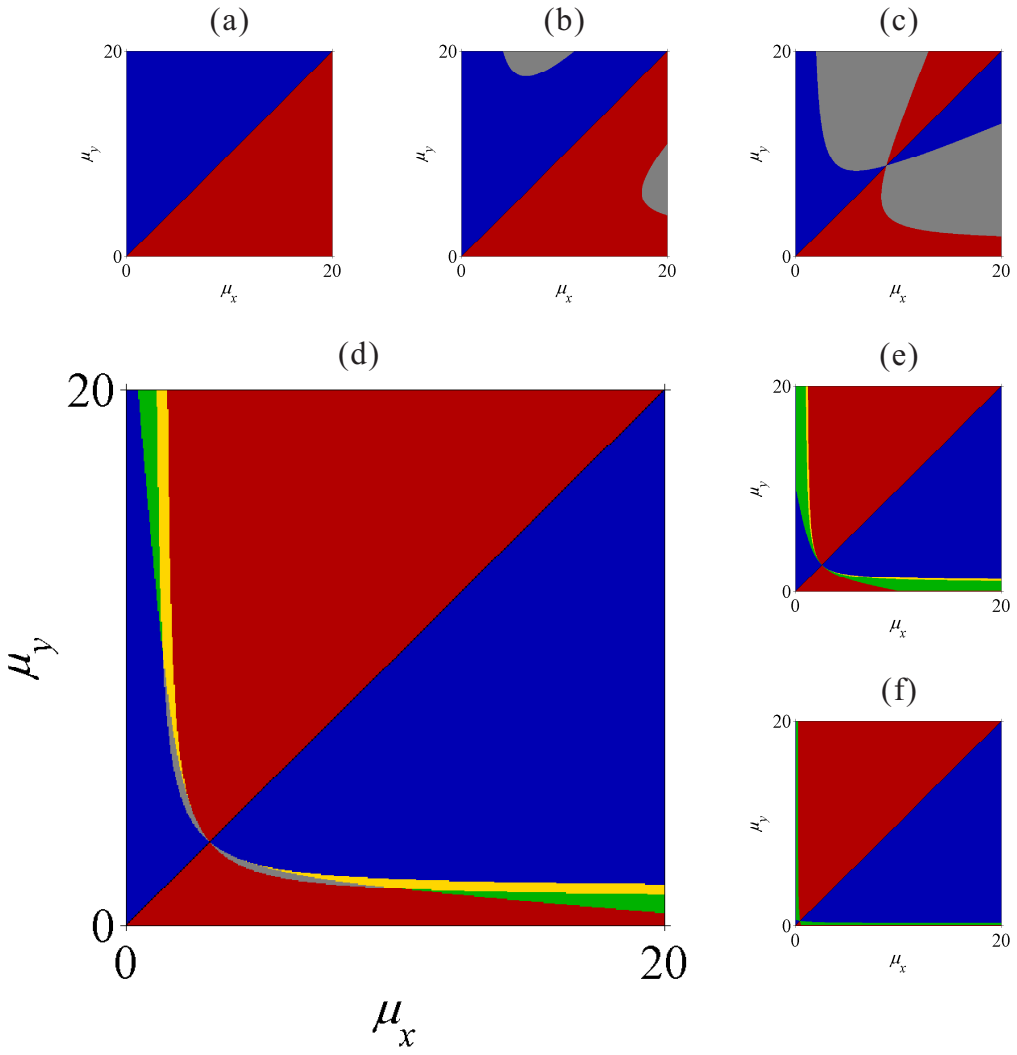

Figure 6.6: Landscape of evolutionary advantage in (μ_x, μ_y) for $\rho = 2$. As $t \to \infty$, the space is separated by red: X always wins the competition, blue: Y always wins the competition, grey: there exists a single unstable coexistence state, green: there exists a single stable coexistence state, gold: there exist a pair of stable/unstable states, and black ($\mu_x = \mu_y$): the system is everywhere stable. (a) $\xi = 1.0$, (b) $\xi = 1.5$, (c) $\xi = 1.7$, (d) $\xi = 2.08$, (e) $\xi = 2.2$, and (f) $\xi = 6$.

to exhibit stable coexistent states and a pair of stable and unstable coexistent states. See Fig. 6.6 (d).

- Further increasing ξ leads the evolutionarily stable dispersal rate μ_* to decrease. The large ξ corresponds to large population limit or large environmental variance. As $\xi \to \infty$, μ_* converges to 0, where the dynamics of the system approaches to the dynamics discussed in Chapter IV. See Fig. 6.6 (e-f).

Finally, we perform the transformation (6.3) to the conclusion to predict the "landscape of selection" of the 2-patch model. Fig. 6.7 resembles the "analytical landscape" parallel to the "numerically computed landscape" Fig. 5.8. It is clear that Fig. 6.7 resembles the qualitative behavior of Fig. 5.8. In addition, it shows our analysis predicts an evolutionarily stable dispersal rate accurately. We remark that in the numerical computations of the 2-patch model, we do not observe the coexistent state for the obvious reason: the effective diffusion creates random motion in the center manifold, and as a consequence drives the system to the absorption states $\langle z \rangle = \pm 1$.

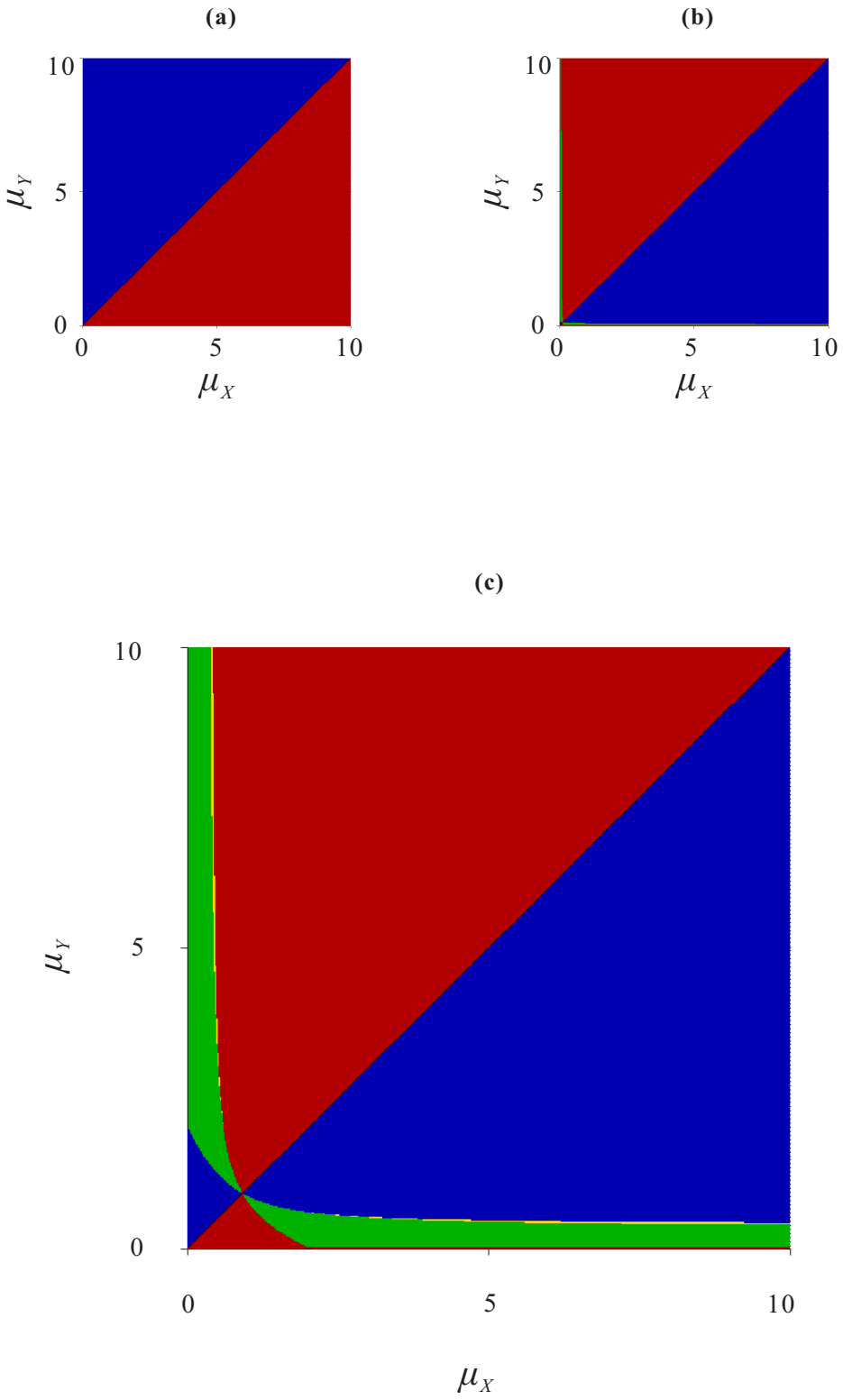

Figure 6.7: Bifurcation analysis of the stochastic 2-patch model for comparison to Fig. 5.8. Color codes are the same to Fig. 6.6. $\rho = 2$ and $\Lambda(= K) = 500$. (a) $\sigma = 0.01$, (b) $\sigma = 0.05$, and (c) $\sigma = 0.10$.

CHAPTER VII

Discussion, Conclusion, and Future Work

7.1 Discussion and conclusion

Development of the physical asymptotic analysis

The two individual-based models discussed in Chapter V were originally developed with the motivation to investigate the dynamical mechanism of the regime shift in competitive dispersal problems reported by Kessler and Sander [25]. Soon after the models were formulated, we realized that owing to the complexity of the dynamics, it is beyond our ability to analytically solve the models. In fact, even in the very special case that the space is homogeneous, i.e. the models discussed in Chapter III, it is apparently infeasible to directly solve the master equations.

One of the key properties of the models in homogeneous environments is that they are *degenerate* in the infinite population limit. That is, the rate equations describing the population dynamics, which are a set of ordinary differential equations, have an infinite number of solutions—as long as the spatial distribution of the population of each species is homogeneous and the total population on each patch is equal to the carrying capacity, the species could coexist in an arbitrary distribution. Due to the simplicity of the solutions, such models are often neglected in the research literature.

Nevertheless, in the individual-based simulations where the population scale is

large but finite, we observed that the degeneracy is broken and that the *fast dispersers* always enjoy an advantage—in the 2-patch model, the winning probability of the fast dispersers is higher, and in the many patch model, the fast dispersers wins the competition almost with probability 1. The fundamental difference between the models describing by (deterministic) rate equations and the models with individual setting is that the latter includes **demographic stochasticity**.

The observation raised a natural question: how does the demographic stochasticity break the degeneracy of the deterministic dynamics? The answer of the question also seemed to be the hinge to understand how the demographic stochasticity enhance the survival of the fast species in the dispersal problems [25].

In order to proceed analytical inquiries, we performed a model reduction to develop the competitive population dynamics in Chapter II where the competition is between two species with different lifespans on a single patch. The objective was to reduce the dimensionality of the dynamics while preserving the degenerate feature observed in the dispersal problems in homogeneous environments. It is worth mentioning that the model does have biological application [31, 32] even though our original motivation was simply to "mimic" the dynamics of the model in Chapter III. See section 2.5 for a more thorough discussion.

With the reduced dimensionality we were able to develop a special technique, the "*physically motivated asymptotic analysis*", to solve the problem analytically. The analysis is developed based on the idea of separation of time scales, which is the common feature of the models in Chapters II–V. The prediction of the physical asymptotic analysis was verified by conventional asymptotic analysis and numerical simulations. Compared to conventional asymptotic analysis, the physically motivated approximate approach is much more intuitive. As a consequence it was straightforward to generalize the analysis to models with higher dimensionality, models for which it is much more difficult to perform a conventional asymptotic analysis. In addition,

the analysis provides insight of the dynamics: the combined effects of stochasticity and nonlinearity cause a effective motion in the degenerate manifold (the center manifold). Most importantly, the *time scale* of the effective motion is immediately revealed to be proportional to the population scale of the system.

Demographic stochasticity in passive dispersal models

After introducing the model and its approximations in Chapter II, the physically motivated asymptotic analysis was soon generalized and applied to the passive dispersal models with homogeneous environments in Chapter III. By combining the physical asymptotic analysis and regular perturbation calculations, features of the dispersal problem were discovered. The degeneracy in the dynamics is broken due to the individual-level symmetry braking, i.e., one of the species moves faster than the other. In the stochastic models, there exists effective motions in the degenerate manifolds in a time scale which is again proportional to the total population (per patch), and the effective dynamics weakly select the fast dispersers. The stochastic models exhibit *singular limits* such that the fast species always enjoys the advantage when the population scale is finite (but the time scale of the selection diverges as the population scale goes to infinity), and in the infinite-population limit the models are degenerate and neutral.

These discoveries in the passive dispersal models with homogeneous environments show that demographic stochasticity is essential in the competitive population dynamics. The observation that the stochasticity favors the *fast species* in homogeneous environments was reported by Travis [37], followed by several studies in adaptive population dynamics with various model settings [33, 1]. In this line of research, even though the mechanisms of dispersion are usually more complicated, fundamentally they are identical to passive diffusion: the migration rates of the individuals only depend on the population density of their current habitat.

A reasonable conjecture is that the demographic stochasticity still favors the *fast species* in passive dispersal models with *heterogeneous* environments. On the other hand, in the deterministic limit it has been observed and proven that the *slow species* always enjoys the advantage in patchy-like systems by Gadgil [13], Comins [4], Holt [20], and in the systems with continuous space by Hastings [19], and Dockery *et al.*. [8].

This raised an interesting question: how does stochastic dynamics, favoring the fast species, interact with the deterministic and nonlinear dynamics that favors the slow species? What are the critical parameters that determine which species is more likely to win the competition? The answer was partially answered by Kessler and Sander [25], but we had not identified the dynamical mechanism of the selection.

With the derived time scale of the weak selection in the stochastic models with homogeneous environment, a reasonable step to move forward is to investigate the time scale of the deterministic dynamics in heterogeneous environments and to compare the two time scales. Most of the analysis developed in literature [16, 4, 19, 30, 1, 33] are fundamentally stability analyses which do not reveal the time scale of the detail dynamics. In order to proceed, we performed the standard asymptotic analysis in Chapter IV, in which the time scale of the dynamics is identified to be proportional to the inverse of the environmental variance. In addition, we discovered that the selection is also weak when the environmental variance is small.

A time scale argument was presented in Chapter V—whichever effect has significantly shorter time scale than the other will prevail and dominate the final dynamics. Therefore the derived time scales, $\mathcal{O}(K)$ and $\mathcal{O}(\sigma^{-2})$, provide an objective measure to determine which effect is "more important": when $\mathcal{O}(K\sigma^2)$ is much greater than unity, we should treat the models as if they are deterministic ones with heterogeneous environments, and when $\mathcal{O}(K\sigma^2)$ is much smaller than unity, the corresponding stochastic models with homogeneous environments are more adequate. Furthermore,

$\mathcal{O}(K\sigma^2) \approx 1$ corresponds to the critical cases where these two effects are equally important, which correspond to the boundaries of the parameter regime shifts.

In Chapter V we applied the physically motivated asymptotic analysis to the models with heterogeneous environments when $\mathcal{O}(K\sigma^2) \approx 1$. In contrast to the tedious computations, the results turned out to be rather simple: when $\mathcal{O}(K\sigma^2) \approx 1$, the overall effective drift in the slow manifold is the linear superposition of the effective drifts in Chapter III and IV.

Demographic stochasticity and the emergence of a evolutionarily stable dispersal rate

The combined effective drift in the slow manifold could support distinct solutions depending on the values of the parameters. All possible scenarios were presented in Chapter VI, in which the bifurcation analysis shows only four free parameters are involved in predicting the fate of the dynamics

- ξ: the product of the harmonic mean population K and the environmental variance, σ^2,
- μ_x: the dispersal rate of species X,
- μ_y: the dispersal rate of species Y, and
- ρ: the birth-to-death rate ratio at low population level.

The emergence of ξ suggests a system with a larger population scale and a system with more environmental variations could result in the same dynamical outcome as time goes to infinity (but the time scale of the system with larger population is longer.)

The bifurcation analysis also shows that the parameter ξ plays an essential role in the qualitative transitions of the selection landscape in (μ_x, μ_y) space. When the parameter ξ is below a certain critical value, the dynamics behave as if the systems

are in homogeneous environments, that is, the fast dispersers enjoy the advantage. When ξ is higher than another critical value, the dynamics of both models exhibit evolutionarily stable dispersal rates which depend solely on the parameters ρ and ξ. As ξ increases to infinity, the evolutionarily stable dispersal rates (in both models) decrease to zero. In this limit the prediction converges to the conclusion in deterministic models, that is, slower dispersers always win the competition.

The conclusions analytically explain many common features of general stochastic population dynamics. First, the prediction of the hyperbolic regime shift boundary in the (K, σ^2) space explains Kessler's and Sander's observation in a model with distinct geometry (a one dimensional lattice space with nearest neighbor connection) [25]. Second, with large ξ values, the predicted selection landscape in the (μ_x, μ_y) space shows similar features to Holt's numerical results of a model with chaotic population dynamics on connected patches [21]. Third, the predicted existence of evolutionarily stable dispersal rate was numerically reported by Comins *et al.* [4], Cadet *et al.* [1] and Waddell *et al.* [41] in various model settings. Lastly, our model is a natural generalization of the model in Waddell *et al.* [41].

Analytical predictions

In addition to explaining the known facts, our analysis yields the following predictions.

One feature of our global analysis is the prediction of the dynamics with arbitrary initial conditions, in contrast to the stability analysis of stochastic systems [16, 4, 30, 1, 33]. With this approach, the analysis also identifies the niche space that supports polymorphism (coexistence). It is clear that the coexistent state is noise-driven, since without demographic stochasticity only the slow species endures. Numerical evidence is presented in Fig. 7.1. The example in Fig. 7.1 may be biologically unrealistic because the dramatic distinction of dispersal rates ($\mu_x = 40, \mu_y = 0.1$),

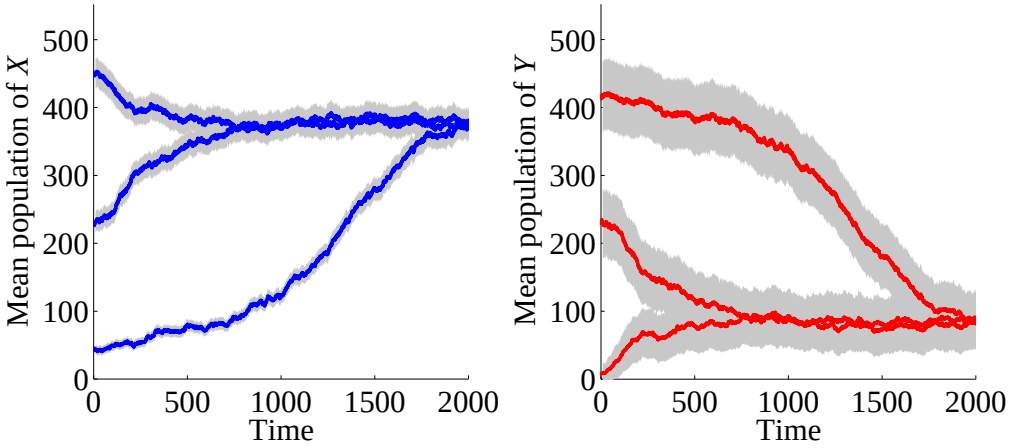

Figure 7.1: Three sample paths of the stochastic many patch model with different initial conditions $z_0 = 0.98$, 0.5, and 0.1. $\mu_x = 40$, $\mu_y = 0.1\,\rho = 2$, and $\Lambda(= K) = 460$. The number of patches is $N = 500$. Solid lines are the mean population among patches, and the grey bands represent the mean populations plus / minus one standard deviation of the distribution. The numerical results confirms the predicted coexistence from asymptotic analysis.

nevertheless, the phenomenon is certainly of interest in general stochastic dynamical system research.

Next, a common rationale reported in the literatures is that demographic stochasticity creates variability that enhances the survival of fast species because the fast ones are more capable of finding an unsaturated spot efficiently. This rationale is only partially accurate. The weak selection is a second-order effect, and to lowest order the fast and the slow species have the same fitness in models with passive dispersal. The fast dispersers can indeed more efficiently spot an "oasis" caused by demographic stochasticity, but they also leave established resident patches more frequently. Since the demographic dynamics of both the fast and the slow species are identical, the resulting vacancy will be filled by both fast and slow species and the final portion of the fast species on the abandoned patch is lowered. The gain and loss of having fast dispersal balance each other; to the lowest order, being fast or slow does not

determine the evolutionary fate.

The analysis highlights what really matters is the *curvature* of the trajectories—which is model dependent—of the nonlinear demographic dynamics near the slow manifold. A concrete counterexample is provided by solely changing the form of mutual competition: consider a model where the death rate

$$\delta\left[1-\frac{(X_i+Y_i)}{K_i}\right]\to\delta\left\{1-\left[\frac{X_i+Y_i}{K_i}\right]^{\lambda}\right\}. \tag{7.1}$$

When λ is sufficiently large, the curvature of the trajectories near the slow manifold could reverse. Our numerical analysis of the trajectories (Fig. 7.2) predicts that the slow dispersers have evolutionary advantage on a majority portion of the slow manifold. The results of the continuous time Markov chain simulation in the 2-patch geometry (Fig. 7.3) supports the prediction—the slower dispersers have a better chance to win the competition, as opposed to the claim that the demographic stochasticity always enhances passive dispersal. That is, to accurately predict the evolutionary outcome of the dynamics, one has to consider the detail interactions between the demographic stochasticity and the nonlinearity near the slow manifold.

7.2 Future Work

In this section, we list prospective future work in three fields of research—population dynamics, general mathematical biology, and applied mathematics.

Population dynamics

- Mutation—one of the essential driving forces in biology—is not considered in these models. Mutation could be modeled as stochastic processes as well. A general question to ask is, what happens if mutation is taken into account in these models?

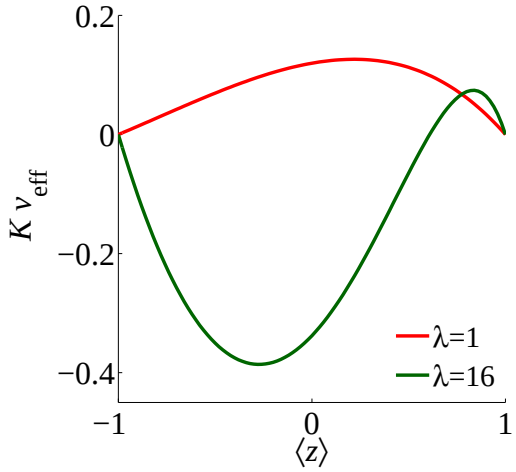

Figure 7.2: Numerical computed effective drift in the slow manifold. $\mu_x = 5$, $\mu_y = 1$, and $\sigma = 0$. We have shown when $\lambda = 1$ the dynamics always favors the fast dispersers in Chapter III. When $\lambda = 16$, in a majority portion of the domain ($z \lesssim 0.7$) the dynamics favors the species with slow dispersal rate. Then one expects that in a head-to-head competition, the slower dispersers prevail.

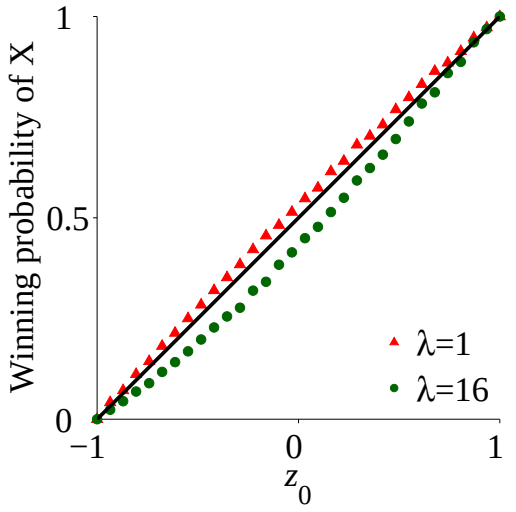

Figure 7.3: Winning probability of faster species X measured in the continuous time Markov chain simulations. $\mu_X = 5$, $\mu_Y = 1$, $\rho = 2$, $\Lambda = 200$, $\sigma = 0$, and $\lambda = 16$. 5×10^5 sample paths are performed. The black diagonal line denotes the winning probability of species X in a degenerate case $\mu_X = \mu_Y > 0$. $\lambda = 16$ shows with initial conditions $z \lesssim 0.8$, the slower dispersers have advantage.

- All the models discussed in the dissertation have only two parties of competitors. The motivation for such models originate from the objectives of early studies [16, 19, 30] where models were developed to determine whether an infinitesimal amount of mutants are able to invade a well-established population. In this context, pairwise-competition models are enough. With a global picture of the dynamics, it is now possible to utilize the analytical tools to investigate more generic competitive population dynamics with more species.

- Combining two points above, a model with mutation in a continuous phenotypical space is adequate for the dispersal problems. The idea is that the species should be able to mutate in an array or a continuous spectrum of migration rates. One of the essential question is whether the competition among multiple species (with mutation) also exhibits the evolutionarily stable dispersal rate we predicted in Chapter VI.

- Recent interest in population ecology and evolutionary dynamics have shifted from passive diffusion to more complicated dispersal mechanisms [6, 2]. Some of the models exhibit degenerate "strategies" of dispersion—that is, the fitness of several strategies are the same. How demographic stochasticity affects the system is also an emerging field of research.

General mathematical biology

- It is suggested that the demographic stochasticity plays important role in virology [38, 39, 23, 7, 40, 34]. Can we develop a similar tool to analytically compute the asymptotic behaviors of the models?

- It has been recently proposed that stochasticity may be an important factor in early cell differentiations and stem cell differentiations[42, 22]. Nevertheless, the noise is often "put in by hand". Can we utilize the knowledge gained in this

dissertation and develop better models and analytical tools for these problem?

Applied mathematics

- A theoretic inquiry into the "physically motivated asymptotic analysis" is the next task. Even with the success to explain numerical results, we do not know how the original stochastic processes converge to the reduce dynamics as the scale of the fluctuation converges to 0. In addition, error estimation must be performed. In addition, our results seem to suggest the large carrying capacity limit ($K \to \infty$) commutes with ordinary perturbation calculations. If one can prove the commutativity of these limits, our asymptotic approach could be widely applied to many stochastic dynamical systems with such separation of time scales.

- A natural conjecture is that effective drifts should be determined *only* by the curvature of the local trajectories. An operational approach may be constructed by directly computing the local curvature along a specific direction (i.e. the direction of the slow manifold in the phase space) from the rate equations. This approach therefore avoid the necessity of knowing detail trajectories of the "kick-out-flow-back" processes in the physical asymptotic analysis. If such operational approach works, we may be able to apply it to multiple-species problem mentioned above, where due to high dimensionality it is difficult to describe the deterministic trajectories via perturbation theory.

APPENDICES

APPENDIX A

Single Species in a Power-law Distributed Environment

In this dissertation, we performed an asymptotic calculation with respect to the expansion of the "spatial variance" σ^2 in the heterogeneous models. One of the assumption of the expansion is the existence of every moment of environmental distribution. Nevertheless, when the environments have power-law distributions, some of the higher moments may not exists.

In this appendix we examine a simple version of the problem: given a power-law distributed environment, what are the resulting *stationary population distributions* of a *single species* with dispersion in deterministic settings? We will show that, when the power-law distribution has a small exponent, a single species with higher dispersal rates will have higher average population. On the other hand, if the exponent is higher than a critical value, there exists an "optimal dispersal rate", and the species with such dispersal rate can achieve maximum average population.

We must point out that the model only considers the stationary distribution of a *single* species, so the analysis does not imply that the "optimal rate" is evolutionarily stable. In fact, as we addressed in this dissertation, in pairwise competitions zero dispersal has been proven to be optimal [19, 8]. The motivation to perform analysis

in this appendix is a pure mathematical inquiry, rather than being motivated by realistic biological phenomenon.

A.1 The problem

We consider a single species with logistic-like demographic dynamics. Let the position be denoted by x and the carrying capacity at $x \in \Omega$ be denoted by $k(x)$, where Ω is the domain. The population $u(x,t)$ is a function of position x and time t. The mean-field dynamics is described by logistic-like equation of motion

$$\frac{\partial u(x,t)}{\partial t} = u(x,t)\left[k(x) - u(x,t)\right] - \mu\left[u(x,t) - \langle u\rangle\right], \tag{A.1}$$

where μ is the dispersal rate and $\langle u(x,t)\rangle$ denotes the 'average population over the domain"

$$\langle u\rangle \equiv \frac{\int_{x\in\Omega} u(x)}{\int_{x\in\Omega} 1} \tag{A.2}$$

We shall denote $\langle\cdot\rangle$ to be the spatial average of the observables as we did in the many patch model.

We are interested in the stationary distribution $u(x)$ which satisfies

$$0 = u(x)\left[k(x) - u(x)\right] - \mu\left[u(x) - \langle u\rangle\right]. \tag{A.3}$$

Note that in this model, due to the dispersal rate $\mu > 0$, $u(x)$ must be nonzero unless it is zero everywhere (that is, no population in the domain). By dividing u to (A.3) we obtain

$$u = k + \mu\left(\langle u\rangle \frac{1}{u} - 1\right). \tag{A.4}$$

After taking average on both side and using Cauchy-Schwartz inequality,

$$\langle u\rangle = \langle k\rangle + \mu\left(\langle u\rangle\left\langle\frac{1}{u}\right\rangle - 1\right) \geq \langle k\rangle\,, \tag{A.5}$$

which implies

$$\langle u\rangle \geq \langle k\rangle\,. \tag{A.6}$$

That is, for any distribution, it is better to disperse (and therefore increase the total population) than to stay at home in the single-species setting. (Nevertheless in the pairwise-competition setting, it is better to be slow in the deterministic limit [8].)

In addition, u can be implicitly solved by

$$u = \frac{k-\mu}{2} + \sqrt{\left(\frac{k-\mu}{2}\right)^2 + \mu\,\langle u\rangle}. \tag{A.7}$$

Taking average on both sides, we obtain

$$\langle u\rangle = \left\langle \frac{k-\mu}{2} + \sqrt{\left(\frac{k-\mu}{2}\right)^2 + \mu\,\langle u\rangle}\right\rangle. \tag{A.8}$$

The key to solve this implicit equation is to find the $\langle u\rangle$ that solve this equation because $\langle k\rangle$ is given. Furthermore, the solution hinges on the term $\left\langle\sqrt{(k-\mu)^2 + 4\mu\,\langle u\rangle}\right\rangle$. Note that this equation always has only one solution. Suppose

$$f_1\left(\langle u\rangle\right) \equiv \langle u\rangle\,, \tag{A.9}$$

$$f_2\left(\langle u\rangle\right) \equiv \left\langle \frac{k-\mu}{2} + \sqrt{\left(\frac{k-\mu}{2}\right)^2 + \mu\,\langle u\rangle}\right\rangle \tag{A.10}$$

then f_2 is increasing but concave in $\langle u\rangle$, and $f_2(0) > f_1(0) = 0$. Since f_1 is linear in $\langle u\rangle$ there must exist a single positive solution for $f_1\left(\langle u\rangle\right) = f_2\left(\langle u\rangle\right)$.

Specifically, we assume the environmental distribution has a power-law tail, that is, the frequency of the occurrence of a "patch" with carrying capacity k is proportional $\mathbf{1}_{\{k\geq 1\}}/k^n$ with indicator function $\mathbf{1}_{\{\cdot\}}$. In addition we assume the mean carrying capacity exists, that is, $n > 2$. The following sections aim to evaluate the upper bound of $f_2(x) \equiv \left\langle \sqrt{(k-\mu)^2 + 4\mu\, x} \right\rangle$ to establish an upper bound for $\langle u \rangle$ as $\mu \to \infty$.

A.2 Change of variable

To find an upper bound for

$$f_2(x) \equiv \left\langle \sqrt{(k-\mu)^2 + 4\mu\, x} \right\rangle = (n-1) \int_1^\infty \sqrt{(k-\mu)^2 + 4\mu\, x}\, \frac{1}{k^n} dk, \tag{A.11}$$

we consider to map the problem to a bounded set. Let $y \in (0,1)$ which satisfies

$$k = \frac{1}{y^\alpha} \tag{A.12}$$

and make a change of variable to (A.11). Then after some algebra,

$$f_2 = \alpha\,(n-1) \int_0^1 \sqrt{(1 - y^\alpha \mu)^2 + 4\mu x y^{2\alpha}}\, \frac{y^{n\alpha}}{y^{2\alpha+1}} dy. \tag{A.13}$$

We choose α such that $n\alpha = 2\alpha + 1$ (that is, $\alpha = 1/(n-2)$), therefore

$$\alpha\,(n-1) = \frac{n-1}{n-2} \equiv \langle k \rangle \tag{A.14}$$

and

$$f_2 = \mu\, \langle k \rangle \int_0^1 \sqrt{\left(y^\alpha - \frac{1}{\mu}\right)^2 + \frac{4x}{\mu} y^{2\alpha}}\, dy. \tag{A.15}$$

For simplicity, we define

$$g(y) \equiv \sqrt{\left(y^{\alpha} - \frac{1}{\mu}\right)^2 + \frac{4x}{\mu} y^{2\alpha}} \tag{A.16}$$

in the following sections, we are using Lebesgue's measure $m(\cdot)$. In addition, a special value $\alpha = 1$ can be integrated analytically; for the asymptotic upper bound analysis we only consider $\alpha \neq 1$.

A.3 Asymptotic upper bound

A.3.1 Set 1 $S_1 := \{y : y^{\alpha} < 2/\mu\}$

Consider the set $\{y : y^{\alpha} < 2/\mu\}$. We will prove that the integration on such set is bounded.

Claim A.1. *One observation is that*

$$g(y) \leq \frac{1}{\mu}\sqrt{1 + 16\frac{x}{\mu}} \tag{A.17}$$

Proof. Trivial. □

Claim A.2. *Integration on this set is bounded by $O(1/\mu^{n-2})$ as $\mu \to \infty$.*

Proof.

$$\begin{aligned} \mu \langle k \rangle \int_{S_1} g(y)\, dy &\leq \mu \langle k \rangle \frac{1}{\mu}\sqrt{1 + 16\frac{x}{\mu}} \int_{S_1} dy \\ &= \langle k \rangle \sqrt{1 + 16\frac{x}{\mu}}\, m(S_1) \\ &= \langle k \rangle \left(\frac{2}{\mu}\right)^{n-2} \left(1 + \frac{16x}{\mu}\right)^{\frac{1}{2}}. \end{aligned} \tag{A.18}$$

□

A.3.2 Set 2 $S_2 := \{y : y^\alpha > 2/\mu\}$

Note that

$$g(y) = y^\alpha \sqrt{1 + \frac{4x}{\mu}} \sqrt{1 + \frac{-2\frac{1}{\mu y^\alpha} + \left(\frac{1}{\mu y^\alpha}\right)^2}{1 + \frac{4x}{\mu}}}. \tag{A.19}$$

Define

$$\epsilon := \frac{-2\frac{1}{\mu y^\alpha} + \left(\frac{1}{\mu y^\alpha}\right)^2}{1 + \frac{4x}{\mu}}, \tag{A.20}$$

notably, $0 < |\epsilon| < 1$ on S_2. Then we expand g by Taylor series expansion and obtain

$$g(y) = y^\alpha \sqrt{1 + \frac{4x}{\mu}} \left[1 + \frac{\epsilon}{2} - \frac{\epsilon^2}{8} + \frac{\epsilon^3}{16}\right] + \mathcal{O}\left(\epsilon^4\right) \tag{A.21}$$

which suggests

$$g(y) < y^\alpha \sqrt{1 + \frac{4x}{\mu}} \left[1 + \frac{\epsilon}{2} - \frac{\epsilon^2}{8}\right] \tag{A.22}$$

since $\epsilon^3/16 < 0$.

Next we compute $\epsilon/2 - \epsilon^2/8$. For simplicity let $\beta \equiv 1/\mu y^\alpha < 1/2$, $\gamma \equiv (1 + 4x/\mu)^{-1}$, then $\epsilon = \gamma\left(-2\beta + \beta^2\right)$ and

$$\begin{aligned} \frac{\epsilon}{2} - \frac{\epsilon^2}{8} &= \frac{1}{2}\gamma\left(-2\beta + \beta^2\right) - \frac{1}{8}\gamma^2\left(-2\beta + \beta^2\right)^2 \\ &= \frac{1}{2}\gamma\left[\left(-2\beta + \beta^2\right) - \frac{1}{4}\gamma\left(4\beta^2 - 4\beta^3 + \beta^4\right)\right] \\ &< -\beta\gamma + \frac{1}{2}\gamma\beta^2(1 - \gamma) + \frac{1}{2}\gamma^2\beta^3. \end{aligned} \tag{A.23}$$

The integration of β^m with respect to y when $m < 2$ is

$$\int_{S_2} \beta^m \, dy = \frac{n-2}{n-2-m} \left[\frac{1}{\mu^m} - \frac{1}{2^m} \left(\frac{2}{\mu} \right)^{n-2} \right]. \tag{A.24}$$

Note that when $m\alpha = 1$, i.e., $m = n-2$, then

$$\int_{S_2} \beta^m \, dy = \frac{n-2}{\mu^{n-2}} \log \left| \frac{\mu}{2} \right| \tag{A.25}$$

Overall, the integrand is bounded by

$$\begin{aligned} g(y) &< y^\alpha \sqrt{1 + \frac{4x}{\mu}} \left(1 + \frac{\epsilon}{2} - \frac{\epsilon^2}{8} \right) \\ &< y^\alpha \sqrt{1 + \frac{4x}{\mu}} \left(1 - \beta\gamma + \frac{1}{2}\gamma\beta^2 (1-\gamma) + \frac{1}{2}\gamma^2\beta^3 \right) \\ &= \frac{1}{\mu} \sqrt{1 + \frac{4x}{\mu}} \left[\frac{1}{\beta} - \gamma + 2x\gamma^2\beta\frac{1}{\mu} + \frac{1}{2}\gamma^2\beta^2 \right] \\ &< \frac{1}{\mu} \sqrt{1 + \frac{4x}{\mu}} \left(\frac{1}{\beta} - \gamma + 2x\beta\frac{1}{\mu} + \frac{1}{2}\beta^2 \right) \end{aligned} \tag{A.26}$$

Integrating both sides yields

$$\begin{aligned} f_2 < \langle k \rangle \sqrt{1 + \frac{4x}{\mu}} \Bigg\{ & \frac{n-2}{n-1} \left[\mu - 2 \left(\frac{2}{\mu} \right)^{n-2} \right] \\ & -\gamma \left(1 - \left(\frac{2}{\mu} \right)^{n-2} \right) + 2xA_1 \frac{1}{\mu} + \frac{1}{2} A_2 \Bigg\} \end{aligned} \tag{A.27}$$

where A_1 and A_2 are defined to be

$$A_1 = \int_{S_2} \beta \, dy > 0, \tag{A.28}$$

$$A_2 = \int_{S_2} \beta^2 \, dy > 0. \tag{A.29}$$

Depending on the value of n, A_2 and A_3 may take different forms. For example, for

$n = 3$,

$$A_1 = \frac{1}{\mu} \log \left| \frac{\mu}{2} \right| \tag{A.30}$$

$$A_2 = -3 \left[\frac{1}{\mu^2} - \frac{1}{2} \left(\frac{2}{\mu} \right) \right] \tag{A.31}$$

and for $n = 4$,

$$A_1 = 2 \left[\frac{1}{\mu} - \frac{1}{2} \left(\frac{2}{\mu} \right)^{n-2} \right], \tag{A.32}$$

$$A_2 = \frac{2}{\mu^2} \log \left| \frac{\mu}{2} \right|. \tag{A.33}$$

otherwise,

$$A_1 = \frac{n-2}{n-3} \left[\frac{1}{\mu} - \frac{1}{2} \left(\frac{2}{\mu} \right)^{n-2} \right] \tag{A.34}$$

$$A_2 = \frac{n-2}{n-4} \left[\frac{1}{\mu^2} - \frac{1}{2^2} \left(\frac{2}{\mu} \right)^{n-2} \right] \tag{A.35}$$

A.4 Asymptotic solution of the upper bound as $\mu \to \infty$

Recall that we are trying to solve the equation

$$\langle u \rangle = \frac{\langle k \rangle - \mu}{2} + \left\langle \sqrt{\left(\frac{k-\mu}{2} \right)^2 + \mu \langle u \rangle} \right\rangle. \tag{A.36}$$

We now use the upper bound in (A.27) to substitute the $f_2(x) \equiv \left\langle \sqrt{(k-\mu)^2 + 4\mu x} \right\rangle$ and solve for x asymptotically as $\mu \to \infty$. Therefore, we solve for (from the upper

bounds of (A.18) and (A.27))

(A.37)
$$x = \frac{\langle k\rangle - \mu}{2} + \frac{\langle k\rangle}{2}\left(\frac{2}{\mu}\right)^{n-2}\left(1+\frac{16x}{\mu}\right)^{1/2} + \frac{\langle k\rangle}{2}\sqrt{1+\frac{4x}{\mu}}\left\{\frac{n-2}{n-1}\left[\mu - 2\left(\frac{2}{\mu}\right)^{n-2}\right] - \gamma\left[1-\left(\frac{2}{\mu}\right)^{n-2}\right] + 2xA_1\frac{1}{\mu} + \frac{1}{2}A_2\right\}.$$

By dividing $\mu/2$ to both sides and plugging in the identity $\langle k\rangle = (n-1)/(n-2)$, we obtain

(A.38)
$$\frac{2x}{\mu} = \frac{1}{\mu}\frac{n-1}{n-2} - 1 + \frac{n-1}{n-2}\frac{1}{\mu}\left(\frac{2}{\mu}\right)^{n-2}\left(1+\frac{16x}{\mu}\right)^{\frac{1}{2}} + \frac{n-1}{n-2}\sqrt{1+\frac{4x}{\mu}}\left\{\frac{n-2}{n-1}\left[\mu - 2\left(\frac{2}{\mu}\right)^{n-2}\right] - \gamma\left[1-\left(\frac{2}{\mu}\right)^{n-2}\right] + 2xA_1\frac{1}{\mu} + \frac{1}{2}A_2\right\}$$

Next, we take the ansatz

(A.39)
$$x \equiv x_b + x_p$$

where x_b is an order 1 number and $x_p \in o(\mu^0)$. Then

(A.40)
$$\frac{2x_b + 2x_p}{\mu} = \frac{1}{\mu}\frac{n-1}{n-2} - 1 + \frac{n-1}{n-2}\frac{1}{\mu^\eta}\left(1+\frac{16x_b+16x_p}{\mu}\right)^{\frac{1}{2}} + \frac{n-1}{n-2}\sqrt{1+\frac{4x_b+4x_p}{\mu}}\times \left\{\frac{n-2}{n-1}\left(1-\frac{2}{\mu^\eta}\right)\left[-\frac{1}{1+\frac{4x_b+4x_p}{\mu}}\left(\frac{1}{\mu}-\frac{1}{\mu^\eta}\right) + 2xA_1\frac{1}{\mu^2} + \frac{1}{2\mu}A_2\right]\right\}.$$

It is algebraically hard to solve for the equation. Therefore, we will perform asymptotic computation to solve the equation as $\mu \to \infty$ for different n. We will need to expand $\sqrt{1+4x_b+4x_p/\mu}$ and $1/(1+4x_b+4x_p/\mu)$ in the following sections

$$\sqrt{1+\frac{4x_b+4x_p}{\mu}} = 1+\frac{2x_b+2x_p}{\mu} - 2\frac{x_b^2+2x_bx_p+x_p^2}{\mu^2} + o\left(\frac{1}{\mu^2}\right), \tag{A.41a}$$

$$\frac{1}{1+\frac{4x_b+4x_p}{\mu}} = 1-\frac{4x_b+4x_p}{\mu} + 16\frac{x_b^2+2x_b\,x_p+x_p^2}{\mu^2} + o\left(\frac{1}{\mu^2}\right). \tag{A.41b}$$

A.4.1 $n > 4$

In the case where $n > 4$, we have $\eta > 3$. Take the ansatz x_p is of $\mathcal{O}(1/\mu)$. We expand (A.40) with (A.41), and after some computation, we arrive at the trivial $\mathcal{O}(1)$ equation

$$0 = -1 + 1, \tag{A.42}$$

and the following $\mathcal{O}(\mu^{-1})$ has

$$\frac{2x_b}{\mu} = \frac{1}{\mu}\frac{n-1}{n-2} + \frac{2x_b}{\mu} - \frac{1}{\mu}\frac{n-1}{n-2}. \tag{A.43}$$

At the $\mathcal{O}(\mu^{-2})$,

$$\frac{2x_p}{\mu} = \frac{2x_p}{\mu} - 2\frac{x_b^2}{\mu^2} - \frac{n-1}{n-2}\frac{2x_b}{\mu^2} \tag{A.44}$$

derives a nontrivial equation

$$\frac{2x_p}{\mu} = \left[\frac{2x_p}{\mu} - 2\frac{x_b^2}{\mu^2} - \frac{1}{\mu}\frac{n-1}{n-2}\frac{2x_b}{\mu} + \frac{1}{\mu^2}\frac{n-1}{n-2}4x_b\right] \tag{A.45}$$

which implies

$$x_b = \frac{n-1}{n-2} = \langle k \rangle. \tag{A.46}$$

To obtain the correction to the trivial mean $\langle k \rangle$, we have to go to $\mathcal{O}(\mu^{-3})$. After some computations we obtain the equation at this order

$$0 = 4x_p \frac{n-1}{n-2}\frac{1}{\mu^2} - 16\frac{n-1}{n-2}\frac{x_b^2}{\mu^3} + 2\,x_b\,\frac{n-1}{n-3}\frac{1}{\mu^3} + \frac{1}{2}\frac{n-1}{n-4}\frac{1}{\mu^3} + \frac{2x_b}{\mu}4x_b\frac{n-1}{n-2}\frac{1}{\mu^2} - \frac{2x_p}{\mu}\frac{n-1}{n-2}\frac{1}{\mu} - 2\frac{x_b^2}{\mu^2}\left(-\frac{n-1}{n-2}\frac{1}{\mu}\right) - 4\frac{x_b x_p}{\mu^2} + 4\frac{x_b^3}{\mu^3}. \tag{A.47}$$

Substituting x_b by (A.46) yields

$$x_p = \left(\frac{(n-1)}{(n-2)^2\,(n-3)} + \frac{1}{4}\frac{n-2}{n-4}\right)\frac{1}{\mu} \tag{A.48}$$

which demonstrates the solution is bounded by $1/\mu$ as $\mu \to \infty$.

A.4.2 $n = 3$

When $n < 3$, as $\mu \to \infty$, numerical computations suggest $\langle u \rangle$ diverges as $\mu \to \infty$. $n = 3$ is the critical point for convergence to $\langle k \rangle$ at large μ. In this section we compute the large μ behavior to the first order and compare to the numerical computations.

When $n = 3$ (consequently, $\alpha = 1$ and $\langle k \rangle = 2$), $g(x)$, Eq.(A.19), can be integrated analytically:

(A.49)

$$g(x) = \langle k \rangle - \mu + 2\mu\left[\frac{2\left(1+\frac{4x}{\mu}\right) - \frac{2}{\mu}}{4\left(1+\frac{4x}{\mu}\right)}\sqrt{1 - \frac{2}{\mu} + \frac{4x}{\mu} + \frac{1}{\mu^2}} + \frac{2}{\mu}\frac{1}{\mu}\left(\frac{1}{4\left(1+\frac{4x}{\mu}\right)}\right)\right] + \frac{2x\mu}{\mu^3\left(1+\frac{4x}{\mu}\right)^{3/2}}\log\left|\left(2+\frac{8x}{\mu}\right) - \frac{4}{\mu} + 2\sqrt{1 + \frac{4x-2}{\mu} + \frac{1}{\mu^2}} + 2\sqrt{\left(1+\frac{4x}{\mu}\right)\left(\frac{1}{\mu^2}\right)}\right|$$

It is again unlikely to solve the equation analytically. Instead, we consider the large

μ approximation, by taking the ansatz

$$x = x_b + x_p + \mathcal{O}\left(\mu^{-2}\right). \tag{A.50}$$

and expand the square roots and logarithm in (A.49). Again we assume x_p is of $\mathcal{O}\left(\mu^{-1}\right)$

The objective of the following analysis is to show that x_b is *not* $\langle k \rangle$. To $\mathcal{O}\left(\mu\right)$, we have a consistent relation

$$0 = -\mu + 2\mu\left(\frac{2}{4}\right), \tag{A.51}$$

and so is to $\mathcal{O}\left(1\right)$,

$$2x_b = 2 + \mu \times \left[-\frac{1}{\mu} - \frac{1}{\mu} + \frac{2x_b}{\mu}\right]. \tag{A.52}$$

To the order $\mathcal{O}\left(\mu^{-1}\right)$, we obtained

$$2x_p = \mu\left[\frac{2x_p}{\mu} + \frac{1}{2\mu^2} - \frac{1}{2\mu^2}\left(2x_b - 1\right)^2 - \frac{1}{\mu}\left(-\frac{1}{\mu} + \frac{2x_0}{\mu}\right) + \frac{4x_0}{\mu^2} + 1\right], \tag{A.53}$$

which can be simplified after cancel x_p out:

$$0 = x_b^2 - 2x_b - 1, \tag{A.54}$$

and the (positive) solution is $x_b = 1 + \sqrt{2}$.

A.5 Small μ approximation

When μ is small, it is elementary to expand series expansion with respect to μ^n

$$
\begin{aligned}
2x &= \langle k\rangle - \mu + \langle k\rangle \int_0^1 \sqrt{\left(1-\mu y^{\alpha}\right)^2 + 4x\mu y^{2\alpha}}dy \\
&= \langle k\rangle - \mu + \langle k\rangle \int_0^1 \sqrt{1 - 2\mu y^{\alpha} + \mu^2 y^{2\alpha} + 4x\mu y^{2\alpha}}dy \\
&= \langle k\rangle - \mu + \langle k\rangle \int_0^1 \left(1 - \mu y^{\alpha} + 2x\mu y^{2\alpha}\right) dy + \mathcal{O}\left(\mu^2\right) \\
&= \langle k\rangle - \mu + \langle k\rangle - \mu + 2x\left\langle k\right\rangle \, \mu\frac{n-2}{n} + \mathcal{O}\left(\mu^2\right),
\end{aligned}
\tag{A.55}
$$

and to the order $\mathcal{O}(\mu)$, and to the order $\mathcal{O}(\mu)$,

$$
2x - 2x\left\langle k\right\rangle \, \mu\frac{n-2}{n} = 2\left[\langle k\rangle - \mu\right] \tag{A.56}
$$

which implies

$$
x = \langle k\rangle + \frac{\mu}{n(n-2)}. \tag{A.57}
$$

A.6 Numerical verification and discussion

The implicit equation (A.8) can be solved numerically by root-finding after the mapping defined in A.2. We develop a set of standard root-finding algorithm to find the root $\langle u\rangle$. Fig. A.1 shows the numerical solution in log-log scale to verify our asymptotic analyses.

In the small μ region, the asymptotic analysis in section A.5 provides accurate description. In the large μ region, the asymptotic computation in section A.4.1 shows $\langle u\rangle - \langle k\rangle$ behaves like μ^{-1} when $n > 4$. When $n = 3$, the numerical solution also

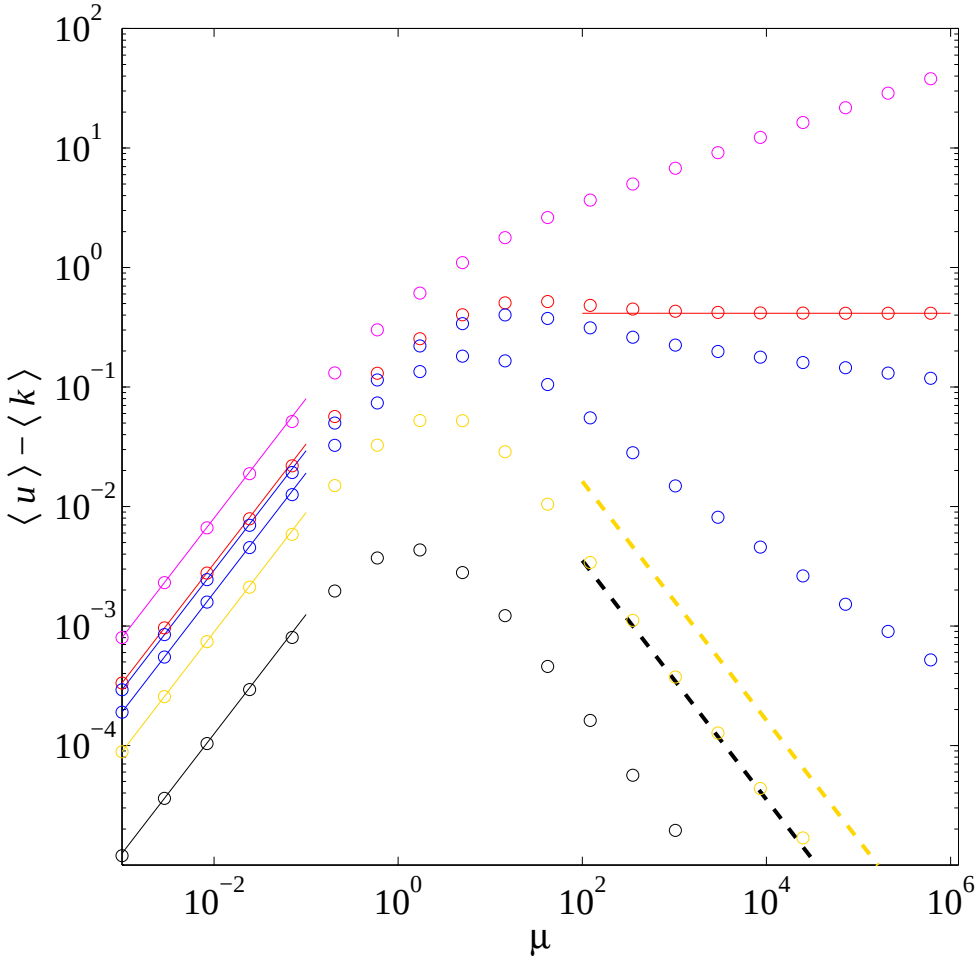

Figure A.1: The solutions of $\langle u \rangle - \langle k \rangle$ from Eq.(A.8). From top to bottom, $n = 2.5, 3.0, 3.1, 3.5, 4.5, 10$. Discrete markers are from directly solving Eq.(A.8) numerically. At low μ, continuous line are asymptotes, Eq.(A.57). At large μ, continuous red line is the asymptotic behavior $\langle u \rangle \to 1 + \sqrt{2}$ from section A.4.2, and the dashed lines are the upper bounds, (A.48), derived in section A.4.1.

verifies our large-μ analysis: $\langle u \rangle = 1 + \sqrt{2}$.

Recall that for our distribution $\mathbf{1}_{\{k \geq 1\}}/k^n$, the moments of k^m

$$\langle k^m \rangle = \int_1^\infty k^{m-n} dk \tag{A.58}$$

diverge when $m \geq n-1$. What is interesting about this system is that the solution of $\langle u \rangle$ has some sort of "critical transitions" at $n = 4$ and $n = 3$. When $n > 4$, the large-μ behavior of $\langle u \rangle$ is $\sim \mu^{-1}$, consistent with the results for a well-behaved environmental distribution which is discussed in Appendix B. This implies when the environmental distribution has a finite *third* moment, our small-σ expansion is good enough to predict the large-μ behavior (at least for this single-species problem.) Between $n = 4$ and $n = 3$ the asymptotic (large-μ) power-law exponent, $d \log(\langle u \rangle - \langle k \rangle)/d\mu$, gradually changes from -1 to 0. When $n = 3$, a sharp critical transition occurs—when $n > 3$, $\lim_{\mu \to \infty} \langle u \rangle = \langle k \rangle$, but at $n = 3$, $\lim_{\mu \to \infty} \langle u \rangle = \langle k \rangle - 1 + \sqrt{2}$. On the other hand, when $n < 3$ numerical solution shows the exponent is positive.

In conclusion, this analysis suggests in an isolated universe with only one species, it is more favorable to disperse and explore the "fat tail" if the environment has a divergent *second* moment. Otherwise, there exists an optimal dispersal rate, indicated by the maxima showed in Fig. A.1

APPENDIX B

Single Species in an Environments with Well-Behaved Distribution

In this Appendix, we briefly document our analysis of the stationary distribution of a single species living in an environment whose every moment exists. The model setting is the same as Appendix A. We seek for the stationary distribution $u(x)$ which satisfies

$$0 = u\left(x\right)\left[k\left(x\right) - u\left(x\right)\right] - \mu\left[u\left(x\right) - \langle u\rangle\right], \tag{B.1}$$

by expanding the "small" environmental variance.

Starting with the assumption

$$k(x) \equiv \langle k\rangle + \tilde{k}(x) \tag{B.2}$$

such that $|\tilde{k}| \ll 1\,\forall x \in \Omega$, we take the ansatz

$$u(x) \equiv u_0 + u_1(x) + u_2(x) + \ldots \tag{B.3}$$

where $\mathcal{O}(u_n) = \mathcal{O}(|\tilde{k}^n|)$ for $n = 1, 2, \ldots$. A straightforward asymptotic computation

shows

$$u_0 = \langle k \rangle, \tag{B.4}$$

$$u_1 = \frac{\langle k \rangle}{\langle k \rangle + \mu} \tilde{k}, \tag{B.5}$$

$$u_2 = \frac{\mu}{(\langle k \rangle + \mu)^3} \left(\langle k \rangle \tilde{k}^2 + \mu \langle \tilde{k}^2 \rangle \right), \tag{B.6}$$

$$u_3 = \frac{\mu(\mu - \langle k \rangle)}{(\langle k \rangle + \mu)^4} \left[\mu \langle \tilde{k}^3 \rangle + \langle k \rangle \tilde{k}^3 + \mu \langle \tilde{k}^2 \rangle \tilde{k} \right], \tag{B.7}$$

Since we are interested in the "average population" in the domain, with

$$\langle \tilde{k} \rangle = 0 \tag{B.8}$$

we arrive at

$$\langle u_0 \rangle = \langle k \rangle, \tag{B.9}$$

$$\langle u_1 \rangle = 0, \tag{B.10}$$

$$\langle u_2 \rangle = \frac{\mu}{(\langle k \rangle + \mu)^2} \langle \tilde{k}^2 \rangle, \tag{B.11}$$

$$\langle u_3 \rangle = \frac{\mu(\mu - \langle k \rangle)}{(\langle k \rangle + \mu)^3} \langle \tilde{k}^3 \rangle. \tag{B.12}$$

Note that the functional form at the lowest order approximation is universal after scaling μ by $\langle k \rangle$. The "optimal" rate of this problem, which can be obtained after elementary calculation, is $\mu_{\text{opt}} = \langle k \rangle$, and the maximum average population the environment is able support (to the lowest order approximation) is $\langle k^2 \rangle / (4 \langle k \rangle)$.

A final remark is that as $\mu \to \infty$, the first order correction scales $1/\mu$ and the prefactor is proportional to the environmental variance $\langle \tilde{k}^2 \rangle$. The scaling relation also shows in Appendix A when $n > 4$.

Numerical verification

We perform numerical simulations to verify our analysis. To justify the robustness of our analysis, in addition to the one we presented in Appendix A, we develop an alternative patch-like setting for the problem in this section.

In the simulation there are 500 patches. The carrying capacities $\{K_i\}_{i=1}^{500}$ of the patches are generated by a certain distribution. Similar to Chapter IV, we choose (1) bounded uniform distribution, (2) bounded normal distribution, and (3) bounded Laplace (double-exponential) distributions. Then we simulate the *dynamics* by integration the following equations of motion

$$\frac{du_i(t)}{dt} = u_i(t)\left(K_i - u_i\right) + \mu\left(\sum_{j=1}^{500} u_j - u_i\right),\ i = 1, 2, \ldots, 500, \tag{B.13}$$

until each patch reaches to the stationary equilibrium. Then we plot the average population among the patches

$$\langle u \rangle \equiv \frac{1}{500}\sum_{j=1}^{500} u_j(t \to \infty) \tag{B.14}$$

versus the control parameter μ.

Note that since the final results of the asymptotic analysis only depend on the moments, to the lowest order, distributions with identical mean and variance will have the identical results. Such prediction is verified by large-scale simulations.

Figs. B.1 and B.2 present the results of bounded uniform distribution and the bounded normal distribution.

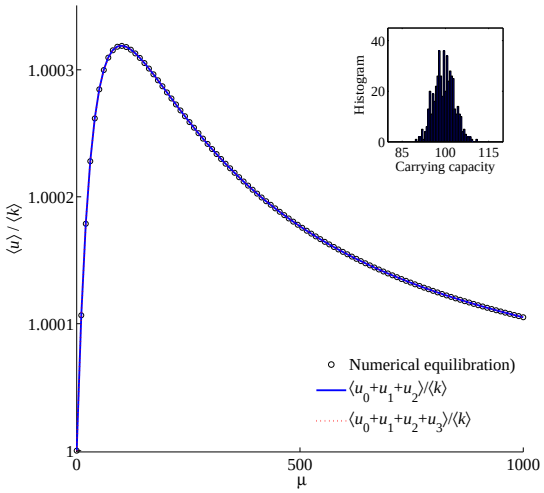

Figure B.1: Results of numerical simulations and asymptotic analysis when the environment is normal distributed.

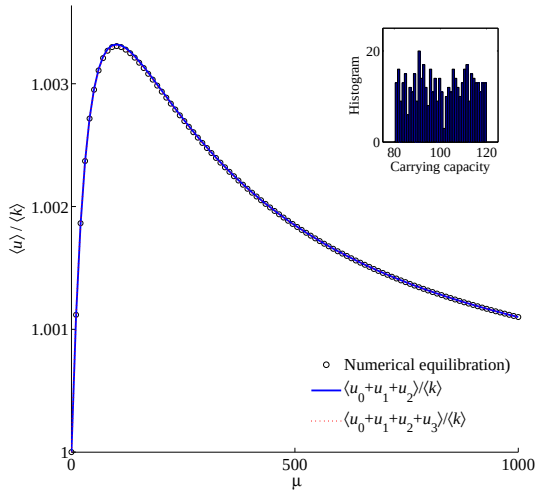

Figure B.2: Results of numerical simulations and asymptotic analysis when the environment is uniformly distributed.

APPENDIX C

Deterministic Competitive Dynamics between Multiple Species

In this Appendix we generalize the (deterministic) pairwise competitive dynamics to (deterministic) competitive dynamics between multiple species with mutation. The analysis is parallel to the analysis in Chapter IV, so we list only the essential steps for documentation. The analysis of this Appendix will serve as a guide when we investigate stochastic competitive dynamics between multiple species in the future.

C.1 The model

We consider an infinite number of species living on two patches. The competition on each patch is still logistic-like, and the species differ only in their diffusion rate. In addition, any individual can possibly mutate. The objective is to find the stationary distribution of the population in the phenotypic space.

Specifically, assume the passive diffusion constant of the species is in some interval:

$$\mu \in \mathcal{D} \equiv (\mu_1, \mu_2)\,, \tag{C.1}$$

with $\mu_2 > \mu_1 > 0$. We will refer such interval $\mathcal{D}$ to be the "phenotypic space" of the system. Let the population density distribution of the species to be $\rho(\mu, t)$ with time variable t, and the strength of unbiased mutation to be α. The reaction–diffusion type of the dynamics on a patch is

$$\frac{\partial \rho}{\partial t} = \alpha \nabla^2 \rho + \rho \left[1 - \int \rho(\mu', t)\, d\mu'\right], \tag{C.2}$$

where the integration is over the phenotypic space $\mathcal{D}$. In this Appendix, we will adopt the notation that any integration symbol represents an integration over the entire phenotypic space $\mathcal{D}$. Now we put in the spatial distribution (see Chapter IV for reference),

$$\frac{\partial \rho_1}{\partial t} = \alpha \nabla^2 \rho_1 + \rho_1 \left[1 - (1+\sigma) \int \rho_1(\mu', t)\, d\mu'\right] + \mu(\rho_2 - \rho_1), \tag{C.3a}$$

$$\frac{\partial \rho_2}{\partial t} = \alpha \nabla^2 \rho_2 + \rho_2^{(0)} \left[1 - (1-\sigma) \int \rho_2(\mu', t)\, d\mu'\right] + \mu(\rho_1 - \rho_2). \tag{C.3b}$$

where $\rho_i(\mu, t)$ is the population density of the species with dispersal rate μ on patch i at time t.

The intuition behind the following analysis is that, in Chapter IV we gained the knowledge that the population of the slower dispersers gradually increases on a time scale $\mathcal{O}(K)$. It is a natural conjecture that the slower species still enjoy the advantage in the multiple-species competition, and as $t \to \infty$, the slowest species dominates the entire population. Mathematically it means there exists a slow drift in the phenotypic space. On the other hand, the mutation forbids the possibility of a single species dominating the entire space. Since the effect of mutation is modeled as a "diffusion" in the phenotypic space, the entire dynamics should have a certain "fluctuation-dissipation" balance in the phenotypic space when the strengths of the effects are comparable. We will show in section

C.1.1 Asymptotic analysis when $\alpha = 0$

We begin with the assumption $\mu = 0$. That is, we consider the problem *without* any mutation. In this case, the dynamics can be described by

$$\frac{\partial \rho_1^{(0)}}{\partial t} = \rho_1^{(0)} \left[1 - (1+\sigma) \int \rho_1^{(0)}(\mu', t)\, d\mu'\right] + \mu\left(\rho_2^{(0)} - \rho_1^{(0)}\right), \tag{C.4}$$

$$\frac{\partial \rho_2^{(0)}}{\partial t} = \rho_2^{(0)} \left[1 - (1-\sigma) \int \rho_2^{(0)}(\mu', t)\, d\mu'\right] + \mu\left(\rho_1^{(0)} - \rho_2^{(0)}\right). \tag{C.5}$$

Following the analysis in Chapter IV, we take the ansatz

$$\rho_1^{(0)} := \rho_1^{(0)} + \sigma\rho_1^{(1)}(t) + \sigma^2\rho_1^{(2)}(t) + ...$$
$$\rho_2^{(0)} := \rho_2^{(0)} + \sigma\rho_2^{(1)}(t) + \sigma^2\rho_2^{(2)}(t) + ...$$

and then we expand the equations of motion with respect to "small" σ.

Now we prove the following theorem which is parallel to Thm. III.1.

Theorem C.1. *To $\mathcal{O}(1)$, the stationary solution is $\rho_1^{(0)} = \rho_2^{(0)}$.*

Proof. Prove by contradiction. At this order $\mathcal{O}(1)$, we have

$$0 = \rho_1^{(0)} \left[1 - \int \rho_1^{(0)}(\mu', t)\, d\mu'\right] + \mu\left(\rho_2^{(0)} - \rho_1^{(0)}\right), \tag{C.6}$$

$$0 = \rho_2^{(0)} \left[1 - \int \rho_2^{(0)}(\mu', t)\, d\mu'\right] + \mu\left(\rho_1^{(0)} - \rho_2^{(0)}\right). \tag{C.7}$$

Suppose for any $\mu \in \mathcal{D}$, $\rho_1^{(0)}(\mu) > \rho_2^{(0)}(\mu)$. This implies

$$\int \rho_1^{(0)}(\mu', t)\, d\mu' < 1, \tag{C.8a}$$

$$\int \rho_2^{(0)}(\mu', t)\, d\mu' > 1. \tag{C.8b}$$

Furthermore, by Eq.(C.6), $\rho_1^{(0)} > \rho_2^{(0)}$ for every $\mu \in \mathcal{D}$. Since $\forall \mu \in \mathcal{D}$, $\rho_1^{(0)}(\mu) > \rho_2^{(0)}(\mu)$, we have a contradiction to the inequalities (C.8). □

Therefore, at order $\mathcal{O}(1)$, we have the trivial constraint:

$$0 = \rho_1^{(0)} \left[1 - \int \rho_1^{(0)}(\mu', t)\, d\mu'\right] \Rightarrow \int \rho_1^{(0)}(\mu', t)\, d\mu' = 1. \tag{C.9}$$

That is, the total population on each patch is the harmonic mean carrying capacity, as expected.

Moving on to order $\mathcal{O}(\sigma)$, we have the equations of motion

$$\sigma\dot{\rho}_1^{(1)} = \left(\rho_1^{(0)} + \sigma\rho_1^{(1)}\right)\left[1 - (1+\sigma)\int\left(\rho_1^{(0)} + \sigma\rho_1^{1}\right)\right] + \sigma\mu\left(\rho_2^{(1)} - \rho_1^{(1)}\right), \tag{C.10}$$

$$\sigma\dot{\rho}_2^{(1)} = \left(\rho_2^{(0)} + \sigma\rho_2^{(1)}\right)\left[1 - (1-\sigma)\int\left(\rho_2^{(0)} + \sigma\rho_2^{(1)}\right)\right] + \sigma\mu\left(\rho_1^{(1)} - \rho_2^{(1)}\right). \tag{C.11}$$

We plug in $\int \rho_1^{(0)} = \int \rho_2^{(0)} = 1$ to obtain

$$\sigma\dot{\rho}_1^{(1)} = \left(\rho_1^{(0)} + \sigma\rho_1^{(1)}\right)\left[1 - (1+\sigma)\left(1 + \sigma\int\rho_1^{(1)}\right)\right] + \sigma\mu\left(\rho_2^{(1)} - \rho_1^{(1)}\right), \tag{C.12}$$

$$\sigma\dot{\rho}_2^{(1)} = \left(\rho_2^{(0)} + \sigma\rho_2^{(1)}\right)\left[1 - (1-\sigma)\left(1 + \sigma\int\rho_2^{(1)}\right)\right] + \sigma\mu\left(\rho_1^{(1)} - \rho_2^{(1)}\right). \tag{C.13}$$

To order $\mathcal{O}(\sigma)$,

$$\dot{\rho}_1^{(1)} = -\rho_1^{(0)}\left(1 + \int\rho_1^{(1)}\right) + \mu\left(\rho_2^{(1)} - \rho_1^{(1)}\right), \tag{C.14}$$

$$\dot{\rho}_2^{(1)} = -\rho_2^{(0)}\left(-1 + \int\rho_2^{(1)}\right) + \mu\left(\rho_1^{(1)} - \rho_2^{(1)}\right). \tag{C.15}$$

Define the total population density (among patches) at this order to be $\rho^{(1)}(\mu, t) \equiv \rho_1^{(1)}(\mu, t) + \rho_2^{(1)}(\mu, t)$, and by Thm. C.1, we define $\rho_0 = \rho_1^{(0)} = \rho_2^{(0)}$. Then

$$\dot{\rho}^{(1)}(\mu, t) = -\rho_0(\mu)\int\rho(\mu', t)\, d\mu'. \tag{C.16}$$

If we integrate over the phenotypic space to get the total population of all the species

$\Theta^{(1)}(t) \equiv \int \rho^{(1)}(\mu', t)\, d\mu'$, then $\Theta^{(1)}(t)$ has simple dynamics

$$\dot{\Theta}^{(1)} = -\Theta^{(1)}. \tag{C.17}$$

Clearly the solution is

$$\Theta^{(1)}(t) = e^{-t}\Theta^{(1)}(0), \tag{C.18}$$

and the dynamics of $\rho^{(1)}$ is

$$\dot{\rho}^{(1)}(\mu, t) = -e^{-t} \times \rho_0(\mu) \times \Theta^{(1)}(0). \tag{C.19}$$

The intuition behind the equation above is clear: at this order, the total populations of the system exponentially decays to 0. For each species, the fraction of its population to the total population is determined by the initial distribution $\rho_0(\mu)$. The convergence rate to 0 is uniformly e^{-t} among species. With the initial condition $\rho^{(1)}(\mu, t = 0) = \rho^{(1)}(\mu)$, then the solution of (C.19) is

$$\rho^{(1)}(\mu, t) = \rho_0(\mu) \times \Theta^{(1)}(0) \times e^{-t} + \left(\rho^{(1)}(\mu, 0) - \rho_0(\mu) \times \Theta^{(1)}(0)\right), \tag{C.20}$$

which implies as $t \to \infty$ the population density of such species is equal to

$$\rho(\mu) = \rho^{(1)}(\mu, 0) - \rho_0(\mu) \times \Theta^{(1)}(0) \tag{C.21}$$

with the constraint of $\Theta^{(1)}(0)$

$$\Theta^{(1)}(0) = \int \rho^{(1)}(\mu', 0)\, d\mu'. \tag{C.22}$$

On the other hand, the dynamics of the *difference* of the population densities

$\delta^{(1)} := \rho_1^{(1)} - \rho_2^{(1)}$ are

$$\dot{\delta}^1 = -\rho_1^{(0)} \left(2 + \int \delta^1\right) - 2\mu\delta^{(1)}. \tag{C.23}$$

The "fixed-point distribution" δ_*^1 satisfies:

$$0 = -\frac{\rho_1^{(0)}}{2\mu} \left(2 + \int \delta_*^{(1)}\right) - \delta_*^{(1)}. \tag{C.24}$$

After integrating over $\mathcal{D}$, and defining

$$\Phi \equiv \int \frac{\rho_1^{(0)}(\mu')}{2\mu'} d\mu', \tag{C.25}$$

$$\Delta_*^{(1)} \equiv \int \delta_*^{(1)}, \tag{C.26}$$

we have

$$0 = -\Phi \left[2 + \Delta_*^{(1)}\right] - \Delta_*^{(1)}, \tag{C.27}$$

as a consequent,

$$\Delta_*^{(1)} = \frac{-2\Phi}{1 + \Phi}. \tag{C.28}$$

For the distribution of a single species, we have

$$\delta_*^{(1)}(\mu) = \frac{-\rho_0(\mu)}{\mu} \left(\frac{1}{1 + \Phi}\right) \tag{C.29}$$

We ignore the transient dynamics. The intuition behind this is that the transient parts are not involved in determining the effective drifts, as depicted in Chapter IV.

To sum up, at $\mathcal{O}(\sigma^1)$, we have the difference of the population density among

patches will converged to a metastable distribution

$$\delta^{(1)}(\mu, t \to \infty) = \frac{-\rho_0(\mu)}{\mu}\left(\frac{1}{1+\Phi}\right), \tag{C.30}$$

and we will assume to this order the total population density among patches is 0 (see Chapter V):

$$\rho^{(1)}(\mu, t \to \infty) = 0. \tag{C.31}$$

Therefore, we arrive at the metastable distribution $\rho_{i*}^{(1)}$

$$\rho_{1*}^{(1)}(\mu, t \to \infty) = -\frac{1}{2}\frac{\rho_0(\mu)}{\mu}\left(\frac{1}{1+\Phi}\right), \tag{C.32}$$

$$\rho_{2*}^{(1)}(\mu, t \to \infty) = \frac{1}{2}\frac{\rho_0(\mu)}{\mu}\left(\frac{1}{1+\Phi}\right). \tag{C.33}$$

Next, consider $\mathcal{O}(\sigma^2)$. The dynamics are

$$\dot{\rho}_1^{(2)} = -\rho_0\left(\int \rho_1^{(1)} + \int \rho_1^{(2)}\right) - \rho_1^{(1)}\left(1 + \int \rho_1^{(1)}\right) + \mu\left(\rho_2^{(2)} - \rho_1^{(2)}\right), \tag{C.34}$$

$$\dot{\rho}_2^{(2)} = -\rho_0\left(-\int \rho_2^{(1)} + \int \rho_2^{(2)}\right) - \rho_2^{(1)}\left(-1 + \int \rho_2^{(1)}\right) + \mu\left(\rho_2^{(1)} - \rho_1^{(2)}\right). \tag{C.35}$$

The total population density among patches at this order, defined to be $\rho^{(2)} \equiv \rho_1^{(2)} + \rho_2^{(2)}$, has equation of motion

$$\dot{\rho}^{(2)} = -\rho_0\left(\int \delta^{(1)} + \int \sigma^{(2)}\right) - \delta^{(1)} - \left[\rho_1^{(1)}\int \rho_1^{(1)} + \rho_2^{(1)}\int \rho_2^{(1)}\right]. \tag{C.36}$$

Parallel computations to the ones in Chapter IV yield

$$\dot{\Theta}^{(2)} = -\Delta^{(1)} - \Theta^{(2)} - \Delta^{(1)} - \left[\left(\int \rho_1^{(1)}\right)^2 + \left(\int \rho_2^{(1)}\right)^2\right]. \tag{C.37}$$

The stationary solution is

$$\Theta_*^{(2)} = -2\Delta_*^{(1)} - \left[\left(\int \rho_{1*}^{(1)}\right)^2 + \left(\int \rho_{2*}^{(1)}\right)^2\right]$$

As a consequence, as $t \to \infty$, the total population density of species with dispersal rate μ changes with a rate

(C.38)
$$-\rho_0\left\{-\Delta_*^{(1)} - \left[\left(\int \rho_{1*}^{(1)}\right)^2 + \left(\int \rho_{2*}^{(1)}\right)^2\right]\right\} - \delta_*^1 - \left(\rho_{1*}^{(1)}\int \rho_{1*}^{(1)} + \rho_{2*}^{(1)}\int \rho_{2*}^{(1)}\right).$$

Observe that

$$\int \rho_{1*}^{(1)} = \int -\frac{1}{2}\frac{\rho_0}{\mu}\left(\frac{1}{1+\Phi}\right) = -\frac{\Phi}{1+\Phi}, \tag{C.39}$$

$$\int \rho_{2*}^{(1)} = \int \frac{1}{2}\frac{\rho_0}{\mu}\left(\frac{1}{1+\Phi}\right) = \frac{\Phi}{1+\Phi}, \tag{C.40}$$

which implies

$$\left[\left(\int \rho_1^{(1)}\right)^2 + \left(\int \rho_2^{(1)}\right)^2\right] = 2\frac{\Phi^2}{(1+\Phi)^2}, \tag{C.41}$$

and straightforward computations establish the following "effective dynamics"

$$\dot{\rho}^{(2)} = \frac{\rho_1^{(0)}}{(1+\Phi)^2}\left(\frac{1}{\mu} - 2\Phi\right). \tag{C.42}$$

This analysis provides us a detail picture, similar to the picture in Chapter IV. Starting from any initial condition, in a short amount of time, the "distribution" of the species converges to a metastable distribution—on patch 1 and 2, the population densities of the species with dispersal rate μ is equal to $\rho_0 - \rho_{1*}^{(1)}$ and $\rho_0 + \rho_{1*}^{(1)}$ respectively. The total population on each patch of such metastable distributions fail to match the carrying capacities, and consequentially they produce a higher order

($\mathcal{O}(\sigma^2)$) nonlinear motion, which is effectively described by Eq.(C.42).

C.2 Effective dynamics, dynamical interpretation, and numerical simulations

The analysis in section C.1.1 shows that the phenotypic difference in the propensity to relocate in the space results in a effective drift that drives the slower species to increase their population. The critical measure is the *weighted harmonic mean*, 2Φ in (C.25), of the dispersal rates; that is, if a species is moving *slower* than $1/(2\Phi)$, its population increases.

Most importantly, the time scale of the dynamics reveals in the analysis; similar to the problem in Chapter IV, the time scale is $\mathcal{O}(\sigma^{-2})$.

The mutation, on the other hand, is modeled by pure diffusion in the phenotypic space $\mathcal{D}$. It is well-known that the such process has a time scale that is proportional to its diffusivity, i.e., α^{-1}. An time-scale argument similar to the one we made in Chapter V can be made as follows. When $\mathcal{O}(\sigma^2/\alpha) \gg 1$, the distribution is rather singular: the population will be dominated by the species with the slowest rate. On the other hand, when $\mathcal{O}(\sigma^2/\alpha) \ll 1$ the distribution in the phenotypic space is uniform due to strong effect of mutation. Only when $\mathcal{O}(\sigma^2/\alpha) \approx 1$, we have a nontrivial interaction between the "selection by nonlinear dynamics" and the "mutation". In such scenario, we will have a nontrivial distribution with finite width determined by the "fluctuation"—the mutation—and the "dissipation"—the effective drift due to nonlinear dynamics—in the phenotypic space. This reminds us the famous fluctuation–dissipation theorem in statistical mechanics.

Finally, the "effective" dynamics in this scenario can be formulated as

$$\frac{\partial \rho(\mu,t)}{\partial t} = \nabla^2 \rho(\mu,t) + \sigma^2 \frac{\rho(\mu,t)}{(1+\Phi)^2}\left(\frac{1}{\mu} - 2\Phi\right), \tag{C.43}$$

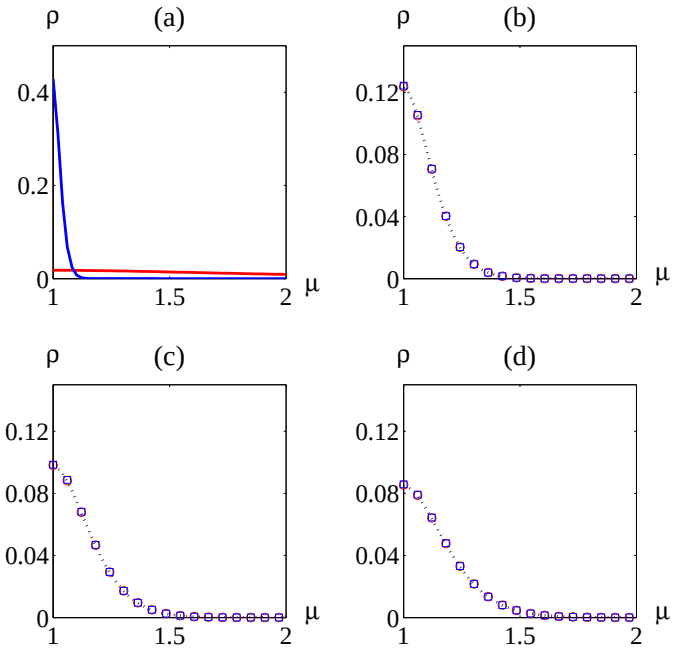

Figure C.1: Numerical simulation of a system with $\mu_1 = 1$ and $\mu_2 = 3$. (a) The stationary solution obtained by direct simulation of Eqs.(C.3). The blue distribution demonstrates when $\alpha = 10^{-4}$ and $\sigma = 0.1$, the effective drift due to nonlinear demographic dynamics dominates the dynamics. The distribution is therefore sharply peaked at the slowest species $\mu = \mu_1 = 1$. The red distribution demonstrates when $\alpha = 1$ and $\sigma = 0.1$, mutation dominates and we observe a rather uniform distribution. In (b-d), α/σ^2 is 0.5, 1, and 1.5 respectively. The dotted lines are the stationary distributions which satisfy the effective dynamics (C.44), and the discrete markers are from direct simulation of Eqs.(C.3). The circles and the squares, represent $\sigma = 0.05$ and $\sigma = 0.1$ respectively.

and the stationary distribution $\rho_*(\mu)$, which satisfies

$$0 = \nabla^2 \rho_*(\mu) + \sigma^2 \frac{\rho_*(\mu)}{(1+\Phi)^2} \left(\frac{1}{\mu} - 2\Phi \right), \tag{C.44}$$

can be obtained by numerical solution (or alternatively, standard asymptotic analysis.) Fig. C.1 shows numerically obtained distributions from simulating the dynamics, Eqs.(C.3), and the solutions of Eq.(C.44) agree, and the results verify our dynamical interpretations.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] C. Cadet, R. Ferriere, J. A. J. Metz, and M. van Baalen. The evolution of dispersal under demographic stochasticity. *American Naturalist*, 162(4):427–441, 2003.

[2] R. S. Cantrell, C. Cosner, D. L. Deangelis, and V. Padron. The ideal free distribution as an evolutionarily stable strategy. *Journal of Biological Dynamics*, 1(3):249–271, 2007.

[3] D. Cohen and S. A. Levin. Dispersal in patchy environments: The effects of temporal and spatial structure. *Theoretical Population Biology*, 39(1):63 – 99, 1991.

[4] H. N. Comins, W. D. Hamilton, and R. M. May. Evolutionarily stable dispersal strategies. *Journal of Theoretical Biology*, 82(2):205–230, 1980.

[5] H. N. Comins, M. P. Hassell, and R. M. May. The spatial dynamics of host parasitoid systems. *Journal of Animal Ecology*, 61(3):735–748, 1992.

[6] R. Cressman, V. Krivan, and J. Garay. Ideal free distributions, evolutionary games, and population dynamics in multiple-species environments. *American Naturalist*, 164(4):473–489, 2004.

[7] N. Dalal, D. Greenhalgh, and X. Mao. A stochastic model for internal hiv dynamics. *Journal of Mathematical Analysis and Applications*, 341(2):1084–1101, 2008.

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

[9] C. R. Doering, K. V. Sargsyan, and L. M. Sander. Extinction times for birth-death processes: Exact results, continuum asymptotics, and the failure of the Fokker-Planck approximation. *Multiscale Modeling & Simulation*, 3(2):283–299, 2005.

[10] C. R. Doering, K. V. Sargsyan, L. M. Sander, and E. Vanden-Eijnden. Asymptotics of rare events in birth-death processes bypassing the exact solutions. *Journal of Physics-Condensed Matter*, 19(6), 2007.

[11] R. Durrett and L. Popovic. Degenerate diffusions arising from gene duplication models. *Annals of Applied Probability*, 19(1):15–48, 2009.

[12] S. D. Fretwell and H. L. Jr Lucas. On territorial behavior and other factors influencing habitat distribution in birds. i. theoretical development. *Acta Biotheoretica*, 19(1):16–36, 1969.

[13] M. Gadgil. Dispersal - population consequences and evolution. *Ecology*, 52(2):253–261, 1971.

[14] C. W. Gardiner. *Handbook of stochastic methods for physics, chemistry, and the natural sciences*. Springer-Verlag, 1983.

[15] W. D. Hamilton. Extraordinary sex ratios. *Science*, 156(3774):477–488, 1967.

[16] W. D. Hamilton and R. M. May. Dispersal in stable habitats. *Nature*, 269(5629):578–581, 1977.

[17] M. P. Hassell, H. N. Comins, and R. M. May. Spatial structure and chaos in insect population-dynamics. *Nature*, 353(6341):255–258, 1991.

[18] M. P. Hassell, H. N. Comins, and R. M. May. Species coexistence and self-organizing spatial dynamics. *Nature*, 370(6487):290–292, 1994.

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

[20] R. D. Holt. population-dynamics in 2-patch environments - some anomalous consequences of an optimal habitat distribution. *Theoretical Population Biology*, 28(2):181–208, 1985.

[21] R. D. Holt and M. A. McPeek. Chaotic population dynamics favors the evolution of dispersal. *American Naturalist*, 148(4):709–718, 1996.

[22] S. Huang. The molecular and mathematical basis of waddington's epigenetic landscape: a framework for post-darwinian biology? *Bioessays*, 34(2):149–57, 2012.

[23] A. Kamina, R. W. Makuch, and H. Y. Zhao. A stochastic modeling of early hiv-1 population dynamics. *Mathematical Biosciences*, 170(2):187–198, 2001.

[24] G. S. Katzenberger. Solutions of a stochastic differential-equation forced onto a manifold by a large drift. *Annals of Probability*, 19(4):1587–1628, 1991.

[25] D. A. Kessler and L. M. Sander. Fluctuations and dispersal rates in population dynamics. *Physical Review E*, 80(4):041907, 2009.

[26] E. Khain, Y. T. Lin, and L. M. Sander. Fluctuations and stability in front propagation. *Europhysics Letters*, 93(2), 2011.

[27] T. G. Kurtz. Solutions of ordinary differential equations as limits of pure jump markov processes. *Journal of Applied Probability*, 7(1):49–58, 1970.

[28] T. G. Kurtz. Limit theorems for sequences of jump markov processes approximating ordinary differential processes. *Journal of Applied Probability*, 8(2):344–356, 1971.

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

[30] J. A. J. Metz and M. Gyllenberg. How should we define fitness in structured metapopulation models? including an application to the calculation of evolutionarily stable dispersal strategies. *Proceedings of the Royal Society B-Biological Sciences*, 268(1466):499–508, 2001.

[31] T. L. Parsons and C. Quince. Fixation in haploid populations exhibiting density dependence ii: The quasi-neutral case. *Theoretical Population Biology*, 72(4):468–479, 2007.

[32] T. L. Parsons, C. Quince, and J. B. Plotkin. Absorption and fixation times for neutral and quasi-neutral populations with density dependence. *Theoretical Population Biology*, 74(4):302–310, 2008.

[33] K. Parvinen, U. Dieckmann, M. Gyllenberg, and J. A. J. Metz. Evolution of dispersal in metapopulations with local density dependence and demographic stochasticity. *Journal of Evolutionary Biology*, 16(1):143–153, 2003.

[34] J. M. Read and M. J. Keeling. Stochasticity generates an evolutionary instability for infectious disease. *Ecology Letters*, 10(9):818–827, 2007.

[35] Russell Schwartz. *Biological modeling and simulation: a survey of practical models, algorithms, and numerical methods*. MIT Press, 2008.

[36] Steven H. Strogatz. *Nonlinear dynamics and chaos : with applications to physics, biology, chemistry, and engineering*. Westview Press, 2000.

[37] J. M. J. Travis and C. Dytham. The evolution of dispersal in a metapopulation: a spatially explicit, individual-based model. *Proceedings of the Royal Society B-Biological Sciences*, 265(1390):17–23, 1998.

[38] H. C. Tuckwell and E. Le Corfec. A stochastic model for early hiv-1 population dynamics. *Journal of Theoretical Biology*, 195(4):451–463, 1998.

[39] H. C. Tuckwell and F. Y. M. Wan. First passage time to detection in stochastic population dynamical models for hiv-1. *Applied Mathematics Letters*, 13(5):79–83, 2000.

[40] T. G. Vaughan, P. D. Drummond, and A. J. Drummond. Within-host demographic fluctuations and correlations in early retroviral infection. *Journal of Theoretical Biology*, 295:86–99, 2012.

[41] J. N. Waddell, L. M. Sander, and C. R. Doering. Demographic stochasticity versus spatial variation in the competition between fast and slow dispersers. *Theoretical Population Biology*, 77(4):279–286, 2010.

[42] J. Wang, L. Xu, E. Wang, and S. Huang. The potential landscape of genetic circuits imposes the arrow of time in stem cell differentiation. *Biophys J*, 99(1):29–39, 2010.