MULTIPLE TIMESCALES IN MICROBIAL INTERACTIONS

Thi Minh Thao Le

UNIVERSITÉ DE TOURS

ÉCOLE DOCTORALE : MIPTIS

Institut Denis Poisson

THÈSE présentée par :

Thi Minh Thao LE

soutenue le : **10 Décembre 2021**

pour obtenir le grade de : **Docteur de l'université de Tours**

Discipline/ Spécialité : Mathématiques

ÉCHELLES DE TEMPS MULTIPLES DANS LES INTERACTIONS MICROBIENNES

THÈSE dirigée par :

M. ANDREIANOV Boris	Professeur à l'Université de Tours
Mme. GJINI Erida	Chargée de recherche à l'Université de Lisbonne
M. MADEC Sten	Maître de conférence à l'Université de Tours

RAPPORTEURS :

M. POGGIALE Jean-Christophe	Professeur à Aix Marseille Université
M. RAPAPORT Alain	Chargé de recherche à INRAE, Montpellier

JURY :

M. ANDREIANOV Boris	Professeur à l'Université de Tours
Mme. COBBOLD Christina	Professeur à l'Université de Glasgow
Mme. EFTIMIE Raluca	Professeur à l'Université de Franche-Comté
Mme. GJINI Erida	Chargée de recherche à l'Université de Lisbonne
M. MADEC Sten	Maître de conférence à l'Université de Tours
M. POGGIALE Jean-Christophe	Professeur à Aix Marseille Université
M. RAPAPORT Alain	Chargé de recherche à INRAE, Montpellier
Mme. RIBOT Magali	Professeur à l'Université d'Orléans- *(Président du jury)*

"Do not mistake the finger pointing to the moon for the moon;
but to truly see the moon, such is unavoidable."
-Requote from ***Lord Buddha Shakyamuni***-

Acknowledgements

First and foremost, I would like to thank my supervisors for giving me this opportunity to work with this thesis for three years.

I thank Professor Boris Andreianov, my master and Ph.D. thesis supervisor, for his extraordinary knowledge and viewpoint in analysis (PDEs and dynamical systems), which gives me great intuition and motivation to deeply comprehend and connect my fundamental theory and ideas. I am also grateful to him for his enthusiastic help and support when I first came to France and my mathematical problems, even on weekends or vacations.

I thank Professor Sten Madec, the main supervisor for my Ph.D. thesis, for his exceptional comprehension and viewpoint in mathematics, not only in the problems of my thesis but also in many related aspects. Under his supervision, I can reinforce and expand my knowledge in the theory of dynamical systems. Furthermore, I owe a lot to his patience with my shortcoming at the beginning and mistakes throughout my works.

I thank Professor Erida Gjini, my Ph.D. thesis supervisor, for her prodigious expertise. I am always filled with gratitude for her precise and apparent explanations about mathematical biology. Under her passionate supervision, I can understand and improve my biology sense, which will help my career. Alternatively, I owe very much to her sharing and encouragement through our talks or lunches together.

I also would like to express my grate to other professors at the University of Tours. I thank Professor Marc Peigné, who kindly helped me very much when I came to France. Moreover, I thank him for his course "Introductory Markov chain" in my master's degree and his enthusiastic help with clear explanations for my probability problems. I thank Professor Guy Barles, who kindly helped me with my PDEs problems with very detailed explanations and his enthusiastic suggestion of references.

I thank Professor Jean-Christophe Poggiale and Professor Alain Rapaport, rapporteurs of my thesis, who gave me remarks and comments that I can improve my manuscript. I thank the faculty thesis committee, Professor Christina Cobbold, Professor Raluca Eftimie and Professor Magali Ribot, for spending their time on my thesis defense. I also thank Professor Magali Ribot for her very nice and warm treat to me.

Professors, I'm always thankful for your talents and generosities.

I thank the University of Tours and the Laboratory of Mathematics and Theoric Physics (Institute Denis Poisson), including all of the Professors, Ms. Elodie, Ms. Anouchka, Mrs. Sandrine, Ms. Oriane Mousset (Mobility Center), etc., for theirs support that has enabled a foreigner student like me to carry out my research.

I thank Dr. Thi Da Cam Pham, an elder sister, who always cares, supports, and helps me with my troubles. Thank you, sister Cam, for your help and the very delicious food you cook for me.

I thank my other friends in France and Vietnam, my former teachers and lecturers in Vietnam, who supported and helped me.

I want to dedicate this thesis to my parents, who may understand nothing of it, but they always care for me, protect me and support me from my childhood until now. I thank my father, who always cares for me in his

very own way, always scares that I get hungry and tries to cook any food for me. Especially, I want to thank my mother, who always loves me unconditionally and gives me all the best despite our condition was not well-off in the past. She encouraged me to study from my childhood and always makes the best conditions for my study and my future, although she may forget to care for herself. She happily ate the cheap food to spend money on tasty treats for me. She may spend thriftily for herself, but she willingly pays for all of my study costs. Without her care, I could never grow up and study. Thank you, my parents, for giving me this life and always being my family and motivation.

I thank my former mathematics teacher, Mr. Van Quy Nguyen, who passed away last year because of cancer. I thank him for his help from my high school time until later and for his encouragement to follow my degree.

I thank my boyfriend's parents for theirs cares when I was in Vietnam and even here. They always support me to follow my degree. I thank them for the presents and food from Vietnam they sent me when I could not go to my home country because of the pandemic.

Last but not least, I thank my boyfriend, who may be called "my fiancé," Mr. Ta Dang Khoa Le – doctoral student at Sorbonne University. Since our first meeting nine years ago, thank you for always showing me beauties and treasures that I never knew before. Thank you for our hour-lasting talks, in which you tell me about the intuition of many things in life and mathematics, about the inspiring stories and people. Thank you for the care, protection, encouragement, support, and help that you do for me every time I need it, no matter what situation you are in, even when you are at your work's deadline or sick. Without your love, I could never have come so far and become a better version of myself. Thank you, in all languages and billions of times. *"Tu es le soleil"*.

Thank you again, for all!

Abstract

The purpose of this thesis is the theoretical and numerical study of an epidemiological model of multi-strain co-infection. Depending on the situation, the model is written as ordinary differential equations or reaction-advection-diffusion equations.

In all cases, the model is written at the host population level on the basis of a classical susceptible-infected-susceptible system (SIS).

The infecting agent is structured into N strains, which differ according to 5 traits: transmissibility, clearance rate of single infections, clearance rate of double infections, probability of transmission of strains, and co-infection rates.

The resulting system is a large system (N^2+N+1 equations) whose complete theoretical study is generally inaccessible. This thesis is therefore based on a simplifying assumption of trait similarity - the so-called quasi-neutrality assumption. In this framework, it is then possible to implement Tikhonov-type time scale separation methods. The system is thus decomposed into two simpler subsystems. The first one is a so-called neutral system - i.e., the value of the traits of all the strains are equal - which supports a detailed mathematical analysis and whose dynamics turn out to be quite simple. The second one is a "replication equation" type system that describes the frequency dynamics of the strains and contains all the complexity of the interactions between strains induced by the small variations in the trait values.

The first work explicitly determines the slow system in an aspatial framework for N strains using a system of ordinary differential equations and justifies that this system describes the complete system well. This system is a replication system that can be described using the $N(N-1)$ fitnesses of interaction between the pairs of strains. It is shown that these fitnesses are a weighted average of the perturbations of each trait.

The second work consists in using explicit expressions of these fitnesses to describe the dynamics of the pairs (i.e. the case $N=2$) exhaustively.

This part is illustrated with many simulations, and applications on vaccination are discussed.

The last work consists in using this approach in a spatialized framework. The SIS model is then a reaction-diffusion system in which the coefficients are spatially heterogeneous. Two limiting cases are considered: The case of an asymptotically small diffusion coefficient and the case of an asymptotically large diffusion coefficient. In the case of slow diffusion, we show that the slow system is a system of type "replication equations", describing again the temporal but also spatial evolution of the frequencies of the strains. This system is of the reaction-advection-diffusion type, the additional advection term explicitly involving the heterogeneity of the associated neutral system. In the case of fast diffusion, classical methods of aggregation of variables are used to reduce the spatialized SIS problem to a homogenized SIS system on which we can directly apply the previous results.

Keywords: multiple timescales, microbial, interactions, replicator, ODEs, PDEs

Résumé

Cette thèse a pour objet l'étude théorique et numérique d'un modèle épidémiologique de co-infection multi-souche. Selon la situation considéré, le modèle s'écrit sous la forme d'un système d'équations différentielles ordinaires ou d'équations de réaction-advection-diffusion.

Dans tous les cas, le modèle s'écrit à l'échelle de la population hôte sur la base d'un classique système susceptibles-infectés-susceptibles (SIS).

L'agent infectueux est structuré en N souches, qui différent selon 5 traits : la transmissibilité, le taux de clairance des infections simples, le taux de clairance des infections doubles, la probabilité de transmission des souches et les taux de co-infection.

Le système obtenu est un système de grande taille (N^2+N+1 équations) dont l'étude théorique complète est inaccessible en général. Cette thèse se fonde donc sur une hypothèse simplificatrice de similarité des traits - que l'on nomme hypothèse de quasi-neutralité. Dans ce cadre, il est alors possible de mettre en oeuvre des méthodes de séparations des échelles de temps de type Tikhonov. Le système est ainsi décomposé en deux sous-systèmes plus simple. Le premier est un système dit neutre - c'est-à-dire dans lequel la valeur des traits de toutes les souches sont égales - qui supporte une analyse mathématique détaillée et dont la dynamique s'avère assez simple. Le second se trouve être un système de type "équations de réplication" qui décrit la dynamique en fréquence des souches et contient toute la complexité des interactions entre souche qu'induit les petites variations dans les valeurs des traits.

Le premier travail consiste à déterminer explicitement le système lent dans un cadre aspatial pour N souches faisant intervenir un système d'équations différentielles ordinaires et à justifier, que ce système décrit bien le système complet. Ce système est un système de réplication qui peut être décrit à l'aide des $N(N-1)$ fitnesses d'interaction entre les paires de souche. Il est montré que ces fitnesses sont une moyenne pondéré des perturbations de chaque traits.

Le second travail consiste a utiliser les expression explicite de ces fitnesses pour décrire exhaustivement la dynamique des paires (c'est-à-dire le cas $N=2$).

Cette partie est illustré à l'aide de beaucoup de simulations des application sur la vaccination sont discutées.

Le dernier travail consiste à reprendre cette approche dans un cadre spatialisé. Le modèle SIS est alors un système de réaction-diffusion dans lequelle les coefficients sot spatiallement hétérogènes. Deux cas limites sont considérés : Le cas d'un coefficient de diffusion asymptotiquement petit et celui d'un coefficient de diffusion asymptotiquement grande. Dans lcas de la diffusion lente on montre que le système lent est un système de type "équations de réplication" , décrivant à nouveau l'évolution temporelles mais également spatiale des fréquences des souches. Ce système est de type réaction-advection-diffusion, le terme d'advection additionelle faisant intervenir explicitement l'hétérogénéité du système neutre associé. Dans le cas de la diffusion rapide, l'utilisation de méthodes classiques d'aggrrégation des variables sont utilisées pour ramener le problème SIS spatialisé à un système SIS homogénéisé sur lequel les résultats précédents peuvent directement s'appliquer.

Mots clés: échelles de temps multiples, microbiennes, interactions, replication, EDO, EDP

Table of contents

List of figures

List of tables

Chapter 0

General introduction

Scientists use models to approximate reality, but reality is a complicated and unpredictable place, rife with uncertainty. There is no such thing as a model exactly representing the world or which can possibly capture everything of the dynamics in the real world. One reason is that small differences in the parameters are amplified as the dynamics of the system progresses. The butterfly effect, an underlying principle of deterministic chaos, describes how a small change in one state of a deterministic nonlinear system can result in large differences in a later state (meaning that there is sensitive dependence on initial conditions) [133]. So, lightly different initial conditions may lead to a very different outcome, which becomes a spreading cone of possible trajectories of the future. However, we may know this quote: "All models are wrong, but some are useful" (by the British Statistician George E. P. Box - who is called "one of the great statistical minds of the 20th century"), which is a common aphorism in statistics. It is usually considered to be applicable to not only statistical models, but to scientific models generally. The aphorism recognizes that statistical or scientific models always fall short of the complexities of reality but can still be of use. Instead of predictions, something that modelers can offer us is projections. These are glimpses into a range of possible futures that we have the power to change. This may be thought as we drive a car, and by the rearview mirror, we may see what may happen and avoid accidents. This is what a useful model helps.

Because there are quantifiable aspects of life science, mathematics plays a critical role in better understanding the natural world. Mathematical biology is a field of research that examines mathematical representations of biological systems. Rich ecosystems comprise many species interacting together in various ways and on multiple temporal and spatial scales. Understanding the scope and consequences of such interactions has been the focus of countless theoretical ecology studies, starting with the seminal work by Lotka and Volterra ([136, 203]) on mathematical models of the population dynamics of interacting species. One key role of math in biology is the creation of mathematical models. These equations or formulas can predict or describe natural occurrences, such as organism behavior patterns or population changes over time. For scientists, mathematical models make it so much easier to view and describe a measurable phenomenon without having to stay stuck in the raw, numerical data. Most fields of medicine are also very dependent upon mathematical models, especially concerning the frequencies of gene expression and the spreading rates of diseases.

In this thesis, we are interested in models describing the competition between strains. These activities are frequently found in the reality, for instant, in the epidemiology, the models describing lives or migration of species (worms, mosquitoes, etc.), the activity of antibiotic and resistance, etc. We propose and study a type of SIS system with similar strains. Using method of different timescale, we obtains some results in approximating

and analyzing the ultimate outcome of models.

This thesis is organized as follows. The **Chapter 1** recall us a brief history of dynamical system and mathematical biology, which is followed by a preliminary of SIS (Susceptible-Infected-Susceptible) - model. We make some general analysis for this type of model, including behaviour of the solution and some remarkable notes on the famous ratio - the basic reproduction ratio R_0. We introduce the model used throughout this thesis with al the details here. There is an intuitive scheme for our model and unified conventions for parameters and variables in this section. The main tool of this thesis - the slow-fast decomposition is defined with its related results, which is applied throughout. The slow dynamics is proven to follow a replicator system. Next, we introduce the replicator system and its main properties. A concise definition and some examples are given, allows us better to understand this type of famous equations in game theory. Interestingly, the replicator system and the Lotka-Volterra can be inferred from each other. We end this chapter by a section presenting some remarkable points about reaction-diffusion equations. In this section, we mention the traveling wave's effects in open domains and Turing instability with weak Allée effects in compact domains. Moreover, we explain the possibility of applying slow-fast approach for our reaction-diffusion equations in this thesis.

In the **Chapter 2**, we propose the general type of our main model in ordinary differential equations form for N strains. Step by step, we completely study it when the variations happen in all traits, including transmission rates, singled and doubled infection clearance rates, transmission capacity of a strain by a co-colonized host containing that strain and relative factor of altered susceptibility to co-colonization between a colonizing strain to an other co-colonizing strain. The slow-fast approximation and in particular the Tikhonov's theorem, is the main tool in proofs, come with precise transformations and arguments. It is shown in this chapter that the completed slow dynamics is given by a (linear) combination of the different elementary blocks, each block depends on the perturbation from single trait. After all, the replicator system is deduced, in which, the variables are the frequencies of strains. For a better intuition of the longtime behavior, we also presents the simulations in many particular cases. Accordingly, it is more convenient to track the different ultimate outcomes in the various cases of trait perturbations.

In the **Chapter 3**, we go deeper in the two-strain system. Based on the general results in the chapter 2, it is necessary to well understand the two strain system because the competition of two is the elementary component in a whole system with many strains. In this part, a concrete picture is presented, which allows us to see how traits affect to the behaviours, for examples, how outcomes depend on the infection-to-coinfection ratio μ in particular cases, how the outcomes depend on the differences of transmission rates and infection clearance rates, etc. Moreover, there are some claims on the possible outcomes with proofs. Changing in values of traits are considered, gives us some applications in antibiotic treatment and vaccination. Thank to the simplicity of number two, all the formulas in this chapter are explicit, makes the precise computations and easy simulations with verifying. It is remarkable to note that, even in a two strains system, the strain-specific R_0 does not determine the winner in the competition, which is against our common sense.

The **Chapter 4** is a study of a similar model in Chapter 2 with a spatial structure, when all of the parameters depend on the space. In this chapter, we propose two cases, including diffusion rates in type of small $\varepsilon\Delta$ and large $\frac{d}{\varepsilon}\Delta$. These additions make the original system even harder to compute, and is not yet seen in the literature. First, for the case of slow diffusion, we can apply the previous method, but with adaptations to the partial differential equations (PDEs), to transform the system into the slow-fast form. The slow system is then proved to be governed by a replicator system with diffusion. The rest part is similar to the Chapter 2, but with adaptations to PDEs. The case of fast diffusion $\left(\frac{1}{\varepsilon}\Delta\right)$ can be studied by applying the Central Manifold Theorem directly. We obtain a system similar to the main system in the Chapter 2, in which, the variables can

be calculated using the mean values of variables in the original system with fast diffusion. Under an assumption for the capacity transmission from a co-colonized host, we then use the result before to have the replicator equations as the reduced model. We close this chapter by a comparison section between two cases, which summarizes the two reduced system in slow manifold and displays some examples to see the distinguished effects of singular and large diffusivity.

In the **Appendices**, we show some theories used in the main flows of this thesis and present numerical simulations. Concretely, the Appendix A presents the definition of the Lyapunov function and the well-known theorem LaSalle's invariance principle, which is applied widely in dynamical systems. The Appendix B shows us the whole proof for the Tikhonov's theorem from the original paper [196], the tool used in the main problem of this thesis. We think it is useful to put the proof here for whom may have interest in it. Last but not least, the Appendix C show us many simulations, allow us to have an intuitive view of the approximation of replicator system to the original model, the space trajectories of a multiple strain system. It is also interesting to see the outcome of a two-strain model in a large samples and compute the probability of the exclusion of either strains. Moreover, we study a simulation for the outcome of a system with large number of strains, to have some comparison to the two-strain model case. The last part in the Appendices considers the intermediate diffusion case $(d\Delta)$ for quasi-neutral model, beside two cases already studied in Chapter 4. It is even impossible to calculate the explicit equilibrium for the neutral system or applying the same method as before. Hence, for this section, we make some analysis and solve the problem when the neutral system's parameters do not depend on the space.

Chapter 1

Models, approaches and an over view on replicator system

1.1 An overview: History, a simple SIS model and some analyzing

Dynamics, which was originally a branch of physics, began in the mid-1600, when Newton invented differential equations, discovered his laws of motion and universal gravitation [191]. He solved the two-body problem-the problem of calculating the earth around the sun. Then, many mathematicians and physicists tried to extend Newton's method to the three body problem (for example: sun, moon and earth), but it was impossible to solve for an explicit formulas after very long time. In the late 1800s, Henri Poincaré made a breakthrough by introducing a new point of view that emphasized qualitative rather than quantitative questions, [23]. For example, instead of asking for the exact positions of three bodies at all times, he wondered whether the solar system was stable forever or some planets eventually flew flew off to infinity. Poincaré developed a powerful geometric approach to these questions, which lead to the modern subject of dynamics.

Many mathematicians were working on dynamics, with numerous contributions in both mathematical theory and applications. In mathematical aspects, in particular, there are works about dynamical system focusing on ordinary differential equations (ODEs) [34, 44], and we can not dismiss the famous Lyapunov stability theory by Aleksandr Lyapunov [138]. In partial differential equations (PDEs), there is well-known KAM theory (Kolmogorov–Arnold–Moser) of Andrey Kolmogorov, Jürgen Moser and Vladimir Arnold [111, 153, 17] and Henri Poincaré in bifurcation theory [177, 19]. Beside, we also have very big names such as John von Neumann, Birkhoff in Ergodic Theory [31, 161, 162], Lorenz and Mandelbrot in Chaos Theory [134, 133, 142, 45], etc. Moreover, people use dynamical ideas in various applications— in , classical mechanics [18, 101], biology [11, 40, 107], chemical kinetics [68, 185, 188], population biology [102, 164, 170], etc. Under the perspective of dynamics, all of these subjects can be placed in a common framework and many are solved by mathematical tools.

It is known that both physics and biology now use mathematical model to deal with problems. However, [132] illustrates a comparison of fundamental principles for theory construction from physics to biology. These are summarized by a very relevant conceptual duality: the genericity of physical objects and the specificity of their trajectories, in contrast to the specificity of biological objects and the genericity of their possible trajectories. Mathematical models which are necessary to understand complex, non-linear interactions need to be grounded on robust biological principles. More clearly, the target of theoretical modeling and experimental

investigation in the biomedical sciences is to show the underlying biological processes that result in a particular observed phenomenon [157, 158, 166].

This thesis studies a kind of population dynamics, that has traditionally been the dominant branch of mathematical biology. Population dynamics overlap with another active area of research in mathematical biology: mathematical epidemiology, the study of infectious disease affecting populations [36–38]. Various models of viral spread have been proposed and analysed, and provide important results that may be applied to health policy decisions [33, 60, 110, 141, 150]. The beginning of population dynamics is widely regarded as the work of Malthus in 1798, formulated as the Malthusian growth model. According to Malthus, assuming that the conditions (the environment) remain constant, a population will grow (or decline) exponentially [197]. In 1838, the logistic equation was originally derived by Pierre François Verhulst [201]. In 1910, the Lotka–Volterra predator-prey equations was initially proposed by Alfred J. Lotka in the theory of autocatalytic chemical reactions, then it was extended by him in 1920 to 1925. In 1926, the same set of equations was published independently in 1926 by Vito Volterra, [26, 91, 102, 135, 193]. An alternative to the Lotka–Volterra predator–prey model, Arditi–Ginzburg equations was in 70 years later [2, 16]. A more general model formulation was proposed by F. J. Richards in 1959 [178], further expanded by Simon Hopkins, in which the models of Gompertz, Verhulst and also Ludwig von Bertalanffy are covered as special cases of the general formulation. The considered population dynamics are often studied in epidemiology, to predict the population or prevalence of infected strains. Models of disease transmission, or epidemic models for short, have been an integral part of the epidemiological toolkit, dating back from pioneer models of [106]. The main goal of epidemic models can be summarized as the ability to accurately predict spreading patterns of a given communicable disease afflicting a specific population. These models allow decision makers to assess the various intervention strategies available to them and to plan accordingly.

Before introducing our models which is of type SIS, let us make some points about the SIR model - whose variations include SIS. The compartmental SIR model simplify the mathematical modelling of infectious diseases and it is more realistic in general as well as applicable to the disease awareness framework. The origin of such models is the early 20th century, with an important work being that of Ross [180] in 1916, Ross and Hudson in 1917 [181], Kermack and McKendrick [106] in 1927. The SIR model [99] is one of the simplest compartmental models, and many models are derivatives of this basic form. Meanwhile, the SIS model is more useful when infections do not give immunity upon recovery from infection, and individuals become susceptible again. It allows us to examine how the infection spreads (and is potentially reintroduced) over time .

In these types of models, the population is assigned to compartments with labels, in which individuals may progress between. The order of the labels usually shows the flow patterns between the compartments; in particular, we have the following conventions in the model SIR:

- S: The number of susceptible individuals. When a susceptible and an infectious individual come into "infectious contact", the susceptible individual contracts the disease and transitions to the infectious compartment.
- I: The number of infectious individuals. These are individuals who have been infected and are capable of infecting susceptible individuals.
- R: for the number of removed (and immune) or deceased individuals. These are individuals who have been infected and have either recovered from the disease and entered the removed compartment, or died. It is assumed that the number of deaths is negligible with respect to the total population. This compartment may also be called "recovered" or "resistant".

For a clear view, we the the following scheme of SIR model in Figure 1.1

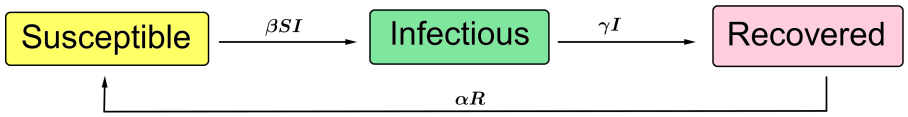


Fig. 1.1 A simple diagram for SIR model

The explicit corresponding system is given by

$$\begin{cases} \dfrac{dS}{dt} = -\dfrac{\beta}{N}SI + \alpha R \\ \dfrac{dI}{dt} = \dfrac{\beta}{N}SI - \gamma I \\ \dfrac{dR}{dt} = \gamma I - \alpha R \end{cases} \tag{1.1.1}$$

with the total population N.

These variables (S, I, and R) represent the proportion of population in each compartment at a particular time. To represent that the number of susceptible, infectious and removed individuals may vary over time (even if the total population size remains constant), we make the precise numbers a function of t (time): $S(t)$, $I(t)$ and $R(t)$. For a specific disease in a specific population, these functions may be worked out in order to predict possible outbreaks and bring them under control. This model SIR is reasonably predictive for infectious diseases that are transmitted from human to human, and where recovery confers lasting resistance, for examples, measles [32], mumps [123] and rubella [182].

Moreover, we have some other examples of compartmental models derived from SIR, such as, SEIS means susceptible, exposed, infectious, then susceptible again, or the SIRD model means the Susceptible-Infectious-Recovered-Deceased model, which differentiates between Recovered (meaning specifically individuals having survived the disease and now immune) and Deceased.

Now, we introduce the SIS model, that describes the dissemination of a single communicable disease in a susceptible population of size N. The transmission of the pathogen occurs when infectious hosts transmit the disease pathogen to healthy susceptible individuals. The infectious period extends throughout the whole course of the disease until the recovery of the patient, warranting a two-stage model: either infected or susceptible [159]. Some infections, for example, those from the common cold and influenza, do not confer any long-lasting immunity. Such infections do not give immunity upon recovery from infection, and individuals become susceptible again. This is an example where we can apply the type SIS model to figure out the epidemic. For a clear glance, the scheme of SIS model is presented in figure 1.2 below.

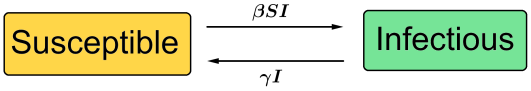


Fig. 1.2 A simple diagram for SIS model

The explicit corresponding model is given by

$$\begin{cases} \dfrac{dS}{dt} = -\dfrac{\beta}{N}SI + \gamma I \\ \dfrac{dI}{dt} = \dfrac{\beta}{N}SI - \gamma I \end{cases} \tag{1.1.2}$$

Note that denoting with N the total population it holds that for all time $t > 0$:

$$\frac{dS}{dt} + \frac{dI}{dt} = 0 \Longrightarrow S + I = N.$$

Before analyzing this system (1.1.2), let's briefly explore the meaning of these terms.

- The βSI term: An average infected individual makes contact sufficient to infect βN others per unit time. Also, the probability that a given individual that each infected individual comes in contact with is susceptible is S/N. Thus, each infected individual causes $(\beta N)(S/N) = \beta S$ infections per unit time. Therefore, I infected individuals cause a total number of infections per unit time of βSI.
- The γI term is even simpler to understand: γ is the fraction of infected individuals who recover (and re-enter the susceptible class) per unit time.

Back to model (1.1.2), we have that

$$\frac{dI}{dt} = (\beta - \gamma)I - \frac{\beta}{N}I^2$$

i.e. the dynamics of infectious is ruled by a logistic function, so that for all $I(0) > 0$,

$$\begin{aligned} &\frac{\beta}{\gamma} < 1 \Longrightarrow \lim_{t\to\infty} I(t) = 0 \\ &\frac{\beta}{\gamma} > 1 \Longrightarrow \lim_{t\to\infty} I(t) = \left(1 - \frac{\gamma}{\beta}\right)N \end{aligned}.$$

It is possible to find the explicit formula for solution of (1.1.2), which is

$$I(t) = \frac{\left(1 - \frac{\gamma}{\beta}\right)N}{1 + C\exp\left(-(\beta - \gamma)t\right)}, \qquad \text{where} \quad C = \frac{\left(1 - \frac{\gamma}{\beta}\right)N}{I(0) - 1}$$

and $S(t) = N - I(t)$.

Dividing two sides of each equation in (1.1.2) to N, we have the same system for variables S/N and I/N. In this thesis, we can assume that $N = 1$, which means variables S, I are population frequencies.

Before, we have two cases of value β/γ for the long time behavior of infectious $I(t)$. This kind of value is often reference as **basic reproductive ratio** or **basic reproductive number**. Generally in epidemiology, people define the basic reproductive ratio to be the expected number of cases directly generated by one case in a population where all individuals are susceptible to infection [80]. This definition assumes that no other individuals are infected or immunized (naturally or through vaccination). R_0 is not a biological constant for a pathogen as it is also affected by other factors such as environmental conditions and the behaviour of the infected population. According to this definition, in this thesis, we find that

$$R_0 = \frac{\beta}{\gamma}, \tag{1.1.3}$$

R_0 values are usually estimated from mathematical models, and the estimated values are dependent on the model used and values of other parameters. Thus values given in the literature only make sense in the given context and it is recommended not to use obsolete values or compare values based on different models [64]. It is important to note that R_0 does not by itself give an estimate of how fast an infection spreads in the population.

The most important uses of R_0 are to determine if an emerging infectious disease can spread in a population and what proportion of the population should be immunized through vaccination to eradicate a disease. In commonly used infection models, when $R_0 > 1$ the infection will be able to start spreading in a population, but not if $R_0 < 1$. Generally, the larger the value of R_0, the harder it is to control the epidemic. For simple models, the proportion of the population that needs to be effectively immunized (meaning not susceptible to infection) to prevent sustained spread of the infection has to be larger than $1 - 1/R_0$ [76]. Conversely, the proportion of the population that remains susceptible to infection in the endemic equilibrium is $1/R_0$. This fact holds in our quasi-neutral model in chapter 2. To end this part, we display some values of basic reproductive number R_0 of well-known infectious diseases.

Disease	Measles [96]	Pertussis [112]	Mumps [15]	COVID-19 [28]	HIV-AIDS [13]	SARS [207]	Influenza [50] (seasonal strains)	Dengue [129]
Value of R_0	12-18	5.5	10-12	3	2-5	2-4	1.3 (1.2-1.4)	4.74

Table 1.1 **Values of R_0 of well-known infectious diseases**

To have a glance on how R_0 can be estimated in reality, we consider the example of COVID-19. In [131], authors estimate an average R_0 in Western Europe of 2.2, by combining data from several countries. They thus apply the principle of "borrowing strength", trying to consolidate the partly unreliable data collected in some countries by pooling them with better data collected in other (albeit similar) countries. To do this, they robustly estimate the exponential growth rate of the disease in 15 Western European countries based on the number of reported daily incidences (new cases) and then average them to obtain an overall estimate of the exponential growth rate for Western Europe. To obtain an estimated value of R_0, the latter estimation is combined with an estimation of the parameters of the generation interval distribution, i.e., the time needed for an infected person (primary case) to infect another person (secondary case), as reconstructed from the literature.

Estimating R_0 in Western Europe is necessary although there are many R_0-estimation reports using Chinese data. However, as explained e.g. by Delamater et al. [64], the value of R_0 is essentially the combination of three factors: the (average) number of daily contacts that one contagious person has, the probability of transmission of the disease during such a contact, and the (average) number of days that an infected person is contagious. While the latter factor mainly depends on the biological characteristics of the disease, the first two factors strongly depend on the social habits of a given population. Since these habits may vary considerably in countries with different cultures, an estimated value of R_0 in China is not necessarily valid in Europe.

1.2 The models in this thesis and preliminaries

Host infections by more than one parasite strain or species are ubiquitous in the wild [171, 58]. Multiple-infections, also referred to as mixed infections, diverse infections or poly-microbial infections receive many interests because they tend to worsen human health compared with single infections [94]. Hence, well understanding how multiple infections shape virulence may have direct applications for health. However, problems of multi-infections are more complex than single-infection. Since the force of infection-governing the frequency of multiple infection-depends on the density of infected hosts, which in turn is determined by the virulence of

the pathogen population, we are confronted with a problem in which evolution and population dynamics are closely intertwined, [199]. With singly infected hosts, virulence (i.e., disease-induced host mortality) is the outcome of the struggle between single hosts and single pathogen clones. When only single infections occur, knowledge of neither population dynamics nor the resident host exploitation strategy is required to find the optimal host-exploitation strategy; it is purely a matter of "prudent" host exploitation [199]. However, multiple infection means, in the first place, that the conflict is not only between host and pathogens but among different pathogen clones as well [69, 122]. But in the second place, population dynamics can no longer be ignored, because the frequency of multiple infection and the strategies adopted by different pathogen clones will depend on the state of the entire system [199].

This thesis consider an SIS model with multiple infections. In particular, we focus on modeling the host-to-host transmissions of different strains, using the SIS (susceptible - infected - susceptible) modeling approach with quasi-neutral assumptions. This structure of multi-type interactions rises in the epidemiological dynamics of co-colonizing strains of the same microbial species, [199, 6, 69]. Actually, the simplest model that incorporates the effect of multiple infection is a model in which hosts can become infected twice. Though in principle there is no limit to the number of infections a host can receive,we restrict our analysis to at most two infections per host because it reduces the complexity of the model, while it still allows us to analyze how optimal virulence depends on exclusive versus shared host exploitation. For instance, if there are n strains in the population, following the dynamics of each strain requires of the order of n^2 equations. Furthermore, if we want to account for the order of infection, this means that the number of parameters also becomes huge. Then, in our system, microbial strains can infect a host simultaneously and we concentrate only on the case of up to two strains co-colonizing a host. This is an assumption most models make, [6].

Single colonization by a resident strain alters the susceptiblity to incoming strains, increasing or decreasing it, by a factor k_{ij}, relative to unifected hosts, without acquired immunity. The transmission and clearance rates of all strains, singled/ doubled infected hosts, are assumed to depend on the strains and to be slightly different. Initially, we establish our model for the case of 2 strains, which can be seen in the figure 1.3 below. Concrete examples of such a situation can include bacteria of the respiratory tract [83] such as Streptococcus pneumoniae and Haemophilus influenzae, displaying several genetic and antigenic variants.

We tract the proportion of hosts in seven compartments:

- S: susceptibles
- I_i, for $i \in \{1,2\}$: hosts colonized by one strain 1st or 2nd
- I_{ij}, for $i, j \in \{1,2\}$: hosts co-colonized by one strain i initially then infected by strain j (the order of their acquisition is considered).

We denote the forces F_1 and F_2 of infection for strain 1 and 2, respectively, as follows

$$F_1 = \beta_1 \left(I_1 + p_{12}^1 I_{12} + p_{21}^1 I_{21}\right), \qquad F_2 = \beta_2 \left(I_2 + p_{12}^2 I_{12} + p_{21}^2 I_{21}\right).$$

We make the assumption that, susceptibles are recruited at constant rate r, equal to the per-capita departure rate. Upon exposure, a host can become colonized by strain 1 or 2. Single and dual carriers contribute equally to the force of infection for each group of types, and hosts carrying two pathogen clones transmit either with probability p_{ij}^s. β_i's, $i \in \{1,2\}$ denote the per-capita transmission rate, while γ_i's and γ_{ij}, $i, j \in \{1,2\}$ denote the clearance rates for infected hosts and co-infected hosts, respectively, assuming no immune memory.

Co-infections by the same (resident) strain can seem counter-intuitive. As Van Baalen and Sabelis in [199] and Alizon in [6] argue, this solves the neutrality problem and avoids giving a frequency-dependent advantage to the mutant. One could wonder what the biological basis for these hosts infected by the same strain is. If we are dealing with macro-parasites, the interpretation is intuitive as, for example I_{11}, can be seen as hosts with a double parasite load. The reasoning is the same for micro-parasites but, given their underlying biology, I_{11} hosts resemble singly infected host.

Moreover, we assume the direct clearance of colonization and co-colonization, back to the susceptible state reflects train-transcending immunity as the main player in pathogen clearance. Single carriers I_i can acquire an additional pathogen clone j at a rate modified by a relative factor k_{ij}. The case $k_{ij} < 1$ corresponds to antagonism/competition between types from strain i to j. Meanwhile, $k_{ij} \geq 1$ means the facilitation between clones and enhancement of host susceptibility.

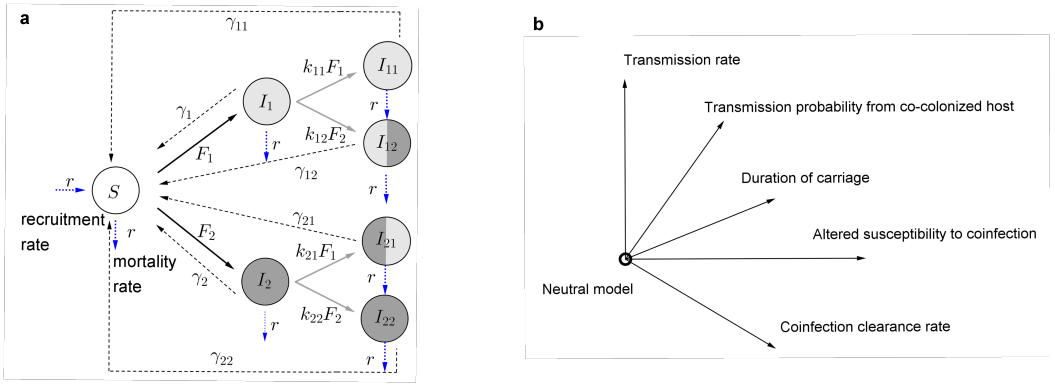


Fig. 1.3 (a) The scheme of two-strain model. The black arrows refer to acquisition of a first clone with forces of infection F_1 and F_2. The grey arrows refer to altered acquisition of a secondary clone $k_{ij}F_i$ in an already colonized host, where clone interactions can range from competition ($k_{ij} < 1$) to cooperation ($k_{ij} \geq 1$). The dashed arrows depict colonization clearance rate γ_i γ_{ij} accordingly. The dotted arrows reflect host demographic processes: birth and death at equal rate r. (b) The considered dimensions of model.

Generally, we can describe and study a general system for epidemiological dynamics of similar co-infecting strains of the same infectious agent.

$$\begin{cases} \dfrac{dS}{dt} &= r(1-S) + \displaystyle\sum_{i=1}^{N} \gamma_i I_i + \sum_{i,j=1}^{N} \gamma_{ij} I_{ij} - S \sum_{i=1}^{N} F_i, \\ \dfrac{dI_i}{dt} &= \lambda_i S - (r+\gamma_i) I_i - I_i \displaystyle\sum_{j=1}^{N} k_{ij} F_j, \qquad 1 \leq i \leq N, \\ \dfrac{dI_{ij}}{dt} &= k_{ij} I_i F_j - (r+\gamma_{ij}) I_{ij}, \qquad 1 \leq i, j \leq N \end{cases} \tag{1.2.1}$$

where

$$F_i = \beta_i \left(I_i + \sum_{j=1}^{N} \left(p_{ij}^{i} I_{ij} + p_{ji}^{i} I_{ji} \right) \right). \tag{1.2.2}$$

Note that, from now on, we use N with the meaning number of pathogens/number of strains, which is different from the meaning of total population (assumed to be constant) in (1.1.1) and (1.1.2).

The parameters are described in the following table

Table 1.2 **Conventions and notations of parameters and variables, where we assume the traits are numerically close for closely-related strains of the same infectious agent.** This similarity assumption ($0 < \varepsilon \ll 1$, small) forms the basis for dynamic decomposition into fast and slow components.

Parameter	Interpretation
Original system	
$\beta_i = \beta(1+\varepsilon b_i)$	Strain-specific transmission rates
$\gamma_i = \gamma(1+\varepsilon v_i)$	Strain-specific clearance rates of single colonization
$\gamma_{ij} = \gamma(1+\varepsilon v_{ij})$	Clearance rates of co-colonization with i and j
$k_{ij} = k+\varepsilon\alpha_{ij}$	Relative factor of altered susceptibility to co-colonization between colonizing strain i and co-colonizing strain j
$p_{ij}^s = \frac{1}{2} + \varepsilon\omega_{ij}^s$, $s \in \{i,j\}$	Transmission capacity of a host co-colonized by strain-i then strain-j, $\left(p_{ij}^i + p_{ij}^j = 1\right)$
Neutral system	
r	Susceptible recruitment rate (Equal to natural mortality)
m	Net host turnover rate

According to our assumption of altered transmission rates and clearance rates, our model captures the possibility that R_0 may change or fluctuate over time, for example due to weather conditions, seasonality or antibiotic use. It is possible that parallel changes in overall transmission, during such targeted vaccination program might boost the decline of vaccine subtypes, or counteract it, depending on the underlying interaction strengths between vaccine and non-vaccine strains [86].

In this thesis, we make the very strong assumption that all the trait are close, that means ε in Table 1.2 is small.

1.3 The slow-fast time scale and its application to our models

Before presenting the slow-fast time scale method, we introduce shortly about the multiple-scale method, which is mentioned in [104]. In dynamics problems, people may uses matched asymptotic expansions, the solution is constructed in different regions that are then patched together to form a composite expansion. The method of multiple scales differs from this approach in that it essentially starts with a generalized version of a composite expansion. In doing this, one introduces coordinates for each region (or layer), and these new variables are considered to be independent of one another. The history of multiple scales is more difficult to track than, say, boundary-layer theory. This is because the method is so general that many apparently unrelated approximation procedures are special cases of it. One might argue that the procedure got its start in the first half of the 19th century. For example, Stokes (1843) [190] used a type of coordinate expansion in his calculations of fluid flow around an elliptic cylinder. Most of these early efforts were limited, and it was not until the latter half of the 19th century that Poincaré (1886) [176], based on the work of Lindstedt (1882) [125], made more extensive use of the ideas underlying multiple scales in his investigations into the periodic motion of planets. He found that the approximation obtained from a regular expansion accurately described the motion for only a few revolutions of a planet, after which the approximation becomes progressively worse. The error was due, in part, to the contributions of the second term of the expansion. He referred to this difficulty as the presence of a secular term.

To remedy the situation, he expanded the independent variable with the intention of making the approximation uniformly valid by removing the secular term. This idea is also at the heart of the modern version of the method. What Poincaré was missing was the introduction of multiple independent variables based on the expansion parameter. This step came much later; the most influential work in this regard is by Kuzmak (1959) [115] and Cole and Kevorkian (1963) [54].

In particular, in many cases of ordinary differential equations, see Chapter 3 of [104], it is useful to introduce the two time-scales $t_1 = t$ and $t_2 = \varepsilon t$ to construct the solution well. This implies that the multi-scale method is used only with the time variable in these cases. Intuitively, we may see that the solution varies on a "fast" time scale, i.e., one that is smaller than $O(1)$. The relevant scale here is $O(\varepsilon)$. It will therefore be convenient to introduce an additional time variable, which splits the variables into two groups.

We state the following definition.

Definition 1.1. A **slow-fast system** is a system of ordinary differential equations taking the form

$$\begin{cases} \varepsilon \dfrac{dx}{d\tau} &= f(x,y,\varepsilon) \\ \dfrac{dy}{d\tau} &= g(x,y,\varepsilon) \end{cases}. \tag{1.3.1}$$

where $f : \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^m$ and $g : \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ and $0 < \varepsilon \ll 1$.
Furthermore, the x variables are called **fast variables** and the y variables are called **slow variables**. Setting $t = \dfrac{\tau}{\varepsilon}$ gives the equivalent form

$$\begin{cases} \dfrac{dx}{dt} &= f(x,y,\varepsilon) \\ \dfrac{dy}{dt} &= \varepsilon g(x,y,\varepsilon) \end{cases}. \tag{1.3.2}$$

We refer to t as the **fast time scale** or **fast time** and to τ as the **slow time scale** or **slow time**.

The parameter ε can be thought of as the "separation" of time scales and is sometimes called the **time-scale parameter**. If the notation $0 < \varepsilon \ll 1$ follows an equation, it indicates that we are interested only in the case that ε is small and positive. If it appears in a statement or theorem, then it indicates that ε is sufficiently small, i.e., $0 < \varepsilon \ll 1$ means that there exists some $\varepsilon_0 > 0$ such that for all $\varepsilon \in (0, \varepsilon_0]$, the statement of the theorem holds.

In many situations, f and g are independent of ε, and we can write the fast–slow system as

$$\begin{cases} \varepsilon \dfrac{dx}{d\tau} &= f(x,y) \\ \dfrac{dy}{d\tau} &= g(x,y) \end{cases}$$

with the analogous obvious reformulation on the fast time scale.

There are two compelling facts that should encourage one to study fast–slow systems. The first is that the mathematical structure of fast–slow systems is very intricate due to the small parameter ε. Many techniques ranging from classical asymptotic methods and nonstandard analysis to geometric methods and numerical treatments have been used successfully alongside the toolbox of dynamical systems theory, [114]. Although the first reason alone would clearly suffice to command further investigation, fast–slow systems also regularly appear in mathematical modeling across many areas of the natural sciences.

Let us mention the solution of slow-fast time scale. Solutions of slow-fast systems generally consist of slow and fast time, i.e. long periods of slow change interspersed by short periods of fast change. As $\varepsilon \to 0$, the solution of (1.3.2) approach the solutions of **fast subsystem** or **layer equations**

$$\begin{cases} \dfrac{dx}{dt} &= f(x,y,0) \\ \dfrac{dy}{dt} &= 0 \end{cases} \tag{1.3.3}$$

During slow time τ, trajectories of (1.3.1) are well described by the solutions of **slow subsystem** or **reduced system**

$$\begin{cases} 0 &= f(x,y,0) \\ \dfrac{dy}{d\tau} &= g(x,y,0) \end{cases} \tag{1.3.4}$$

which is a differential-algebraic equation. The algebraic equation in (1.3.4) defines the **critical manifold**

$$C = \{(x,y) \in \mathbb{R}^m \times \mathbb{R}^m | f(x,y,0) = 0\}$$

which is the phase space of the reduced problem (1.3.4) and is nothing but the set of equilibrium points of the layer equation (1.3.3). One major goal of singular perturbation theory is to use these lower dimensional subsystems (1.3.3) and (1.3.4), to understand the dynamics of the full system (1.3.2) when $0 < \varepsilon \ll 1$.

The geometric language is useful to describe the model. Roughly speaking, the real dynamics fast follow the fast dynamics (1.3.3) near the set C. In this timescale, the slow dynamics has no effect. Once near C, we have that $\frac{dy}{dt} = 0(\varepsilon)$, the system is described at the slow timescale τ by the slow system (1.3.4).

The system (1.3.3) allows us to classify the singularly perturbed systems. A compact sub-manifold $C_h \subset C$ is **normally hyperbolic** if all $(x,y) \in C_h$ are hyperbolic equilibria of the layer problem, that is, the Jacobian $(D_y f)(x,y,0)$ has no eigenvalues with zero real part. Then, Fenichel theory [71] guarantees the existence of a perturbed locally invariant manifold C_ε called (Fenichel) slow manifold the original system $0 < \varepsilon \ll 1$. Fenichel slow manifolds are typically non-unique but they are exponentially close to each other (away from their boundary). They inherit the regularity of the vector filed, they are $O(\varepsilon)$-close to the unperturbed manifold C_h (for the Hausdorff distance) and the flow on them converges to the slow flow as $\varepsilon \to 0$. If the unperturbed manifold C_h belongs to an attracting (resp. a repelling) sheet of C – the real part of the eigenvalues of the Jacobian $(D_y f)(x,y,0)$ being strictly negative (resp. strictly positive) — then the corresponding slow manifolds are called *attracting* (resp. *repelling*) *slow manifolds*. If C_h is normally hyperbolic and neither attracting nor repelling, it is of saddle type, and so are the corresponding slow manifolds. In this present thesis, the result with strictly negative-real-part eigenvalues are restated in the form of the Tikhonov's theorem [196]. For a simulation of this, see the figure 1.4 as follows. In this thesis, the slow-fast dynamics decomposition offers a useful tool to provide insight into microbial interactions. Although we analyze an endemic multi-strain pathogen and its dynamics on the epidemiological scale, our results have implications beyond this immediate system, contributing to novel conceptual unification in community ecology. We show that neutral dynamics between interacting strains occurs on a fast timescale, whereas the non-neutral stabilizing forces act on a slow timescale, dependent on trait differences manifested upon co-colonization. We quantify exactly which asymmetries matter for multi-strain coexistence, and how overall transmission intensity affects stabilization of diversity. Together, our results define a promising analytical approach to better understand microbial ecosystem.

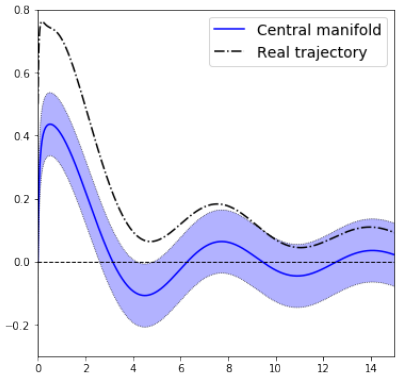


Fig. 1.4 The system trajectories (black dashed arrows) reach an ε-neighborhood of the slow manifold (in blue) in a time of the order $O(\varepsilon \ln(1/\varepsilon))$. Once in this neighbourhood, the trajectories stay ε-close to the slow manifold and follow the slow dynamics, taking a time of order $O(\varepsilon)$ to reach the equilibria.

1.4 The replicator system

We show in this thesis that the system in slow manifold is in the form of replicator system. To introduce this famous kind of equations, we first recall the Lotka-Vottera system, which reads as follows

$$\frac{dx_i}{dt} = x_i f_i(x_1, x_2, \ldots, x_n) = x_i((Ax)_i + c_i),$$

with f_i's are linear for all i.

A well-known particular case of the Lotka–Volterra equations, the predator-prey system is a pair of first-order nonlinear differential equations, ussually used to describe the dynamics of biological systems in which two species interact, one as a predator and the other as prey. This type of model was initial introduced in [135], and the populations change through time and are described as follows

$$\begin{cases} \dfrac{dx}{dt} &= \alpha x - \beta xy \\ \dfrac{dy}{dt} &= \delta xy - \gamma y \end{cases} \tag{1.4.1}$$

where x and y are the densities of preys and predators in the population, respectively. α, β, δ, γ are positive real parameters describing the interaction of the two species.

Concretely, the prey are assumed to have an unlimited food supply and to reproduce/given birth exponentially, unless subject to predation; this exponential growth is represented in the equation above by the term αx. The rate of predation upon the prey is assumed to be proportional to the rate at which the predators and the prey meet, this is represented above by βxy. If either x or y is zero, then there can be no predation. With these two terms the equation above can be interpreted as follows: the rate of change of the prey's population is given by its own growth rate minus the rate at which it is preyed upon.

On the other side, in the equation for the predators, δxy represents the growth of the predator population. Note the similarity to the predation rate; however, a different constant δ is used, as the rate at which the predator population grows is not necessarily equal to the rate at which it consumes the prey. The term γy represents the mortality rate of the predators, which leads to an exponential decay in the absence of prey. Hence the equation expresses that the rate of change of the predator's population depends upon the rate at which it consumes prey, minus its intrinsic death rate.

Setting

$$V(x,y) = \delta x - \gamma \ln(x) + \beta y - \alpha \ln(y),$$

then $\nabla V \cdot \begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix} = 0$ which means $G(t) = V(x(t), y(t))$ is a constant if $(x(t), y(t))$ is the solution of (1.4.1). We note that V only depends on the initial values $(x(0), y(0))$, see figure 1.5 below.

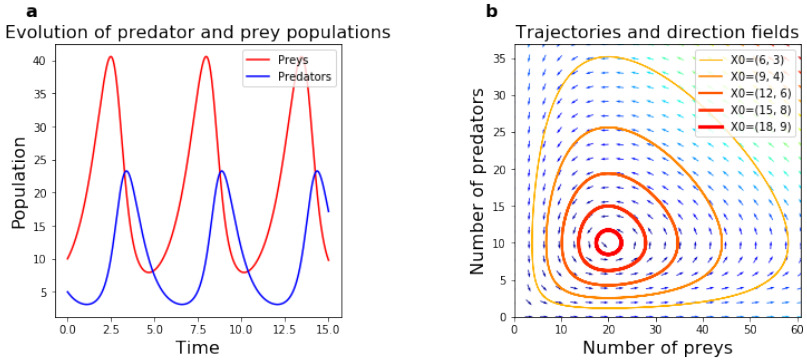


Fig. 1.5 **A simulation for solution of the Lotka-Volterra of 2 strains.** In this simulations, we consider the system (1.4.1) with the parameters $\alpha = 1$, $\beta = 0.1$, $\delta = 0.075$, $\gamma = 1.5$. The figure **(a)** plots the populations of prey and predator in time with initial values $(x(0), y(0)) = (10, 5)$. The figure **(b)** is the phase-space plotting for the predator prey problem for various initial conditions of the predator population. It can be proved that these trajectories are all closed and this holds true for general case of two-strain Lotka-Volterra system.

Now, we consider a system of N populations and denote by x_i the densities of i-th species $(i = 1, 2, \ldots, N)$. Then the growth rates of i-th species an be determined by the total effects from itself and from the other ones. This can be read as

$$\frac{dx_i}{dt} = x_i \left(b_i + \sum_{j=1}^{N} a_{ij} x_j \right), \qquad i = 1, 2, \ldots, N. \tag{1.4.2}$$

This system is called a generalized Lotka-Volterra system for N species and broadly investigated, see [193, 102]. The parameters b_i, $i = 1, 2, \ldots, N$ is intrinsic growth or decay rate for i-th strain and the a_{ij} represents the effect of the j-th strain upon the i-th one. Similarly as in the case two strain, the terms b_i and a_{ii}, $i = 1, 2, \ldots, N$, describes the growth rates of i-th strain in the absence of the others, whereas the terms a_{ij}'s are the effects on the grwoth of i-th species from j-th one.

The matrix $A = (a_{ij})_{1 \le i,j \le N}$ is called the **interaction matrix**. If the i-th and j-th strains compete to each other then the sign pattern of (a_{ij}, a_{ji}) becomes $(-,-)$. Meanwhile, if they cooperates together, the sign pattern now is $(+,+)$. They are the predator-prey relationship means $(+,-)$ for the case i-th species is the predator, j-th strain is the prey and vice versa. The element on the principle diagonal are usually non-positive for all i sine they represent the relationship among the same strains and reflect the limited resource. The existence and uniqueness of solution for (1.4.2) in the nonnegative space $\mathbb{R}^n_+$ is showed in [193].

The replicator dynamics, which is studied much in evolutionary system and game theory, see [102, 59], describes the evolution of the frequencies of strategies in a population. The replicator equation allows the fitness function to incorporate the distribution of the population types rather than setting the fitness of a particular type constant. This important property allows the replicator equation to capture the essence of selection. A disadvantage is that, the replicator equation does not incorporate mutation and so is not able to innovate new types or pure strategies. Interestingly, we can transfer the replicator equation to the Lotka-Volterra equation

and conversely, which we prove later in this subsection. Initial, we assume that the considered population has n types 1st, 2nd, etc., $n-th$, corresponding n frequencies $(x_1, x_2, \ldots, x_n)$. The fitness f_i of type i-th will be a function of the composition of the population. The rate of increase $\frac{\dot{x}}{x}$ of type i-th is a measure of its evolutionary success which can be expressed as the difference between the fitness $f_i(x)$ of type i-th and the average fitness $\bar{f}(x) = \sum x_i f_i(x)$ of the population. Then we observe the most general continuous form of the replicator equation as follows

$$\frac{dx_i}{dt} = x_i \left(f_i(x) - \bar{f}(x) \right), \qquad 1 \leq i \leq n, \tag{1.4.3}$$

where $x = (x_1, x_2, \ldots, x_n)$ is the vector of the distribution of types in the population.

The set $S_n = \{x \in \mathbb{R}^n : x_1 + x_2 + \cdots + x_n = 1\}$ is invariant under (1.4.3), i.e. if $x \in S_n$ then $x(t) \in S_n$ for all $t \in \mathbb{R}$. Indeed, according to (1.4.3), the sum $S = x_1 + x_2 + \cdots + x_n$ verifies $\frac{dS}{dt} = (1-S)\bar{f}$, in which $S(t) = 1$ is an equilibrium. From now on, we should only consider the restriction of (1.4.3) on S_n.

It can be shown that the change in the ratio of two proportions x_i/x_j with respect to time is

$$\frac{d}{dt}\left(\frac{x_i}{x_j}\right) = \frac{x_i}{x_j}\left(f_i(x) - f_j(x)\right).$$

In other words, the change in the ratio is driven entirely by the difference in fitness between types. Moreover, if we add an arbitrary function $\phi : S_n \to \mathbb{R}$ to all the f_i, this does not affect (1.4.3) on S_n. Indeed, setting $g_i(x) = f_i(x) + \phi(x)$ for all i then we obtain that $\bar{g}(x) = \sum x_i g_i(x) = \bar{f}(x) + \phi(x)$ (on S_n), which implies $g_i(x) - \bar{g}(x) = f_i(x) - \bar{f}(x)$ for all i.

Let us now prove that the replicator equation in n dimensions is equivalent to the generalized Lotka-Volterra equation in $n-1$ dimensions. Indeed , setting that

$$x_i = \frac{y_i}{1 + \sum_{j=1}^{n-1} y_j}, \qquad 1 \leq i \leq n-1; \qquad\qquad x_n = \frac{1}{1 + \sum_{j=1}^{n-1} y_j} \tag{1.4.4}$$

where y_i, $1 \leq i \leq n$ are the variables of the Lotka-Volterra equation in $n-1$ dimensions of (1.4.2). By direct validation, $(x_1, x_2, \ldots, x_n)$ verifies the replicator equation (1.4.3).

In particular, we are interested in the case linear f_i for all i. Then, there exists an $n \times n$ matrix $A = (a_{ij})$ such that $f_i(x) = (Ax)_i$. Hence, when the fitness is assumed to depend linearly upon the population distribution, which allows the replicator equation to be written in the form

$$\frac{dx_i}{dt} = x_i \left((Ax)_i - x^T Ax \right), \qquad 1 \leq i \leq n \tag{1.4.5}$$

where the *pay-off matrix* A holds all the the fitness information for the population: the expected payoff can be written as $(Ax)_i$ and the mean fitness of the population as a whole can be written as $x^T Ax$. We note that adding a same constant to all the entries a_{ij}'s of matrix A will not change (1.4.5). This gives a freedom of the definition of the f_i's in (1.4.3). In particular, it can be assumed that $a_{ii} = 0$ for all i.

The equilibrium of (1.4.5) verifies

$$\begin{cases} (Ax)_1 = (Ax)_2 = \cdots = (Ax)_n \\ x_1 + x_2 + \cdots + x_n = 1 \end{cases}$$

satisfying $x_i > 0$ for all i. In general, there exists at most one such solution. Indeed, we can compute the difference between two successive lines i-th and $i+1$-th which leads to the equivalent system

$$\begin{cases} \sum_{j=1}^{n} \left(a_{ij} - a_{i+1,j}\right) x_j = 0 \\ x_1 + x_2 + \cdots + x_n = 1 \end{cases}$$

which is nothing but the system $Bx = (0,\ldots,0,1)^T$ with $B = (b_{ij})$ and $b_{ij} = a_{ij} - a_{i+1,j}$ for $1 \leq i \leq n-1$ and $b_{nj} = 1$ for all j. This system have generically (if $\det B \neq 0$) exactly one solution on $\mathbb{R}^n$.

An interesting and intuitive way to interpret the replicator system is through game theory. We consider a large population, N, of players Each period, a player is randomly matched with another player and they play a two-player game. Each player is assigned a strategy and they cannot choose their strategies. We can think of this in a few ways, for example:

- Players are "born with" their mother's strategy (ignore sexual reproduction).
- Players "imitate" others' strategies.

We note that rationality and consciousness do not enter the picture. Now, denoting the proportion of the population playing strategy A and B as x_A and x_B, respectively, so:

$$x_A = \frac{N_A}{N}, \qquad x_B = \frac{N_B}{N},$$

where N_A and N_B are the populations playing strategy A and B, respectively. The state of the population is given by (x_A, x_B) where $x_A \geq 0$, $x_B \geq 0$, and $x_A + x_B = 1$.

Since players interacts with another randomly chosen player in the population, a player's *expected payoff* is determined by the payoff matrix and the proportion of each strategy in the population. Payoff for player who is playing A is f_A. Since f_A depends on x_A and x_B we write $f_A(x_A, x_B)$ with

$$f_A(x_A, x_B) = (\text{probability of interacting with } A \text{ player}) \times U_A(A,A) + (\text{probability of interacting with } B \text{ player}) \times U_A(A,B)$$

which implies that

$$f_A(x_A, x_B) = x_A \times U_A(A,A) + x_B \times U_A(A,B).$$

In which, $U_A(A,A)$ and $U_A(A,B)$ are the payoff for players playing strategy A when they meet the other player playing strategy A and B, respectively. We interpret payoff as *rate of reproduction* (fitness). The average fitness, $\bar{f}$- the weighted average of the two fitness values now becomes

$$\bar{f}(x) = x_A f_A(x_A, x_B) + x_B f_B(x_A, x_B).$$

Now, we wonder that how fast x_A and x_B grow. We argue for the case x_A first and do the same for x_B. Recalling that $x_A = \frac{N_A}{N}$ and we need to know how fast N_A grows. We have that each individual reproduces at a rate f_A, and there are N_A of them, so

$$\frac{dN_A}{dt} = N_A f_A(x_A, x_B).$$

Similarly we deduce that $\frac{dN_B}{dt} = N_B f_B(x_A, x_B)$ and $\frac{dN}{dt} = N\bar{f}(x_A, x_B)$. By the quotient rule, and combining with a little simplification, we obtain the replicator system

$$\begin{cases} \frac{dx_A}{dt} = x_A\left(f_A(x_A, x_B) - \bar{f}(x_A, x_B)\right) \\ \frac{dx_B}{dt} = x_B\left(f_B(x_A, x_B) - \bar{f}(x_A, x_B)\right) \end{cases}$$

In this deduced system, let us consider the equation for x_A, we have the meanings that the growth rate of A is equal to the product of current frequency of strategy A with its own fitness relative to the average. The term x_A represents how many A's can reproduce and the term $f_A(x_A, x_B) - \bar{f}(x_A, x_B)$ implies that the more successful strategies will grow faster. For example, if $x_A > 0$ and $f_A(x_A, x_B) > \bar{f}(x_A, x_B)$ then we have that the proportion of A's is non-zero and the fitness of A is above average, which yields to $\frac{dx_A}{dt} > 0$, i.e., A will be increasing in the population, which makes sense in reality. We skip the theory about the Nash equilibria and evolutionarily stable states in this section. These important parts can be seen in chapter 7 of [102].

To close this part, let us now generalize the replicator equation's example to a 3-player game which is very familiar with us: Rock, Paper, Scissor, whose strategies are denoted R, P and S, respectively. Supposing N_R, N_P, N_S are the numbers of players playing R, P, S then x_R, x_P, x_S are the proportions of population playing R, P, S. The payoff matrix of this game is

$$A = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}$$

and we note that this is a zero-sum game. The explicit dynamics for this case is

$$\begin{cases} \frac{dR}{dt} = R(-P+S) \\ \frac{dP}{dt} = P(R-S). \\ \frac{dS}{dt} = S(-R+P) \end{cases} \tag{1.4.6}$$

There exist only four rest points for (1.4.6), one in the center, $m = (1/3, 1/3, 1/3) \in \text{int}S_3$ with $S_3 = \{R+P+S = 1\}$, and the other three at the vertices e_i.

Let us consider the functions $V_1 = R+P+S$ and $V_2 := RPS$. By (1.4.6), we can calculate directly that $\frac{dV_1}{dt} = 0$ and $\frac{dV_2}{dt} = 0$. By visualizing the level sets of the functions V_1 and V_2 in the three-dimensional (R,P,S) space, we can see that all orbits in $\text{int}S_3$ are closed orbits surrounding m.

We can change the payoff matrix to $A = \begin{pmatrix} 0 & -1 & 2 \\ 2 & 0 & -1 \\ -1 & 2 & 0 \end{pmatrix}$ then by the similar argument in page 42 of [187], we can show that interior fixed point attracts all trajectories on the interior of (R,P,S) We can see these phenomena in the following figures 1.6.

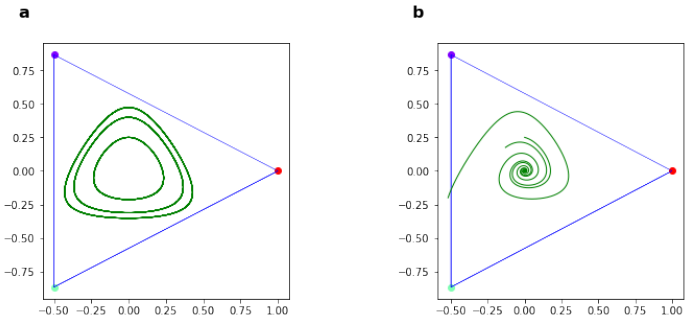


Fig. 1.6 **Phase space for Rock-Paper-Scissors game. (b)** For the generalized case with matrix $\begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}$. The phase space now is always cycle around the interior fixed point for all initial value. **(a)** For the normal case with matrix $\begin{pmatrix} 0 & -1 & 2 \\ 2 & 0 & -1 \\ -1 & 2 & 0 \end{pmatrix}$. The phase space now is spiral turning to the attractor.

1.5 Reaction-Diffusion Equations

Reaction-diffusion systems are nonlinear parabolic equations or systems of the form

$$\begin{cases} \dfrac{\partial u}{\partial t} - M\Delta u = f(u) \quad \text{in} \quad \Omega \times (0,T) \\ + \text{ boundary condtions and initial data,} \end{cases} \tag{1.5.1}$$

where $u(x,t)$ takes its values in $\mathbb{R}^m$, M is an $m \times m$ (diagonal) matrix, and f is a nonlinear map from $\mathbb{R}^m$ into $\mathbb{R}^m$.

These system are used to model phenomena occurring in various fields: chemistry, biology, epidemiology, population genetics, ecology, etc., see [75]. The solutions of reaction-diffusion equations display a wide range of behaviors, including the formation of traveling waves and self organized patterns, [39].

For reaction-diffusion equations in unbounded domains, we may have a typical question: Is it possible for a species to invade into new habitats and how does this work? This question is often studied by the traveling wave approach. A *traveling wave* is a wave that advances in a particular direction, with the addition of retaining a fixed shape. Moreover, a traveling wave is associated to having a constant velocity throughout its course of propagation. Such waves are observed in many areas of science, like in combustion, which may occur as a result of a chemical reaction [202]. In mathematical biology, the impulses that are apparent in nerve fibres are represented as traveling waves [158]. Furthermore, the structures present in solid mechanics are typically modelled as standing waves [46]. From the seminal work of Fisher [77], the reaction-diffusion systems are know to often exhibit traveling wave solutions. These traveling wave solution can describe an invasion into a new habitat, where this move appears with a constant speed c (the so-called wave speed), i.e.

$$u(x,t) = \phi(x - ct). \tag{1.5.2}$$

For a clearer view, the solution can be illustrated in Figure 1.7.

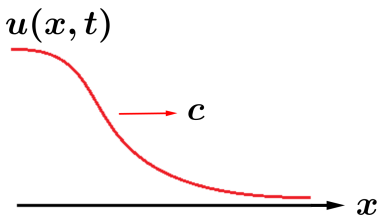


Fig. 1.7 Traveling wave solution

The new variable $z := x - ct$ denotes the wave variable and $\phi(z)$ is the wave profile. The travelling wave ansatz includes conditions at $\pm\infty$ instead of "classical" boundary conditions:

$$u(x,t) = \phi(x-ct), \quad \phi(-\infty) = 1, \quad \phi(+\infty) = 0 \tag{1.5.3}$$

which means that the population has reached its capacity for $x \to -\infty$ (normalised to 1), and no population has arrived yet for $x \to +\infty$, (1.5.3) leads to:

$$\frac{\partial u(x,t)}{\partial t} = -c\phi', \qquad \frac{\partial^2 u(x,t)}{\partial x^2} = \phi'' \tag{1.5.4}$$

Such waves do not propagate, and are typically observed when inducing a fixed boundary. In fact, we can categorise traveling waves into forms that are attributed to having certain properties. For a traveling wave that approaches constant states given by $\phi(-\infty) = u_l$ and $\phi(+\infty) = u_r$, with $u_l \neq u_r$, we have what we call a *wave front* see Figure 1.7 for an example. However, if the constant states are equal with $u_l = u_r$, the corresponding wave is known as a *pulse wave*. If a wave exhibits periodicity with $\phi(z+T) = \phi(z)$, where $T > 0$, the wave is called a *spatially periodic wave*.

The first and simplest example is the Fisher equation:

$$u_t = Du_{xx} + \mu u(1-u) \tag{1.5.5}$$

together with the travelling wave ansatz (1.5.3) and (1.5.4), this yields

$$-c\phi' = D\phi'' + \mu\phi(1-\phi)$$

which corresponds to a second order ODE and we can transform it into a 2D system of first order ODEs by introducing a new variable $v := \phi'$, we write again u instead of ϕ for reasons of simplicity:

$$\begin{cases} u' = v \\ v' = -\dfrac{c}{D}v - \dfrac{\mu}{D}u(1-u) \end{cases}$$

In [77], Fisher showed that, for every wave speed $c \geq 2\sqrt{\mu D}$, (1.5.5) admits traveling wave solution. No such solution exists for $c < 2\sqrt{\mu D}$. The wave shape for a given wave speed is unique. We can find the detailed proof in [77].

Conversely, we do not have traveling wave solution for a reaction-diffusion in a compact domain. Nevertheless, behaviors of dynamics are closed to the case of ordinary differential equations, leading to the similar effects in ODEs including Turing instability and weak Allee effects.

Firstly, we recall the Turing instability. There is compelling evidence that *Turing instabilities* (*diffusion-driven instabilities*) can induce patterns in biological systems [157, 158]. A Turing instability, first proposed by Turing in 1952 [198], asserts that there are real linear constant coefficient linear dynamics

$$X' = \tilde{A}X, \qquad X' = \tilde{B}X$$

both strongly stable in the sense that the eigenvalues of $\tilde{A}$ and $\tilde{B}$ lie in the left half plane and so that combining the effects to yield $X' = \left(\tilde{A}+\tilde{B}\right)X$ yields an unstable equilibrium. In particular, a steady state, which is stable in the absence of diffusion, may become unstable when diffusion is present, and patterns are temporally stable and spatially heterogeneous solutions of the reaction–diffusion system.

Secondly, we recall definition of the Allee effect, which is presented in [57].

Definition 1.2. We have the following terminology related to the Allee effect, see Figure 1.8.

- **Allee threshold:** a critical population size or density below which the per capita population growth rate becomes negative.
- **Demographic Allee effect:** a positive correlation between the overall individual fitness, usually quantified by the per capita population growth rate, and population size or density.
- **Strong Allee effect:** a demographic Allee effect with an Allee threshold.
- **Weak Allee effect:** a demographic Allee effect without an Allee threshold.

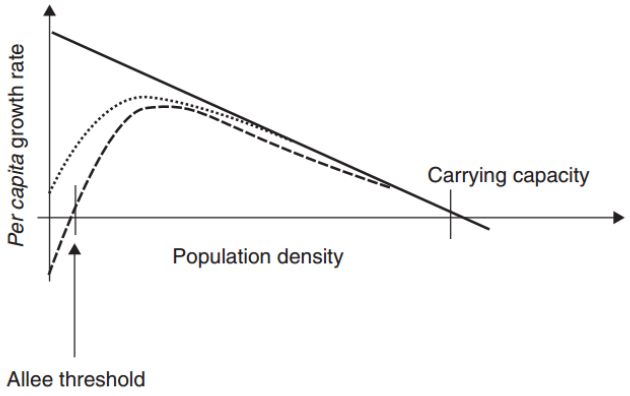


Fig. 1.8 Classical negative density dependence (solid) compared to strong (dashed) and weak (dotted) Allee effects, [57]

A simple mathematical example of an Allee effect is given by the cubic growth model

$$\frac{dN}{dt} = rN\left(\frac{N}{A}-1\right)\left(1-\frac{N}{K}\right) \tag{1.5.6}$$

where the population has a negative growth rate for $0 < N < A$, and a positive growth rate for $A < N < K$ (assuming $0 < A < K$). The equation (1.5.6) has difference comparing with the logistic growth equation

$$\frac{dN}{dt} = rN\left(1-\frac{N}{K}\right) \tag{1.5.7}$$

where N is the population size; r is the intrinsic rate of increase, K is the carrying capacity, A is the critical point, and dN/dt is rate of increase of the population.

After dividing both sides of (1.5.7) by the population size N, left hand side of (1.5.7) represents the per capita population growth rate, which is dependent on the population size N, and decreases with increasing N throughout the entire range of population sizes. In contrast, when there is an Allee effect the per-capita growth rate increases with increasing N over some range of population sizes $[0,N]$.

Spatio-temporal models can take Allee effect into account as well. A 1D simple example is given by the reaction-diffusion model in domain $[0,1]$

$$\frac{\partial N}{\partial t} = D\frac{\partial^2 N}{\partial x^2} + rN\left(\frac{N}{A} - 1\right)\left(1 - \frac{N}{K}\right) \tag{1.5.8}$$

with diffusion coefficient D. For some result of (1.5.8), we can see [84]. The dynamical behavior of (1.5.8) without Allee effect is well-known; meanwhile, when the Allee effect is present, the structure of the set of the steady state solutions is more complicated [186], in which (1.5.8) can have multiple steady state solutions, and extinction and persistence are both possible depending on the initial value.

These three phenomena (traveling wave, Turing instability and weak Allee effect) show that reaction-diffusion systems can exhibit some behaviors which cannot be capture by ODE systems, **even if the parameters are spatially homogeneous**. Moreover, and it is actually what we do in this thesis, such a systems can model explicit spatial heterogeneity by writing a reaction function explicitly dependent on x : $f(u)(x,t) = f(x,u(x,t))$. In this thesis we will not focus on the qualitative behavior of reaction diffusion systems but we will construct several reaction-diffusion systems spatially heterogeneous, by using slow fast techniques.

Chapter 2

The quasi-neutral model for N strains

This chapter is the work in the paper "Quasi-neutral Dynamics in a Coinfection System with N Strains and Asymmetries along Multiple Traits", with Dr. Erida Gjini and Dr. Sten Madec, 2021, see [120].

2.1 Introduction

Multiple infections are ubiquitous in nature [21]. They may occur between pathogen strains of the same species or between different species [51, 29, 206], and have implications for virulence evolution and maintenance of various polymorphisms among infectious agents [199, 154, 6, 7]. The importance of multiple infection for antibiotic resistance and vaccination effects in multi-strain systems has also been increasingly highlighted [127, 30]. Due to its inherent difficulties, multiple infection has only been tackled in a limited manner by mathematical models so far. A majority of studies focus on coexistence and competitive exclusion criteria for coinfection systems with $N = 2$ or $N = 3$ strains [49, 90, 88, 173]. A few studies, using arbitrary system size, derive analytical results for any number of coinfecting strains N [4, 140]. But the vast majority of N-strain coinfection models are entirely based on simulations [51, 61], with limited analytical insight and organic syntheses for the mechanisms of emergent dynamics.

In this article, we uncover the subtle structure of coinfection model with N strains. We introduce a general model to describe the population dynamics of multiple strains circulating in a host population with the possibility of co-infection. In particular, we focus on modeling the host-to-host transmission of different strains, using the SIS (susceptible - infected - susceptible) compartmental framework for endemic diseases. There are two sources of complexity in the model: i) the number of strains, which increases quadratically the dimensionality of the system, and ii) all the fitness dimensions in which the strains may vary. The latter is the main novelty of our framework.

We present a method for approximating the solution of this SIS- N-strain co-infection system, under a quasi-neutral assumption for the strain-defining parameters. To that end, we first analyze multi-strain co-infection system with symmetric traits. Then, based on the theoretical results in [72, 114, 196] and their applications to similar models in [86, 88, 140], we use the slow-fast dynamics approach and the method of multiple timescales to approximate the solution of systems with non-symmetric traits.

Extending the foundational work in [86, 88, 140], this article studies a more general dynamic system, with perturbations in many more dimensions of variation across strains, namely transmission, clearance rates and within-host competitiveness, besides the co-colonization vulnerability parameters (Figure 2.1). The complexity

of this general problem is reduced by the quasi-neutral assumption, with each parameter constrained to be close to its default value, allowing us to leverage the neutral system to approximate the quasi-neutral system. The difficulty lies in reformulating the original system starting from a neutral component plus perturbations, in terms of slow fast dynamics consisting of a fast sub-system and a slow sub-system. Thanks to the singular perturbation theory in [200] and Tikhonov's Theorem, we expect to find explicitly the emergent system describing the slow dynamics.

More precisely, we find how to rewrite the original system in the form $\frac{dx}{dt} = \varepsilon f(x,y,t,\varepsilon)$ and $\frac{dy}{dt} = g(x,y,t,\varepsilon)$ where x describes the slow dynamics and y the fast dynamics. Taking $\varepsilon = 0$ we obtain the degenerate fast system $\frac{dx}{dt} = 0$ and $\frac{dy}{dt} = g(x,y,t,0)$. Under appropriate assumptions, this fast system admit a (degenerate) attractor called the *slow manifold* of the form $y = \phi(x,t)$. Then, at the *slow time scale* $\tau = \frac{t}{\varepsilon}$ we obtain the *slow* dynamics on this slow manifold as $\frac{dx}{d\tau} = f(x,\phi(x,t),\tau,0)$ that needs to be computed explicitly. The singular perturbation theory makes the link between this slow dynamics and the dynamics of the original system for $0 < \varepsilon \ll 1$.

Even though we have an intuition for how the final model approximation in terms of fast-slow dynamics should work, with the neutral model as the organizing centre [93], it is not at all obvious from the start which should be the necessary mathematical steps when multiple perturbations occur and interact at the same time between N strains. In this article we uncover these steps, which ultimately lead us to a similar replicator equation to the one derived in [140] but now more complete because it involves variation among strains along more fitness dimensions. Indeed, we obtain an N dimensional replicator equation for strain frequencies over long time in terms of their pairwise invasion fitness matrix, and this connects our multi-strain coinfection framework in an endemic setting with the work of [102] which extensively researches this well-known model, and shows its contribution to evolution and game theory. With this simplifcation, qualitative and quantitative aspects of the competitive dynamics between N strains, leading to regimes of exclusion, coexistence, multi-stability, family of cycles or chaotic behavior can be investigated, and directly linked to their trait variations.

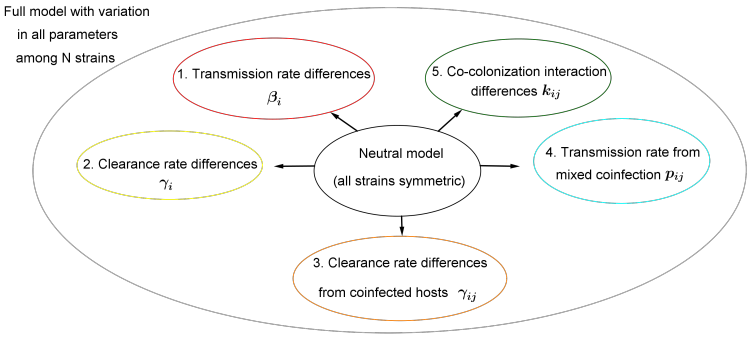


Fig. 2.1 **Schematic description of the spirit of our study.** We study the full N-strain SIS model with coinfection like in [140, 88], but here include variation in several parameters among strains, besides co-colonization interactions. For this, we consider the neutral model as the organizing center of the dynamics, and the slow-fast form for each case of trait variation. Finally, we combine all cases of singular perturbation in each parameter to obtain the general system. Our result is the dynamics in the slow manifold, which corresponds to a replicator system for N strain frequencies, governed by the pairwise invasion fitness matrix.

The paper is organized as follows. Section 2.2 outlines the general systems studied in this paper with corresponding quasi-neutral and neutral models. Then it introduces Tikhonov's theorem and the expansion theorem used to approximate the target model. Section 2.3 presents the main framework used to decompose the dynamics into fast and slow components, accompanied with lemmas and concrete steps. In this section, we state

the main result: the replicator system for strain frequencies, whose coefficients' matrix is defined by pairwise invasion fitnesses. Section 2.4 is devoted to the explicit computations for perturbations in each trait, and ends with the proof for the error estimate between the original system and the slow-fast approximation. In Section 2.5 we provide illustration by numerical simulations about the different regimes of system behaviour, including coexistence, competitive exclusion and more complex dynamics. This helps to contextualize the competitive outcomes between strains as a function of parameters. Finally, in Section 2.6 we close with conclusions and a discussion.

2.2 System, methods and results

This section aims to provide a general description of the dynamics followed by an outline of the analytical framework applied. We introduce the general structure, then subsequently present explicitly the steps of our approach, consisting in the quasi-neutral model, neutral model and slow-fast model. We then present Tikhonov's theorem, which is the key tool we use to approximate the singular perturbation dynamics efficiently. Important lemmas and main results are also stated in this section.

2.2.1 The general SIS coinfection model with Ntrains and some initial analysis

The dynamics studied in this article groups the pathogen types in N subsets, indexed by i, $1 \leq i \leq N$. With a set of ordinary differential equations, we then track the proportion of hosts in $1+N+N^2$ compartments: susceptible: S, hosts colonized by strain-i: I_i, hosts co-colonized by strain-i then strain-j: I_{ij}. Notice that we include also same strain coinfection, as argued in [140]. We formulate the general model based on the same structure as that in [140] but here allow for strains to vary in their transmission rates β_i, clearance rates of single infection γ_i (or duration of carriage $1/\gamma_i$), clearance rates from mixed co-colonization γ_{ij}, within-host competition reflected in relative transmissibilities from mixed coinfected hosts (p^i_{ij} and p^i_{ji}), as well as co-colonization vulnerabilities k_{ij}, already studied in [140].

$$\begin{cases} \dfrac{dS}{dt} &= r(1-S) + \displaystyle\sum_{i=1}^{N} \gamma_i I_i + \sum_{i,j=1}^{N} \gamma_{ij} I_{ij} - S \sum_{i=1}^{N} \beta_i J_i, \\ \dfrac{dI_i}{dt} &= \beta_i J_i S - (r+\gamma_i) I_i - I_i \displaystyle\sum_{j=1}^{N} k_{ij} \beta_j J_j, \qquad 1 \leq i \leq N, \\ \dfrac{dI_{ij}}{dt} &= k_{ij} I_i \beta_j J_j - (r+\gamma_{ij}) I_{ij}, \qquad 1 \leq i,j \leq N \end{cases} \tag{2.2.1}$$

where J_i is proportion of all hosts transmitting strain i, including singly- and co-colonized hosts and has the explicit formula

$$J_i = I_i + \sum_{j=1}^{N} \left(p^i_{ij} I_{ij} + p^i_{ji} I_{ji} \right). \tag{2.2.2}$$

Note that $\beta_i J_i$ is the force of infection of strain i, for all i. All mixed coinfection hosts, harboring strain i (and j), in any order, whether acquired first or second, can transmit strain i and the two probabilities of transmission are denoted by p^i_{ij} and p^i_{ji}. The corresponding probabilities to transmit the other strain for such hosts, is simply $1-p^i_{ij}$ and $1-p^i_{ji}$ respectively. Thus we allow for variation between strains in both transmissibility from mixed coinfection, and in the benefit gained within-host for transmission when landed there first (a precedence effect). In (2.2.1), for $1 \leq i,j \leq N$, parameters are summarized in Table 2.1. Summing up all the equations of (2.2.1)

Table 2.1 **Conventions and notations of parameters defining strains in our model, and host turnover. Under strain similarity assumptions, we can write each trait using a common reference for all strains, and express the variation as a deviation from neutrality, with ε a small number between 0 and 1.**

	Parameter	Interpretation	Strain similarity
1.	β_i	Strain-specific transmission rates	$\beta_i = \beta\,(1+\varepsilon b_i)$
2.	γ_i	Strain-specific clearance rates of single colonization	$\gamma_i = \gamma(1+\varepsilon v_i)$
3.	γ_{ij}	Clearance rates of co-colonization with i and j	$\gamma_{ij} = \gamma(1+\varepsilon u_{ij})$
4.	p^s_{ij}	Transmission probability of strain $s \in \{i,j\}$ from a host co-colonized by strain-i then strain-j, $\left(p^i_{ij} + p^j_{ij} = 1\right)$	$p^s_{ij} = \dfrac{1}{2} + \varepsilon\omega^s_{ij}$
5.	k_{ij}	Relative factor of altered susceptibility to co-colonization by strain j when a host is already colonized by strain i	$k_{ij} = k + \varepsilon\alpha_{ij}$
	r	Susceptible recruitment rate (equal to natural mortality rate)	
	R_0	Basic reproduction number	$R_0 = \frac{\beta}{\gamma+r}$

on both sides yields the equation for total mass

$$\frac{d}{dt}\left(S + \sum_{i=1}^{N} I_i + \sum_{i,j=1}^{N} I_{ij}\right) = r(1-S) - r\left(\sum_{i=1}^{N} I_i + \sum_{i,j=1}^{N} I_{ij}\right), \tag{2.2.3}$$

which leads to $S + \sum_{i=1}^{N} I_i + \sum_{i,j=1}^{N} I_{ij} = 1 - e^{-rt}$. Hence, $S + \sum_{i=1}^{N} I_i + \sum_{i,j=1}^{N} I_{ij}$ tends to 1 as $t \to \infty$.

We want to study a system whose host population is invariant. Such an expectation leads to the assumption that, (2.2.1) has the same recruitment rate of susceptibility host and mortality rate of strains. It is plausible to from now on assume that the total population size is constant and rescaled to unit. We also take the system (2.2.1) as given the initial conditions $S(0) + \sum_{i=1}^{N} I_i(0) + \sum_{i,j=1}^{N} I_{ij}(0) = 1$, which implies that the total population size is always one for any time. Thus our compartmental variables can be taken to reflect proportions of host in different epidemiological states.

2.2.2 Quasi-neutral system and new variables

A straightforward understanding of (2.2.1) is not possible due to its complexity, high-dimensional parameter space and number of equations. However, for indistinguishable strains, i.e. if all the parameters do not depend on the strain i, we obtain the so-called *neutral system* which is analytically tractable (see [86, 88, 140]). In this text, we make a *quasi neutral* assumption by assuming that the parameters are nearly equal, because the strains are similar. Without loss of generality we can take the same epsilon in all parameters with the perturbations written in the form presented in table 2.1. For the sake of simplicity, we denote the inverse duration of a carriage episode by strain i with $m_i = r + \gamma_i$, of a co-carriage episode by strains i and j with $m_{ij} = r + \gamma_{ij}$ and the corresponding inverse duration of carriage if all strains were equivalent with $m = r + \gamma$.

To work on the neutral system, it's useful to denote some new state variables, including the total 'mass' of singly-infected hosts I, the total 'mass' of doubly-infected hosts D, and the total 'mass' of infected hosts $T = I + D$. According to these definitions of T, I, D, we have the formulae:

$$I = \sum_{i=1}^{N} I_i, \quad D = \sum_{i,j=1}^{N} I_{ij}, \quad T = I + D. \tag{2.2.4}$$

It can be easily deduced from (2.2.4) together with $\omega_{ij}^i + \omega_{ij}^j = 0$ that $\sum_{i=1}^N J_i = T$.

Thanks to these new variables, the original system (2.2.1) can be rewritten into the extensive new form

$$\begin{cases}
\frac{dS}{dt} &= r(1-S) + \gamma T + \varepsilon\gamma\left(\sum_{i=1}^N \nu_i I_i + \sum_{i,j=1}^N u_{ij} I_{ij}\right) - \beta S T - \varepsilon\beta S \sum_{i=1}^N b_i J_i \\
\frac{dT}{dt} &= \beta S T - mT + \varepsilon\beta S \sum_{i=1}^N b_i J_i - \varepsilon\gamma\left(\sum_{i=1}^N \nu_i I_i + \sum_{i,j=1}^N u_{ij} I_{ij}\right) \\
\frac{dI_i}{dt} &= \beta J_i S + \varepsilon\beta b_i J_i S - (m + \varepsilon\gamma\nu_i) I_i - \beta I_i \sum_{j=1}^N (k+\varepsilon\alpha_{ij})(1+\varepsilon b_j) J_j \\
\frac{dJ_i}{dt} &= \beta(1+\varepsilon b_i) J_i S - \beta I_i \sum_{j=1}^N (k+\varepsilon\alpha_{ij})(1+\varepsilon b_j) J_j - \varepsilon\gamma\left[\nu_i I_i + \sum_{j=1}^N \left(\left(\frac{1}{2} + \varepsilon\omega_{ij}^i\right) u_{ij} I_{ij} + \left(\frac{1}{2} + \varepsilon\omega_{ji}^i\right) u_{ji} I_{ji}\right)\right] \\
&\quad - mJ_i + \beta \sum_{j=1}^N \left(\left(\frac{1}{2} + \varepsilon\omega_{ij}^i\right)(k+\varepsilon\alpha_{ij})(1+\varepsilon b_j) I_i J_j + \left(\frac{1}{2} + \varepsilon\omega_{ji}^i\right)(k+\varepsilon\alpha_{ji})(1+\varepsilon b_i) I_j J_i\right) \\
\frac{dI}{dt} &= \beta T S + \varepsilon\beta S \sum_{i=1}^N b_i J_i - mI - \varepsilon\gamma \sum_{i=1}^N \nu_i I_i - \beta \sum_{i=1}^N I_i \left(\sum_{j=1}^N (k+\varepsilon\alpha_{ij})(1+\varepsilon b_j) J_j\right) \\
\frac{dI_{ij}}{dt} &= \beta\,(k+\varepsilon\alpha_{ij})\,(1+\varepsilon b_j) I_i J_j - (m + \varepsilon\gamma u_{ij}) I_{ij}.
\end{cases} \tag{2.2.5}$$

This system has the generic form

$$\frac{dX}{dt} = \tilde{F}(X,\varepsilon) \quad \text{where} \quad X = (X_1, X_2, \ldots, X_{\tilde{n}}) \in \mathbb{R}^{\tilde{n}}$$

,for some integer $\tilde{n}$, and is equivalent to $\frac{dX}{dt} = F(X) + O(\varepsilon)$ after some algebraic transformations.

In ours case, the part $\frac{dX}{dt} = F(X)$ is known as the *neutral system*, consistently stays unaltered and be investigated in the subsection 2.2.3. It is important to note that this neutral system is structurally unstable.

Then, the part $O(\varepsilon)$ is a singular perturbation of the neutral system. To treat such an emergence by Tikhonov's theorem, it's essential to rewrite $\frac{dX}{dt} = F(X) + O(\varepsilon)$ into an equivalent slow-fast form

$$\begin{cases}
\frac{dx}{dt} = & \varepsilon\,(f(x,y) + O(\varepsilon)) \\
\frac{dy}{dt} = & \quad g(x,y) + O(\varepsilon)
\end{cases} \tag{2.2.6}$$

where $y \in \mathbb{R}^{n_y}$ is the fast variable and $x \in \mathbb{R}^{n_x}$ is the slow variable (with $n_x + n_y = \widetilde{n}$).

In general, the finding of this slow-fast reformulation is strongly dependent on the specific system. Here, it is achieved thanks to the ansatz (2.2.25) which is yielded from the study of the neutral system.
Hence, we start to study the important *neutral system* which is obtained for $\varepsilon = 0$ in (2.2.5).

This study yields the definition of the appropriate slow and fast variables (v_i, z_i). These variables together with the ansatz (2.2.25) are the key for the slow-fast study of the next section.

2.2.3 Neutral system, $\varepsilon = 0$

Taking $\varepsilon = 0$ in (2.2.5) leads to the so-called *Neutral System*[1] for $S, T, I, I_i, J_i, I_{ij}$ which reads after some simplifications:

$$\begin{cases} \dfrac{dS}{dt} = r(1-S) + \gamma T - S\beta T \\ \dfrac{dT}{dt} = S\beta T - mT \\ \dfrac{dI}{dt} = \beta TS - (m + k\beta T)I \\ \dfrac{dI_i}{dt} = \beta J_i S - mI_i - kI_i\beta T, & 1 \le i \le N \\ \dfrac{dJ_i}{dt} = (\beta S - m)J_i + \dfrac{1}{2}\beta kIJ_i - \dfrac{1}{2}\beta kI_iT, & 1 \le i \le N \\ \dfrac{dI_{ij}}{dt} = k\beta I_iJ_j - mI_{ij}, & 1 \le i, j \le N. \end{cases} \tag{2.2.7}$$

Such a triangular structure of this system enables to successively consider the subsystems for (S,T), I, (I_i, J_i) and I_{ij}.

• *Firstly*, we consider the neutral system for S, T as following

$$\begin{cases} \dfrac{dS}{dt} &= m(1-S) - \beta ST \\ \dfrac{dT}{dt} &= -mT + \beta ST \end{cases} \tag{2.2.8}$$

This system is a classical. As in [139], we define the basic reproduction number as $R_0 = \dfrac{\beta}{m}$. If $R_0 > 1$ then it admits a positive steady state (S^*, T^*) where $S^* = \dfrac{1}{R_0}$ and $T^* = 1 - S^*$.

We now recall a crucial proposition, which follows the definition of S^* and T^*.

Proposition 2.1. Assume that $S(0) > 0$ and $T(0) > 0$. If $R_0 \le 1$ then the solution S, T of system (2.2.8) tends to $(1,0)$. Otherwise, it tends to (S^*, T^*) asymptotically.

Proof. To apply the LaSalle's invariance principle in Appendix A, we need to find a suitable Lyapunov function $V(x)$. Consider that

$$V(S,T) = \frac{1}{2}(S - S^*)^2.$$

It's trivial that $V(S,T) > 0$ for all $(S,T) \neq (S^*, T^*)$ and $V(S^*, T^*) = 0$. We have that

$$\dot{V}(S,T) = (S - S^*)\dot{S} = -\beta(1-S)(S-S^*) \le 0.$$

Applying LaSalle's invariance priciple, we have the conclusion. □

• *Secondly*, we prove that $I(t) \to I^* := \dfrac{mT^*}{m + \beta kT^*}$.

[1] The name *neutral system* comes from the fact that if $\varepsilon = 0$ then the parameters do not depend on the strains as in the neutral theory, and the model describes indistinguishable strains.

Indeed, substitute (S,T) by $(S^*+(S-S^*),T^*+(T-T^*))$ into the equation of I in (2.2.8) then make some manipulations to obtain

$$\frac{dI}{dt} = mT^* - (m+\beta kT^*)I + [\beta S^*(T-T^*) + \beta T^*(S-S^*) + \beta(T-T^*)(S-S^*)]. \tag{2.2.9}$$

Consider the equation

$$\frac{d\tilde{I}}{dt} = mT^* - (m+\beta kT^*)\tilde{I} \tag{2.2.10}$$

which has the explicit solution $\tilde{I}(t) = \dfrac{mT^*}{m+\beta kT^*}\left(1 - \dfrac{m+\beta kT^*}{mT^*}I(0)\exp(-(m+\beta kT^*)t)\right)$. We simultaneously have the equation for $I-\tilde{I}$ as follows

$$\frac{d}{dt}\left(I-\tilde{I}\right) = -(m+\beta kT^*)\left(I-\tilde{I}\right) + [\beta S^*(T-T^*) + \beta T^*(S-S^*) + \beta(T-T^*)(S-S^*)]. \tag{2.2.11}$$

Set $f(t) = \beta S^*(T-T^*) + \beta T^*(S-S^*) + \beta(T-T^*)(S-S^*)$ then $f(t) \to 0$ asymptotically when $t \to \infty$, by Proposition 2.1. It's easy to see $\left(I-\tilde{I}\right) = \exp(-(m+\beta kT^*)t)\left(\int_0^t \exp((m+\beta kT^*)s) f(s)ds + C\right)$, with C is some suitable constant. Hence, $I(t) - \tilde{I}(t) \to 0$ when $t \to \infty$ then leads to $I(t) \to I^*$ as $t \to \infty$.

For later reference, we also write their equilibrium values in the neutral system

$$S^* = \frac{m}{\beta}, \qquad T^* = 1 - \frac{m}{\beta}, \qquad I^* = \frac{mT^*}{m+\beta kT^*}, \qquad D^* = T^* - I^* = \frac{\beta kT^{*2}}{m+\beta kT^*}. \tag{2.2.12}$$

•*Thirdly*, from (2.2.7), we also have the neutral model for I_i, J_i for all $1 \le i \le N$. This is the very important part which gives crucial insight for $0 < \varepsilon \ll 1$ in the next section. For now, $\varepsilon = 0$ and substitute (S,T,I) by the limit (S^*,T^*,I^*), we obtain the (degenerate) linear system

$$\frac{d}{dt}\begin{pmatrix} I_i \\ J_i \end{pmatrix} = \begin{pmatrix} -(m+\beta kT^*) & m \\ -\dfrac{\beta kT^*}{2} & \dfrac{\beta kI^*}{2} \end{pmatrix}\begin{pmatrix} I_i \\ J_i \end{pmatrix} \quad . \tag{2.2.13}$$

Setting $A = \begin{pmatrix} -(m+\beta kT^*) & m \\ -\dfrac{\beta kT^*}{2} & \dfrac{\beta kI^*}{2} \end{pmatrix}$ then from the determinant and the trace of A, we note that A has two eigenvalues 0 and $-\xi$ with $\xi = m + \beta kT^* - \dfrac{1}{2}\beta kI^* > m + \dfrac{1}{2}\beta k(T^* - I^*) > 0$.

Hence, we read the matrix to determine a variable z_i given by the proper direction associated with the value proper zero, therefore constant.

To do this, we set that

$$\begin{gathered} P = \begin{pmatrix} 2T^* & I^* \\ D^* & T^* \end{pmatrix}, \qquad P^{-1} = \frac{1}{|P|}\begin{pmatrix} T^* & -I^* \\ -D^* & 2T^* \end{pmatrix}, \\ \text{and for } i = 1,\cdots,N\,, \quad \text{setting:} \quad \begin{pmatrix} v_i \\ z_i \end{pmatrix} = P^{-1}\begin{pmatrix} I_i \\ J_i \end{pmatrix} \end{gathered} \tag{2.2.14}$$

We have the diagonalization for matrix A, which is $A = P\begin{pmatrix} -\xi & 0 \\ 0 & 0 \end{pmatrix}P^{-1}$ and $|P| = 2T^{*2} - I^*D^* > 0$.

From (2.2.13) and (2.2.14), we infer an equation for $\begin{pmatrix} v_i \\ z_i \end{pmatrix}$ for each $1 \leq i \leq N$:

$$\begin{cases} \dfrac{dv_i}{dt} &= -\xi v_i \\ \dfrac{dz_i}{dt} &= 0. \end{cases} \tag{2.2.15}$$

This step of changing to (v_i, z_i) plays an important role. Since under these new variables, we can rewrite into the slow-fast form. It allows us to apply the Tikhonov's Theorem introduced in the next subsection.
Let us remark that z_i is exactly frequency of strain i in the total of infected, see the proof in [140].

• *Fourthly*, the N^2 last equations for I_{ij} in (2.2.7) yields $1 \leq i \leq N$

$$\frac{dI_{ij}}{dt} = \beta k I_i J_j - m I_{ij}. \tag{2.2.16}$$

Whose dynamics is trivial once I_i and J_i are known. Indeed, assume that for each i, there exists $\left(\tilde{I}_i, \tilde{J}_i\right)$ such that $I_i(t) - \tilde{I}_i(t) = O(\varepsilon)$ and $J_i(t) - \tilde{J}_i(t) = O(\varepsilon)$, then we can rewrite (2.2.16) into

$$\frac{dI_{ij}}{dt} = -mI_{ij} + \beta k \tilde{I}_i \tilde{J}_j + \beta k \left[\left(I_i - \tilde{I}_i\right) \tilde{J}_j + \left(J_j - \tilde{J}_j\right) \tilde{I}_i + \left(I_i - \tilde{I}_i\right) \left(J_j - \tilde{J}_j\right) \right]. \tag{2.2.17}$$

Consider the equation

$$\frac{d\tilde{I}_{ij}}{dt} = -m\tilde{I}_{ij} + \beta k \tilde{I}_i \tilde{J}_j \tag{2.2.18}$$

then we can obtain the differential equation for $I_{ij} - \tilde{I}_{ij}$

$$\frac{d}{dt}\left(I_{ij} - \tilde{I}_{ij}\right) = -m\left(I_{ij} - \tilde{I}_{ij}\right) + \beta k \left[\left(I_i - \tilde{I}_i\right) \tilde{J}_j + \left(J_j - \tilde{J}_j\right) \bar{I}_i + \left(I_i - \tilde{I}_i\right) \left(J_j - \tilde{J}_j\right) \right]. \tag{2.2.19}$$

By our assumption on $\tilde{I}_i, \tilde{J}_j$ and use the same arguments for $I(t) \to I^*$, we deduce that $I_{ij}(t) - \tilde{I}_{ij}(t) = O(\varepsilon)$ on each bounded interval of time.

2.2.4 Tikhonov's Theorem and derivation of the non-neutral dynamics

Using the above idea, we transform the problem into an equivalent slow-fast form which is analyzed through singular perturbations method.

According to previous arguments, our slow-fast form includes variables $(X, Y, \mathbf{L}, \mathbf{v}, \mathbf{z})$. Using (2.2.14), we have the variables $\begin{pmatrix} v_i \\ z_i \end{pmatrix}$ that we determine from $\begin{pmatrix} I_i \\ J_i \end{pmatrix}$.

Proceeding like in (2.2.15), we obtain for $\varepsilon > 0$:

$$\begin{cases} \dfrac{dv_i}{dt} &= -\xi v_i + O(\varepsilon) \\ \dfrac{dz_i}{dt} &= O(\varepsilon). \end{cases} \tag{2.2.20}$$

By setting $\tau = \varepsilon t$, (2.2.20) can be read as the slow time scale:

$$\begin{cases} \varepsilon \dfrac{dv_i}{d\tau} &= -\xi v_i + O(\varepsilon) \\ \dfrac{dz_i}{d\tau} &= O(1). \end{cases} \tag{2.2.21}$$

We need to compute explicitly the perturbation $O(1)$ in (2.2.21). This computation is quite complex especially when involving perturbation in each parameters, so it's worthwhile dividing this progress into five sub-cases wherein only one perturbation at a time occurs.

After that, we will treat the slow-fast form by the Tikhonov's theorem, that is presented as follows.

Theorem 2.2 (Tikhonov, 1952, see [196])**.** *Consider the initial value problem*

$$\begin{cases} \dfrac{dx}{d\tau} &= f(x,y,\tau) + \varepsilon \ldots, \qquad x(0) = x_0, \quad x \in D \subset \mathbb{R}^n, \\ \varepsilon \dfrac{dy}{d\tau} &= g(x,y,\tau) + \varepsilon \ldots, \qquad y(0) = y_0, \quad y \in G \subset \mathbb{R}^n. \end{cases} \tag{2.2.22}$$

For f and g, we take sufficiently smooth vector functions in x, y and t; the dots represent (smooth) higher-order terms in ε.

a. *We assume that a unique solution of the initial value problem exists and suppose this holds also for the reduced problem*

$$\begin{cases} \dfrac{dx}{d\tau} &= f(x,y,\tau), \qquad x(0) = x_0, \\ 0 &= g(x,y,\tau), \end{cases} \tag{2.2.23}$$

with solution $\bar{x}(\tau)$, $\bar{y}(\tau)$.

b. *Suppose that $0 = g(x,y,\tau)$ is solved by $\bar{y} = \phi(x,\tau)$, where $\phi(x,\tau)$ is a continuous function and an isolated root, i.e. there exists a neighbor of $\phi(x,\tau)$ such that there is no other solution for $0 = g(x,y,\tau)$ in this vicinity. Also, suppose that $\bar{y} = \phi(x,t)$ is an asymptotically stable solution [2] of the equation $\dfrac{dy}{dt} = g(x,y,\tau)$, where $\tau = \varepsilon t$, that is uniform in the parameters $x \in D$ and $t \in \mathbb{R}^+$.*

c. *$y(0)$ is contained in an interior subset of the domain of attraction of $\bar{y} = \phi(x,\tau)$ in the case of the parameter values $x = x(0)$, $\tau = 0$.*

Then, we have

$$\begin{aligned} \lim_{\varepsilon \to 0} x_\varepsilon(\tau) &= \bar{x}(\tau), \qquad 0 \le \tau \le T, \\ \lim_{\varepsilon \to 0} y_\varepsilon(\tau) &= \bar{y}(\tau), \qquad 0 < \tau_0 \le \tau \le T, \end{aligned} \tag{2.2.24}$$

with τ_0 and T are constants independent of ε.

[2] Recall that the solution $\bar{y} = \phi(x,\tau)$ is asymptotically stable if for each $\tau_0 > 0$, a $\delta(\tau_0)$ can be found such that: $\|y_0 - \phi(x,\tau_0)\| \le \delta(\tau_0)$ yields $\lim_{\tau \to \infty} \|y(\tau;\tau_0,x_0) - \phi(x,\tau)\| = 0$.

Beside, it needs to use another result that allows us to approximate the original system by the slow-fast form. The following error estimate gives a more precise description of these limits. (theorem 9.1, [200] adapted here for the simple case $m = 0$).

Theorem 2.3. *[see [200]] Consider the initial value problem*

$$\frac{dx}{dt} = f_0(t,x) + \varepsilon R(t,x,\varepsilon) \tag{2.2.25}$$

with $x(t_0) = \eta$ and $|t - t_0| \leq h$, $x \in D \subset \mathbb{R}^n$, $0 \leq \varepsilon \leq \varepsilon_0$. Assume that in this domain we have

a. $f(t,x)$ continuous in t and x, 2 times continuously differentiable in x;

b. $R(t,x,\varepsilon)$ continuous in t,x and ε, Lipschitz-continuous in x.

Let $x_0(t)$ be the solution of

$$\frac{dx}{dt} = f_0(t,x) \tag{2.2.26}$$

with $x_0(t_0) = \eta$ Let $T > 0$ and assume that both x and x_0 are defined on $[0,T]$ for any $\varepsilon \in (0,\varepsilon_0)$. There exist $C > 0$ (depending on T) such that for any $\varepsilon \in (0,\varepsilon_0)$, and $t \in (0,T)$, we have the estimate

$$\|x(t) - x_0(t)\| \leq C\varepsilon \tag{2.2.27}$$

2.3 Integrating many perturbations in the slow-fast approximation

2.3.1 Steps for application of Tikhonov's theorem in our system

Next we develop a lemma showing allowing to linearly combine all the relevant simple cases directly into the slow equation. For this purpose, we use the following notations in system (2.2.5).

$$\begin{aligned} &\beta_i = \beta\,(1+\chi_1\varepsilon b_i)\,; \qquad \gamma_i = \gamma(1+\chi_2\varepsilon v_i)\,; \qquad \gamma_{ij} = \gamma(1+\chi_3\varepsilon u_{ij})\,; \\ &p^s_{ij} = \frac{1}{2} + \chi_4\varepsilon\omega^s_{ij} \quad s\in\{i,j\} \quad \left(\omega^i_{ij} + \omega^j_{ij} = 0\right); \qquad k_{ij} = k + \chi_5\varepsilon\alpha_{ij}; \end{aligned} \tag{2.3.1}$$

where $\chi_d \in \{0,1\}$ for $d = 1,2,3,4,5$.

Any combination of trait variation among strains, can be captured via $\mathscr{A}$ where $\mathscr{A}$ is a subset of $\{1,2,3,4,5\}$ denoting the absence/presence of perturbations in that parameter among strains: for some fixed initial values given, let $C_{\mathscr{A}}$ be the system (2.2.5) with $\chi_d = 1$ if $d \in \mathscr{A}$ and $\chi_d = 0$ if $d \notin \mathscr{A}$. For simplicity, we note also $C_{\{d\}}$ by C_d for $d \in \{1,2,3,4,5\}$.

Remark 2.1. If $\mathscr{A} = \emptyset$ then there is no perturbation and the system $C_\emptyset$ is exactly the *neutral* model (2.2.7). If $\mathscr{A} = \{5\}$ then C_5 is the system with perturbation on the co-colonization interaction parameters k_{ij} only, that has been studied in [86, 88, 140].

In order to capture all the perturbations of order 1 in the equation of the z_i we need these additional changes of variables:

$$\boxed{S(t) = S^* - \varepsilon X(t) + O(\varepsilon^2); \quad T(t) = T^* + \varepsilon X(t) + O(\varepsilon^2); \quad I(t) = I^* + \varepsilon Y(t) + O(\varepsilon^2).} \tag{2.3.2}$$

where S^*, T^* and I^* are defined in (2.2.12), and for $i = 1, \cdots, N$:

$$L_i(t) = \frac{1}{2}\sum_{j=1}^{N}\left(u_{ij}I_{ij}(t) + u_{ji}I_{ji}(t)\right). \tag{2.3.3}$$

With these notations, $C_{\mathscr{A}}$ reads

$$\begin{cases} \dfrac{dX}{dt} &= -\beta T^* X + \chi_1\beta S^*\displaystyle\sum_{i=1}^{N} b_i J_i - \chi_2\gamma\sum_{i=1}^{N}\nu_i I_i - \chi_3\gamma\sum_{i=1}^{N} L_i + O(\varepsilon) \\ \dfrac{dY}{dt} &= \beta(S^* - T^* - kI^*)X - (m+\beta kT^*)Y + \chi_1\beta(S^* - kI^*)\displaystyle\sum_{i=1}^{N} b_i J_i - \chi_2\gamma\sum_{i=1}^{N}\nu_i I_i - \chi_5\beta\sum_{i,j=1}^{N}\alpha_{ij}I_i J_j + O(\varepsilon) \\ \dfrac{dL_i}{dt} &= -mL_i + \chi_3\dfrac{1}{2}\beta\gamma k I_i\displaystyle\sum_{j=1}^{N} u_{ij}J_j + \chi_3\dfrac{1}{2}\beta\gamma k J_i\sum_{j=1}^{N} u_{ji}I_j + O(\varepsilon) \end{cases} \tag{2.3.4}$$

together with (we omit terms of $O\left(\varepsilon^2\right)$)

$$\frac{d}{dt}\begin{pmatrix} I_i \\ J_i \end{pmatrix} = A\begin{pmatrix} I_i \\ J_i \end{pmatrix} - \varepsilon\beta\begin{pmatrix} k & 1 \\ \frac{k}{2} & 1 \end{pmatrix}\begin{pmatrix} I_i \\ J_i \end{pmatrix} X + \varepsilon\frac{\beta k}{2}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} I_i \\ J_i \end{pmatrix} Y + \varepsilon\mathscr{M}_{\mathscr{A}}\begin{pmatrix} I_i \\ J_i \end{pmatrix} - \varepsilon\chi_3\begin{pmatrix} 0 \\ L_i \end{pmatrix} \tag{2.3.5}$$

where A is defined in (2.2.13) and $\mathscr{M}_{\mathscr{A}}$ is the matrix

$$\begin{pmatrix} -\chi_1\beta k\displaystyle\sum_{i=1}^{N} b_i J_i - \chi_2\gamma\nu_i - \chi_5\beta\sum_{j=1}^{N}\alpha_{ij}J_j & \chi_1\beta b_i S^* \\ \beta\displaystyle\sum_{j=1}^{N}\left(\chi_4 k\omega_{ij}^i - \chi_5\frac{\alpha_{ij}}{2}\right)J_j - \chi_1\beta\frac{k}{2}\sum_{i=1}^{N} b_i J_i - \chi_2\gamma\nu_i & \chi_1\beta b_i\left(S^* + \frac{kI^*}{2}\right) + \beta\displaystyle\sum_{j=1}^{N}\left(\chi_4 k\omega_{ji}^i + \chi_5\frac{\alpha_{ji}}{2}\right)I_j \end{pmatrix} \tag{2.3.6}$$

In order to apply the Theorem (2.2), we rewrite system $C_{\mathscr{A}}$ using the changes of variables detailed in (2.2.14). Let us note

$$\mathbf{L} = (L_i)_i, \quad \mathbf{v} = (v_i)_i, \quad \mathbf{z} = (z_i)_i,$$

and $-\xi = -(m+\beta kT^*) + \dfrac{\beta kI^*}{2} < 0$. The system $C_{\mathscr{A}}$ reads now as the slow-fast form

$$\begin{cases} \dfrac{dX}{dt} &= -\beta T^* X + \chi_1 F_X^1(\mathbf{v},\mathbf{z}) + \chi_2 F_X^2(\mathbf{v},\mathbf{z}) + \chi_3 F_X^3(\mathbf{L}) + O(\varepsilon) \\ \dfrac{dY}{dt} &= \beta(S^* - T^* - kI^*)X - (m+\beta kT^*)Y + \chi_1 F_Y^1(\mathbf{v},\mathbf{z}) + \chi_2 F_Y^2(\mathbf{v},\mathbf{z}) + \chi_5 F_Y^5(\mathbf{v},\mathbf{z}) + O(\varepsilon) \\ \dfrac{dL_i}{dt} &= -mL_i + \chi_3 F_{L_i}(\mathbf{v},\mathbf{z}) + O(\varepsilon) \\ \dfrac{dv_i}{dt} &= -\xi v_i + O(\varepsilon) \\ \dfrac{dz_i}{dt} &= \varepsilon\left(F_{z_i}(X,Y,\mathbf{L},\mathbf{v},\mathbf{z}) + O(\varepsilon)\right) \end{cases} \tag{2.3.7}$$

wherein we have replaced I_i and J_i by v_i and z_i though the change of variable (2.2.14), that is:

$$\begin{pmatrix} I_i \\ J_i \end{pmatrix} = P\begin{pmatrix} v_i \\ z_i \end{pmatrix} \text{ with } P = \begin{pmatrix} 2T^* & I^* \\ D^* & T^* \end{pmatrix}.$$

For $i = 1, \cdots, N$, the functions F_X^i, F_Y^i and F_{L_i} are obviously deduced from the right term of (2.3.4) and are linear in theirs variables, X, Y and $\mathbf{L}$ respectively. The function F_Y^4 is quadratic in $(\mathbf{v}, \mathbf{z})$. Finally, F_{z_i} is given by the second line of the right term of (2.3.6) after the linear change of variables (2.2.14):

$$F_{z_i}(X,Y,\mathbf{L},\mathbf{v},\mathbf{z}) = \begin{pmatrix} 0 & 1 \end{pmatrix} P^{-1} \left(\beta \begin{pmatrix} -k & -1 \\ -\frac{k}{2} & -1 \end{pmatrix} X + \frac{\beta k}{2} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} Y + \mathscr{M}_{\mathscr{A}} \right) P \begin{pmatrix} v_i \\ z_i \end{pmatrix} + \begin{pmatrix} 0 & 1 \end{pmatrix} P^{-1} \chi_3 \gamma \begin{pmatrix} 0 \\ L_i \end{pmatrix}. \tag{2.3.8}$$

The next step is to change the time scale. Taking $\tau = \varepsilon t$ in (2.3.7) we obtain[3] the following system which is equivalent to (2.3.7) but in the slow motion τ.

$$\begin{cases} \varepsilon \dfrac{dX}{d\tau} &= -\beta T^* X + \chi_1 F_X^1(\mathbf{v},\mathbf{z}) + \chi_2 F_X^2(\mathbf{v},\mathbf{z}) + \chi_3 F_X^3(\mathbf{L}) + O(\varepsilon) \\ \varepsilon \dfrac{dY}{d\tau} &= \beta(S^* - T^* - kI^*)X - (m + \beta k T^*)Y + \chi_1 F_Y^1(\mathbf{v},\mathbf{z}) + \chi_2 F_Y^2(\mathbf{v},\mathbf{z}) + \chi_5 F_Y^5(\mathbf{v},\mathbf{z}) + O(\varepsilon) \\ \varepsilon \dfrac{dL_i}{d\tau} &= -mL_i + \chi_3 F_{L_i}(\mathbf{v},\mathbf{z}) + O(\varepsilon) \\ \varepsilon \dfrac{dv_i}{d\tau} &= -\xi v_i + O(\varepsilon) \\ \dfrac{dz_i}{d\tau} &= F_{z_i}(X,Y,\mathbf{L},\mathbf{v},\mathbf{z}) + O(\varepsilon) \end{cases} \tag{2.3.9}$$

Using the notation of the Theorem 2.2, we see that the fast variables is $y(\tau) = (X,Y,\mathbf{L},\mathbf{v})$ and the slow variable is $x(\tau) = \mathbf{z}(\tau)$. The first step in applying the Tikhonov theorem is to take $\varepsilon = 0$ in (2.3.9) and to show that the fast variable converge to an attractor $\Phi(\mathbf{z})$ which is parametrized by the slow variable.

Lemma 2.4. Let $\varepsilon = 0$ in (2.3.9). Then there exist a function $\Phi(\mathbf{z}) = (X^*(\mathbf{z}), Y^*(\mathbf{z}), \chi_3 \mathbf{L}^*(\mathbf{z}), 0)$ such that the solution $(X,Y,\mathbf{L},\mathbf{v},\mathbf{z})$ of (2.3.7) with any initial condition

$$(X,Y,\mathbf{L},\mathbf{v},\mathbf{z})(0) = (X_0, Y_0, \mathbf{L}_0, \mathbf{v}_0, \mathbf{z}_0) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$$

verifies $\mathbf{z}(t) = \mathbf{z}_0$ for all $t \geq 0$ and

$$\lim_{t \to +\infty} (X,Y,\mathbf{L},\mathbf{v})(t) = \Phi(\mathbf{z}_0).$$

Moreover, X^* and Y^* are linear function of the χ_i.

Proof. Using the triangular structure of (2.3.9) the idea is to compute the limits step by step of $\mathbf{v}$, $\mathbf{L}$, X and Y in this order. Here we make a quick formal computation by simply plugging the limits obtained at one step into the equation of the next step. It is easy to verified that this computation is justified and we omit it here for clarity. Since (2.3.9) is equivalent to (2.3.7) but in the slow motion, we take $\varepsilon = 0$ in (2.3.7). We have directly $\mathbf{z}(t) = \mathbf{z}_0$ for all $t \geq 0$ and $v_i = e^{-\xi t} v_i(0) \to 0$ asymptotically as $t \to +\infty$. Remark that taking $v_i = 0$ in the others equations leads to the simple change of variables : $I_i = I^* z_i$ and $J_i = T^* z_i$ that we can plug in (2.3.4)-(2.3.5)-(2.3.6) to simplify the explicit computations.
Now we have the following asymptotic limits

$$L_i(t) \to \chi_3 \frac{1}{m} F_{L_i}(0, \mathbf{z}_0) = \chi_3 L_i^*(\mathbf{z}_0).$$

[3] We use the usual notation abuse. Rigorously speaking, we have to define $\widetilde{X}(\tau) = X\left(\frac{\tau}{\varepsilon}\right)$ and the same for each variables. Here we remove the $\widetilde{}$ for simplicity.

Denoting $\mathbf{L}^* = (L_i^*)_i$ and plugging this into the equation of X we have:

$$X(t) \to -\frac{1}{\beta T^*}\left(\chi_1 F_X^1(0,\mathbf{z}_0) + \chi_2 F_X^2(0,\mathbf{z}_0) + \chi_3 F_X^3(\chi_3 \mathbf{L}^*(\mathbf{z}_0))\right) = X^*(\mathbf{z}_0).$$

Remark that by linearity of the F_X^i and the fact that $\chi_d^2 = \chi_d$ for each d, we have the simpler formula

$$X^*(\mathbf{z}_0) = -\frac{1}{\beta T^*}\left(\chi_1 F_X^1(0,\mathbf{z}_0) + \chi_2 F_X^2(0,\mathbf{z}_0) + \chi_3 F_X^3(\mathbf{L}^*(\mathbf{z}_0))\right). \tag{2.3.10}$$

Finally, using the same arguments we get

$$Y(t) \to Y^*(\mathbf{z}_0)$$

wherein we have note

$$Y^*(\mathbf{z}_0) = \frac{1}{m+\beta k T^*}\left(\beta(S^* - T^* - kI^*)X^*(\mathbf{z}_0) + \chi_1 F_Y^1(0,\mathbf{z}_0) + \chi_2 F_Y^2(0,\mathbf{z}_0) + \chi_5 F_Y^5(0,\mathbf{z}_0)\right).$$

□

Now, we take $\varepsilon = 0$ in (2.3.9) and we fix

$$(X,Y,\mathbf{L},\mathbf{v})(\tau) = \Phi(\mathbf{z}(\tau)). \tag{2.3.11}$$

Then the $2+2N$ first equations are satisfied and the N last equations give the *slow system*

$$\frac{dz_i}{d\tau} = F_{z_i}(X^*(\mathbf{z}), Y^*(\mathbf{z}), \mathbf{L}^*(\mathbf{z}), 0, \mathbf{z}). \tag{2.3.12}$$

It's important to note that, since $\mathbf{v} = 0$ (2.3.12) then (2.2.14) gives $\sum_{i=1}^N z_i = 1$ by the formula $I_i = I^* z_i$. Hence z_i reflects the frequency of strain i for all i. Remark that we have also $J_i = T^* z_i$.
The Theorem 2.2 imply that the solutions of (2.3.12) together with (2.3.11) gives a good approximation of the original system (2.3.9) for a small enough but positive ε. Coming back to the original variables of the SIS system, we deduce the following result on error estimate, whose proof will be given in section 2.4.5.

Lemma 2.5. Let $T > 0$ be fixed. There exists $\varepsilon_0 > 0$ and $C_T > 0$ such that for any $\varepsilon \in (0,\varepsilon_0)$ we have for any solution of $(S,(I_i)_i,(I_{ij})_{ij})_{i,j}$ of (2.2.1) and $(z_i)_i$ of (2.3.12)

$$\left|S\left(\frac{\tau}{\varepsilon}\right) - S^*\right| + \sum_{i=1}^N \left|I_i\left(\frac{\tau}{\varepsilon}\right) - I^* z_i(\tau)\right| + \sum_{i,j=1}^N \left|I_{ij}\left(\frac{\tau}{\varepsilon}\right) - k\frac{I^* T^*}{S^*} z_i(\tau) z_j(\tau)\right| \le \varepsilon C_T, \tag{2.3.13}$$

Proof. See section 2.4.5. □

It remains to compute explicitly the slow system (2.3.12). The following lemma shows that it suffices to compute independently the system for each perturbation, that is $\mathscr{A} = \{d\}$ for $d \in \{1,2,3,4,5\}$. The case of a general $\mathscr{A}$ is simply a sum over simple cases thanks to the following result.

Lemma 2.6. Let $\mathscr{A} \subset \{1,\cdots,5\}$. Recall that $\chi_d = 1$ if $d \in \mathscr{A}$ and $\chi_d = 0$ if $d \notin \mathscr{A}$. The functions F_{z_i} for $i = 1,\cdots,N$ in (2.3.12) read

$$F_{z_i}(X^*(\mathbf{z}), Y^*(\mathbf{z}), \mathbf{L}^*(\mathbf{z}), 0, \mathbf{z}) = \sum_{d=1}^5 \chi_d z_i f_{z_i}^d(\mathbf{z}),$$

where the functions $f_{z_i}^d$ do not depend on χ_d.
In particular, if $\mathscr{A} = \{d\}$ for some $d \in \{1,2,3,4,5\}$, then

$$F_{z_i}(X^*(\mathbf{z}), Y^*(\mathbf{z}), \mathbf{L}^*(\mathbf{z}), 0, \mathbf{z}) = z_i f_{z_i}^d(\mathbf{z}).$$

Proof. Taking $v_i = 0$ in (2.3.8) we see that there is two constant C_X and C_Y such that

$$F_{z_i}(X^*(\mathbf{z}), Y^*(\mathbf{z}), \mathbf{L}^*(\mathbf{z}), 0, \mathbf{z}) = z_i \left(C_X X^*(\mathbf{z}), + C_Y Y^*(\mathbf{z}) + \begin{pmatrix} 0 & 1 \end{pmatrix} P \mathscr{M}_{\mathscr{A}} P^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) + \begin{pmatrix} 0 & 1 \end{pmatrix} \chi_3 \gamma P^{-1} \begin{pmatrix} 0 \\ L_i^*(\mathbf{z}) \end{pmatrix}.$$

Firstly, as it is shown in the proof of the lemma 2.4, the expression of X^* and Y^* are both a linear combination of the χ_d.

Secondly, recalling that we have at this step $I_i = I^* z_i$, $J_i = T^* z_i$, $\mathbf{L} = \chi_3 \mathbf{L}^*$ and, in particular, $\chi_3^2 = \chi_3$. Plugging this in (2.3.7), we see that the matrix $\mathscr{M}_{\mathscr{A}}$ is also a linear combination of the χ_d :

$$\mathscr{M}_{\mathscr{A}} = \sum_{d \in \mathscr{A}} \mathscr{M}_{\{d\}} = \sum_{d \in \{1,2,3,4,5\}} \chi_d \mathscr{M}_d.$$

denoting $m_d(\mathbf{z}) = \begin{pmatrix} 0 & 1 \end{pmatrix} P^{-1} \mathscr{M}_{\{d\}} P \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, this yields to:

$$\begin{pmatrix} 0 & 1 \end{pmatrix} P^{-1} \mathscr{M}_{\mathscr{A}} P \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \sum_{d \in \{1,2,3,4,5\}} \chi_d m_d(\mathbf{z}). \tag{2.3.14}$$

Thirdly, plugging $I_i = I^* z_i$ and $J_i = T^* z_i$, for all i in (2.3.4) we prove that

$$L_i^*(\mathbf{z}) = \frac{1}{2m} \beta k I^* T^* z_i \sum_{j=1}^{N} (u_{ij} + u_{ji}) z_j.$$

Actually, this value $L^*(\mathbf{z})$ is exact as in (2.4.23) computed in section 2.4.3.
The result follows directly from the three previous points. □

In the next section 2.4, these functions $f_{z_i}^d$ are explicitly computed for any d.

2.3.2 Main Results

In the earlier study [140], we computed the slow dynamics for $\mathscr{A} = \{5\}$, that is for perturbation in $k_{ij} = k + \varepsilon \alpha_{ij}$ only, i.e. for strains varying only in their co-colonization susceptibility interactions. We found that the slow system obeys a *replicator equation* which has the from

$$\begin{cases} \dot{z}_i = \Theta z_i \left((\Lambda \mathbf{z})_i - \mathbf{z}^T \Lambda \mathbf{z} \right), \ i = 1, \cdots, N, \\ \sum_{i=1}^{N} z_i = 1 \end{cases} \tag{2.3.15}$$

where Θ is a positive constant depending on the parameters of the *neutral system* and $\Lambda = \left(\lambda_i^j\right)_{i,j}$ is the $N \times N$ matrix of pairwise invasion fitness among strains where the term of line i and column j was

$$\lambda_i^j = \frac{I^*}{D^*}(\alpha_{ji} - \alpha_{ij}) + (\alpha_{ji} - \alpha_{jj}).$$

We remark that, the notation "dot"-$\dot{}$ (in $\dot{z}_i$) designates here the derivative temporal for the timescale τ (and this is also in the rest of this chapter).

In this present article, we show that the system (2.3.15) is true for any type of perturbation. The change is that the constant Θ and the pairwise fitness λ_i^j depend on the multiple trait variations which occur in the system. From the Lemma 2.6, we infer in particular that the λ_i^j are just a linear combination of the different perturbations. This implies that the pairwise invasion fitness between any two strains is an explicit weighted sum over all fitness dimensions where the two strains vary. More precisely, the main result of this article is as follows.

Let $\mathscr{A} \subset \{1,2,3,4,5\}$. Using the notations in the previous section, we prove in the 2.4 that (2.3.12) reads.

$$\boxed{\begin{aligned}\frac{dz_i}{d\tau} = \Theta_1 z_i\left(b_i - \sum_{j=1}^N b_j z_j\right) + \Theta_2 z_i\left(-\nu_i + \sum_{j=1}^N \nu_j z_j\right) + \Theta_3 z_i\left[-\sum_{j=1}^N (u_{ij}+u_{ji})z_j + \sum_{j,l=1}^N (u_{jl}+u_{lj})z_l z_j\right] \\ + \Theta_4 z_i\left[\sum_{j=1}^N (\omega_{ij}^i - \omega_{ji}^j)z_j\right] + \Theta_5 z_i\left[\sum_{j=1}^N \left(\frac{T^*}{D^*}\alpha_{ji} - \frac{I^*}{D^*}\alpha_{ij}\right)z_j - \sum_{j,l=1}^N \alpha_{jl} z_j z_l\right]\end{aligned}} \tag{2.3.16}$$

where

$$\Theta_1 = \chi_1 \frac{2\beta S^* T^{*2}}{|P|}, \quad \Theta_2 = \chi_2 \frac{\gamma I^*(I^*+T^*)}{|P|}, \quad \Theta_3 = \chi_3 \frac{\gamma T^* D^*}{|P|}, \quad \Theta_4 = \chi_4 \frac{2mT^*D^*}{|P|}, \quad \Theta_5 = \chi_5 \frac{\beta T^* I^* D^*}{|P|}. \tag{2.3.17}$$

Naturally, if $\mathscr{A} = \emptyset$, (2.3.16) becomes simply $\dfrac{dz_i}{d\tau} = 0$. Otherwise, if $\mathscr{A} \neq \emptyset$, it is useful to rewrite (2.3.16) using the pairwise invasion fitness between strains in (2.3.15). Define

$$\Theta = \Theta_1 + \Theta_2 + \Theta_3 + \Theta_4 + \Theta_5 \quad \text{and } \theta_i = \frac{\Theta_i}{\Theta}. \tag{2.3.18}$$

we see that $\theta_i \geq 0$ for each $i = 1,2,3,4,5$ and $\theta_1 + \theta_2 + \theta_3 + \theta_4 + \theta_5 = 1$. For completeness, if $\mathscr{A} = \emptyset$ then we set $\Theta = 1$. Using these notations, we obtain our main result.

Theorem 2.7. *Consider the system of equations*

$$\boxed{\begin{cases}\dot{z}_i = \Theta z_i\left((\Lambda \mathbf{z})_i - \mathbf{z}^T \Lambda \mathbf{z}\right), \ i = 1, \cdots, N, \\ z_1 + z_2 + \cdots + z_N = 1.\end{cases}} \tag{2.3.19}$$

where Λ is the square matrix of size $N \times N$ whose coefficients $(i;j)$ are the pairwise invasion fitnesses λ_i^j which satisfy

$$\boxed{\lambda_i^j = \theta_1(b_i - b_j) + \theta_2(-\nu_i + \nu_j) + \theta_3(-u_{ij} - u_{ji} + 2u_{jj}) + \theta_4\left(\omega_{ij}^i - \omega_{ji}^j\right) + \theta_5(\mu(\alpha_{ji} - \alpha_{ij}) + \alpha_{ji} - \alpha_{jj}).} \tag{2.3.20}$$

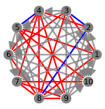
with $\mu = \dfrac{I^*}{D^*}$.
Then, for any initial values of (2.2.1)*, for each* $\tau_0 > 0$*,* $T > \tau_0$ *arbitrarily and independent on* ε*, there is* $\varepsilon_0 > 0$*,* $C > 0$ *and a vector of positive coefficients* $\mathbf{z}_0 \in \mathbb{R}^N$ *verifying* $\sum_{i=1}^N \mathbf{z}_{0,i} = 1$*, such that* $\forall \varepsilon < \varepsilon_0$

$$\left|S\left(\frac{\tau}{\varepsilon}\right) - S^*\right| + \sum_{i=1}^N \left|I_i\left(\frac{\tau}{\varepsilon}\right) - I^* z_i(\tau)\right| + \sum_{i,j=1}^N \left|I_{ij}\left(\frac{\tau}{\varepsilon}\right) - k\frac{I^*T^*}{S^*} z_i(\tau) z_j(\tau)\right| \leq \varepsilon C, \quad \forall \tau \in (\tau_0, T)\,. \tag{2.3.21}$$

where S*,* $(I_1, I_2, \ldots, I_N)$*,* $(I_{ij})_{i,j\in\{1,\ldots,N\}}$ *is the solution of* (2.2.1) *and* $(z_1, z_2, \ldots, z_N)$ *is the solution of reduced system* (2.3.19) *together with* $\mathbf{z}(0) = \mathbf{z}_0$.

This system (2.3.19) is a general replicator system, which is studied in [102].
We have two remarks on λ_i^j in (2.3.20). The first is that, each coefficient θ_i, $i \in \{1,2,3,4,5\}$ measures the weight of each trait perturbation on pairwise invasion fitness. Thus, each λ_i^j is a weighted average of the perturbations. Secondly, the pairwise invasion fitnesses play an important role in predicting collective dynamics, since λ_i^j is the pairwise invasion fitness between strains i and j, describing the quantitative initial growth rate of i invading an equilibrium set by j alone. In a 2-strain system, recall the final outcome results depend on the signs of the these mutual coefficients between the strains (Table 2.2), mentioned and used in [86, 88, 140].

Table 2.2 **From 2-strain invasion dynamics to collective multi-strain dynamics.** Each pair of strains in the system falls in one of 4 classes, according to λ_1^2 and λ_2^1 in (2.3.19): either competitive exclusion of 1, competitive exclusion of 2, coexistence, or bistability. The N-strain mutual invasion network drives competitive dynamics over long time.

Mutual invasion $\left(\lambda_1^2, \lambda_2^1\right)$	Pairwise Outcome	N-strain network	Strain freq.
$(+,+)$	Stable coexistence		$\dot{z}_i = \Theta z_i \left((\Lambda \mathbf{z})_i - \mathbf{z}^T \Lambda \mathbf{z}\right)$ $i = 1 \ldots N$
$(+,-)$	Exclusion of type 1		
$(-,+)$	Exclusion of type 2		
$(-,-)$	Bistability		
$\lambda_i^j = \theta_1 (b_i - b_j) + \theta_2 (-\nu_i + \nu_j) + \theta_3 (-u_{ij} - u_{ji} + 2u_{jj}) + \theta_4 \left(\omega_{ij}^i - \omega_{ji}^j\right) + \theta_5 (\mu (\alpha_{ji} - \alpha_{ij}) + \alpha_{ji} - \alpha_{jj})$			

In the next section, we present explicitly all the necessary computations and we also prove the lemma for the error estimate Lemma 2.5.

2.4 Proofs and explicit computations

Initially, let us recall the following definitions.

- S: total proportion of susceptible hosts
- T: the total proportion of infected hosts (prevalence of colonization)
- I_i: the proportion of hosts singly-colonized by strain-i
- I_{ij}: the proportion of hosts co-colonized by strain-i then strain-j (Including I_{ii}).

2.4.1 $\mathscr{A} = \{1\}$. Perturbations only in transmission rates β_i

Here we compute the functions $f^1_{z_i}$. In (2.3.7), take $\varepsilon = 0$, $\chi_1 = 1$ and $\chi_d = 0$ for $d > 1$. It comes

$$\begin{cases} \dfrac{dX}{dt} &= -\beta T^* X + F^1_X(\mathbf{v},\mathbf{z}) \\ \dfrac{dY}{dt} &= \beta(S^* - T^* - kI^*)X - (m+\beta k T^*)Y + F^1_Y(\mathbf{v},\mathbf{z}) \\ \dfrac{dL_i}{dt} &= -mL_i \\ \dfrac{dv_i}{dt} &= -\xi v_i \\ \dfrac{dz_i}{dt} &= 0 \end{cases} \tag{2.4.1}$$

Following the notation of the lemma 2.4, we obtain that the solution $(X,Y,\mathbf{L},\mathbf{v},\mathbf{z})$ of (2.4.1) with the initial condition $(X,Y,\mathbf{L},\mathbf{v},\mathbf{z})(0) = (X_0,Y_0,\mathbf{L}_0,\mathbf{v}_0,\mathbf{z}_0) \in \mathbb{R}\times\mathbb{R}\times(\mathbb{R}^n)^3$ verifies

$$\lim_{t\to+\infty}(X,Y,\mathbf{L},\mathbf{v})(t) = (X^*(\mathbf{z}_0),Y^*(\mathbf{z}_0),0,0)\,.$$

for some functions $X^*(\mathbf{z})$ and $Y^*(\mathbf{z})$ which remains to be compute.
Replacing $\mathbf{L}$ and $\mathbf{v}$ by 0 in the two first equation of (2.4.1) yields

$$\begin{cases} \dfrac{dX}{dt} &= -\beta T^* X + F^1_X(0,\mathbf{z}) \\ \dfrac{dY}{dt} &= \beta(S^* - T^* - kI^*)X - (m+\beta k T^*)Y + F^1_Y(0,\mathbf{z}) \end{cases} \tag{2.4.2}$$

Note that $\mathbf{v} = 0$ implies that the change of variables (2.2.13) reads simply

$$I_i = I^* z_i, \quad J_i = T^* z_i.$$

The quantities $F^1_X(0,\mathbf{z})$ and $F^1_Y(0,\mathbf{z})$ are then easily deducting from (2.3.4)

$$F^1_X(0,\mathbf{z}) = \beta S^* T^* \sum_{i=1}^N b_j z_j, \qquad F^1_Y(0,\mathbf{z}) = \beta\left(S^* - kI^*\right)T^* \sum_{i=1}^N b_j z_j. \tag{2.4.3}$$

Plugging this in (2.4.2), we obtain

$$X^*(\mathbf{z}) = S^* \sum_{i=1}^N b_j z_j$$

and then

$$Y^*(\mathbf{z}) = \beta\cdot\frac{S^{*2} - kI^*S^* - kI^*T^*}{m+\beta kT^*}\sum_{i=1}^N b_j z_j = \beta\cdot\frac{S^{*2} - kI^*}{m+\beta kT^*}\sum_{i=1}^N b_j z_j.$$

Now, (2.3.8) with the notation of the lemma 2.6 gives

$$\begin{aligned} f^1_{z_i}(\mathbf{z}) = &-\frac{\beta}{|P|}\left(kI^{*2} - T^*D^* + 2T^{*2}\right)S^*\sum_{i=1}^N b_j z_j + \frac{\beta}{|P|}\frac{\beta k T^{*2}}{m+\beta kT^*}\left(S^{*2} - kI^*\right)\sum_{i=1}^N b_j z_j \\ &-\frac{\beta}{|P|}\left(kI^{*2}T^*\sum_{i=1}^N b_j z_j + b_i D^* T^* S^* - 2b_i T^{*2} S^* - k b_i T^{*2} I^*\right). \end{aligned} \tag{2.4.4}$$

Denote

$$\begin{aligned} G &= -D^*T^*S^* + 2T^{*2}S^* + kT^{*2}I^* \\ H &= -\left(kI^{*2} - T^*D^* + 2T^{*2}\right)S^* + D^*\left(S^{*2} - kI^*\right) - kI^{*2}T^* \end{aligned} \tag{2.4.5}$$

then $G = -H = 2T^{*2}S^* > 0$, by straightforward computations. Setting $\Theta_1 = \dfrac{2\beta T^{*2}S^*}{|P|} > 0$, we have

$$f^1_{z_i}(\mathbf{z}) = \Theta_1\left(b_i - \sum_{j=1}^{N} b_j z_j\right). \tag{2.4.6}$$

It follows that the slow system (2.3.12) reads

$$\boxed{\frac{dz_i}{d\tau} = \Theta_1 z_i\left(b_i - \sum_{j=1}^{N} b_j z_j\right), \qquad 1 \le i \le N.} \tag{2.4.7}$$

Now we will show the simple computations showing that this system is exactly on the form of the replicator equation (2.3.19). It is clear that the set $\left\{\mathbf{z} \in [0,1]^N,\ \sum_{i=1}^N z_i = 1\right\}$, is conserved for (2.4.7). Hence, (2.4.6) may be rewrite as

$$f^1_{z_i}(\mathbf{z}) = \Theta_1\left(\sum_{j=1}^{N}(b_i - b_j)z_j\right). \tag{2.4.8}$$

Denoting pairwise invasion fitness between strains i and j, i invading in an equilibrium set by j, $\lambda_i^j = (b_i - b_j)$ and $\Lambda = (\lambda_i^j)$, we have

$$f^1_{z_i}(\mathbf{z}) = \Theta_1\left(\Lambda\mathbf{z}\right)_i. \tag{2.4.9}$$

Finally, from $\Lambda^T = -\Lambda$ we see that $\mathbf{z}^T\Lambda\mathbf{z} = 0$ which leads to the (artificial) representation of (2.4.7) :

$$\boxed{\frac{dz_i}{d\tau} = \Theta_1 z_i\left(\left(\Lambda\mathbf{z}\right)_i - \mathbf{z}^T\Lambda\mathbf{z}\right), \qquad 1 \le i \le N.} \tag{2.4.10}$$

which is nothing but the slow system (2.3.19) with $\lambda_i^j = b_i - b_j$.

2.4.2 $\mathscr{A} = \{2\}$. Perturbations only in clearance rates of single colonization γ_i

Similarly to the case $\mathscr{A} = \{1\}$, we compute the functions $f^2_{z_i}$. In (2.3.7), take $\varepsilon = 0$, $\chi_2 = 1$ and $\chi_d = 0$ for $d \neq 2$. It comes

$$\begin{cases} \dfrac{dX}{dt} &= -\beta T^* X + F^2_X(\mathbf{v},\mathbf{z}) \\ \dfrac{dY}{dt} &= \beta(S^* - T^* - kI^*)X - (m + \beta k T^*)Y + F^2_Y(\mathbf{v},\mathbf{z}) \\ \dfrac{dL_i}{dt} &= -mL_i \\ \dfrac{dv_i}{dt} &= -\xi v_i \\ \dfrac{dz_i}{dt} &= 0 \end{cases} \tag{2.4.11}$$

Following the notation of the lemma 2.6, we obtain that the solution $(X,Y,\mathbf{L},\mathbf{v},\mathbf{z})$ of (2.4.11) with the initial condition $(X,Y,\mathbf{L},\mathbf{v},\mathbf{z})(0) = (X_0,Y_0,\mathbf{L}_0,\mathbf{v}_0,\mathbf{z}_0) \in \mathbb{R}\times\mathbb{R}\times(\mathbb{R}^n)^3$ verifies

$$\lim_{t\to+\infty}(X,Y,\mathbf{L},\mathbf{v})(t)=(X^*(\mathbf{z}_0),Y^*(\mathbf{z}_0),0,0)\,.$$

for some functions $X^*(\mathbf{z})$ and $Y^*(\mathbf{z})$ which remains to be compute.
Replacing $\mathbf{L}$, $\mathbf{K}$ and $\mathbf{v}$ by 0 in the two first equation of (2.4.11) yields

$$\begin{cases}\dfrac{dX}{dt} &= -\beta T^*X+F_X^2(0,\mathbf{z})\\ \dfrac{dY}{dt} &= \beta(S^*-T^*-kI^*)X-(m+\beta kT^*)Y+F_Y^2(0,\mathbf{z})\end{cases} \tag{2.4.12}$$

The quantities $F_X^2(0,\mathbf{z})$ and $F_Y^2(0,\mathbf{z})$ are then easily deducting from (2.3.4)

$$F_X^2(0,\mathbf{z})=-\gamma I^*\sum_{i=1}^N \nu_i z_i, \qquad F_Y^2(0,\mathbf{z})=-\gamma I^*\sum_{i=1}^N \nu_i z_i. \tag{2.4.13}$$

Plugging this in (2.4.12), we obtain

$$X^*(\mathbf{z})=-\frac{\gamma I^*}{\beta^2 T^*}\sum_{i=1}^N \nu_i z_i$$

and then

$$Y^*(\mathbf{z})=\frac{\gamma I^*(kI^*-S^*)}{T^*(m+\beta kT^*)}\sum_{i=1}^N \nu_i z_i.$$

Now, (2.3.8) with the notation of the lemma 2.6 gives

$$f_{z_i}^2(\mathbf{z})=\frac{\gamma}{|P|}\left[-\nu_i I^*(I^*+T^*)+\frac{I^*(kI^*T^*-D^*+2T^*)}{T^*}\sum_{j=1}^N \nu_j z_j\right]. \tag{2.4.14}$$

By straightforward computations we can verify that

$$(-kI^*T^*+D^*-2T^*)+T^*(I^*+T^*)=0. \tag{2.4.15}$$

Setting $\Theta_2=\dfrac{\gamma I^*(I^*+T^*)}{|P|}>0$, we have

$$f_{z_i}^2(\mathbf{z})=\Theta_2\left(-\nu_i+\sum_{j=1}^N \nu_j z_j\right). \tag{2.4.16}$$

It follows that the slow system (2.3.12) reads

$$\boxed{\frac{dz_i}{d\tau}=\Theta_2 z_i\left(-\nu_i+\sum_{j=1}^N \nu_j z_j\right), \qquad 1\le i\le N.} \tag{2.4.17}$$

By the same arguments in section 2.4.1, we can show the simple computations showing that this system is exactly on the form of the replicator equation (2.3.19). Denoting the pairwise invasion fitness

$$\lambda_i^j=(-\nu_i+\nu_j)$$

and $\Lambda = (\lambda_i^j)$, we have

$$f_{z_i}^2(\mathbf{z}) = \Theta_2 \left(\Lambda \mathbf{z}\right)_i . \tag{2.4.18}$$

Finally, from $\Lambda^T = -\Lambda$ we see that $\mathbf{z}^T \Lambda \mathbf{z} = 0$ which leads to the (artificial) representation of (2.4.17) :

$$\boxed{\frac{dz_i}{d\tau} = \Theta_2 z_i \left(\left(\Lambda \mathbf{z}\right)_i - \mathbf{z}^T \Lambda \mathbf{z} \right), \qquad 1 \leq i \leq N.} \tag{2.4.19}$$

which is nothing but slow system (2.3.19) with $\lambda_i^j = -\nu_i + \nu_j$.

2.4.3 $\mathscr{A} = \{3\}$. Perturbations only in clearance rates of co-colonization γ_{ij}

Similarly to the case $\mathscr{A} = \{1\}$, we compute the functions $f_{z_i}^3$. In (2.3.7), take $\varepsilon = 0$, $\chi_3 = 1$ and $\chi_d = 0$ for $d \neq 3$. It comes

$$\begin{cases} \dfrac{dX}{dt} &= -\beta T^* X + F_X^3 \left(\mathbf{L}\right) \\ \dfrac{dY}{dt} &= \beta (S^* - T^* - kI^*) X - (m + \beta k T^*) Y \\ \dfrac{dL_i}{dt} &= -mL_i + F_{L_i} \left(\mathbf{v}, \mathbf{z}\right) \\ \dfrac{dv_i}{dt} &= -\xi v_i \\ \dfrac{dz_i}{dt} &= 0 \end{cases} \tag{2.4.20}$$

Following the notation of the lemma 2.6, we obtain that the solution $(X, Y, \mathbf{L}, \mathbf{v}, \mathbf{z})$ of (2.4.20) with the initial condition $(X, Y, \mathbf{L}, \mathbf{v}, \mathbf{z})(0) = (X_0, Y_0, \mathbf{L}_0, \mathbf{v}_0, \mathbf{z}_0) \in \mathbb{R} \times \mathbb{R} \times (\mathbb{R}^n)^3$ verifies

$$\lim_{t \to +\infty} (X, Y, \mathbf{L}, \mathbf{v})(t) = (X^*(\mathbf{z}_0), Y^*(\mathbf{z}_0), \mathbf{L}^*(\mathbf{z}_0), 0, 0) .$$

for some functions $X^*(\mathbf{z})$, $Y^*(\mathbf{z})$ and $\mathbf{L}^*(\mathbf{z}_0)$ which remains to be compute.
Replacing $\mathbf{v}$ by 0 in the two first equation of (2.4.20) yields

$$\begin{cases} \dfrac{dX}{dt} &= -\beta T^* X + F_X^3 \left(\mathbf{L}\right) \\ \dfrac{dY}{dt} &= \beta (S^* - T^* - kI^*) X - (m + \beta k T^*) Y \\ \dfrac{dL_i}{dt} &= -mL_i + F_{L_i} \left(0, \mathbf{z}\right) \end{cases} \tag{2.4.21}$$

The quantities $F_{L_i} \left(0, \mathbf{z}\right)$ and $F_X^3 \left(\mathbf{L}\right)$ are then easily deducting from (2.3.4).

$$F_X^3 \left(\mathbf{L}\right) = -\gamma \sum_{i=1}^N L_i, \qquad F_{L_i} \left(0, \mathbf{z}\right) = \frac{1}{2} \beta k I^* T^* z_i \sum_{j=1}^N \left(u_{ij} + u_{ji}\right) z_j. \tag{2.4.22}$$

Plugging this in (2.4.21), we obtain

$$L_i{}^*(\mathbf{z}) = \frac{1}{2m} \beta k I^* T^* z_i \sum_{j=1}^N \left(u_{ij} + u_{ji}\right) z_j, \tag{2.4.23}$$

then we deduce that

$$X^*(\mathbf{z}) = -\frac{\gamma k I^*}{2m} \sum_{i,j=1}^{N} (u_{ij} + u_{ji}) z_i z_j$$

and

$$Y^*(\mathbf{z}) = -\frac{\beta\gamma k I^*(S^* - T^* - kI^*)}{2m(m+\beta k T^*)} \sum_{i,j=1}^{N} (u_{ij} + u_{ji}) z_i z_j.$$

Now, (2.3.8) with the notation of the lemma 2.6 gives

$$f^3_{z_i}(\mathbf{z}) = \frac{\gamma}{|P|} \left[\frac{\beta k I^* T^{*2}}{m} \sum_{j,l=1}^{N} (u_{jl} - u_{lj}) z_l z_j + \frac{\beta k I^* T^{*2}}{m} \sum_{j=1}^{N} (u_{ij} + u_{ji}) z_j \right]. \tag{2.4.24}$$

It's trivial to see that $\dfrac{\beta k I^* T^{*2}}{m} = T^* D^*$. Setting $\Theta_3 = \dfrac{\gamma T^* D^*}{|P|} > 0$, we have

$$f^3_{z_i}(\mathbf{z}) = \Theta_3 \left[-\sum_{j=1}^{N} (u_{ij} + u_{ji}) z_j + \sum_{j,l=1}^{N} (u_{jl} + u_{lj}) z_l z_j \right]. \tag{2.4.25}$$

It follows that the slow system (2.3.12) reads

$$\boxed{\frac{dz_i}{d\tau} = \Theta_3 z_i \left[-\sum_{j=1}^{N} (u_{ij} + u_{ji}) z_j + \sum_{j,l=1}^{N} (u_{jl} + u_{lj}) z_l z_j \right], \qquad 1 \le i \le N.} \tag{2.4.26}$$

By the same arguments in section 2.4.1, we can show the simple computations showing that this system is exactly on the form of the replicator equation (2.3.19). Denoting pairwise invasion fitness

$$\lambda_i^j = -u_{ij} - u_{ji} + 2u_{jj}$$

and $\Lambda = (\lambda_i^j)$, we have

$$f^3_{z_i}(\mathbf{z}) = \Theta_3 \left((\Lambda \mathbf{z})_i - \mathbf{z}^T \Lambda \mathbf{z} \right). \tag{2.4.27}$$

Finally, we see the (artificial) representation of (2.4.26) :

$$\boxed{\frac{dz_i}{d\tau} = \Theta_3 z_i \left((\Lambda \mathbf{z})_i - \mathbf{z}^T \Lambda \mathbf{z} \right), \qquad 1 \le i \le N.} \tag{2.4.28}$$

which is nothing but replicator system (2.3.19) with $\lambda_i^j = -u_{ij} - u_{ji} + 2u_{jj}$.

2.4.4 $\mathscr{A} = \{4\}$. Perturbations only in transmission coefficients from mixed co-colonization p^i_{ij}

Similarly to the case $\mathscr{A} = \{1\}$, we compute the functions $f^4_{z_i}$. In (2.3.7), take $\varepsilon = 0$, $\chi_4 = 1$ and $\chi_d = 0$ for $d \neq 4$. It comes

$$\begin{cases} \dfrac{dX}{dt} &= -\beta T^* X \\ \dfrac{dY}{dt} &= \beta(S^* - T^* - kI^*)X - (m + \beta k T^*)Y \\ \dfrac{dL_i}{dt} &= -mL_i \\ \dfrac{dv_i}{dt} &= -\xi v_i \\ \dfrac{dz_i}{dt} &= 0 \end{cases} \tag{2.4.29}$$

Following the notation of the lemma 2.6, we obtain that the solution $(X, Y, \mathbf{L}, \mathbf{v}, \mathbf{z})$ of (2.4.29) with the initial condition $(X, Y, \mathbf{L}, \mathbf{v}, \mathbf{z})(0) = (X_0, Y_0, \mathbf{L}_0, \mathbf{v}_0, \mathbf{z}_0) \in \mathbb{R} \times \mathbb{R} \times (\mathbb{R}^n)^3$ verifies

$$\lim_{t \to +\infty} (X, Y, \mathbf{L}, \mathbf{v})(t) = (X^*(\mathbf{z}_0), Y^*(\mathbf{z}_0), 0, 0)\,.$$

for some functions $X^*(\mathbf{z})$ and $Y^*(\mathbf{z})$ which remains to be compute.
The two first equation of (2.4.29) reads

$$\begin{cases} \dfrac{dX}{dt} &= -\beta T^* X \\ \dfrac{dY}{dt} &= \beta(S^* - T^* - kI^*)X - (m + \beta k T^*)Y \end{cases} \tag{2.4.30}$$

So $X^*(\mathbf{z}_0) = 0$ and ,$Y^*(\mathbf{z}_0) = 0$. Now, (2.3.8) with the notation of the lemma 2.6 gives

$$f^4_{z_i}(\mathbf{z}) = \frac{1}{|P|}\left[2(T^*)^2 I^* \beta k \sum_{j=1}^{N} (\omega^i_{ij} + \omega^i_{ji}) z_j\right]. \tag{2.4.31}$$

Note that $2\beta k T^{*2} I^* = 2mT^* D^*$. From $\omega^i_{ji} = -\omega^j_{ji}$ we see that the slow system (2.3.12) reads

$$\boxed{\frac{dz_i}{d\tau} = \Theta_4 z_i \sum_{j=1}^{N} \left(\omega^i_{ij} - \omega^j_{ji}\right) z_j, \qquad 1 \leq i \leq N} \tag{2.4.32}$$

with $\Theta_4 = \dfrac{2mT^* D^*}{|P|}$.
Denote the $N \times N$ matrix $\Omega = (\omega^i_{ij})_{i,j}$ and $\Lambda = \Omega - \Omega^T$. We may rewrite this equation as

$$\boxed{\frac{dz_i}{d\tau} = \Theta_4 z_i \left((\Lambda \mathbf{z})_i\right), \qquad 1 \leq i \leq N.} \tag{2.4.33}$$

Finally, noting that $\Lambda = -\Lambda^T$ is skew symmetric, we have $\mathbf{z}^T\Lambda\mathbf{z} = 0$ so the slow equation reads

$$\boxed{\frac{dz_i}{d\tau} = \Theta_4 z_i\left((\Lambda\mathbf{z})_i - \mathbf{z}^T\Lambda\mathbf{z}\right), \qquad 1 \le i \le N.} \tag{2.4.34}$$

which is nothing but (2.3.19) with $\lambda_i^{\ j} = \omega_{ij}^i - \omega_{ji}^j$.

Remark that, this system leads to family of closed trajectories of an odd number $\tilde{N}$ of persistent strains but it is structurally unstable (except if $\tilde{N} = 1$), see [47]. Hence, in this case $\mathscr{A} = \{4\}$ we need to compute the term in ε^2 in the expansion, which we do not do in this text. However, when there are perturbations in other terms then the deviation in this trait conducts to interesting non trivial dynamics, which is shown in sections 2.5.3 and 2.5.3.2. This is similar to the case of large μ with perturbation in co-colonization interaction factor k_{ij}, i.e. $\mathscr{A} = \{5\}$, see [88]. We find that for $\mu \to 0$ and random α_{ij}, we have a case of Generalized Lotka-Volterra (GLV) dynamics with constant growth rates and random interactions. Meanwhile, if $\mu \gg 1$, dynamics converge to hyper-tournament dynamics studied by [10] for anti-symmetric matrix of interaction W with $\mathsf{W}_{ij} = \pm 1$; and by [95] for the case in which all the eigenvalues of $\mathsf{W} + \mathsf{W}^T$ are negative.

2.4.5 Proof of lemma 2.5 of error estimate

Lemma 2.8. The solution $(z_i)_{i=1,\dots,N}$ of the slow-fast form system (2.3.9) tends to the solution of the slow system (2.3.12) as $\varepsilon \to 0$ locally uniformly in time on $[\tau_0, T]$, with $\tau_0 > 0$, $T > \tau_0$ arbitrarily and independent on ε.

Proof. It suffices to verify the conditions for Tikhonov's theorem, see Theorem (2.2).
• Firstly, we prove that (2.3.9) with initial values possesses the unique solution.
The system (2.3.9) with initial values can be rewritten into

$$\frac{dx}{d\tau} = f(x), \qquad x(0) = x_0, \tag{2.4.35}$$

where $x = (X, Y, \mathbf{L}, \mathbf{v}, \mathbf{z})$, then $x(\tau) \in \mathbb{R}^{3N+2}$. We note that the function f of (2.4.35) is a vector function with all the components are polynomial of variables $(X, Y, \mathbf{L}, \mathbf{v}, \mathbf{z})$ (explicitly computed in sections 2.2.4, 2.4.3 and 2.4.4)and we work in the bounded set $[0, T]$ of time where all the functions $(X, Y, \mathbf{L}, \mathbf{v}, \mathbf{z})$ are differentiable. Hence, f is global Lipschitz and the uniqueness of solution for (2.3.9) follows, according to the Picard-Lindelof Theorem, see Theorem 2.2 in [194].
Implement analogously for (2.3.12), we acquire the same conclusion for the uniqueness of solution.

• Secondly, by the proof of lemma 2.4, we have that the solution $(X, Y, \mathbf{L}, \mathbf{v}, \mathbf{z})$ of (2.3.7) with any initial condition

$$(X, Y, \mathbf{L}, \mathbf{v}, \mathbf{z})(0) = (X_0, Y_0, \mathbf{L}_0, \mathbf{v}_0, \mathbf{z}_0) \in \mathbb{R} \times \mathbb{R} \times (\mathbb{R}^n)^3$$

verifies $\mathbf{z}(t) = \mathbf{z}_0$ for all $t \ge 0$ and

$$\lim_{t \to +\infty} (X, Y, \mathbf{L}, \mathbf{v})(t) = \Phi(\mathbf{z}_0)$$

asymptotically, in which, $\Phi(\mathbf{z}) = (X^*(\mathbf{z}), Y^*(\mathbf{z}), \chi_3 \mathbf{L}^*(\mathbf{z}), 0)$ satisfy the system (2.3.9) in slow timescale, with $\varepsilon = 0$ as follows

$$\begin{cases} 0 &= -\beta T^* X + \chi_1 F_X^1(\mathbf{v},\mathbf{z}) + \chi_2 F_X^2(\mathbf{v},\mathbf{z}) + \chi_3 F_X^3(\mathbf{L}) + O(\varepsilon) \\ 0 &= \beta(S^* - T^* - kI^*)X - (m + \beta k T^*)Y + \chi_1 F_Y^1(\mathbf{v},\mathbf{z}) + \chi_2 F_Y^2(\mathbf{v},\mathbf{z}) + \chi_5 F_Y^5(\mathbf{v},\mathbf{z}) \\ 0 &= -mL_i + \chi_3 F_{L_i}(\mathbf{v},\mathbf{z}) \\ 0 &= -\xi v_i \end{cases} \tag{2.4.36}$$

Applying Tikhonov's Theorem, we have the required conclusion. □

Let us now approximate the solution of the original dynamics (2.2.1) using the solution of slow-fast form 2.3.9, when ε is small enough.

Lemma 2.9. Under our assumptions, for any initial values of (2.2.1), there exists $\tau_0 > 0$ and initial value $\mathbf{z}(\tau_0)$ of (2.3.9), such that for any $T > \tau_0$, there are $\varepsilon_0 > 0$ and $C_T > 0$ satisfies $\forall \varepsilon < \varepsilon_0$

$$\left| S\left(\frac{\tau}{\varepsilon}\right) - S^* \right| + \sum_{i=1}^{N} \left| I^* z_i(\tau) - I_i\left(\frac{\tau}{\varepsilon}\right) \right| + \sum_{i=1}^{N} \left| T^* z_i(\tau) - J_i\left(\frac{\tau}{\varepsilon}\right) \right| \leq \varepsilon C_T, \tag{2.4.37}$$

for all $\tau_0 \leq \tau \leq T$, where $(S, I_i, J_i)_{i=1,\dots,N}$ verifies (2.2.1) and $(z_1, \dots, z_N)$ is the solution of (2.3.9).

Proof. To prove this lemma, we make two steps, one is to prove the error estimate between S^*, T^*, I^* and the solution (S,T,I) of (2.2.5), the other one is approximating the solutions of (2.2.5) using the solution of (2.3.9).

• *First step*, we wish to apply the Expansion Theorem 2.3. Note that, if (2.2.5) satisfies the conditions of Theorem 2.3 because of the property of global Lipschitz, then it will also fulfill the conditions of the Picard-Lindelof, see Theorem 2.2 in [194]. Thus, if that, for each initial value, (2.2.5) always has the unique solution. Therefore, it's suffices to verify the two conditions mentioned in Theorem 2.3, including the global Lipschitz properties.

Denote $x = (S, I_1, I_2, \dots, I_N, J_1, J_2, \dots, J_N)$. By the extract of (2.2.5) for S, I_i, J_i, $1 \leq i \leq N$, we write the system for (S, T, I_i, J_i), $i = 1, \dots, N$ in (2.2.5) into the following form

$$\frac{dx}{dt} = f_0(t,x) + \varepsilon f_1(t,x) \tag{2.4.38}$$

and in any bounded domain $|t - t_0| \leq h$ we have

1. $f_0(t,x)$ is continuous in t, continuously differentiable in x;
2. $f_1(t,x)$ continuous in t,x, Lipschitz-continuous in x.

According to this extraction, $f_0(t,x)$ and $f_1(t,x)$ are well-defined. Note that the function $f_0(t,x)$ has the $(f_S(t,x), f_{I_1}(t,x), \dots, f_{I_N}(t,x), f_{J_1}(t,x), \dots, f_{J_N}(t,x))$ for f_S, f_{I_i}, f_{J_i} are functions $\mathbb{R}^{2N+1} \to \mathbb{R}$, for all $1 \leq i \leq N$. The function $f_1(t,x)$ has the same form as well.
It's easy to see that $f_0(t,x)$ is continuous in t, continuously differentiable in x and the function $f_1(t,x)$ continuous in t,x. It remains to prove that $f_1(t,x)$ is Lipschitz-continuous in x in each bounded domain $|t - t_0| \leq h$, for all $h \in \mathbb{R}^+$. Indeed, $f_1(t,x)$ is a polynomial in multi variables (S,T,I_i,J_i), $i = 1, \dots, N$, and note that $S + T = 1$. In consequence, it is Lipschitz-continuous.

By the earlier arguments, if $x^r = (S^r, T^r, I_i^r, J_i^r)$ satisfies the neutral system (2.2.7) and $x = (S,T,I_i,J_i)_{1\le i\le N}$ satisfying (2.2.5) then $\|x - x^r\|_{\mathbb{R}^{2N+2}} = O(\varepsilon)$.

Therefore, note that $I = \sum_{i=1}^N I_i$, we deduce the solution of (2.2.5) can be approximated using neutral system. Combine with the arguments in section 2.2.3, the approximation of solution (S,T,I) of (2.2.5) by (S^*,T^*,I^*) is accordingly plausible in the sense of $O(\varepsilon)$. We have done our first step.

• *Second step*, we claim that all the algebraic and linear transformations from (2.2.5) to (2.3.9) are equivalent with error estimate $O(\varepsilon)$, including changing (S,T,I) to (X,Y) using S^*,T^*,I^* (proved in the first part), changing $\begin{pmatrix} I_i \\ J_i \end{pmatrix}$ to $\begin{pmatrix} v_i \\ z_i \end{pmatrix}$ (linear operator) and changing to time scale $\tau = \varepsilon t$ with re-denote $z(\tau)$ (see argument in (2.3.9)). We follow the steps of the preceding proof, that are verifying the conditions, and using Expansion Theorem 2.3 once again (note that $v(\tau) \to 0$ asymptotically), we have that

$$\sum_{i=1}^N \left| I^* z_i(\tau) - I_i\left(\frac{\tau}{\varepsilon}\right)\right| + \sum_{i=1}^N \left| T^* z_i(\tau) - J_i\left(\frac{\tau}{\varepsilon}\right)\right| = O(\varepsilon),$$

for all $\tau_0 \le \tau \le T$, where $(I_i,J_i)_{i=1,\dots,N}$ verify (2.2.1) and $(z_1,\dots,z_N)$ is the solution of (2.3.9).

Combining two parts, we have the conclusion for this lemma. □

By two lemmas 2.8 and 2.10, we have that

$$\left| S\left(\frac{\tau}{\varepsilon}\right) - S^*\right| + \sum_{i=1}^N \left| I^* z_i(\tau) - I_i\left(\frac{\tau}{\varepsilon}\right)\right| + \sum_{i=1}^N \left| T^* z_i(\tau) - J_i\left(\frac{\tau}{\varepsilon}\right)\right| \le \varepsilon C_T, \tag{2.4.39}$$

for all $\tau_0 \le \tau \le T$, where $(S,I_i,J_i)_{i=1,\dots,N}$ verifies (2.2.5) and $(z_1,\dots,z_N)$ is the solution of (2.3.12).

Finally, we will find an approximation of I_{ij}, $1 \le i \le N$ and estimate the error. Indeed, according to (2.4.39), we substitute $I_i(t)$ by $I^* z_i(\tau) + O(\varepsilon)$ and $J_j(t)$ by $T^* z_j(\tau)$ in all of the equations for $I_{ij}(t)$, $1 \le i,j \le N$ we have the equations

$$\frac{dI_{ij}(t)}{dt} = \beta_i k_{ij}(I^* z_i(\tau) + O(\varepsilon))(T^* z_j(\tau) + O(\varepsilon)) - m_{ij} I_{ij}(t), \qquad 1 \le i,j \le N, \tag{2.4.40}$$

which becomes

$$\frac{dI_{ij}(t)}{dt} = -mI_{ij} + \beta k I^* T^* z_i(\tau) z_j(\tau) + O(\varepsilon), \qquad 1 \le i,j \le N. \tag{2.4.41}$$

Now we formulate and prove the result for approximations of I_{ij}, $1 \le i,j \le N$, then deduce the approximation and error estimate for the whole initial system (2.2.1).

Lemma 2.10. Under our assumptions, for any initial values of (2.2.1), there exists $\tau_0 > 0$ and initial value $\mathbf{z}(\tau_0)$ of (2.3.12), such that for any $T > \tau_0$, there is $\varepsilon_0 > 0$ and $C_T > 0$ satisfies $\forall \varepsilon < \varepsilon_0$

$$\sum_{i,j=1}^N \left| I_{ij}\left(\frac{\tau}{\varepsilon}\right) - k\frac{I^* T^*}{S^*} z_i(\tau) z_j(\tau)\right| \le \varepsilon C_T, \tag{2.4.42}$$

for all $\tau_0 \leq \tau \leq T$, $(I_{ij})_{1\leq i,j\leq N}$ satisfying (2.2.1) and $(z_1,\ldots,z_N)$ is the solution of reduced system (2.3.12).

Proof. Assume $\left(I^r_{ij}\right)_{1\leq i,j\leq N}$ to be the solution of

$$\frac{dI_{ij}(t)}{dt} = -mI_{ij}(t) + \beta k I^* T^* z_i(\varepsilon t) z_j(\varepsilon t), \tag{2.4.43}$$

$1 \leq i,j \leq N$. Then, for each $\tau_0 > 0$ and $T > \tau_0$, after the changing time scale $\tau = \varepsilon t$, we have $\sum\limits_{i,j=1}^{N} \left| I_{ij}\left(\frac{\tau}{\varepsilon}\right) - I^r_{ij}\left(\frac{\tau}{\varepsilon}\right)\right| = O(\varepsilon)$ for any $\tau \in [\tau_0, T]$. Indeed, from (2.2.1) and (2.4.43), we have that

$$\begin{aligned} \frac{dI_{ij}}{dt}\left(\frac{\tau}{\varepsilon}\right) &= -m_{ij} I_{ij}\left(\frac{\tau}{\varepsilon}\right) + \beta_j k_{ij} I_i\left(\frac{\tau}{\varepsilon}\right) J_j\left(\frac{\tau}{\varepsilon}\right) \\ \frac{dI^r_{ij}}{dt}\left(\frac{\tau}{\varepsilon}\right) &= -mI^r_{ij}\left(\frac{\tau}{\varepsilon}\right) + \beta k I^* T^* z_i(\tau) z_j(\tau) \end{aligned} \tag{2.4.44}$$

which implies

$$\frac{d}{dt}\left(I_{ij}\left(\frac{\tau}{\varepsilon}\right) - I^r_{ij}\left(\frac{\tau}{\varepsilon}\right)\right) = -m\left(I_{ij}\left(\frac{\tau}{\varepsilon}\right) - I^r_{ij}\left(\frac{\tau}{\varepsilon}\right)\right) - \varepsilon\gamma u_{ij} I_{ij}\left(\frac{\tau}{\varepsilon}\right) + \left(\beta_j k_{ij} I_i\left(\frac{\tau}{\varepsilon}\right) J_j\left(\frac{\tau}{\varepsilon}\right) - \beta k I^* T^* z_i(\tau) z_j(\tau)\right). \tag{2.4.45}$$

By lemma 2.9, we have that $\left|\beta_j k_{ij} I_i\left(\frac{s}{\varepsilon}\right) J_j\left(\frac{s}{\varepsilon}\right) - \beta k I^* T^* z_i(s) z_j(s)\right| = O(\varepsilon)$ uniformly for $s \in [\tau_0, T]$. It is trivial to note that, since $\left|I_{ij}\right| \leq 1$, $\varepsilon\gamma u_{ij}\left|I_{ij}\left(\frac{\tau}{\varepsilon}\right)\right| = O(\varepsilon)$. Then, for all $1 \leq i,j \leq N$, using the expansion theorem- Theorem 2.3, we observe that

$$\left|I_{ij}\left(\frac{\tau}{\varepsilon}\right) - I^r_{ij}\left(\frac{\tau}{\varepsilon}\right)\right| = O(\varepsilon). \tag{2.4.46}$$

We then compute the solution $\left(I^r_{ij}\right)_{1\leq i,j\leq N}$ of (2.4.43) to be

$$I^r_{ij}(t) = e^{-mt}\left(\beta k I^* T^* \int_0^t e^{ms} z_i(\varepsilon s) z_j(\varepsilon s)\, ds + C\right), \qquad C \in \mathbb{R}. \tag{2.4.47}$$

For any fixed time T and $\tau_0 \leq t \leq T$, when $\varepsilon \to 0$ we can regard $z_i(\varepsilon t)$ invariant. Hence, for all $1 \leq i,j \leq N$, we have $\left|I^r_{ij}(t) - k\frac{I^* T^*}{S^*} z_i(\varepsilon t) z_j(\varepsilon t)\right| = O(\varepsilon)$, which implies $\left|I_{ij}\left(\frac{\tau}{\varepsilon}\right) - k\frac{I^* T^*}{S^*} z_i(\tau) z_j(\tau)\right| = O(\varepsilon)$. □

Combining Lemmas 2.8,2.9 and 2.10, we have the Lemma 2.5.

Thanks to this section, we now have the main result for the error estimate, that allows us to approximate the solution of the original system (2.2.1) using the solution of slow system (2.3.19). The original system (2.2.1) now formally reduces to the slow system (replicator system) (2.3.19), which leads to many advantages in analysis and prediction. The massive number of equations in (2.2.1) now decreases from N^2+N+1 to N equations of (2.3.19), which helps in computation and time. Thus, we may not need to compute the whole original model (2.2.1) to make prediction but only the replicator equations (2.3.19). The main result in section 2.3.2 also has biological meaning, when the coefficients of slow system (2.3.19) are pairwise invasion fitness, giving information about survival outcome of 2-strain system as in table 2.2. Furthermore, λ_i^j's give us the meaning and effects of each trait perturbation on the system and its long time behavior, which can not be seen directly in the (2.2.1).

2.5 Illustrations of the model and interpretations

In this section, we present some results and simulations about survival outcome of model based on the replicator system (2.3.19). Initially, we recall the definition of basic reproduction ratio of strain i that is the expected number of secondary cases produced by a single (typical) infection of strain i in a completely susceptible population and computed by $R_{0,i} = \dfrac{\beta_i}{m_i}$. If there is only variation in transmission rates among strains, then $R_{0,i}$'s fully determine the unique winner in the system. Yet, in cases of variation in transmission and clearance rates, it can be shown that $R_{0,i}$'s alone do not determine the survival outcome anymore because of the feedbacks induced by persistence in the coinfection compartment. These phenomena are illustrated in proofs and numerical simulations as follows in this section.

2.5.1 Competitive exclusion due to variation in transmission and infection clearance rates only, $\mathscr{A} = \{1\}$, $\mathscr{A} = \{2\}$, and $\mathscr{A} = \{1,2\}$

2.5.1.1 Variation in transmission rate or infection clearance rates $\mathscr{A} = \{1\}$ or $\mathscr{A} = \{2\}$

Now, we show the competitive exclusion principle in this case $C_{\mathscr{A}}$ with $\mathscr{A} = \{1\}$. In these cases, the competitive exclusion principle holds: the species with the largest $R_{0,i}$ is the only survivor.

Theorem 2.11. *Assume that $\mathscr{A} = \{1\}$ and $b_1 > b_2 \geq \cdots \geq b_N$. Then $E_1 = (1,0,\ldots,0)$ is globally stable in $(0,1)\times[0,1]^{N-1} \cap \{u \in \mathbb{R}^N : \sum_{i=1}^N z_i = 1\}$.*

This result means that, the strain with the largest basic reproduction number is the best competitor. However, in general, this fact does not always occurs, which we will illustrate in later subsection.

Proof. For simplicity, denote $D = (0,1)\times[0,1]^{n-1} \cap \{u \in \mathbb{R}^N : \sum_{i=1}^N z_i = 1\}$.
We aim to use LaSalle's invariant principle. Consider $V(u) = -\ln z_1$. Since we are consider the coexistence in D, then

$$\frac{dV(u)}{d\tau} = -\Theta_1\left(b_1 - \sum_{j=1}^N b_j z_j\right) = -\Theta_1\left(b_1\sum_{j=1}^N z_j - \sum_{j=1}^N b_j z_j\right) = -\Theta_1\sum_{j=1}^N (b_1 - b_j)z_j. \tag{2.5.1}$$

It's straightforward that $V(u) > 0$ because $0 < z_1 < 1$ in D. Because of the assumption $b_1 = \max\{b_i; 1 \leq i \leq N\}$ then $b_1 - b_j$ must be positive for all $j \neq 1$. Recall that $\Theta_1 > 0$ then, $\dfrac{dV(u)}{d\tau} \leq 0$. We have that

$$\frac{dV(u)}{d\tau} = 0 \Leftrightarrow (b_1 - b_j)z_j = 0, \quad \forall j \quad \Leftrightarrow \begin{cases} z_j &= 1, \quad j = 1 \\ z_j &= 0, \quad j \neq 1. \end{cases} \tag{2.5.2}$$

Thus, $V(u)$ is a Lyapunov function associated to $u(\tau)$. Applying LaSalle's invariant principle, we obtain our solution u tends to E_1 asymptotically. □

Analogously, we have a similar result for $\mathscr{A} = \{2\}$, that states that, the strain with smallest single infection clearance rate (longest duration of carriage) is the unique survivor.

Theorem 2.12. *Assume that $\mathscr{A} = \{2\}$ and $\nu_1 < \nu_2 \leq \cdots \leq \nu_N$. Then $E_1 = (1,0,\ldots,0)$ is globally stable in $(0,1)\times[0,1]^{N-1} \cap \{u \in \mathbb{R}^N : \sum_{i=1}^N z_i = 1\}$.*

The proof for this result uses the same argument in the theorem 2.11 so we do not present it.

2.5.1.2 Variation in transmission and single infection clearance rates, $\mathscr{A} = \{1,2\}$

In this subsection, it is shown that $R_{0,i}$'s do not determine the unique survivor anymore when $\mathscr{A} = \{1,2\}$ by constructing a counterexample. Firstly, we need an auxiliary lemma. With system $C_{\mathscr{A}}$ with $\mathscr{A} = \{1,2\}$, we try to make a result similar to Theorem 2.11 about the longtime scenarios of competition for

$$\begin{cases} \dfrac{dz_i}{d\tau} = \Theta_1 z_i \left(b_i - \sum\limits_{i=1}^{N} b_j z_j \right) + \Theta_2 z_i \left(-\nu_i + \sum\limits_{i=1}^{N} \nu_j z_j \right) \\ z_1 + z_2 + \cdots + z_N = 1. \end{cases} \tag{2.5.3}$$

Recalling that $\Theta_1, \Theta_2 > 0$ by definitions, we can prove the following theorem stating that the competitive exclusion occurs again but depends on the parameters of the neutral model though the quantity $\Theta_1 b_j - \Theta_2 \nu_j$ which characterizes the unique survivor.

Theorem 2.13. *Assume in* (2.5.3) *with N strains, there exists a strain, namely* 1*, satisfies* $\Theta_1 b_1 - \Theta_2 \nu_1 = \max\limits_{1 \le j \le N} \{\Theta_1 b_j - \Theta_2 \nu_j\}$*. Then* $E_1 = (1,0,\ldots,0)$ *is globally stable in* $D = (0,1) \times [0,1]^{N-1} \cap \{u \in \mathbb{R}^N : \sum_{i=1}^{N} z_i = 1\}$.

Proof. Analogously to the earlier result in section 2.5.1, we want to apply LaSalle's invariant principle. Consider the function $V(u) = -\ln z_1$ then by our hypothesis, it's easy to see that $V(u) > 0$ and

$$\frac{dV(u)}{d\tau} = -\sum_{j=1}^{N} (\Theta_1 b_1 - \Theta_2 \nu_1 - \Theta_1 b_j + \Theta_2 \nu_j) z_j \le 0. \tag{2.5.4}$$

Hence, $V(u)$ is an association Lyapunov function. The equation $\dfrac{dV(u)}{d\tau} = 0$ is equivalent to

$$\begin{cases} [(\Theta_1 b_i - \Theta_2 \nu_i) - (\Theta_1 b_j - \Theta_2 \nu_j)] z_j & = 0, \qquad 1 \le j \le N \\ z_1 + z_2 + \cdots + z_N & = 0 \end{cases} \tag{2.5.5}$$

which is equivalent to $(z_1, z_2, \ldots, z_N) = (1,0,\ldots,0)$. By LaSalle's invariant, E_1 is globally stable in D. □

We next come to see how this result is used in the forthcoming examples. We then compare the results with relations of $R_{0,i}$ to see how basic reproduction numbers affect the final competitive outcomes. Firstly, with the perturbations existing in clearance rates, the $R_{0,i}$ now becomes

$$R_{0,i} = \frac{\beta_i}{m_i} = \frac{\beta + \varepsilon b_i}{m + \varepsilon \nu_i} = \frac{\beta}{m}(1 + \frac{\varepsilon}{\beta} b_i) \left(1 - \frac{\varepsilon}{m} \nu_i\right) + O(\varepsilon^2), \tag{2.5.6}$$

which is equivalent to $R_{0,i} = \dfrac{\beta}{m} + \varepsilon \dfrac{\beta}{m} \left(\dfrac{b_i}{\beta} - \dfrac{\nu_i}{m} \right) + O(\varepsilon^2)$.

Hence, note that $R_0 = \dfrac{\beta}{m}$ we have that $R_{0,i} \le R_{0,j}$ if and only if $b_i - b_j \le R_0 (\nu_i - \nu_j)$ when $\varepsilon \to 0$.

Example 2.14. Consider the system (2.5.3).
Initially, we can directly apply Lemma 2.13 and infer that the strain, called 1, satisfying $\Theta_1 b_1 - \Theta_2 \nu_1 = \max\limits_{1 \le j \le N} \{\Theta_1 b_j - \Theta_2 \nu_j\}$ will be the winner.

Yet, unlikely to such result in section 2.5.1, according to the explicit calculation on $R_{0,i}$, we can construct so

that this strain 1 may not have the biggest basic reproduction number. Indeed, $\Theta_1 b_i - \Theta_2 \nu_i \geq \Theta_1 b_j - \Theta_2 \nu_j$ is equivalent to

$$b_i - b_j \geq \frac{1}{2}\frac{\gamma}{\beta} R_0 \frac{1}{1+k(R_0-1)} \left(\frac{1}{1+k(R_0-1)} + 1 \right) (\nu_i - \nu_j). \tag{2.5.7}$$

We can choose b_i, ν_i, $1 \leq i \leq n$ and $\gamma > 0$, $r > 0$, $k > 0$ and $R_0 > 1$ such that for $j \neq 1$,

$$\begin{cases} b_1 - b_j \geq \frac{1}{2}\frac{\gamma}{\beta} R_0 \frac{1}{1+k(R_0-1)} \left(\frac{1}{1+k(R_0-1)} + 1 \right) (\nu_1 - \nu_j), \\ b_1 - b_j \leq R_0 (\nu_1 - \nu_j), \end{cases} \tag{2.5.8}$$

then strain 1 has $\Theta_1 b_1 - \Theta_2 \nu_1 = \max\limits_{1\leq j\leq N} \{\Theta_1 b_j - \Theta_2 \nu_j\}$ and $R_{0,1} = \min\limits_{1\leq i\leq N} \{R_{0,i}\}$.
It is possible because

$$\frac{1}{2}\frac{\gamma}{\beta} R_0 \frac{1}{1+k(R_0-1)} \left(\frac{1}{1+k(R_0-1)} + 1 \right) < R_0$$

and we can pick, for instance, $\nu_1 = \max\limits_{1\leq j\leq N} \{\nu_j\}$, then easily find satisfactory b_i, ν_i.
This example shows us that, even a strain i with smallest basic reproduction number $R_{0,i}$ can be the single competitively exclusive strain if there is variation in both transmission and clearance rates in a system with co-infection. Explicitly, the strain 1 is the only survivor but it has the smallest R_0.

Hence, we can see that, even when there is competitive exclusion, $R_{0,i}$ alone still do not determine the winner if there are perturbations in the transmission rates and clearance rates. More detailed consideration of such effects and interplay between parameters for the 2-strain general system is provided in [119]. To close this subsection, we present simulations in figure 2.2 of competitive exclusion to illustrate claims in sections 2.5.1.1 and 2.5.1.2. We choose the 10-strain system and plot frequencies of strains in two cases: perturbation in only transmission rates β_i; and, perturbation in transmission rates β_i and in clearance rates of single colonization γ_i.

2.5.2 Variation in clearance rate of co-colonization may yield coexistence

2.5.2.1 Variation in clearance rate of co-colonization only, $\mathscr{A} = \{3\}$

In this case, the very first claim about competitive outcomes of the system is that, in contrast to the above cases $\mathscr{A} \subset \{1,2\}$, there can be coexistence of strains. Indeed, in this case the system can be rewritten on the form of a replicator system with a symmetric matrices for which several results exists (see in particular [102]). In particular we have :

Theorem 2.15. *Let $\mathscr{A} = \{3\}$ which means variation in coinfection clearance rates only. The system* (2.3.19) *may be rewritten as*

$$\begin{cases} \dot{z}_i = 2\Theta_3 \left(((-\bar{U})z)_i - z^T (-\bar{U}) z \right), \quad 1 \leq i \leq N \\ z_1 + \cdots + z_n = 1. \end{cases}$$

where the symmetric matrix $\bar{U} = \frac{1}{2}\left(U + U^T\right)$ is symmetric part of the matrix of perturbation $U = (u_{ij})_{1\leq i,j\leq N}$. In particular, the function $z \mapsto z^T \bar{U} z$ is a strict Lyapunov function and any positive asymptotic equilibria z^* are globally stable.

Proof. We refer here to the theorem 7.8.1 page 82 of [102] for the results about a replicator system with a symmetric matrix A. Then we only have to prove that the system (2.3.19) may be rewritten in terms of the symmetric matrix $-\bar{U}$.

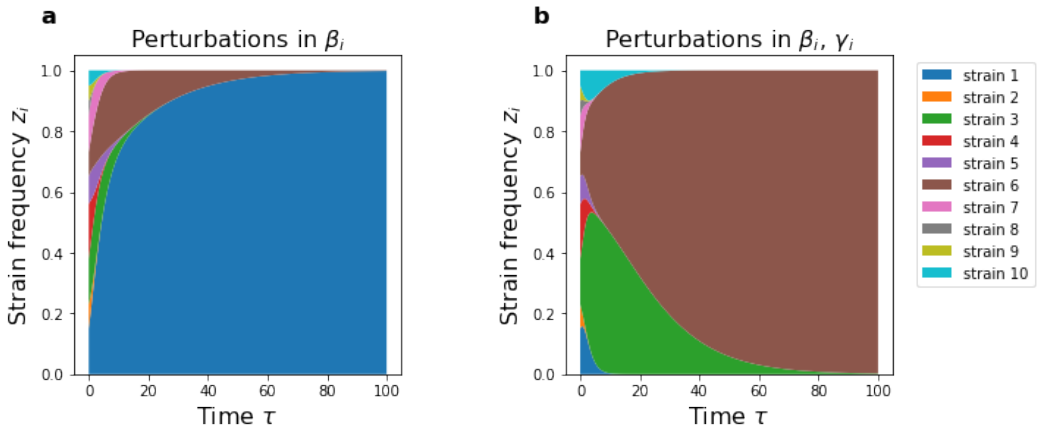


Fig. 2.2 **Illustration of competitive exclusion dynamics for $N = 10$ when strains vary in transmission and clearance rates.** We choose the parameter values of the neutral system $\beta = 4$, $m = 2$, $\gamma = 1$ and $k = 1.5$. The variation of β is given by $b = \begin{pmatrix} b_1 & b_2 & \dots & b_N \end{pmatrix}$ and is set to be the equal in both cases and equals $b = \begin{pmatrix} 0.25 & -0.2 & 0.125 & -0.125 & 0.075 & 0.225 & 0.05 & -0.5 & -0.175 & 0 \end{pmatrix}$. The matrix of ν_i in **(b)** is chosen to be $\nu = \begin{pmatrix} 1 & 0.8 & -1.5 & -0.5 & 0.3 & -1 & 1.2 & -2 & 0.7 & -2 \end{pmatrix}$. **(a)** Strains vary only in transmission rates β_i: $\mathscr{A} = \{1\}$. **(b)** Strains vary in transmission and clearance rates β_i, γ_i: $\mathscr{A} = \{1,2\}$. We can see that competitive exclusion is the only outcome in either case. However in **(a)** the strain with the highest reproduction number will persist while all other strains will go extinct. In contrast, in **(b)** the coinfection parameters matter, and it is not true that the strain with highest R_0 will persist. In this example strain 10 has highest basic reproduction number but strain 6 is the ultimate winner, because of its exact advantage in clearance rate (as explained in Example 2.14).(Data & Codes)

This comes from the following general fact in the replicator equation. Let $x = (x_j)_{1\le j\le N}$ be a vector and $A = (a_{ij})_{1\le i,j\le N}$ and $C = (c_{ij})_{1\le i,j\le N}$ be two $N\times N$ matrix such that $c_{ij} = a_{ij} + x_j$. For every $z = (z_1, \cdots, z_n)$ we have

$$(Cz)_k - z^T Cz = \sum_{j=1}^{N} c_{kj} z_j - \sum_{i,j} c_{ij} z_j z_i = \sum_{j=1}^{N} a_{kj} z_j + \sum_{j=1}^{N} x_j z_j - \sum_{i,j} a_{ij} z_j z_i - \sum_{i,j} x_j z_j z_i, \quad 1 \le k \le N.$$

If $\sum_{i=1}^{N} z_i = 1$ then $\sum_{i,j} x_j z_j z_i = \sum_{j=1}^{N} x_j z_j$ which yields

$$(Cz)_k - z^T Cz = (Az)_k - z^T Az, \quad 1 \le k \le N.$$

The proof follows from the explicit expression of (2.3.19) when $\mathscr{A} = \{3\}$ and by taking $a_{ij} = \frac{1}{2}\lambda_i^j = u_{jj} - \frac{1}{2}(u_{ij} + u_{ji})$, $x_j = -u_{jj}$ and $c_{ij} = -\frac{1}{2}(u_{ij} + u_{ji})$. □

Two important features of the dynamics in the case $\mathscr{A} = \{3\}$ are:

- Large possibilities of stable coexistence steady states.
- The parameters of the neutral models affect only the speed of the dynamics, given by Θ_3, but not the qualitative behavior. The latter depends only on the symmetric part of the perturbation $U = (u_{ij})$.

For an illustration of this case, we take the following example.
We consider a system of $N = 10$ strains with $\mathscr{A} = \{3\}$. In figure 2.3, we plot strains frequencies for multiple values of k showing that the *same* coexistence equilibrium of 3 strains is achieved with a speed dependent on k. Note that a similar effect would hold if we vary R_0.

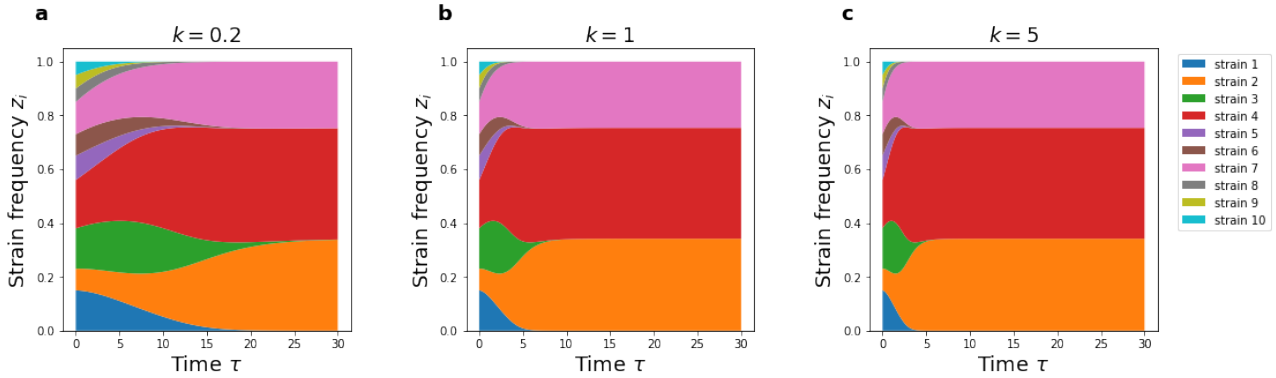


Fig. 2.3 **Strain coexistence is possible when there is variation in coinfection clearance rate and the speed of the dynamics depends on the parameters of the neutral model.** Here, we illustrate coexistence dynamics under the effect of k for $k = 0.2$ **(a)**, $k = 1$ **(b)** and $k = 5$ **(c)**. In the top sub-panels we show the dynamics of 10 strain frequencies. We choose $\beta = 4$, and basic reproduction number $R_0 = 2$. It can be seen that as k increases, the system tends to its stable state faster. In figures **(a, b, c)**, three strains 2, 4, 7 coexist after a long time. (Data & Codes)

We note that the speed of the dynamics is given by

$$\Theta_3 = \frac{\gamma T^* D^*}{|P|} = \frac{\gamma T^*}{2T^* + \frac{S^*}{k}\left(1 + \frac{m}{m+\beta k T^*}\right)},$$

which increases with k. Thus, in this case, increasing k only multiplies whole matrix Λ by a factor, which increases the speed of the convergence to the stable state of coexistence.

2.5.2.2 Variation in transmission and coinfection clearance rates, $\mathscr{A} = \{1,3\}$

When $\mathscr{A} \subset \{1,2,3\}$, the perturbations occur both on traits $\{1\}$ and $\{2\}$ leading on competitive exclusions and on $\{3\}$ leading on coexistence. Thus the relative weights of the perturbation, depending on the parameters on the neutral model, will affect the *qualitative* outcomes of the dynamics among strains.

Hence, unlike in section 2.5.2.1, the qualitative behavior does not depend only on the pertubations b_i, ν_i and u_{ij} but also on the values of the parameters of the neutral model. A precise generic result is out of the scope of this paper.
For simplicity, consider the case $\mathscr{A} = \{1,3\}$. From the formula $\lambda_i^j = \Theta_1(b_i - b_j) + \Theta_3(-u_{ij} - u_{ji} + 2u_{jj})$, we infer that the larger the ratio $\frac{\Theta_3}{\Theta_1}$ is the more chance a coexistence may happen. From

$$\frac{\Theta_3}{\Theta_1} = \frac{R_0}{2}\frac{\gamma}{\frac{m}{k} + \beta T^*},$$

we see in particular that this ratio increases with k. This is illustrated in the figure 2.4. We see in this figure that shifting k alters qualitatively the dynamics and the ultimate outcome among strains. In figure 2.4**(a)** $k = 0.1$ and the only winner is strain 8, whereas for $k = 1$, figure 2.4**(b)**, then the winners turn to strain 3 and 6. Finally, for $k = 3$, the outcome in figure 2.4 **(c)** is the coexistence of strains 2, 4 and 7.
Note that the short explanation above, gives only an overview of the phenomena and do not explain all the details. For instance, we observe that the set of coexistent species depends on the value of k in a complex maner.

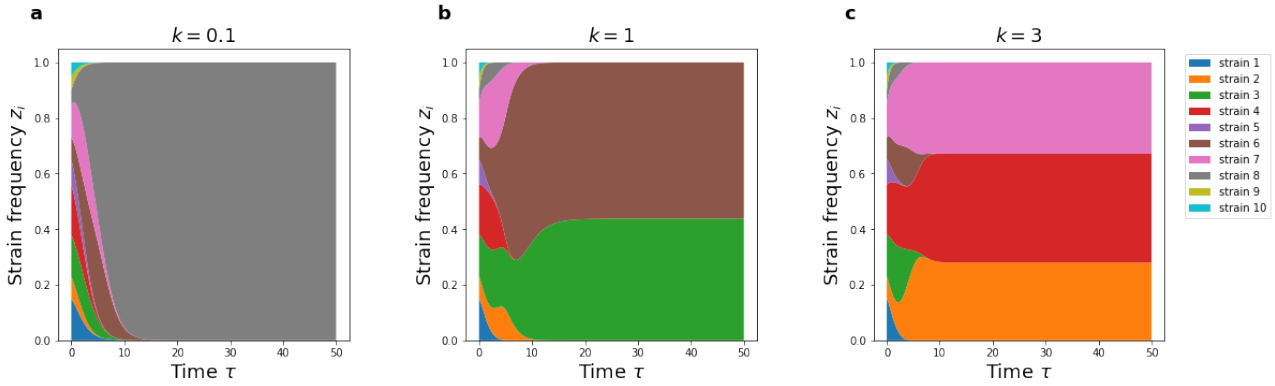


Fig. 2.4 **The final ecological outcome can shift with changing vulnerability to coinfection, when strains vary in transmission and coinfection clearance rates.** We illustrate coexistence dynamics for $k = 0.1$ **(a)**, $k = 1$ **(b)**, $k = 3$ **(c)**. In the top sub-panels we show the dynamics of 10 strain frequencies. We choose $\beta = 4$, $R_0 = 5$ and $\gamma = 0.5$. We keep the initial values in 2.3 and the matrix of value's b_i as follows, in which b_i is in cell i-th $\begin{pmatrix} 0 & -0.2 & 0.125 & -0.125 & 0.225 & 0.75 & 0.5 & 1.25 & -0.175 & 0 \end{pmatrix}$. We plot for multiple values of k respectively equal to 0.1, 1 and 3, to show effects of k to transient phenomena. It can be seen that as k increases, changes the survival strains. (Data & Codes)

2.5.3 Variation in transmission probability from mixed carriage may lead to cycles among strains.

In this subsections, we make simulations in which variation at least in transmission probability from mixed carriage, $4 \in \mathscr{A}$. Despite of the antisymmetric matrix of pairwise invasion fitness $\Lambda = \left(\lambda_i^j\right)_{i,j}$ as in cases $\mathscr{A} = \{1\}$ and $\mathscr{A} = \{2\}$, there are many long time behaviors that may occur in this case. In [47], one proves that there can be coexistence with higher possibility than competitive exclusive of one strain. However, when there are combinations with other trait perturbation, the outcome survival can shift due to neutral parameters, which will be presented in the next subsections 2.5.3.1 and 2.5.3.2.

2.5.3.1 Variation in transmission rates and transmission probability from mixed carriage, $\mathscr{A} = \{1,2,4\}$

We make simulations when perturbations in transmission rates β_i and transmission capacity of a strain by a host co-colonized. From (2.3.19) when $\mathscr{A} \subset \{1,2,4\}$, the equations for this case can be written as

$$\frac{dz}{d\tau} = z \cdot (\Lambda z)$$

where anti-symmetric matrix Λ is the invasion fitness matrix with

$$\Lambda_i^j = \Theta_1 (b_i - b_j) + \Theta_2 (\nu_j - \nu_i) + \Theta_4 \left(\omega_{ij}^i - \omega_{ji}^j\right).$$

this type of replicator equation is known as a zero sum games tournaments from which several results are known (see [47]). In particular the set E of persistent strains is unique, regardless the initial values, and the number of persistent strains is odd.

- If this odd number is 1, then the competitive exclusion principle occurs, as we saw above in the particular case $\emptyset \neq \mathscr{A} \subset \{1,2\}$.

- If this odd number is above 1, the system is conservative and structurally unstable: there is a family of cycles around a single steady states of these strains E. This is possible thus the effect of a perturbation in ω_{ij}^{i} (i.e. $4 \in \mathscr{A}$).

As in the section 2.5.2.2, the parameters of the neutral models affect the relative weight of the pertubation and therefore the type of outcomes.

In figure 2.5, we take $\mathscr{A} = \{1,4\}$. We have

$$\frac{\Theta_4}{\Theta_1} = \frac{k(R_0-1)}{1+k(R_0-1)} = \frac{1}{\mu+1}.$$

Hence, changing $\mu = \dfrac{1}{k(R_0-1)}$ shift the dynamics outcome. When $\mu = 0.6$, i.e. small enough, makes $\dfrac{\Theta_1}{\Theta_4}$ large yielding to a cycle of 3 persistent strains. Conversely, $\mu = 1.2$ large enough leads to the competitive exclusion.

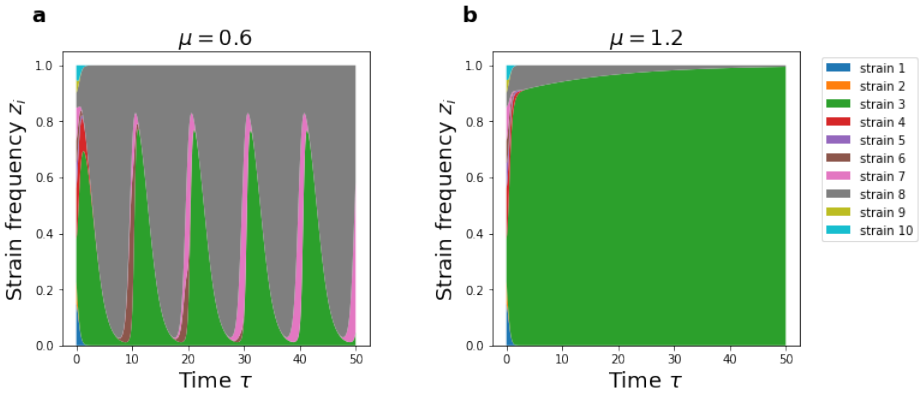


Fig. 2.5 **The long time behavior can shift with changing the co-infection prevalence $\left(\mu = \frac{I}{D}\right)$ rate, when strains vary in transmission rate and transmission coefficients from mixed carriage.** We illustrate co-existence dynamics for $\mu = 0.6$ **(a)** and $\mu = 1.2$ **(b)**. We choose $\beta = 3$, $\gamma = 1.2$, $R_0 = 2$ and $b = \begin{pmatrix} 0.3 & -0.8 & 2.4 & -0.5 & 0.9 & 2 & 1.2 & 1 & -0.7 & 0.5 \end{pmatrix}$. It can be seen in this case that an increase in μ (reducing co-infection prevalence), shifts the cycle of persistent strains in **(a)**, to the competitive exclusion of strain 3-with biggest transmission rate β_i in **(b)**. (Data & Codes)

2.5.3.2 Variation in coinfection clearance rates and transmission probability from mixed carriage, $\mathscr{A} = \{3,4\}$

When there are perturbations in coinfection clearance rates and transmission probability from mixed carriage, pairwise invasion fitness matrix Λ is not anti-symmetric anymore. The analysis of the sections 2.5.2.1 and 2.5.3.1 suggest that, depending on the ratio $\dfrac{\Theta_4}{\Theta_3}$, we may observe coexistence through stable steady states if $\dfrac{\Theta_4}{\Theta_3} \ll 1$ and through cycles if $\dfrac{\Theta_4}{\Theta_3} \gg 1$. We have the explicit formula

$$\frac{\Theta_4}{\Theta_3} = \frac{2m}{\gamma} = 2\left(1+\frac{r}{\gamma}\right),$$

then, depending on the values of r and γ, we can have other interesting phenomena.
We make simulations for two cases of r, susceptible host recruitment rate. When $r = 0.2$ small enough, we obtain the coexistence of 3 strains, that is structurally stable, although it oscillates in a first period of time. When $r = 3$ large enough, the coexistence of strains becomes structurally unstable. It can be seen that, the number of coexistent strains is 3, which is odd as mentioned.

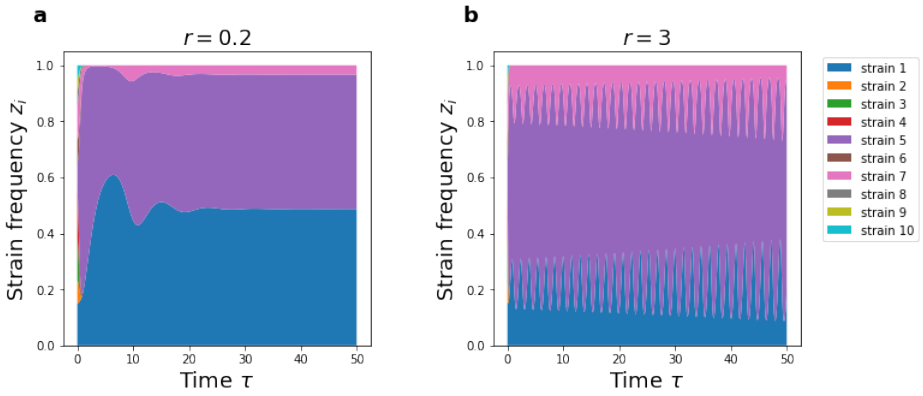


Fig. 2.6 **The long time behavior can shift with changing neutral transmission rate, when strains vary in transmission rate and transmission coefficients from mixed carriage.** We illustrate coexistence dynamics for $r = 0.2$ **(a)** and $r = 3$ **(b)**. We choose $k = 3$, $R_0 = 2$, $\gamma = 1$ and reuse the initial values in figure 2.3. As r increases, the stable state coexistence of 3 strains in **(a)**, shifts to the unstable trajectory of strains in **(b)**. (Data & Codes)

2.5.4 Summary of multi-strain outcomes by studying the slow system

In general, when there are many traits varying among similar strains, the long time behaviour may lead to complex outcomes. However, in cases of single trait perturbations only $\mathscr{A} = \{j\}$, the outcome is often easier to understand.

- If $j \in \{1,2,3,4\}$ (the cases explored within this section), then we can prove or refer to existing results to predict the dynamics. In particular, in these cases, the values of the pairwise invasion fitness Λ_{ij} do not depend on the parameters of the neutral system and then
  If $\mathscr{A} = \{j\}$ with $j \in \{1,2,3,4\}$ the qualitative outcomes do not depend on the parameters of the Neutral model.

- If $j = 5$ (perturbation in k_{ij} only) the outcomes are more complex, and an introduction to the phenomena is given in [88]. In particular, the pairwise invasion fitness reads $\lambda_i^j = \alpha_{jj} - \alpha_{ji} + \mu(\alpha_{ji} - \alpha_{ij})$ and depends on the parameter $\mu = \dfrac{I^*}{D^*} = \dfrac{1}{k(R_0 - 1)}$. It follows that
  if $A = \{5\}$ the qualitative outcome does not depend on the parameters of the Neutral model.

In the Table 2.3 we give a summary of the multi-strain system behavior when there are perturbations in only one trait.

In general, when there are perturbations in several traits, the qualitative outcomes result in a complex manner from each single case. The weight of each perturbation in the λ_i^j's, and thus on the qualitative dynamics, is governed exactly by the Θ_i's, which are explicit functions of the parameters of the neutral system. Hence, if

the ratio between the Θ_i is changing we may observe a change in the qualitative dynamics. This implies that a change in any of the global parameters of the neutral model (k, R_0, r, γ, β) may affect not only the speed of the dynamics, but also, and in a complex manner, the type of the dynamics.

Table 2.3 **Summary of multi-strain outcomes for each case of a single trait varying.** Note that all the Θ_i admits the same denominator $|P| = \left(2T^{*2} - I^*D^*\right) > 0$. Since, only the ratio between the Θ_i impact the qualitative behavior, we represent the values of $|P|\Theta_i$.

	Trait varying	**Formula of $\|P\|\Theta_i$**	**Type of dynamics**
1.	Transmission rates β_i	$2\beta S^* T^{*2}$	Competitive exclusion
2.	Single infection clearance rates γ_i	$\gamma I^* (I^* + T^*)$	Competitive exclusion
3.	Co-infection clearance rates γ_{ij}	$\gamma T^* D^*$	Possibility of Coexistence
4.	Transmission probability from mixed carriage p^s_{ij}	$2mT^* D^*$	Family of cycles
5.	Co-colonization interaction factor via altered susceptibilities, k_{ij}	$\beta T^* I^* D^*$	Anything

2.6 Concluding remarks

This mathematical study provides a fundamental advance in understanding analytically quasi-neutral dynamics between multiple strains in a co-infection system. Until now, explicit and general derivations of coinfection dynamics among N strains are very rare in the literature [4, 140]. Previous studies have considered $N = 2$, $N = 3$ or N-strain dynamics without coinfection, typically with variation in just one fitness dimension. Others have sketched the conceptual framework linking neutrality with non-neutral dynamics [128]. Here, we go beyond the state of the art, and provide a full analytical characterization of the coinfection dynamics among N strains that vary along multiple fitness dimensions, under the assumption that such variation is relatively small. We complete a series of studies based on slow-fast dynamics, made explicit, for linking neutral and non-neutral dynamics in interacting multi-strain pathogens [86, 88, 140].

Naturally in this endemic compartmental model, infectious strains compete for susceptible and singly-colonized hosts, which are the only resources that can favour their growth and propagation. The different traits provide each strain with variable fitness advantages or disadvantages in exploiting such dynamic resources in the system, and interact together to shape multi-strain selection. We establish some remarkable results by simplifying the dynamics when small perturbations arise in the clearance rates, transmission rates, within-host competitiveness coefficients, as well as co-colonization susceptibility interaction factors between strains. We derive the corresponding slow-fast form for the global dynamics, with the system of strain frequencies completely

explicit, and provide the formal approximation for solutions of all epidemiological variables by quantifying error estimates. We reduce the complexity of N^2+N+1 equations of the original SIS compartmental model to the N-equations of replicator dynamics, which reduces substantially time for computation.

Instead of studying concurrently all compartmental variables, our approach separately considers the neutral system and the perturbation components, then integrates them at the final stage. It would be possible to obtain a solution immediately for the whole emergence within perturbations in all traits. Nevertheless, such an undertaking in our view would involve many massive and complicated manipulations, and hence constitute a more difficult route than the one chosen here. This difficulty led us to the main lemma, Lemma 2.6. This result enables us to integrate all particular cases for the most general problem. It only leaves us concrete special cases, with the same structure, but simpler.

As a first step, we comprehend the neutral model and deduce the globally asymptotically stable state of variables (S,T,I) by (S^*,T^*,I^*), which give us a conservation law for global quantities in the co-infection system: susceptibles prevalence, total colonization and single colonization prevalence, reached in a fast time scale. The expansion theorem in [200] plays a role as the first step in bridging between neutrality and the slow-fast system. Thanks to new variables z_i, denoting strain frequencies, and the new time-scale $\tau=\varepsilon t$, understanding the emergent model now becomes an exploration of the so-called replicator system for $\{z_i\}_{1\leq i\leq N}$. This derivation makes sense, in light of Tikhonov's theorem. The model with perturbations is consequently well approximated by the slow-fast formulation, which helps us to explicitly demonstrate error estimates in term $O(\varepsilon)$ as well.

Concerning the system of strain frequencies, we find out and work in the invariant set $\{u\in\mathbb{R}^{N+}: z_1+\cdots+z_N=1\}$. In general, by interpreting pairwise invasion fitness numbers in each singular perturbation case, the closing equations at each section become special instances of the same replicator system of $\{z_i\}_{1\leq i\leq N}$. This enables us to study the relative dominance of strains, longtime scenarios of dynamics and other important properties of coexistence and competition. Such a replicator system always admits the competitive exclusion equilibria $E_i=(0,0,\ldots,1,\ldots,0)$. An essential and sufficient condition for the linear asymptotic stability of such trivial steady state E_i, is that the pairwise invasion fitness numbers λ_j^i (for all $j\neq i$) must be negative. Another remarkable sequel is that there is at most one $\mathscr{I}$-coexistence solution for any nonempty subset $\mathscr{I}$ given. Other well-known results for the replicator equation can be applied directly to our case, and linked with epidemiological processes and dynamics.

Whereas many deterministic multi-strain models rely on the presence of immune interactions to generate oscillations in multi-strain dynamics [62, 92, 73] typically focusing on a small number of strains (up to $N=4$), other factors such as natural heterogeneities or stochasticity have also been implicated in regular or chaotic oscillations in single-strain and ecological predator-prey systems [24, 147, 12], including spatial effects in multi-strain contexts [137]. However, as argued by [113], the precise assumptions that lead to oscillations in different strain models are yet to be established, and hence represent an area of constant investigation. In our study, we have contributed to close further the gap in this area, by providing a direct link between SIS multi-strain epidemiology with coinfection, and the replicator equation for N strains, where oscillatory regimes are naturally expected, and have a substantial history of study [102]. We know that in our model, for any parameter combinations that make the pairwise invasion matrix $\left(\lambda_i^j\right)$ approach the anti-symmetric numerical structure, multi-strain relative frequencies will tend to display oscillations similar to dynamics in zero-sum games ([89, 140]). Remarkably, with the λ_i^j entirely explicit in this model, one can envision several mechanisms and trait trade-offs by which oscillations, whether regular or irregular, are the rule and not the exception. Inevitably, this will require an in-depth investigation, links with existing models, and biological interpretation in the future.

It is exciting to envision how this approach could be extended to other epidemiological models of multi-strain dynamics. An essential requirement is that their embedded neutral system admits a central manifold which is globally stable. The challenge would then be to identify the equations governing slow motion on this manifold in each specific model. Until now we have not considered a spatial component to the multi-strain dynamics. A further perspective is considering space and a diffusion model for the replicator equation (e.g see [35]). Many more extensions and model applications to data in an explicit manner should be now within reach in the near future. As argued in [140, 89], this coinfection model and its dynamics could also be translated by analogy to other biological scales, e.g. the colonization dynamics of multi-species communities [10, 95] or gut microbiota within host [70], which would open new frontiers for application, interpretation and computational tool development.

Chapter 3

Disentangling two-strain system

This chapter 3 is the work in the paper "Disentangling how multiple traits drive 2 strain frequencies in SIS dynamics with coinfection" with Dr. Erida Gjini and Dr. Sten Madec, 2021, see [119].

3.1 Introduction

Epidemiological models of coinfection have a long history of study [122, 4, 165, 146, 199, 154, 145, 195, 6]. Examples of multi-strain infectious agents where coinfection processes appear and shape epidemiology include *Streptococcus pneumoniae* bacteria [127, 90], *Bordetella pertussis* [163], *Mycobacterium tuberculosis* [53], *Staphylococcus Aureus*, [173] and many others, comprising plants [192, 97], and also inter-species co-colonization such as between *Haemophilus Influenzae* and pneumococcus serotypes [144, 52] and coinfection with different viruses [81]. Typically the strain-defining parameters vary much less within than between species. However, until now, models have not leveraged the conceptual and analytic advantages of strain similarity to the full extent, except for the classical comparisons between strain-specific basic reproduction numbers.

While it has long been recognized that in coinfection systems, basic reproduction numbers alone do not determine strain competitive dynamics [165, 199], a generic framework to integrate variation among strains along several phenotypic axes and coinfection, and map these directly to strain frequencies, has not been developed. Moreover, until now simplified versions of traits involved in SIS dynamics between two strains have been addressed: either modeling vulnerability to coinfection as a single parameter [6], or focusing just on cross-strain competition [127], four-way competitive interactions via altered susceptibilities to coinfection [90], exclusively cooperative dynamics [49], or focusing on transmission and clearance rate variation [145, 195].

A key level of strain interactions in coinfection is the within-host level, where the order and timing of arrival, can matter for onward transmission or clearance. Such interactions when studied empirically have revealed strong priority effects, where the first arriving genotype has an advantage over later arriving ones [63, 97]. Independently of the underlying mechanisms, whether via host immunity, resource overlap within host or others, priority effects have repercussions on disease dynamics and parasite assemblage dynamics at higher scales. Yet, the full extent of the inter-dependence between this trait and other traits involved in epidemiological dynamics remains poorly understood. Thus, although several aspects of coinfection have been studied, typically with simulations or analyses restricted to special cases and particular models, a comprehensive and concise theoretical framework for how coinfection prevalence broadly interplays with multiple traits between 2 coinfecting agents in endemic systems is still missing.

In this paper, we describe and study a general system for epidemiological dynamics of similar co-infecting entities (e.g. discrete strains of the same infectious agent or similar species) that comprise a rich ecological and epidemiological phenomenology. Such a system could apply, but not be limited to, polymorphic *Streptococcus pneumonia* bacteria or other commensal bacteria [127, 51, 90, 61]. We generalize a previously- introduced quasi-neutral *Susceptible-Infected-Susceptible* (SIS) framework for 2 circulating strains and co-infection, where we showed that asymmetries in pairwise susceptibilities to co-infection create frequency-dependent advantage for one of the strains and can give rise to coexistence, bistability as well as competitive exclusion between strains [90, 86]. Here, we study additional variation between two microbial strains, namely, in other traits besides vulnerability to co-infection, including transmission rate and duration of carriage, two classical traits that are known to vary among colonizing pneumococcus serotypes [1], but in general can also vary among two arbitrary infectious agents, e.g. an antibiotic-resistant and an antibiotic-sensitive strain. Furthermore, we also allow for variation in transmission biases from co-infected hosts carrying a mixture of two strains, and in duration of coinfection episodes, where priority effects can play a role, adding new layers where competitive abilities and asymmetries can manifest. Until now analytic solutions for how such systems behave in time have not been obtained, although theoretical models have studied the conditions for coexistence vs. competitive exclusion or used numerical simulations as an approach in specific cases [127, 209, 90, 5, 98].

The novelty of our approach lies in applying singular perturbation theory to a quasi-neutral model, whereby we obtain a timescale separation, similar to [86], in order to express the total dynamics as a fast plus a slow component, related to broken symmetries along 5 traits between strains: transmission rate β_i, clearance rate γ_i, coinfection duration γ_{ij}, pairwise vulnerabilitites to coinfection k_{ij}, and transmission priority effects from coinfection p^i_{ij}. Net competition dynamics can be complex because all traits interact nonlinearly to determine final strain fitness at the host population level, but here we make such selective process entirely explicit. Moreover, we show how the strain-transcending parameters, defining the neutral model at the center, feed back on the strain dynamics on the slow timescale, and tune the net importance of each phenotypic axis.

The bigger N-strain SIS model with coinfection is derived in [120], where all the formal mathematical details leading to a replicator equation that governs multi-strain frequency dynamics are provided. Here, we apply this framework to the simpler $N = 2$ case, and study in depth the emergent ecology mediated by different epidemiological traits. While in the general $N-$ strain case analytical expressions for equilibrium frequencies can only be obtained in special cases, as shown by [140, 120], in the $N = 2$ system, considered here, equilibrium frequencies are entirely explicit. This allows us to provide exact analytical results on qualitative and quantitative shifts in system behavior (coexistence vs. exclusion) with coinfection prevalence across a wide spectrum of biological scenarios.

As in [86, 140], we assume only up to 2 strains may co-infect a host (MOI=2). We model how primary infection by one strain alters host susceptibility to secondary strains, (increasing or decreasing it) by a factor k_{ij}, relative to uninfected hosts, without acquired immunity. The altered susceptibilities to co-infection, given by a 2×2 matrix in the case of 2 strains, can comprise antagonistic or facilitative interactions ($k_{ij} < 1$ or ≥ 1). Beyond enabling competition and cooperation to be studied under the same framework, our model allows also for any asymmetries in this coinfection susceptibility matrix, as for the coinfection clearance rate matrix, depending on strain composition, and for transmission biases from coinfected hosts, depending on order of strain arrival.

Considering the complex epidemiological multi-strain dynamics in fast and slow components has many analytic and computational advantages as argued in [140, 120]. Our neutral model satisfies the criteria for ecological and population-genetic neutrality discussed in the context of 'no coexistence for free' [128], but much more than a neutral null model, our approach highlights the neutral model as the core organizing centre

of the multi-strain dynamics. This is made entirely explicit via the slow-fast timescale separation formalism [140, 120], linking neutral and non-neutral dynamics in an 'organic' manner, and showing how parameters of the neutral model impact significantly on the slow frequency-dependent selection dynamics between strains, and tune the ecological feedbacks among different traits.

While the multi-strain SIS model with coinfection that we analyze here is deterministic, it should lend itself easily to later implementations of stochasticity. Although stochasticity has been rarely studied for coinfection models, the role of stochasticity in epidemiological models in general has a long history of investigation, involving analyses of pathogen transmission in finite or structured populations [25, 20, 160, 109, 8], simulations of pathogen diversity [74], simulations of explicit host contact structure [100, 173], individual-based simulations [61], studies of the role of environmental fluctuations in global parameters and seasonality for disease dynamics [117, 179], up to statistical inference and parameter estimation [108]. While it is well-known that stochastic models can differ in outcomes and are more challenging to analyze than deterministic counterparts, having a solid analytical grip on the dynamic behavior of a deterministic coinfection model for multi-strain dynamics, should be considered a fundamental theoretical step for subsequent advances toward stochasticity.

Our paper is organized as follows. First, we describe the epidemiological framework. Secondly, we expose and elaborate on a closed and generic analytic solution for 2-strain frequency dynamics over the slow timescale, in a changing fitness landscape shaped by multi-trait variation. This solution coincides with a version of the classical replicator equation in 2 dimensions [103], but with an explicit payoff matrix derived from 5-dimensional trait variation between strains relatively weighted in the overall pairwise invasion fitness [120]. Third, we analyze why and how coexistence, bistability or competitive exclusion of either strain may occur between any two strains, for a fixed given trait variation between them. Fourth, we focus our attention on an in-depth analytic investigation of how strain-transcending mean-field gradients can shift the same system across these regimes, for different values of global R_0 or coinfection prevalence, detailing the context-dependence of net outcomes. Finally, we conclude with a roadmap for biological applications. We believe our analysis and approach offer a fresh perspective, to quantify and predict how multiple traits together shape coinfecting strain dynamics and final equilibria via joint and nonlinear population feedbacks.

3.2 The modeling framework

3.2.1 The SIS model with coinfection

We study an infectious agent transmitted in a host population following Susceptible-Infected-Susceptible dynamics, where there are two co-circulating strains (denoted by 1 and 2). Susceptible hosts S can acquire any strain i, by which they enter the single infection (colonization) compartment I_i. Singly-infected hosts I_i can acquire any secondary strain j, leading them to enter the coinfection (co-colonization) compartment I_{ij}. As in the pioneering model by [199], an important epidemiological feature here is that hosts can be coinfected twice by the same strain (I_{11} and I_{22} compartments). Without this assumption, a rare strain always has an advantage: it can infect hosts already infected by the common strain while the common strain has few hosts to coinfect. Co-colonized and singly-colonized hosts transmit at equal total rate, and hosts carrying a mixture of two different strains transmit any strain i with a given probability p^i_{ij} which can be different from 1/2 and may depend on the order of arrival within-host (e.g. $p^i_{ij} \neq p^i_{ji}$). The model follows the structure in [90, 86], but here we have a more general model, allowing for more trait variation between strains. In [86], only pairwise susceptibilities to co-infection were modeled as different between strains (2-by-2 matrix of k_{ij} coefficients), and this was sufficient to generate stabilizing mechanisms for coexistence. Here, in addition to k_{ij},

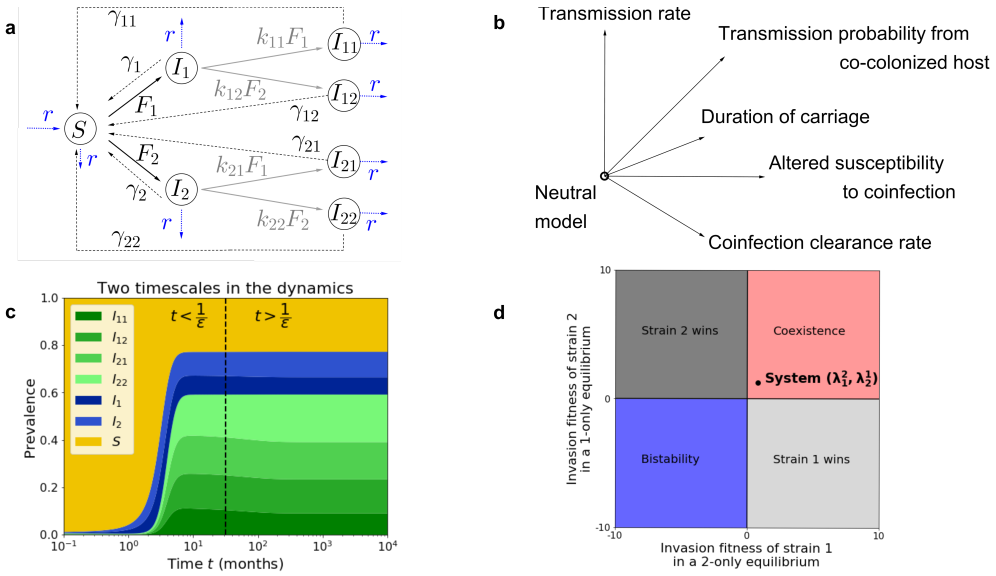


Fig. 3.1 **SIS coinfection model diagram for two strains ($N = 2$) and multivariate selection dynamics. a.** The model follows the structure in [86] but here two strains can differ in: transmission rate β_i, clearance rate γ_i, co-colonization clearance rate γ_{ij}, altered susceptibilities to co-colonization k_{ij} and transmission biases from coinfection p^i_{ij}. Thus any combination of relative fitness costs and advantages can be encapsulated, provided that their variation is not too big, as expected for similar conspecific strains, or similar infectious co-circulating 'species'. Non-carriers (S) become carriers of either strain 1 or 2 (I_i) with force of infection $F_i = \beta_i \left(I_i + I_{ii} + \sum_{ij} \left(p^i_{ij} I_{ij} + p^i_{ji} I_{ji} \right) \right)$ where the mixed carriage compartment (I_{ij}) may transmit either strain with a slightly biased probability away from 1/2 depending on the order of arrival (see [120]). Here $1/\gamma_i$ is the strain-specific duration of single colonization, $1/\gamma_{ij}$ are the composition-specific durations of co-colonization, which can vary for all four I_{ij} classes. The coefficients k_{ij} capture the altered relative susceptibilities to co-colonization between strains, when a host is already colonized, and transitions from primary colonization to co-colonization. The parameter r is the natural birth/death rate of the host. **b.** Assuming strain similarity, the epidemiological dynamics in such an SIS model with coinfection, can be decomposed into a fast (neutral) component and slow (non-neutral) component. The slow dynamics are shown to follow an explicit replicator equation which includes in the net payoff matrix variation across 5 dimensions of fitness for each strain [120]. This equation allows to predict analytically the entire temporal dynamics of two strains as a function of their epidemiological phenotypes. **c.** We simulate an example of 2-strain system in two timescales. On the fast time-scale ($o(1/\varepsilon)$), strains follow neutral dynamics, driven by mean-field parameters, where total prevalence of susceptibles, single infection and co-infection stabilize. On a slow time-scale, εt, within conserved global epidemiological compartments, complex non-neutral dynamics between strains takes place, depicted here by the blue and green shadings. **d.** Each system can be in one of four scenarios between 2 strains, depending on the signs of mutual invasion fitnesses (e.g. dynamics in **c** corresponds to the black point in the coexistence region). We find that frequency dynamics are explicitly governed by the λ^j_i. In our model, invasion fitnesses are explicit functions of strain variability along different traits and global mean-field parameters.

we model strain-specific transmission rate, β_i, and clearance rate γ_i, as well as coinfection clearance rates γ_{ij} and transmission biases from coinfection p^i_{ij}, depending on strain composition in coinfection. Thus we describe two types of priority effects: at the between-host level in terms of the k_{ij}, and at the within-host level in terms of the p^i_{ij}. Recruitment of susceptibles happens at per-capita rate r, assumed equal to the natural mortality rate. The scheme of the model for two strains is given in Figure 3.1. The explicit dynamical system of equations for

the N-strain version of this epidemiological model is derived in [120], and given by:

$$\begin{cases} \dfrac{dS}{dt} & = r(1-S) + \displaystyle\sum_{i=1}^{N} \gamma_i I_i + \sum_{i,j=1}^{N} \gamma_{ij} I_{ij} - S \sum_{i=1}^{N} F_i, \\ \dfrac{dI_i}{dt} & = F_i S - (r+\gamma_i) I_i - I_i \displaystyle\sum_{j=1}^{N} k_{ij} F_j, \qquad 1 \le i \le N, \\ \dfrac{dI_{ij}}{dt} & = k_{ij} I_i F_j - (r+\gamma_{ij}) I_{ij}, \qquad 1 \le i,j \le N, \end{cases} \tag{3.2.1}$$

where in our case for $N = 2$, $i, j \in \{1,2\}$, and the force of infection for each strain is given by:

$$F_i = \beta_i \left(I_i + \sum_{j=1}^{N} \left(p_{ij}^i I_{ij} + p_{ji}^i I_{ji} \right) \right), \tag{3.2.2}$$

3.2.2 The similarity assumption

Some conventions and notations for the parameters in a multiple-trait model, under the similarity assumption between strains, are given in Table 3.1, where strain-specific parameters are defined in terms of relative variation from a common reference. For example, strain-specific transmission rates are described as $\beta_i = \beta(1+\varepsilon b_i)$. In particular, in this approach ε gives the scale of the perturbation around neutrality, which must be relatively small for the quasi-neutral approximation to hold, whereas the magnitudes of the perturbations in different traits are captured by $\Delta b = b_i - b_j$, etc. which may be large. Furthermore, it is important to note that there is not a unique representation in terms of '*scale* $\times$ *magnitude*' of perturbations for a particular system. For example there are many possibilities preserving $\varepsilon \Delta b = \frac{\beta_1 - \beta_2}{\beta}$, for fixed β_1, β_2 close to each other. In our framework, the magnitudes and directions of the perturbations in transmission rates, given by b_1, b_2 provide a description of how far proportionally from the common reference β are the two respective transmission rates, when measured in the scale of ε. Since we have many traits varying simultaneously, the choice of the appropriate ε must be convenient in order to describe all parameters under a common scale. In practice, capturing multivariate dissimilarity between two strains under the same scale, will involve an intermediate choice of ε balancing the requirement of relatively small for the approximation to hold, and sufficient for the slow dynamics to be of significant ecological/observational relevance. However, mathematically speaking, for the slow-fast method to be applied, the only requirement is that ε has to be as small as needed and that the perturbations b_i, ν_i etc. are fixed.

To obtain the fast-slow decomposition, we rewrite the system 2.2.1 in terms of new aggregate variables as in [140], such as the total prevalence of colonized hosts T, the total prevalence of hosts transmitting either strain J_i, the total prevalence of single colonization I, and co-colonization D:

$$T = \sum_{i=1}^{2} I_i + \sum_{i,j=1}^{2} I_{ij}, \quad J_i = I_i + \sum_{j=1}^{2} \left(p_{ij}^i I_{ij} + p_{ji}^i I_{ji} \right), \quad I = \sum_{i=1}^{N} I_i, \quad D = T - I. \tag{3.2.3}$$

Following similar technical steps as in [140], and applying Tikhonov's theorem [196], we have derived for the N strain model in [120] that during the fast timescale, strains behave as neutral (all parameters are identical between them) and each global aggregated variable tends to its equilibrium: $S \to S^*$, $T \to T^*$, $I \to I^*$ and $D \to D^*$ as $\varepsilon \to 0$. With the basic reproduction number in this system denoted $R_0 = \dfrac{\beta}{m}$, this equilibrium is

Table 3.1 **Conventions and notations of parameters and variables, where we assume the traits are numerically close for closely-related strains, or similar infectious entities.** This similarity assumption ($0 < \varepsilon \ll 1$, small) forms the basis for dynamic decomposition into fast and slow components [120].

Parameter	Interpretation	Specification	Features
Original system (Quasi-neutral)			
$\beta_i = \beta\,(1+\varepsilon b_i)$	Strain-specific transmission rates	$\Delta b = b_1 - b_2$	Favours 1 if $\Delta b > 0$
$\gamma_i = \gamma(1+\varepsilon \nu_i)$	Strain-specific clearance rates of single colonization	$\Delta \nu = \nu_2 - \nu_1$	Favours 1 if $\Delta \nu > 0$
$\gamma_{ij} = \gamma(1+\varepsilon u_{ij})$	Clearance rates of co-colonization with i and j	$\Delta_i u = 2u_{jj} - (u_{ij} + u_{ji})$	Favours i if $\Delta_i u > 0$
$p_{ij}^s = \frac{1}{2} + \varepsilon \omega_{ij}^s$	Transmission probability [1] of $s \in \{i, j\}$; from a host co-colonized by strain-i then -j (priority effects).	$\Delta\omega = \omega_{12}^1 - \omega_{21}^2$	Favours 1 if $\Delta\omega > 0$
$k_{ij} = k + \varepsilon \alpha_{ij}$	Relative factor of altered susceptibility to co-colonization between colonizing strain i and co-colonizing strain j	$\Delta_i \alpha = \alpha_{ji} - \alpha_{jj} + \mu(\alpha_{ji} - \alpha_{ij})$	Favours i if $\Delta_i \alpha > 0$
Embedded neutral system (Fast):			
r	Susceptible recruitment rate	Equal to natural mortality	Host turnover $\uparrow$
β	Transmission rate (infectiousness)	$\beta > 0$	Transmission $\uparrow$
γ	Clearance rate of single infection	Equal to clearance rate of co-infection, $\gamma > 0$	Transmission $\downarrow$
m	Net infected host turnover rate	$m = r + \gamma$	$R_0 \downarrow$
R_0	Basic reproduction number	$R_0 = \beta/m > 1$	Colonization $\uparrow$
k	Altered susceptibility to co-infection when infected	$k > 0$	$k \geq 1$: facilitation, $k < 1$: competition
μ	Single to co-infection ratio	$\mu = \frac{1}{(R_0-1)k}$	Monotone in R_0 and k
Non-neutral system (Slow):			
θ_1	Weight of transmission rate axis β_i	$(\mu+1)^2$	Depends on μ
θ_2	Weight of clearance rate axis γ_i	$\frac{\gamma}{2(\gamma+r)}\left(2\mu^2+\mu\right)$	Depends on μ, γ, r
θ_3	Weight of co-infection clearance rate axis γ_{ij}	$\frac{\gamma}{2(\gamma+r)}\left(\mu+1\right)$	Depends on μ, γ, r
θ_4	Weight of transmission priority effects from co-infection axis p_{ij}^i	$\mu + 1$	Depends on μ
θ_5	Weight of susceptibilities to co-infection axis k_{ij}	$\frac{1}{2k}$	Depends on k

given by:

$$\boxed{S^* = \frac{1}{R_0}, \qquad T^* = 1 - S^*, \qquad I^* = \frac{T^*}{1 + k(R_0 - 1)}, \qquad D^* = T^* - I^*.} \tag{3.2.4}$$

and the ratio of single infection to co-infection is given by $\mu = \dfrac{I^*}{D^*} = \dfrac{1}{k(R_0-1)}$.

Further, during the slow time scale εt, strains are not equivalent, their differences in fitness start to get manifested, and what follows is non-neutral dynamics at the level of strain frequencies z_i. For a 2-strain system, there are four possibilities for the equilibrium: i) coexistence, ii-iii) exclusion of each strain, and iv) bistability of competitive exclusion states, also known as a priority effect. Below, these outcomes and their dynamics are shown to depend explicitly on mutual invasion growth rates between two strains.

3.2.3 Pairwise invasion fitness and replicator dynamics via timescale separation

In this more complex model, we follow the same reasoning as in [140], focusing on mutual invasion fitnesses, to express the selective dynamics occurring on the slow time scale. We will define λ_i^j to be invasion fitness of strain i in an equilibrium set by strain j alone, a classical approach in adaptive dynamics [85]. Initially, based on [120] and Chapter 3 equation (2.3.17), we redefine θ_i via the global quantities and parameters of the neutral model:

$$\theta_1 = \left(\frac{T^*}{D^*}\right)^2, \quad \theta_2 = \frac{\gamma I^*(I^*+T^*)}{2mD^{*2}}, \quad \theta_3 = \frac{\gamma T^*}{2mD^*}, \quad \theta_4 = \frac{T^*}{D^*}, \quad \theta_5 = \frac{\beta I^* T^*}{2mD^*}. \tag{3.2.5}$$

As derived in detail in [120] and Chapter 3 equation (2.3.20), we have that, in our model with multi-trait variation between strains, for $i, j \in \{1,2\}$, the mutual invasion fitnesses are given by:

$$\boxed{\lambda_i^j = \theta_1(b_i - b_j) + \theta_2(-\nu_i + \nu_j) + \theta_3(-u_{ij} - u_{ji} + 2u_{jj}) + \theta_4\left(\omega_{ij}^i - \omega_{ji}^j\right) + \theta_5(\mu(\alpha_{ji} - \alpha_{ij}) + \alpha_{ji} - \alpha_{jj})}. \tag{3.2.6}$$

This analytic expression sums the relative contributions of multiple trait variations at the same time, with the weighting constants θ_i defined above and given explicitly in Table 3.1. By the notations of λ_i^j, setting $\Lambda = \left(\lambda_i^j\right)_{i,j\in\{1,2\}}$, system (3.2.1) on the slow timescale εt, can be approximated by the replicator equation

$$\boxed{\begin{cases} \dfrac{dz}{d\tau} = \Theta z\left(\Lambda z - z^T \Lambda z\right) \\ z_1 + z_2 + \cdots + z_N = 1 \end{cases}} \tag{3.2.7}$$

for variables $z = (z_1, z_2)$ denoting strain frequencies, where the overall speed of dynamics Θ is given by:

$$\Theta = \frac{2mD^{*2}}{2T^{*2} - I^*D^*} = \frac{2m}{2\mu^2 + 3\mu + 2}. \tag{3.2.8}$$

When there is just variation in co-infection susceptibility coefficients k_{ij}, we recover the model and the Θ in [140]. Further, recall in [140] the term $z^T\Lambda z$ is denoted as Q and referred to as mean invasibility of the system, capturing resistance of the system to invasion by outsiders [89]. In the 2-strain system, more explicitly we have: $Q = z^T\Lambda z = (\lambda_1^2 + \lambda_2^1)z_1z_2$, and this quantity is positive only in the case of coexistence between two strains.

In this slow-fast derivation, epidemiological variables of the original model (system 2.2.1) are then a function of strain frequencies of the slow system:

$$\boxed{I_i(\tau) = I^* z_i(\tau) \quad \text{and} \quad I_{ij}(\tau) = D^* z_i(\tau) z_j(\tau),} \tag{3.2.9}$$

where I^* and D^* give the overall prevalence of single infection and co-infection in the endemic system (neutral model), and z_i and $z_i z_j$ give the proportions occupied by strain i and the pair of strains i and j, among singly infected hosts and co-infected hosts respectively.

3.3 General outcomes of the 2-strain system

3.3.1 Equilibria of the system

Denote by $z^* = (z_1^*, z_2^*)$ the nonzero equilibrium state of (3.2.7), where strain frequencies are given by:

$$(z_1^*, z_2^*) = \left(\frac{\lambda_1^2}{\lambda_1^2 + \lambda_2^1}, \frac{\lambda_2^1}{\lambda_1^2 + \lambda_2^1} \right).$$

Depending on the signs of both invasion growth rates, we therefore have conditions for λ_1^2 and λ_2^1, leading to four ecological scenarios between two strains as in [140] (see Table 3.2). Thus, to investigate the equilibria

Table 3.2 **System equilibria for 2-strain dynamics according to λ_1^2 and λ_2^1** as expected from the replicator equation (3.2.7)

Mutual invasion $(\lambda_1^2, \lambda_2^1)$	Outcome	Strain frequencies	Quadratic form $Q = z^T \Lambda z$
$(+,+)$	Stable coexistence	$z_1^* > 0$, $z_2^* > 0$: $\frac{1}{\lambda_1^2+\lambda_2^1}\begin{pmatrix}\lambda_1^2\\ \lambda_2^1\end{pmatrix}$	$Q > 0$ and $Q \to Q^*$
$(-,+)$	Exclusion of type 1	$z_1^* = 0$, $z_2^* = 1$	$Q \to 0$
$(+,-)$	Exclusion of type 2	$z_1^* = 1$, $z_2^* = 0$	$Q \to 0$
$(-,-)$	Bistability	$z_1^* = 1$,or $z_2^* = 1$	$Q < 0$ and $Q \to 0$

and their stability in this 2-strain system, it suffices to study the values and signs of pairwise invasion fitness coefficients $\left(\lambda_1^2, \lambda_2^1\right)$, given explicitly as follows:

$$\begin{aligned}
\lambda_1^2 &= \theta_1 (b_1 - b_2) + \theta_2 (\nu_2 - \nu_1) + \theta_3 (-u_{12} - u_{21} + 2u_{22}) + \theta_4 \left(\omega_{12}^1 - \omega_{21}^2\right) + \theta_5 (\mu (\alpha_{21} - \alpha_{12}) + \alpha_{21} - \alpha_{22}) \\
\lambda_2^1 &= -\theta_1 (b_1 - b_2) - \theta_2 (\nu_2 - \nu_1) + \theta_3 (-u_{21} - u_{12} + 2u_{11}) - \theta_4 \left(\omega_{12}^1 - \omega_{21}^2\right) + \theta_5 (\mu (\alpha_{12} - \alpha_{21}) + \alpha_{12} - \alpha_{11}),
\end{aligned} \tag{3.3.1}$$

while their sum is $\lambda_1^2 + \lambda_2^1 = 2\theta_3 (u_{11} + u_{22} - u_{12} - u_{21}) + \theta_5 (\alpha_{12} + \alpha_{21} - \alpha_{11} - \alpha_{22})$. It's not easy to determine the exact long time scenario or the winner in a two-strain system because parameters with perturbations affect all together the dynamics. The table 3.2 gives us criteria to determine the long time behavior. However, instead of computing the fitness coefficients explicitly in each case, we can base on (3.3.1) to determine quickly and for a various range of cases. Inspecting closely equations (3.3.1), we can see two parts:

1. The part $\theta_1 \Delta b + \theta_2 \Delta \nu + \theta_4 \Delta \omega + \theta_5 \mu (\alpha_{21} - \alpha_{12})$, which keeps the $\left(\lambda_2^1, \lambda_1^2\right)$ close to the line $\lambda_1^2 + \lambda_2^1 = 0$, i.e. the dynamics tend to exclusion of one strain.

2. The part $\theta_3 \Delta_2 u + \theta_5 (\alpha_{21} - \alpha_{22})$, which pulls $\left(\lambda_2^1, \lambda_1^2\right)$ away from the line $\lambda_1^2 + \lambda_2^1 = 0$, i.e. driving the dynamics toward coexistence or bi-stability.

This makes it easy to see that variation in transmissibility (Δb) or duration of infection between strains ($\Delta \nu$), and the precedence effect in transmission from mixed coinfection ($\Delta \omega$), always promotes competitive exclusion

in the system, whereas variation in coinfection parameters (susceptibilities and clearance rates) can oppose competitive exclusion.

3.3.2 An overview on four system outcomes dependent on R_0 and k

What determines competitive exclusion?

The competitive exclusion occurs if and only if $\lambda_1^2 > 0$ and $\lambda_2^1 < 0$ or reversely, $\lambda_1^2 < 0$ and $\lambda_2^1 > 0$. By (3.3.1), taking $R_0 \to 1^+$ or $k \to 0^+$, which implies $\mu \to \infty$, makes the nonlinear part tends to 0, which leads to the competitive exclusion, in the general case when variations occur in all traits. This remark coincides with the result about $\mu = \frac{1}{(R_0-1)k}$ in [88].

We note that a biologically feasible range for R_0 and k is: $1 \leq R_0 \leq 10$ and $0 \leq k \leq 10$, thus we use values of these parameters in such range to illustrate our model behavior through simulations. However, the model is general to accommodate any other positive values of such parameters.
Next, we will consider the case if we take $\mu \to \infty$, which implies $R_0 \to 1^+$ or $k \to 0^+$, and determine the strain winning in competitive exclusion.

As mentioned, we can rewrite pairwise invasion fitness so that we highlight two opposing terms, where the first one, is completely anti-symmetric in the reverse λ_2^1, thus contributes only to competitive exclusion. Whereas, the second term in the square bracket captures the trait variation that may lead to outcomes beyond exclusion. Here, recall that, if we impact on the system so that $R_0 \to 1$ or $k \to 0$ to get the phenomenon of exclusion, then $\dfrac{\theta_2}{\theta_1} \to \dfrac{\gamma}{\gamma+r}$, and $\dfrac{\theta_3}{\theta_1}$, $\dfrac{\theta_5}{\theta_1}$ go to 0. Hence, the second part tends to 0 and determination of winner/extinct strain depends on the sign of

$$\theta_1 (b_1 - b_2) + \theta_2 (\nu_2 - \nu_1) + \theta_4 \left(\omega_{ij}^i - \omega_{ji}^j\right) + \theta_5 \mu (\alpha_{21} - \alpha_{12}). \tag{3.3.2}$$

If the sign of this expression is positive then strain 1 will be the winner strain and vice versa. Generally, using these arguments, it is still hard to consider exactly the single winner without computing explicitly the term (3.3.2). The final answer will depend on the advantage in terms including transmission Δb, duration of carriage $\Delta\nu$, transmission probability of a strain from a co-colonized host $\omega_{12}^1 - \omega_{21}^2$ and susceptibility to co-colonization $(\alpha_{21} - \alpha_{12})$. We will study particular cases in the next sections, but an overview of the range of possible scenarios is given in Table 3.3.

What determines coexistence?

By the previous arguments, in order to have the coexistence of two strains, in our model with coinfection, the essential condition is $R_0 > 1$ and $k > 0$ large enough. Coexistence opportunities can only come from advantages that may arise in coinfection. In other words, this requires the ratio of single to co-colonization μ tend to 0. In conclusion, from the analysis until here, we have that:

1. The perturbations only in β_i, γ_i and p_{ij}^i lead to the competitive exclusion. Thus strain-specific transmission and/or clearance rates, and the strain-specific transmission biases from mixed co-colonized hosts only create forces favouring exclusion in the system.

2. The perturbations in co-colonization clearance rates and susceptibilities, γ_{ij} and k_{ij}, create more complex scenarios including exclusion of each strain, coexistence or bistable exclusion steady states. Thus,

Table 3.3 **Which trait variation between 2 strains leads to which final outcome?** Scenarios of variation in biological parameters analyzed with the quasi-neutral coinfection SIS model for 2-strains, and the final ecological outcome of their dynamics. There can be exclusion of strain 1 (E1), exclusion of strain 2 (E2), coexistence (C), and bistability (B). Not all trait variations lead to coexistence. In a majority of scenarios, for a given set of trait variation on ≥ 2 trait axes, the final outcome between two strains may shift with coinfection prevalence in the system.

Ecology	**Exclusion Axis**			**Coexistence-Bistability Axis**		**Final outcome for 2 strains**
Nr. traits varying	β_i	γ_i	p^i_{ij}	γ_{ij}	k_{ij}	*Possible equilibria with given trait variation*
1	✓	×	×	×	×	E1 or E2 (depends on R^i_0)
1	×	✓	×	×	×	E1 or E2 (depends on R^i_0)
1	×	×	✓	×	×	E1 or E2 (depends on $\omega^1_{12} - \omega^2_{21}$)
1	×	×	×	✓	×	E1, E2, C, B (depends on u_{ij})
1	×	×	×	×	✓	E1, E2, C, B (depends on α_{ij},$(**)$
2	•	•	•	×	×	E1 or E2 $(*)$
3	✓	✓	✓	×	×	E1 or E2 $(**)$
≥ 2	•	•	•	✓	×	E1, E2, C, B (#)
≥ 2	•	•	•	×	✓	E1, E2, C, B (#)
≥ 2	•	•	•	✓	✓	E1, E2, C, B (#)

The role of co-infection (co-colonization) as a potential gradient modulator for system behavior:

$^{(*)}$ In this case, for any combination of *fixed* trait variation between two strains (with traits denoted by • varying or not varying), there is only competitive exclusion and up to one shift is possible as a function of μ (I^*/D^*). Thus the system can shift from *Exclusion of 1* $\rightarrow$ *Exclusion of 2* or vice versa, only once, because of a particular overall coinfection prevalence ($D/T = 1/(\mu+1)$).

$^{(**)}$ In this case, for any combination of *fixed* trait variation between two strains, only competitive exclusion is possible, but now up to 2 shifts are possible as a function of μ (I^*/D^*). Thus the system can shift three times between opposite exclusion equilibria, because of overall coinfection prevalence.

$^{(\#)}$ In all these cases, (with traits denoted by • varying or not varying), for any combination of *fixed* trait variation between two strains, there can be more qualitative shifts between scenarios as a function of μ (I^*/D^*, i.e. R_0 and k). For example, by varying coinfection prevalence, the system can shift: *Exclusion* $\rightarrow$ *Coexistence* $\rightarrow$ *Exclusion*; or *Exclusion* $\rightarrow$ *Bistability* $\rightarrow$ *Exclusion* (see Figs. S3-S4). In very specific cases, for a given combination of *fixed* trait variation between two strains, the system may even shift 4 times to different regimes as a function of μ, depending on how the λ^2_1 and λ^1_2 intersect the x-axis and each-other (see Supplementary proofs in Text S3.4).

only through the possibility of asymmetries in co-colonization (co-infection) parameters can the strains mediate their mutual coexistence.

In the next section, we will consider the phenomena: exclusion, coexistence or bistable exclusion of either strain according to μ.

3.4 Effect on each trait variation on final outcome

In this very general version of the model, two strains vary along several fitness dimensions: transmission, clearance rate, co-infection susceptibilities, and possible biases in the clearance of co-infection and transmission from the co-infection compartment. In the following we will explore these dimensions in detail, and what is their effect on the competitive dynamics for $N = 2$.

In the subsections 3.4.1, 3.4.2, 3.4.3, and 3.4.4, for each case considered, we study the values of mutual invasion fitnesses λ^2_1 and λ^1_2 as a function of the ratio of single to co-colonization μ. Without loss of generality, in computation we assume that $\nu_1 < \nu_2$, hence strain 1 is cleared more slowly than strain 2, giving it an

advantage in duration of carriage ($\Delta\nu > 0$). First, we recall the relative weights (θ_i) of each trait in terms of μ, $\mu \in [0, +\infty)$ in the Table 3.1.

3.4.1 Definite drivers of competitive exclusion

In this model, exclusion always results from strain-specific transmission and clearance rate of single infection, and transmission probability from coinfected hosts, if other parameters are equal. Below we explore these three axes of trait variation in more detail. First, we consider variation only in β_i, γ_i and p^i_{ij}. We assume equal parameters for co-colonization clearance $\gamma_{ij} = \gamma$ and interaction coefficients between strains $k_{ij} = k$. The two invasion fitnesses (Figure 3.2) are:

$$\begin{cases} \lambda_1^2 = \dfrac{\gamma}{2(\gamma+r)}\mu(2\mu+1)\Delta\nu + (\mu+1)^2\Delta b + (\mu+1)\Delta\omega, \\ \lambda_2^1 = -\lambda_1^2. \end{cases} \tag{3.4.1}$$

These fitness coefficients are completely anti-symmetric, implying competitive exclusion as the only outcome. Thus it doesn't matter that there is co-infection in the system ($k > 0$). For promoting coexistence, this is not sufficient by itself. Variable co-infection susceptibilities or traits between strains would be an additional requirement. When strains behave equally in all processes related to coinfection, coinfection cannot rescue them from the destiny of competitive exclusion. However, as we explore below, overall prevalence of coinfection can actually shift between the winning and losing strain. This can happen only if there is variation in duration of carriage. For example noticing that when β_i is the only trait varying between two strains, k does not appear in θ_1, this indicates that transmissibility's variation, uniquely determines the winner between two strains; its relative contribution to λ_i^j cannot be altered by coinfection.

If variation is only in transmission rates β_i, then the strain with bigger β_i excludes the other one, see figure 3.2 **(a)**. This fact holds for N strains in general and is proved in [120].

If a strain is superior in both fitness dimensions, which means it has smaller clearance rate and greater transmission rate ($\Delta b > 0, \Delta\nu > 0$), then for all value of μ or R_0, it surely will be the winner par excellence, which can be easily seen from (3.4.1). This is unsurprising and naturally expected.

However, if a strain is better in one trait but worse in another, for example if the strain with longer duration of carriage (lower clearance) also has smaller transmission rate, the determination of the winning strain depends on value of μ (and in general also R_0), see Figure 3.2b.

Since we have fixed here $\gamma = 1.5$, $r = 0.5$ by convention, if we fix neutral transmission rate β, the value of μ now depends only on the mean interaction coefficient in co-colonization k. This means, when k is high, thus when strains tend to allow each-other more in co-colonization, μ is sufficiently close to 0, strain 2, which has the smaller transmission rate but longer duration of carriage is the winner. In contrast, when μ is larger, thus when hosts are less vulnerable to co-colonization, the strain 1, which has bigger transmission rate and smaller duration of carriage is the winner.

This illustrates how the relative advantage between two strains, differing in two traits, depends on coinfection prevalence.

It is interesting to note that, if variations are only in two of which including transmission rates β_i, clearance rates γ_i and transmission probability from coinfected hosts p^i_{ij}, we can have at most one shifting outcome, i.e. shifting once from the exclusion of one strain to the exclusion of the other one, see proof in S3.4.1.

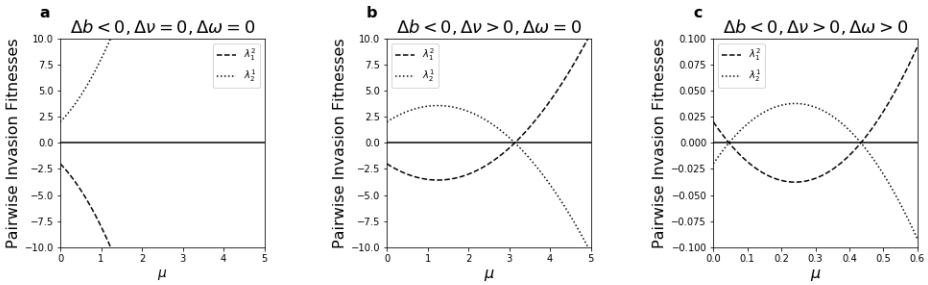


Fig. 3.2 **Pure competitive exclusion from variation in three traits, but the winner strain may depend on relative coinfection prevalence (single-to-co-infection ratio $\mu = I^*/D^*$).** Competitive exclusion is the only scenario when two strains vary only in transmission rate β_i and/or infection clearance rate γ_i and/or transmission probability from coinfected hosts p_{ij}^i (See (3.4.1)). Completely anti-symmetric mutual invasion leads to competitive exclusion. The strain with the positive invasion fitness excludes the one with the negative invasion fitness. Here we illustrate the values of λ_1^2 and λ_2^1 as a function of the ratio of single to co-colonization in the system $\mu = 1/(k(R_0-1))$, with $\gamma = 1.5$, $r = 0.5$, in three cases of (**a**) Variations in transmission rates β_i only with $\Delta b = -2$, (**b**) Variations in transmission rates β_i and clearance rates γ_i with $\Delta\nu = 4$ and $\Delta b = -2$, and (**c**) Variations in transmission rates β_i, clearance rates γ_i and transmission probability from coinfected hosts p_{ij}^i with $\Delta\nu = 4$, $\Delta b = -2$ and $\Delta\omega = 2.02$.

Following the same analysis, we can study the model in which transmission probability from co-colonized hosts, denoting within-host advantage, p_{ij}^i displays strain-specific perturbation. We note that if there is variation in p_{ij}^i only, if $\omega_{12}^1 - \omega_{21}^2 > 0$, strain 1 excludes strain 2, and vice versa. This can be understood via the precedence advantage that one strain has from mixed coinfected hosts if it arrives first, and thus gets transmitted more.

If there are combinations of variations in within-host transmission advantage p_{ij}^i, as well as other traits β_i and/or γ_i, the final outcome is more complex (see Table 3.3 and Figure 3.2 **(c)**). However, the long time competitive result is always exclusion of one strain from the system. If perturbations occur in transmission rate β_i and transmission probabilities from mixed coinfection p_{ij}^i, the winning strain depends on coinfection prevalence in the system, ($\mu = I^*/D^*$) because $\dfrac{\theta_4}{\theta_1} = \dfrac{1}{1+\mu}$. If perturbations occur in duration of single carriage (i.e. strain-specific clearance rates γ_i) and p_{ij}^i, the winning strain depends on more strain-transcending parameters: μ, γ and r. This can be explicitly observed in the ratio $\dfrac{\theta_4}{\theta_2} = \dfrac{2(\gamma+r)}{\gamma}\dfrac{\mu+1}{\mu(2\mu+1)}$, which ultimately affects the signs of pairwise invasion fitnesses. Figure 3.2 **(c)** shows us a special example in which the exclusion of either strains shifts twice when μ varies from 0 to ∞.

3.4.2 Four scenarios possible with variable co-infection clearance rates γ_{ij}

As we can see from the fully explicit expression of pairwise invasion fitness, in this model, the coinfection clearance rate axis contributes to λ_i^j with a term $2(u_{jj} - \frac{u_{ji}+u_{ij}}{2})$. Thus what matters is the comparison between clearance rate of same strain coinfection vs. the mean coinfection clearance rate of mixed-strain coinfection. While there are no restrictions for how these can vary, in the following we consider three special cases where the variation in coinfection duration depends on variation in single infection duration. We also assume no variation in transmission probability from coinfected hosts ($p_{ij}^i = 1/2$) but the results for the general case can be easily derived using Eqs.(3.3.1).

3.4.2.1 Case 1: Unbiased clearance in mixed carriage $u_{12} = u_{21} = \dfrac{\nu_1 + \nu_2}{2}$

In this first case, we consider that the clearance rate of mixed carriage is unbiased and equal to the mean of the two clearance rates of single colonization. We have $u_{11} = \nu_1$, $u_{22} = \nu_2$ and the invasion fitnesses between two strains are anti-symmetric:

$$\begin{cases} \lambda_1^2 = \dfrac{\gamma}{2(\gamma+r)}\left(2\mu^2+2\mu+1\right)\Delta\nu + (1+\mu)^2\Delta b, \\ \lambda_2^1 = -\lambda_1^2. \end{cases} \tag{3.4.2}$$

This case leads again to the pure competitive exclusion. Whichever strain has positive λ_i^j will be the winner.

Similar to the case in section 3.1, if a strain is superior in both transmission β_i and clearance ν_i, it will be the winner.

However, if $\Delta\nu$ and Δb have opposite sign, meaning one strain has advantage in clearance and the other has advantage in transmission, the final winner will depend on coinfection prevalence, hence on μ. From the formula of the invasion fitness, it can be seen that in this system, the clearance rate differential (i.e. in duration of carriage) has more important role than the transmission rate difference in helping an inferior strain overcome and overturn its fitness disadvantage as μ increases.

3.4.2.2 Case 2: Decreased clearance in mixed carriage $u_{12} = u_{21} = \min\{\nu_1, \nu_2\} = \nu_1$

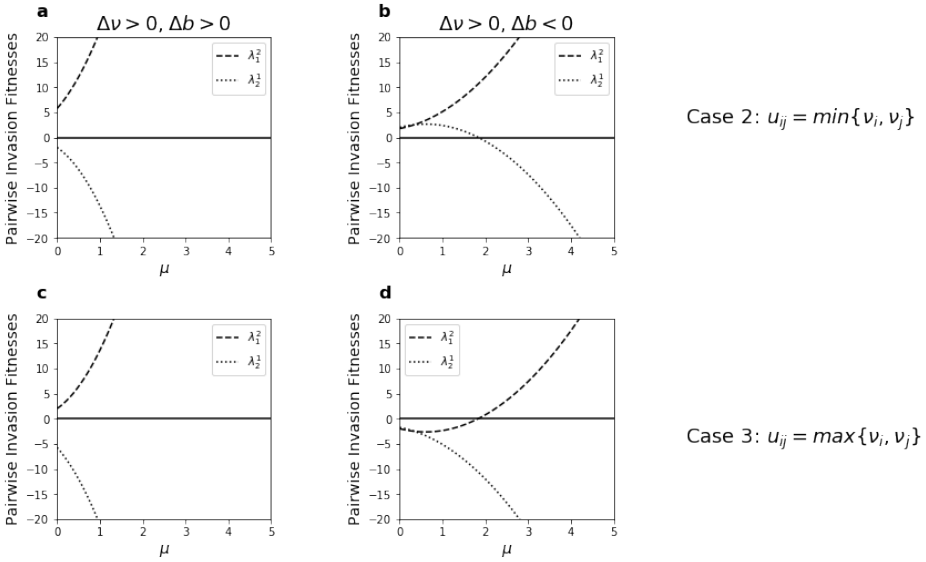


Fig. 3.3 **Competitive exclusion can be broken with variable coinfection clearance rates, and the result may depend on μ.** We illustrate possible scenarios resulting from deviation from symmetry in the mixed coinfection clearance rate γ_{ij} as a function of the ratio of single to coinfection μ. **a-b** Coinfection clearance equals the minimum clearance rate of either strain: $u_{12} = u_{21} = \min\{\nu_1, \nu_2\} = \nu_1$ (see (3.4.3)). **c-d** Coinfection clearance equals the maximum clearance rate of either strain: $u_{12} = u_{21} = \max\{\nu_1, \nu_2\} = \nu_2$ (see (3.4.4)). We choose $\gamma = 1.5$ and $r = 0.5$ as in figure 3.2. For each sub case, we plot the mutual invasion fitnesses for transmission advantage and disadvantage of strain 1, respectively: $\Delta b > 0$ (**a,c**) and $\Delta b < 0$ (**b,d**). In particular, in the first column $\Delta b = 2$, and in the second column $\Delta b = -2$. The clearance rate differential $\Delta\nu$ is assumed $\Delta\nu = 5$ attributing higher duration of carriage to strain 1. In the first column, strain 1 is superior in all fitness dimensions, and coinfection clearance cannot overturn the result. In the second column, strain 1 is not superior in all fitness dimensions, and coinfection matters for the final result.

Here, we explore the case when mixed co-infection clearance rate corresponds to the minimum of the two single infection clearance rates. Without loss of generality, we assume that $\Delta\nu \geq 0$ i.e. $\nu_1 \leq \nu_2$. We still have $u_{11} = \nu_1$, $u_{22} = \nu_2$, but in hosts carrying a mixture of strain 1 and strain 2, this creates an advantage in co-infection for the opponent strain (the one with the faster strain-specific clearance). The two invasion fitness coefficients are not anti-symmetric anymore, hence allowing for more scenarios beyond exclusion (see figures 3.3 **(a,b)**):

$$\begin{aligned} \lambda_1^2 &= \frac{\gamma}{2(\gamma+r)}\left[(2\mu+1)\mu+2(\mu+1)\right]\Delta\nu \quad +(\mu+1)^2\Delta b \\ \lambda_2^1 &= -\frac{\gamma}{2(\gamma+r)}(2\mu+1)\mu\Delta\nu \qquad\qquad\qquad\quad -(\mu+1)^2\Delta b. \end{aligned} \tag{3.4.3}$$

If one strain, denoted to be strain 1 in Figure 3.3a, has larger transmission rate and lower clearance rate (case when $\Delta b > 0, \Delta\nu > 0$), it will again be the only survivor for all μ.

However if the advantage is only in one of the two traits (case when $\Delta b, \Delta\nu$ have opposite signs), for example the strain with smaller clearance rate has lower transmission rate, as in Figure 3.3b, then coexistence can occur. However, even in this situation, coexistence can only be possible for sufficiently low μ, i.e. k large enough for fixed R_0 (we already fix γ and r). This means that increasing the relative prevalence of co-colonization or coinfection in the system, via higher facilitative interactions, can promote coexistence of two strains. This links back to the arguments at the end of section 2.

The phenomenon of coexistence arising here is similar to what has been found before in the context of virulence evolution [5], where decreased clearance in coinfection was observed to promote coexistence, hence persistence of more virulent strains (here strain 2 if $\Delta\nu > 0$, and in the limit of strain 2-R_0 below 1).

3.4.2.3 Case 3: Increased clearance in mixed co-infection $u_{12} = u_{21} = \max\{\nu_1, \nu_2\} = \nu_2$

The co-infection clearance rate in mixed carriage here is assumed to be equal to the maximum value of the strain-specific clearance rates in single infection. As in the previous section, we still assume that $\Delta\nu \geq 0$ i.e. $\nu_1 \leq \nu_2$. We have $u_{11} = \nu_1$, $u_{22} = \nu_2$ and the pairwise invasion fitness coefficients are (see figures 3.3 **(c,d)**):

$$\begin{aligned} \lambda_1^2 &= \frac{\gamma}{2(\gamma+r)}(2\mu+1)\mu\Delta\nu \qquad\qquad\qquad\quad +(\mu+1)^2\Delta b \\ \lambda_2^1 &= -\frac{\gamma}{2(\gamma+r)}\left[(2\mu+1)\mu+2(\mu+1)\right]\Delta\nu \quad -(\mu+1)^2\Delta b \end{aligned} \tag{3.4.4}$$

If strain 1 has smaller clearance rate and larger transmission rate, it is again the superior strain in the system, independently of coinfection parameters, like previous cases (Figure 3.3c).

However, if strain 1 has smaller clearance rate but also lower transmission rate, bistability of exclusion can occur when μ is small enough. When μ becomes larger and tends to infinity, we obtain only competitive exclusion, as mentioned earlier and proven in Section 2. In that extreme, strain 1 is the only persistent strain over long time (Figure 3.3d).

In conclusion, by the explicit formulae of $\left(\lambda_1^2, \lambda_2^1\right)$ in the cases above, we can also prove that for μ large enough, the strain which has smaller clearance rate will be the only strain persisting in the system.

3.4.3 Four scenarios from variation in pairwise co-colonization susceptibilities k_{ij}

Another fitness dimension is how the strains facilitate or compete in altered susceptibilities to co-infection via the coefficients k_{ij}. Above we assumed they are all equal to the reference k. But when variation in this parameter

is allowed, as shown already in [86], all four ecological scenarios are possible, and thus the effect is to open up space for coexistence and bistability among two strains, when competitive exclusion is expected from other parameters. According to the derivation of the reduced model in [140], the perturbations in co-colonization interaction matrix for $N=2$ satisfy $\sum_{i,j=1}^{2}\alpha_{ij}=0$ when k is defined by the mean of k_{ij}. However, without loss of generality, one can shift the α_{ij} by the same constant, without changing the mutual λ_i^j and consequently without changing the dynamics. The explicit formulas for two pairwise invasion fitnesses are (see figure 3.4)

$$\begin{aligned}\lambda_1^2 &= \frac{\gamma}{2(\gamma+r)}(2\mu+1)\mu\Delta\nu+(R_0-1)\mu(\mu(\alpha_{21}-\alpha_{12})+\alpha_{21}-\alpha_{22})+(\mu+1)^2\Delta b,\\ \lambda_2^1 &= -\frac{\gamma}{2(\gamma+r)}(2\mu+1)\mu\Delta\nu+(R_0-1)\mu(\mu(\alpha_{12}-\alpha_{21})+\alpha_{12}-\alpha_{11})-(\mu+1)^2\Delta b.\end{aligned} \tag{3.4.5}$$

In this spirit, below we consider a few special cases of α_{ij} variation between strains, to highlight the effect of co-infection susceptibilities when they provide:

i) **an advantage to strain 2:** $(\alpha_{ij})_{ij}=\begin{pmatrix}-1.5 & 0.5\\ 0.5 & 0.5\end{pmatrix}$, whose effect on λ_i^j is equivalent to $(\alpha_{ij})_{ij}=\begin{pmatrix}-2 & 0\\ 0 & 0\end{pmatrix}$, and increases relatively λ_2^1:

$$\begin{cases}\lambda_1^2=\frac{\gamma}{2(\gamma+r)}(2\mu+1)\mu\Delta\nu+(\mu+1)^2\Delta b,\\ \lambda_2^1=-\frac{\gamma}{2(\gamma+r)}(2\mu+1)\mu\Delta\nu+2(R_0-1)\mu-(\mu+1)^2\Delta b.\end{cases} \tag{3.4.6}$$

ii) **a disadvantage to strain 1:** $(\alpha_{ij})_{ij}=\begin{pmatrix}-0.5 & -0.5\\ -0.5 & 1.5\end{pmatrix}$; equivalent to $(\alpha_{ij})_{ij}=\begin{pmatrix}0 & 0\\ 0 & 2\end{pmatrix}$ when shifted by the appropriate constant, which relatively decreases λ_1^2:

$$\begin{cases}\lambda_1^2=\frac{\gamma}{2(\gamma+r)}(2\mu+1)\mu\Delta\nu-2(R_0-1)\mu+(\mu+1)^2\Delta b,\\ \lambda_2^1=-\frac{\gamma}{2(\gamma+r)}(2\mu+1)\mu\Delta\nu-(\mu+1)^2\Delta b.\end{cases} \tag{3.4.7}$$

iii) **and exactly counterbalanced effects on either strain:** $(\alpha_{ij})_{ij}=\begin{pmatrix}-\sqrt{2} & 0\\ 0 & \sqrt{2}\end{pmatrix}$, whose impact on λ_i^j is to decrease λ_1^2 and increase λ_2^1 by exactly the same amount:

$$\begin{cases}\lambda_1^2=\frac{\gamma}{2(\gamma+r)}(2\mu+1)\mu\Delta\nu-\sqrt{2}(R_0-1)\mu+(\mu+1)^2\Delta b,\\ \lambda_2^1=-\frac{\gamma}{2(\gamma+r)}(2\mu+1)\mu\Delta\nu+\sqrt{2}(R_0-1)\mu-(\mu+1)^2\Delta b.\end{cases} \tag{3.4.8}$$

In Figure 3.4, we consider these cases, where besides k_{ij}, we allow also variation in transmission β and clearance γ of each strain, but assume initially symmetry in other traits. We assume $\Delta\nu>0$ (strain 1 is cleared more slowly). In Fig. 3.4a-c we illustrate competitive outcomes (dependent on mutual signs of λ_i^j), when strain 1 is superior in transmissibility, and in Fig. 3.4d-f we show outcomes when strain 2 is superior in transmissibility instead. Naturally as $\mu\to\infty$ the role of coinfection interaction asymmetries vanishes, and

the system tends to exclusion, but for low values of μ, the structure of the α_{ij} matters. In particular, if it is asymmetric (Figure 3.4,a-b, d-e) there can be at most 3 scenarios as a function of μ: exclusion - coexistence -exclusion, or exclusion - bistability- exclusion. In particular increase in λ_1^2 acts to enable coexistence when $\Delta b > 0$, and a decrease in λ_2^1 acts to enable bistability when $\Delta b < 0$. Whereas, if co-colonization interactions have exactly counterbalanced effects on λ_i^j (Fig. 3.4c,f), there can only be alternating exclusion scenarios as a function of μ.

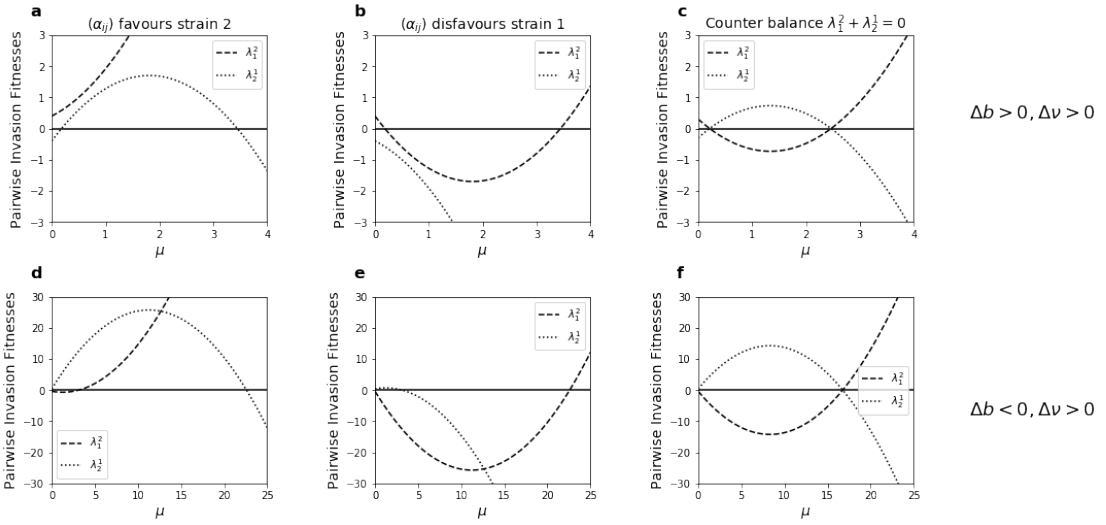


Fig. 3.4 **Breaking the competitive exclusion with co-colonization interactions** k_{ij} (see Eqs (3.4.5)). We compute pairwise invasion fitnesses $\left(\lambda_1^2, \lambda_2^1\right)$according to μ in various cases of co-colonization interaction matrix $\left(\alpha_{ij}\right)$ with $R_0 = 5$, $r = 0.5$ and $\gamma = 1.5$. **(a-c)** We illustrate the cases of transmission superiority of strain 1: $\Delta b > 0$, when $\Delta b = 0.4$, $\Delta \nu = 0.8$. In **(d-f)** we plot 2-strain invasion fitnesses for transmission superiority of strain 2: $\Delta b < 0$, when $\Delta b = -0.4, \Delta \nu = 0.8$, with the same γ, r and R_0 as in **(a, b, c)**. Coinfection clearance rate γ_{ij} is assumed equal to γ and transmission probability from coinfected hosts carrying a mixture of two strains $p_{ij}^i = \frac{1}{2}$. Subplots with the same values $\left(\alpha_{ij}\right)$ lie in the same column. In particular, we consider 3 structures: **(a, d)** $\begin{pmatrix} -2 & 0 \\ 0 & 0 \end{pmatrix}$ (Eqs (3.4.6)); **(b, e)** $\begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$ (Eqs (3.4.7)); **(c, f)** $\begin{pmatrix} -\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{pmatrix}$ (Eqs (3.4.8)) for variation in co-colonization interactions. Except for when the α_{ij} exactly counterbalance effects on λ_i^j (c,f) there is potential for more scenarios beyond competitive exclusion, induced by coinfection susceptibilities between strains.

3.4.4 Adding variation in transmission probability from coinfected hosts

In this case, besides transmission and clearance rates variations $(\Delta b, \Delta \nu)$, and co-colonization susceptibilities (α_{ij}), we add to the system variation in terms of a slight priority effect between strains for transmission from coinfection $\Delta \omega \neq 0$. The explicit formulae for pairwise invasion fitnesses of two strains are as follows:

$$\begin{aligned} \lambda_1^2 &= \frac{\gamma}{2(\gamma+r)}(2\mu+1)\mu\Delta\nu + (\mu+1)\Delta\omega + (R_0-1)\mu(\mu(\alpha_{21}-\alpha_{12}) + \alpha_{21} - \alpha_{22}) + (\mu+1)^2\Delta b \\ \lambda_2^1 &= \frac{\gamma}{2(\gamma+r)}(2\mu+1)\mu\Delta\nu - (\mu+1)\Delta\omega + (R_0-1)\mu(\mu(\alpha_{12}-\alpha_{21}) + \alpha_{12} - \alpha_{11}) - (\mu+1)^2\Delta b, \end{aligned} \tag{3.4.9}$$

Recall that $\Delta\omega = \omega_{12}^1 - \omega_{21}^2$ represents the relative advantage of strain 1 from arriving first within-host, in transmission from mixed coinfection. It is clear from the expression above, that when variation in this within-host advantage is combined with variation in coinfection clearance rates or co-colonization susceptibility factors between strains, coexistence and bistability also become possible. In this case within-host and between-host

competition combine to give rise to different outcomes. In figure S6 we illustrate the effect of $\Delta\omega$ on the baseline outcomes of Figure 3.4a-c. The effect of $\Delta w < 0$ is to increase the potential for strain 2-only competitive exclusion and coexistence with strain 1 in the system, oftentimes overturning the baseline result especially so if μ small. The importance of transmission biases from mixed coinfection Δw is unsurprisingly higher when relative coinfection prevalence is higher in the system.

3.4.5 The qualitative outcome for two strains can shift multiple times with μ

Until now we have seen the important and explicit role of strain-transcending parameters, (e.g. R_0, k, and specifically μ) which define the core neutral system of this model, on the ultimate competitive outcome between strains at the epidemiological level. We have seen that the same relative variation between strains, displayed in $(\Delta b, \Delta v, ...)$ will have a different impact in a system with larger or lower overall prevalence of coinfection, relationships that are completely transparent in the λ_i^j. Sometimes the effect will be quantitative, changing only the speed of dynamics without affecting the λ_i^j signs (see Supplementary figure S2). Other times the effect will be qualitative, changing the signs of the pairwise invasion fitnesses and hence the dynamics. Furthermore, using such full analytic transparency, we can also prove mathematically special results to make the claims about qualitative shifts more precise. Depending on how many traits and which traits vary in the system, we can have at most one, two, or more shifts with μ (see Table 3.2). For example we can prove that with variation only in transmission and clearance rates, there can be at most one shift in final outcome as a function of μ (Text S3.4.1)

A very special case arises, when the same system can shift 4 times as a function of μ. We have proven that a necessary condition for its occurrence is the presence of variation in both coinfection clearance rates γ_{ij} and vulnerabilities to coinfection k_{ij}. If any of these is missing, 4 shifts as a function of μ are impossible (see Text S3.4.2 for the formal proof). An illustration of such a special case is given in Figure 3.5, where the system traces all 4 quadrants as μ goes from 0 to ∞, highlighting an extreme case of the critical role of coinfection, for the relative hierarchical advantages between two strains and their selection dynamics. We see that the system is characterized by competitive exclusion of strain 2 for μ low, then tends to coexistence of both strains for increasing μ, followed by competitive exclusion of strain 1 for even higher μ, until it returns again in the competitive exclusion region of strain 2 as $\mu \to \infty$.

3.4.6 Parameter regions for 4 outcomes in our SIS model with co-infection

Until now, we have considered fixed trait variations between two strains, and varied μ to show how their net competitive dynamics driven by λ_1^2 and λ_2^1 will depend on the ratio of single to coinfection in the neutral system. This context is defined by basic reproduction number R_0, and k, of the neutral model (see Table 1). Next, we consider the distribution of 4 possible ecological outcomes across different systems, as a function of R_0 and k. In figure 3.6 we represent the long time behaviour of a 2-strain model with perturbations in transmission rate β_i, clearance rate of single colonization γ_i and clearance rate of co-colonization γ_{ij} with $u_{11} = u_{12} = u_{21} = v_1$ and $u_{22} = v_2$. Figure 3.6 shows which combinations of Δb and Δv, lead to one of the four scenarios: exclusion of strain 1 or 2, coexistence or bistable state, for each R_0 and k, for assumed symmetry in k_{ij} and in $p_{ij}^i = \frac{1}{2}$.

It can be seen that we can choose suitable values of relative trait differences Δb and Δv to observe a given scenario, typically coexistence and bistability arise for relative advantage in one trait and relative disadvantage in the other, hence a trade-off between transmission and clearance. Notice that when R_0 or k become larger, increasing relative coinfection prevalence in the corresponding neutral system, (i.e. reducing μ), the possibility

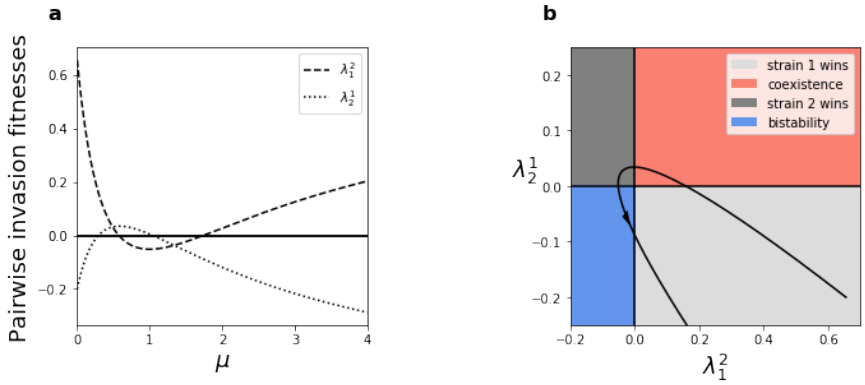


Fig. 3.5 **Four ecological scenarios may happen depending on μ (single-to-co-infection ratio) under fixed trait variation.** Variations here are in transmission rates, infection and coinfection clearance rates, and co-colonization susceptibilities between strains. The strain-transcending parameters are assumed: $\gamma = 2$, $r = 0.2$ and $R_0 = 2.5$. We assume that $\Delta b = 0.2$, $\Delta v = 0.5$, the coinfection clearance rate γ_{ij} with $u_{ij} = \min\{v_i, v_j\}$ and values of (α_{ij}) to be $\begin{pmatrix} -1 & 0 \\ 0 & \sqrt{3} \end{pmatrix}$. The ultimate ecological outcomes when μ goes from 0 to ∞ include: exclusion of strain 2, coexistence, exclusion of strain 1, bistablity, then back to the exclusion of strain 2. Figure **(a)** plots two pairwise invasion fitnesses as function of $\mu \in [0.01, 4]$, corresponding to $k \in [1/6, 66.6]$. Figure **(b)** plots the parametric curve $\left(\lambda_1^2(\mu), \lambda_2^1(\mu)\right)$ for μ shifting in the same range of figure **(a)**. This curve crosses all four quadrants, which means the 2-strain system can display all four qualitative outcomes with changes in coinfection prevalence (i.e. as μ varies).

for coexistence or bistability expands in the system. Near the origin, it's harder to obtain coexistence or bistable state rather than competitive exclusion.

The results of Figure 3.6 can be used to connect our system's behavior to strain-specific basic reproduction numbers $R_0^{(1)}$ and $R_0^{(2)}$. In a model with coinfection, strain-specific R_0 are not sufficient to determine the long-time epidemiological competition between two strains. We show in Text S1 and in Figure S1 that even for the same values of $R_{0,1}$ and $R_{0,2}$ we can have different long-time scenarios between the two strains. The result in coinfection depends on the particular combination of traits, strain-specific transmission and clearance rates. In particular, even for the same strain-specific R_0's, a system with bigger variation in clearance rate between two strains is more sensitive to coinfection, and it is where coinfection parameters can shift the dynamics more easily away from competitive exclusion (Fig. S1a). This is because of the advantage conferred to both strains by staying longer in mixed compartment I_{12}, where from they can have equal chance of transmission.

In Fig. 3.6 coinfection susceptibilities were assumed symmetric and coinfection clearance rate for mixed carriage was assumed biased toward the same clearance rate of single infection by strain 1: $u_{11} = u_{12} = u_{21} = v_1$ and $u_{22} = v_2$. In another case (Figs. S3-S4), when we remove the bias in coinfection clearance assuming $u_{12} = u_{21} = 0$, but allow variation in susceptibilities to coinfection, we see only three scenarios emerge, over all Δb and Δv. In this case, depending on the structure of α_{ij}, we observe either coexistence or bistability as a third possibility flanked by the two opposite exclusion steady states.

3.5 How trait mean and variation impact coexistence frequencies

Next, we zoom in from criteria for coexistence to the details of the coexistence equilibrium in a system with two strains that vary along multiple fitness dimensions. The general formula in terms of pairwise invasion

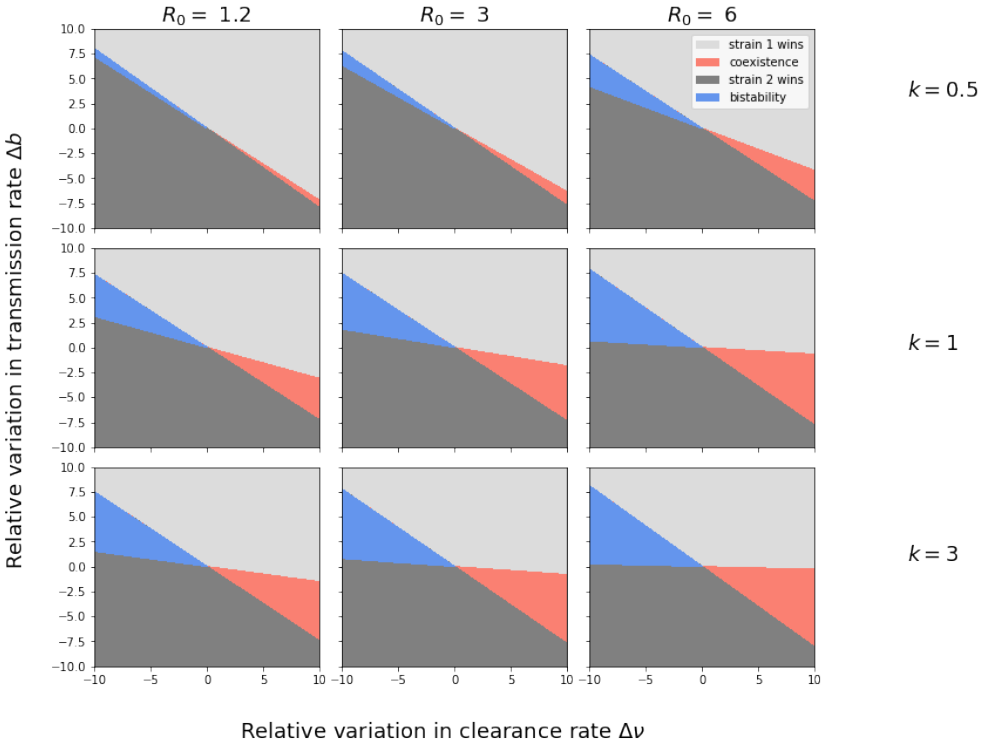


Fig. 3.6 **Illustration of 4 possible outcomes, as a function of relative variation in transmission and clearance rate between two strains, for different values of k and R_0.** We highlight the respective regions in different colors, according to the critical relationship between Δb, $\Delta \nu$, k and R_0 when perturbations happen only β_i, γ_i, and mixed coinfection clearance happens base on strain-specific clearance rates γ_{ij} with $u_{11} = u_{12} = u_{21} = \nu_1$ and $u_{22} = \nu_2$. We choose the values $\gamma = 1$ and $r = 0.2$. Recall that the higher $\Delta \nu > 0$, the higher the advantage of strain 1 in the system; and the higher $\Delta b > 0$, the higher the advantage of strain 1 in the system. We observe coexistence and bistability arise only when the disadvantage in one trait is compensated by an advantage in the other. In particular coexistence is enabled when the disadvantaged strain 2 in clearance rate benefits from reduced clearance in mixed co-colonization with strain 1.

fitnesses in Eqs. (3.3.1), when this equilibrium exists, is given by

$$z_1^* = \frac{\lambda_1^2}{\lambda_1^2 + \lambda_2^1} \quad z_2^* = 1 - z_1^*. \tag{3.5.1}$$

This allows explicit computation for any combination of strain-specific and strain-transcending parameters in the system. Below we focus on the case of variation in transmission rate β_i, clearance rate of infection γ_i and in interaction coefficients via susceptibilities to coinfection k_{ij}. We assume the special case of $u_{ij} = 0$ and initially no transmission biases in coinfection $p_{ij}^i = \frac{1}{2}$ (although we relax this later). We consider only coexistence regimes (Figure 3.7) and illustrate the equilibrium frequency of one of the strains (here z_2^*), as a function of trait mean and variation between strains. Using the explicit formula $z_2^* = \lambda_2^1/(\lambda_1^2 + \lambda_2^1)$, we explore this quantity numerically for two values of mean co-infection susceptibility: $k = 1.5$ and $k = 0.2$, and symmetric cross-strain interactions $\alpha_{12} = \alpha_{21} > \alpha_{11} = \alpha_{22}$. We consider $(\alpha_{ij})_{ij} = \begin{pmatrix} 0 & \sqrt{2} \\ \sqrt{2} & 0 \end{pmatrix}$ but the general formula for other cases is straightforward (see Text S3). The criteria enabling coexistence are derived below.

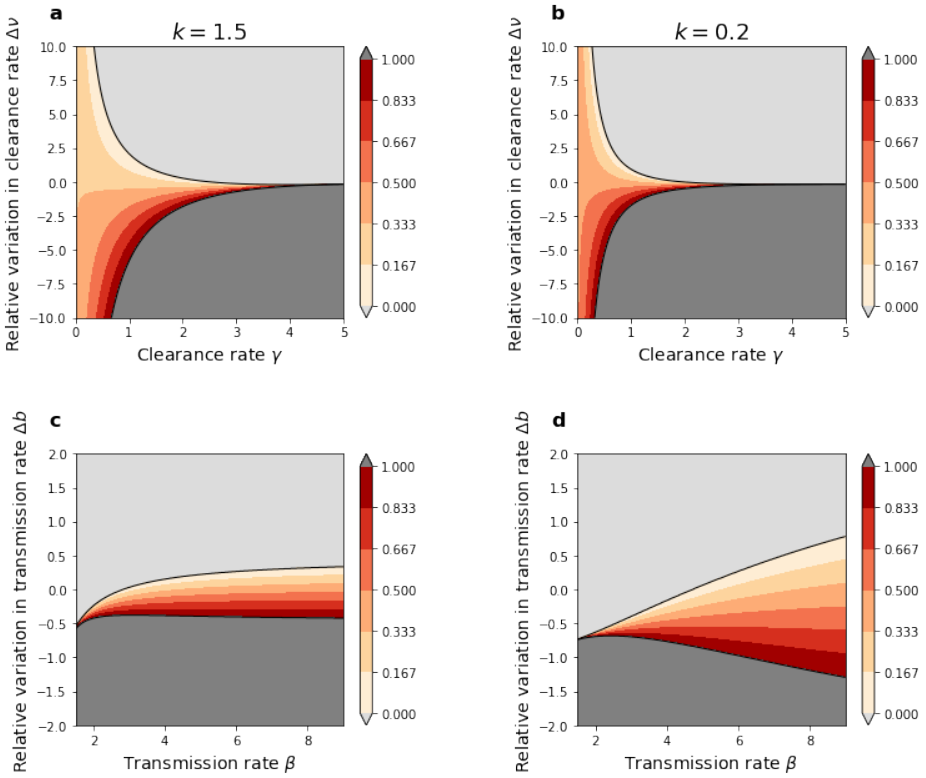


Fig. 3.7 **Strain coexistence frequencies depend on a critical interplay between trait mean and variation.** We plot the coexistence equilibrium frequency of strain 2 (intermediate shading), as a function of $(\beta, \Delta b)$ and $(\gamma, \Delta\nu)$ for higher coinfection prevalence (left) and lower coinfection prevalence (right). **a-b. Effect of clearance rate.** In (a-b) we assume $\beta = 5.3$ and $\Delta b = 0.15$. In these figures, we vary γ in $0 \leq \gamma \leq 5$ and $\beta = 5.3$, to make $R_0 > 1$. **c-d. Effect of transmission rate.** In (c-d) we assume $\gamma = 1$ and $\Delta\nu = 1$ as fixed. We vary β between 1.5 and 9, ensuring $R_0 > 1$. Right-column subplots reflect a system with more strain competition preventing co-infection, hence lower coinfection prevalence (μ higher) than the left-column subplots. The light grey region represents the exclusion of strain 2-$z_2^* = 0$ and the dark grey region is for exclusion of strain 1 from the system i.e. $z_2^* = 1$. We choose $r = 0.3$ and the matrix (α_{ij}) to be $\begin{pmatrix} 0 & \sqrt{2} \\ \sqrt{2} & 0 \end{pmatrix}$ thus unbiased in terms of favouring either strain. The co-infection clearance rate is also assumed unbiased, and equal to the mean. The black lines denote the border lines for which coexistence is no longer possible and the system shifts to either exclusion of strain 1 (yellow) or exclusion of strain 2 (blue). In a-b, the lines are denoted by $T_1(\gamma)$ and $T_2(\gamma)$, and given by (3.5.3) and (3.5.4). In c-d, the graphs of $S_1(\beta)$ and $S_2(\beta)$ are hyperbolic, given by explicit equations (3.5.6) and (3.5.7).

3.5.1 Mean and variation in clearance rate of single infection

By (3.4.3), the equation of boundary $z_2^* = 0$ is equivalent to $\lambda_2^1 = 0$ which can be written explicitly as:

$$\frac{\gamma}{2(\gamma+r)} \frac{1}{1+\frac{1}{\mu}} \left(1 + \frac{1}{1+\frac{1}{\mu}} \right) \Delta\nu = -\Delta b - \frac{\sqrt{2}}{2} (R_0 - 1) \frac{\mu}{(\mu+1)^2} \tag{3.5.2}$$

which, by substituting $\frac{1}{\mu} = k\left(\frac{\beta}{\gamma+r} - 1\right)$, becomes:

$$\Delta\nu = T_1(\gamma), \quad T_1(\gamma) = -\Delta b \frac{2[\gamma+r+k(\beta-\gamma-r)]^2}{\gamma(2\gamma+2r+k(\beta-\gamma-r))} - \sqrt{2}\frac{k}{\gamma} \frac{(\beta-\gamma-r)^2}{2(\gamma+r)+k(\beta-\gamma-r)}. \tag{3.5.3}$$

Analogously, we can compute the equation of boundary $z_2^* = 1$, thus $\lambda_1^2 = 0$, which is

$$\Delta\nu = T_2(\gamma), \quad T_2(\gamma) = -\Delta b \frac{2[\gamma+r+k(\beta-\gamma-r)]^2}{\gamma(2\gamma+2r+k(\beta-\gamma-r))} + \sqrt{2}\frac{k}{\gamma}\frac{(\beta-\gamma-r)^2}{2(\gamma+r)+k(\beta-\gamma-r)}. \tag{3.5.4}$$

Combining (3.5.3) and (3.5.4) we have the condition for 2-strain coexistence between these boundaries, expressed as an inequality for the variation $\Delta\nu$ dependent on the mean γ:

$$\boxed{\text{Stable Coexistence} \iff T_1(\gamma) \leq \Delta\nu \leq T_2(\gamma).} \tag{3.5.5}$$

The figures 3.7a and 3.7b show critical interplay between mean and variation clearance rates of infection γ_i. The light color region in this figure is the region displays coexistence phenomenon, corresponding to the space between two curves which present equations (3.5.3) and (3.5.4). The contour shading corresponds to the relative frequency of strain 2. In these figures, we can observe that by decreasing mean co-infection efficiency k, we can decrease the possibility of coexistence at the same variation in clearance rate $\Delta\nu$ (3.7b compared to 3.7a). For γ small enough, which increases overall prevalence in the system, less values of $\Delta\nu$ lead to coexistence. When $\Delta\nu$ becomes larger, the relative frequency of strain 2, decreases.

3.5.2 Mean and variation in transmission rate

Similarly to (3.5.3) and (3.5.4) for mean and variation in clearance rate (duration of carriage), and the same assumptions about the other parameters, we can find the equation of boundary $z_2^* = 0$ as a function of $(\beta, \triangle b)$ as follows:

$$\Delta b = S_2(\beta), \quad S_2(\beta) = -\Delta\nu \frac{\gamma(2\gamma+2r+k(\beta-\gamma-r))}{2[\gamma+r+k(\beta-\gamma-r)]^2} + \frac{\sqrt{2}}{2}\frac{k(\beta-\gamma-r)^2}{[\gamma+r+k(\beta-\gamma-r)]^2} \tag{3.5.6}$$

The equation of the boundary $z_2^* = 1$ as a function of $(\beta, \triangle b)$ is:

$$\Delta b = S_1(\beta), \quad S_1(\beta) = -\Delta\nu \frac{\gamma(2\gamma+2r+k(\beta-\gamma-r))}{2[\gamma+r+k(\beta-\gamma-r)]^2} - \frac{\sqrt{2}}{2}\frac{k(\beta-\gamma-r)^2}{[\gamma+r+k(\beta-\gamma-r)]^2}. \tag{3.5.7}$$

Combining (3.5.6) and (3.5.7), we have the condition on intermediate values of Δb that allow 2-strain coexistence:

$$\boxed{\text{Stable Coexistence} \iff S_1(\beta) \leq \Delta b \leq S_2(\beta).} \tag{3.5.8}$$

Two figures 3.7c and 3.7d show the value of strain 2 in case of coexistence as a function of mean and variation in transmission rates Δb. In contrast to 3.7a and 3.7b, by decreasing co-infection vulnerability k, we increase the possibility of coexistence at the same variation in transmissibility between strains Δb. This also means that for larger β, more values of Δb lead to coexistence. When Δb becomes larger, increasing the transmission advantage of strain 1, the equilibrium frequency of strain 2, decreases. This is similar to Figure 3.7a and 3.7b.

3.5.3 Coexistence strain frequencies are explicit

While the entire frequency dynamics for each strain and for all time ($z_i(\tau)$) is fully explicit in the system, via the replicator equation, so is also the final equilibrium. The value of the equilibrium frequency of strain 2, z_2^*

under the coexistence regimes studied above is given by:

$$z_2^* = \frac{1}{2} - \gamma \frac{2(\gamma+r)+k(\beta-\gamma-r)}{2\sqrt{2}k(\beta-\gamma-r)^2}\Delta\nu - \frac{[\gamma+r+k(\beta-\gamma-r)]^2}{\sqrt{2}k(\beta-\gamma-r)^2}\Delta b, \tag{3.5.9}$$

and can be seen to be an explicit function of mean parameters (e.g. β, γ) as well as variation between strains $(\Delta b, \Delta\nu)$. Here, because coinfection interactions are assumed symmetric, it becomes obvious that any differences between strains will make the frequency deviate from the expected frequency of 1/2 under balancing selection. If $\Delta\nu, \Delta b > 0$, then strain 1 has an absolute advantage and z_2 will always be inferior than 1/2, if it coexists. However, if the advantage is only in one trait and not in the other ($\Delta\nu$, Δb of different signs), then strain 2 can increase its equilibrium frequency above 1/2.

More generally, for the rescaled co-colonization susceptibilities matrix being symmetric and satisfying: $\alpha_{ii} = \alpha_{jj} = \alpha_{11}$ and $\alpha_{ij} = \alpha_{ji} = \alpha_{12}$, we can write the expression for strain 2 equilibrium frequency, more compactly as a function of μ, k and R_0 constituent parameters, mean transmission rate β, clearance rate γ and average host lifespan $1/r$:

$$z_2^* = \frac{1}{2}\left[1 - \frac{k}{\alpha_{12}-\alpha_{11}}\left(\mu(2\mu+1)\frac{\gamma}{\gamma+r}\Delta\nu + 2(\mu+1)^2\Delta b\right)\right] \tag{3.5.10}$$

One can notice that variation in each trait Δb and $\Delta\nu$ have their own distinct nonlinear scaling factors for how they impact on ultimate strain success at the epidemiological level, depending directly on the prevalence of co-colonization in the system via the parameter $\mu = I^*/D^*$. We can immediately see from this formula, how the predicted coexistence level among two strains is attributable to clear mechanisms and clearly identifiable biological differences between strains, which are explicitly weighted by epidemiologic constants.

Even more generally, if in addition there are biases in strain transmission probabilities from co-colonized hosts carrying a mixture of strains, recalling that $\Delta\omega = \omega_{12}^1 - \omega_{21}^2$ the formula (3.5.10) now reads

$$z_2^* = \frac{1}{2}\left[1 - \frac{k}{\alpha_{12}-\alpha_{11}}\left(\mu(2\mu+1)\frac{\gamma}{\gamma+r}\Delta\nu + 2(\mu+1)^2\Delta b + 2(\mu+1)\Delta\omega\right)\right] \tag{3.5.11}$$

where we can straightforwardly see the contribution of the precedence effect in transmission from co-colonization as a force in coexistence hierarchies between two strains. Indeed, for any value of coinfection prevalence and overall R_0, hence for any μ, the relative contribution of the transmission rate differential between strains (Δb) is higher than that of the transmission bias differential from coinfected hosts ($\Delta\omega$). It is also interesting to notice in the above expression for z_2^* that the relative contribution of the between-strain difference in duration of carriage ($\Delta\nu$) depends on μ but also on the absolute value of the duration of colonization itself γ. Essentially, keeping all else fixed, differences in duration of carriage between two strains matter more for their relative fitness, if the colonization episodes are shorter (γ higher), and if hosts are longer lived (r higher).

Perfectly balancing selection, in this model, under a given $k \neq 0$, would require that the linear combination of $\Delta\nu, \Delta b, \Delta\omega$ in the round brackets be equal to zero. Trait variation possibilities satisfying this requirement are infinite and constitute a plane in 3-d in this case, necessarily encoding trade-offs across different fitness dimensions that would as a whole lead to the same epidemiologic fitness for the two strains. Other particular values of z_2^* can be studied explicitly and investigated similarly in terms of the constraints implied for the linear combination of $\Delta\nu, \Delta b, \Delta\omega$, and other fitness dimensions eventually, if such variation exists.

3.6 Model applications: a roadmap

In earlier theoretical work, interaction among co-infecting agents has been assumed to occur only between different strains, and studied for arbitrary infection multiplicity [4]. Later evolutionary frameworks, based on [199], have considered a full model including same-strain coinfection, but modeling vulnerability to co-infection with a single parameter [6]. This aggregation of within- and between- strain interactions into a net parameter can be found in other co-infection models, considering altered susceptibility to coinfection in the context of disease persistence [82], and diversity in other traits, e.g. virulence [7] and antibiotic resistance [98]. These studies highlight the importance of coinfection and its epidemiological details for persistence and evolution of microorganisms. Sometimes very complex multi-scale models have been invoked to generate coexistence between strains via coinfection, embedding an explicit within-host dynamics framework [61]. We argue that many such coinfection and co-colonization models could be mapped to phenomena in the overarching model proposed here, as special cases, or expansions of a particular parameter.

With the here-proposed explicit framework, the impact of coinfection becomes very easy to understand, via the role of the parameter μ, given by the ratio of single - to co-infection prevalence in the system, which modulates the relative weight of different trait asymmetries (θ_i's) among strains, and even tuning the net asymmetries in some traits, as is the case for k_{ij}. This role of coinfection (in terms of 2 global determining parameters R_0, k) can be studied at a deeper level, at a higher resolution in terms of potential asymmetries within and between strains, and in an entirely analytically-explicit manner which enables precise predictions. These advantages can lead to new applications to study coexistence and vaccination effects in polymorphic systems, going beyond current theoretical insights [127, 90]. Similarly, our modeling framework could also help obtain clear and direct analytical insights into antibiotic resistance evolution, as an alternative or as a complement to the more cumbersome simulation route [61].

Ideally deterministic and stochastic forces shaping population dynamics and evolution should be integrated [56]. Although stochastic effects have been somewhat studied in multi-strain epidemiological models, largely via simulations of ODEs with stochasticity based on the Gillespie Algorithm [3, 179], with the exception of few analytical investigations [48, 109], stochastic simulation approaches in coinfection models are rare [61, 100, 173, 172], with analytical results on stochasticity and coinfection even rarer [22]. Hence, our SIS model with multiple interacting strains in coinfection and emergent replicator dynamics at its deterministic core offers a simple, general, and elegant template on which stochasticity can be straightforwardly built-in, simulated, and moreover studied analytically.

Below we sketch briefly some ways in which the model can be applied.

3.6.1 Antibiotic treatment, fitness costs and competitive release

We can apply this model to understand epidemiological competition dynamics between two strains under antibiotics, co-circulating in a host population with the possibility of coinfection, which constitutes a big study field in the epidemiological literature [156]. In its simplest form, broad-spectrum antibiotic treatment can be modelled as an increase in global clearance rate of colonization γ, keeping fixed all other parameters between strains. Suppose strain 1 is superior in transmission, with strain 2 suffering a relative fitness cost $\Delta b > 0$. For variation in duration of carriage we explore both an advantage to strain 1 $\Delta\nu > 0$ or advantage to strain 2 $\Delta\nu < 0$. Without perturbation in coinfection parameters, under such scenario, there would be competitive exclusion of the less fit strain, seen earlier in Section 3.4.1.

But under interactions through coinfection vulnerabilities, coexistence among the two strains is possible. Thus we may explore, the frequency of strain 2, under such scenario, which would then correspond to the

variable z_2 in our model. We examine how the equilibrium value of the frequency of strain 2, varies as a function of γ (or total broad-spectrum antibiotic treatment), for different values of strain variation in relative duration of carriage (clearance rate) $\Delta\nu$, and relative transmissibility Δb.

Fixing for example the rescaled co-infection susceptibility matrix to $\alpha_{ii} = \alpha_{jj} = 0$ and $\alpha_{ij} = \alpha_{ji} = \sqrt{2}$, corresponds to the case analyzed earlier, where within-strain susceptibilities to coinfection are lower than between-strain susceptibilities, a condition that *a-priori* favours coexistence. Hence, applying the earlier results, we have that for any $(\beta, \gamma, \Delta b, \Delta\nu)$, stable coexistence of two strains is possible only for $T_1(\gamma) \leq \Delta\nu \leq T_2(\gamma)$. The value of z_2^* under such condition is given by Eq. 3.5.9. Thus, in Figure 3.8, we only consider the values of global clearance rate γ guaranteeing coexistence. It is interesting to point out that depending on how the fitness differential between the strain 1 and its competitor strain 2, is manifested $(\Delta b, \Delta\nu)$, increasing antibiotic administration in a population can have opposing effects: it can increase or decrease the prevalence of a focal strain (see Fig.3.8a solid vs. dashed lines). This behavior can be understood in full analytic detail because of the explicit expression for strain frequencies, allowing us to compute and verify directly the first and second derivatives of z_2 with respect to γ (see Supplementary Material S3.6).

To understand the additional effects of possible variations in transmission probability from mixed coinfection between strains, we have repeated the same simulations with $\Delta w < 0$, favouring strain 2 (Supplementary figure S7). It is clear that also this dimension of fitness (within-host advantage) has a substantial effect on the net competitive dynamics between the two strains, and in particular, in this case, enhances the possibility of two-strain coexistence.

3.6.2 Vaccination, coexistence and strain replacement in colonizing bacteria

Similarly, universal vaccination that protects against both strains could be modelled, to a first-order approximation, as a global reduction of β in the system, realized via reduced average susceptibility of all hosts to infection. In figures 3.8c and 3.8d, we explore the effect of a universal reduction in β on the relative prevalence of two strains. As shown earlier, coexistence is possible only for $S_1(\beta) \leq \Delta b \leq S_2(\beta)$. The value of z_2^* under such condition is given by Equation 3.5.9. Plotting this as a function of β in Figures 3.8c and 3.8d, we observe again that changes in strain-transcending transmissibility can have opposing effects on prevalence of a focal strain (here strain 2). They can either favour its increase in prevalence or its decrease, depending on the underlying basic trait variation $(\Delta b, \Delta\nu)$, as well as on coinfection parameters (k_{ij} or others).

Turning the inequalities around, another way to interpret the critical borders for Δb is with regards to a strain-specific vaccine. For example, assuming universal coverage, in order to predict the minimal vaccine efficacy needed to exclude a given strain (or group of strains) from the system, given everything else fixed at pre-vaccine baseline, we can use the model to extract the Δb that violates the coexistence inequality criterion: $S_1(\beta) \leq \Delta b \leq S_2(\beta)$. Notice that this criterion specifically implies that in populations with different overall β a different Δb (targeted vaccine efficacy) may be needed to reach exclusion. In the case of $\Delta b < S_1(\beta)$ we would ensure exclusion of strain 1 from the system, whereas making $\Delta b > S_2(\beta)$ would shift the system to the exclusion of strain 2 steady state.

3.6.3 Effects of host demography on 2-strain epidemiological competition

Changes in natural susceptible recruitment rate (equal host mortality rate) r affect R_0 in the system, which also determines μ, besides appearing itself also explicitly in Eq.3.5.10. So even with everything else fixed (β, γ, k and $\Delta b, \Delta\nu$) it is possible to influence competitive dynamics between strain 1 and strain 2, just via host demography. Increasing host mortality rate, decreases R_0 in the system which increases μ and hence gives a

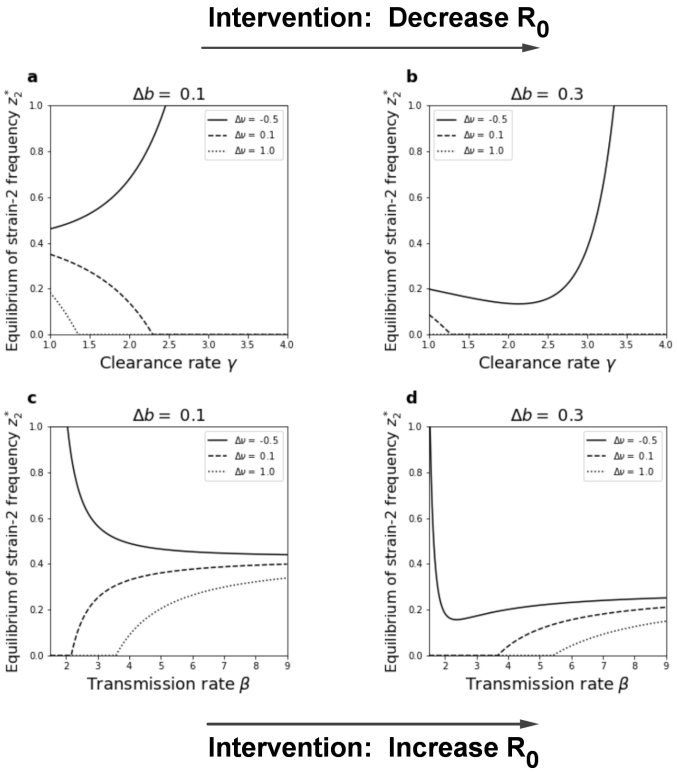


Fig. 3.8 **Increasing global clearance or global transmission can have opposite effects on single strain frequency at the endemic equilibrium, depending on underlying trait variation.** We plot the frequency of strain 2, z_2^* at the coexistence equilibrium (Eq. 3.5.9), as a function of mean transmission rate and mean clearance rate. Variation among 2-strains is encoded in the transmission and clearance rate axes: β_i and γ_i, and co-colonization vulnerabilities k_{ij}. In this simulation, we choose $r = 0.2$, $k = 1$ and the matrix of standardized interactions is assumed symmetric $(\alpha_{ij}) = \begin{pmatrix} 0 & \sqrt{2} \\ \sqrt{2} & 0 \end{pmatrix}$, favouring coexistence with $\alpha_{ij} = \alpha_{ji} > \alpha_{ii} = \alpha_{jj}$. **a-b.** Equilibrium frequency as a function of strain-transcending γ. We plot z_2^* as a function of mean clearance rate γ (varied between 1 and 4) for 2 cases of fitness differentials in transmission Δb and 3 cases of variability in clearance $\Delta\nu$. The global transmission rate is $\beta = 4.5$ to ensure $R_0 \geq 1$. **c-d.** Equilibrium frequency as a function of strain-transcending β. We plot z_2^*, as a function of mean transmission rate β (varied between 1.5 and 9) for 2 cases of different variation Δb and 3 cases of $\Delta\nu$. In these plots, overall clearance rate is held fixed at $\gamma = 1$ to ensure $R_0 \geq 1$.

larger weight to the trait variation in clearance rate of infection $\Delta\nu$. Increasing host turnover rate might then enable coexistence of two strains (maintenance of the less fit strain e.g. strain 2) because it could amplify, and even overturn, the relative advantage in duration of carriage ($\Delta\nu < 0$) versus the disadvantage in transmissibility ($\Delta b > 0$). This would imply that in different populations, with different rates of susceptible host turnover, the dynamics of the same two strains could be different. Thanks to this model, all these mechanisms and special cases in competitive dynamics between two closely-related strains at the epidemiological scale can be studied in a fully parameter-explicit and analytical manner, which should promote easier and more direct testable links with data.

3.6.4 Dynamical transitions: from $N = 2$ to the N-strain ecological network

Throughout this study, we have shown explicitly and illustrated in detail how global parameters of the neutral model, embedded in the center of the dynamics, can shift qualitatively and quantitatively the net competitive

outcome for any given pair of strains. Recall that the $N = 2$ system forms the basic unit in the full competitive network among an arbitrary number N of strains, and that qualitative shifts in each network 'edge', as a function of global parameters, can have far-reaching effects on the collective dynamics among multiple strains, even when strains differ just in co-infection vulnerabilities [88]. Having exposed new and nonlinear gradient effects of μ, R_0, k, γ, r in the more complete 2-strain system with variation along 5 fitness dimensions, opens the way towards deeper analysis of their higher-level effects on the $N-$ strain assembly, dynamics, and coexistence [120]. Studying dynamical transitions mediated via coinfection prevalence and strain-transcending epidemiological parameters, as well as the statistical distributions of trait variations, possibly informed by data [1, 51, 61] in the full system, is the natural and exciting next step.

3.6.5 Extensions of the model towards stochasticity

One could investigate the stability of the global endemic vs. disease-free equilibrium under stochastic noise; which is key for the establishment of the neutral model on the fast time-scale (based on $R_0 > 1$). It has been noted that outcomes predicted by stochastic and deterministic models can differ widely, especially concerning persistence and extinction [48], and, in the case of strains differing in epidemiological traits, in final competitive outcomes [151]. This indicates that the influence of fluctuations on dynamics of infectious agents could be significant and should be studied also in our context of coinfection.

In SIR frameworks with immunity and cross-immunity between strains, simulations have revealed that the dynamics of 'semi-stochastic' multi-strain models, can be strikingly different from those of the corresponding deterministic model, with large amplitude limit cycles typically replaced by oscillatory irregular dynamics with a limited strain diversity circulating at any time [152]. Similar to our key driver of model behavior, μ, dependent on basic reproduction number R_0 and mean susceptibility to coinfection k [88], sharp transitions in other multi-strain studies have been shown as a function of the reproduction number of the pathogen and the intensity of cross–immunity [3]. These could be studied further in stochastic versions of our model with finite host population size, possibly affecting μ-mediated transitions in the $N = 2$ system described here, and likely more relevant and interesting in the full SIS formulation with N strains and coinfection derived in [120].

At their mathematical core, SIS endemic frameworks share many properties with logistic growth, and have stochastic models with well-established phenomena [160, 8]. Possibly, existing or forthcoming analytic results could be translated across this interface, accounting also for multiple strains and coinfection. A clear benefit of the deterministic framework proposed here is the reduction in the number of relevant parameters for predicting dynamics (the two mutual invasion fitnesses in the $N = 2$ system, or N mutual invasion fitnesses in the generic N-strain model [120]) hence enabling easier, more precise and reproducible estimation across countless scenarios of subtle biological trait variations among the co-circulating infectious agents. This means that parametrization of epidemiological coinfection SIS models of this sort with interacting strains, from strain frequency temporal data, could become substantially more efficient, even under stochasticity, observation process and measurement error. Many avenues for links with data can now be explored.

3.7 Discussion

Coinfection is an important aspect of many infectious diseases, and a metaphor applied to model also certain information propagation and species colonization processes. There are substantial modeling efforts dedicated to co-infection in the last decades [4, 127, 90, 7, 98, 61]. Yet, simple and sufficiently general mathematical frameworks to analyze and unify the full spectrum of hierarchical patterns emerging from co-infection interactions

and variation in other fitness dimensions between two strains are missing. Here, we contribute to fill this gap, thanks to a model reduction obtained after assuming strain similarity [120]. Focusing on $N = 2$, here we have modeled simultaneously 5 fitness dimensions where two strains can differ, and used the decomoposition into two timescales to simplify their dynamics: neutral dynamics between types on a fast timescale and non-neutral selective processes on a slow timescale, driven explicitly by trait variation, going beyond [86] where only pairwise vulnerabilities to co-infection (a single 'trait') were studied.

Many studies of coinfection are either totally epidemiological in nature [145, 195, 127], exploring transmission dynamics of infectious agents in a host population, or they focus on evolution of specific pathogen traits (often virulence) [154, 7] using the coinfection framework developed by [199] for microparasites causing persistent infections. The latter group of studies, typically derive the conditions of invasion of a rare mutant in a host population already infected by a resident strain, following adaptive dynamics theory [149], where it is further assumed that the resident strain is at equilibrium, that is, that the densities of susceptible, singly-infected and coinfected hosts have reached their equilibrium values. Invader fitness is then evaluated using the basic reproduction ratio [199], where it becomes clear that the fitness of a mutant strain is the sum of two components: the fitness achieved through the infection of susceptible hosts and the fitness achieved through the infection of hosts already infected by the resident. Typically by analyzing whether the reproduction ratio is greater than or lower than 1, conditions for successful or unsuccessful invasion, and ultimate evolutionary dynamics for the trait in consideration are established. However, sometimes the criteria derived in such models can be model-dependent, involve cumbersome mathematical expressions, and may not provide immediate comprehensive insight into the biological mechanisms. Furthermore, explicit frequency dynamics post-invasion are typically not derived or elaborated upon.

In the present work, we have bridged these fields of study. With a rather generic model, we have revealed coexistence and competition mechanisms in their bare essence, and have integrated, generalized and advanced analytically the epidemiological and adaptive dynamics perspectives on coinfection. We have linked population dynamics of endemic transmission with slow selective dynamics in strain trait space, and shown that such dynamics are given by a replicator equation involving the mutual invasion fitness matrix between strains (Eq.(3.2.7)). We have generalized single trait evolution to multiple trait evolution, exploring phenotypic differentiation along 5 dimensions between two strains: transmission, clearance, vulnerability to coinfection, duration of coinfection, and transmission biases from mixed coinfection, all of which contribute to mutual invasion fitness (Eq.3.3.1). We have illustrated the utility of an analytical expression for explicit frequency dynamics between two strains, under an endemic global equilibrium, which allows to use the well-known replicator equation to make predictions for exclusion vs. coexistence (Section 3, Eq.3.5.1), relative strain abundances over time, the means of different traits, and exposes how the means of different traits act to shape their slow variance dynamics.

By elaborating several systematic examples, we have shown that coinfection effects can be very complex on the epidemiological competition between two strains. Indeed, high coinfection prevalence (small μ here) is not always a promoter of coexistence; instead, its effect to generate or prevent polymorphism is non-monotonic, and crucially depends on the type and level of existing phenotypic differentiation between strains. We have also proven formally what type of trait variation is required for 2, 3, and 4 qualitative shifts in the same system as a function of coinfection prevalence alone (Table 3.3). This may prove extremely relevant when interpreting different ecological outcomes potentially arising between the same strains when they occur together in different geographical locations, environmental settings, temperature, resources or other biophysical gradients that act on R_0 and k.

Although we have illustrated special cases, where the traits are uncorrelated, possible co-variation constraints or trade-offs between different traits (e.g. transmission-clearance, or transmission -competition in co-colonization) can be studied under the same analytically-explicit framework, especially for higher N, provided they do not violate the similarity assumption. One mathematical requirement for the singular perturbation expansion to hold is small ε. Yet, numerically we find that the slow-fast approximation remains valid and quantitatively reasonably close to the original system even for values of ε in the order of 0.2-0.3, expanding its applicability. On the other hand, full characterization of the selective dynamics between two strains for small ε, as provided by our framework, has utility for studying local bifurcations near neutrality and the origin of speciation in such systems. Inevitably, the feature of expressing multivariate phenotypic differentiation into a common currency, ε, is central to the slow-fast representation [120], and restricts somewhat the types of systems that can be modelled to those where multiple traits between strains display similar variance. In practice, there are many ways to define the neutral model for the original system 2.2.1 starting from given parameters. Although they will lead to slightly different fast-slow dynamics, all of them will eventually be $\varepsilon-$ close to the real dynamics and preserve its key features.

It is important to keep in mind that our model is deterministic in nature, and as such it cannot account for the subtleties in selection dynamics induced by demographic stochasticity, possibly even in cases where net invasion fitnesses fully equalize between the two strains. It has been shown using simpler population models that how net fitness is realized in terms of vital rates between two species can have drastic effects on stochastic extinction or mutual coexistence [168, 124, 55]. This aspect in our model remains to be studied in the future with new frameworks that incorporate stochasticity. Stochastic fluctuations could impact strongly the dynamics especially in those cases where bistable outcomes (alternative stable states for general N) are predicted from the deterministic model. Similarly, extensions in model structure, for example sequential instead of direct clearance from coinfection, are likely to require specific investigation to establish whether a similar replicator-like equation is possible to derive using the singular perturbation approach, and determine the eco-evolutionary feedbacks governing strain selection. The very same techniques used here to explore the role of coinfection can have may potential advantages if applied in the context of more complex host-pathogen systems where host population is structured differently, for example heterogeneity in host-specific traits, or contact networks [49, 100]. Another path for future exploration is clearly linking our theoretical results to the broader theoretical developments in quantitative genetics, also based on Taylor expansions around mean traits, for example the oligomorphic approximation by [184] for examining the joint ecological and evolutionary dynamics of populations with multiple interacting morphs, and recent extensions thereof [126].

However, even if arising from a specific epidemiological context, we believe our analytical approach, derived for general number of strains N in [120], starting with a complete coinfection structure, involving both same strain and mixed co-infection as in [199, 6, 90], and made possible via separation of timescales and model reduction under strain similarity [86, 140], can have countless further theoretical and practical applications, not least by harnessing the elegant simplicity and historical study record of the replicator equation [103]. Understanding more deeply and in explicit time-scales multiple-trait selection in systems with co-infecting interacting entities will enable easier and more direct insights on several important health challenges, including antibiotic resistance, virulence evolution, optimal immunization, and patterns of diversity in multi-strain endemic pathogens or colonization-based ecosystems.

Code availability

An illustrative code that simulates the dynamics between two strains in our SIS model with coinfection and multi-trait variation, under the proposed slow-fast approximation and replicator equation, is provided on Github.

S3.1 Supplementary information

S3.1.0.0.1 Disentangling how multiple traits drive 2 strain frequencies in SIS dynamics with coinfection Thi Minh Thao Le [1,*], Sten Madec [1,+], Erida Gjini [2,†]

[1] Laboratory of Mathematics, University of Tours, Tours, France,

[2] Center for Stochastic and Computational Mathematics, Instituto Superior Tecnico, Lisbon, Portugal

*Thi-minh-thao.Le@lmpt.univ-tours.fr,+Sten.Madec@lmpt.univ-tours.fr,† erida.gjini@tecnico.ulisboa.pt

S3.2 Strain-specific R_0 in co-infection vs. actual trait variation

We may consider the relations between the basic reproductive numbers $R_{0,1}$ and $R_{0,2}$ in determining the winner in the case of exclusion. We recall the basic reproductive number of each strain, see [120] as follows

$$R_{0,1} = \frac{\beta}{m} + \varepsilon \frac{\beta}{m}\left(b_1 - \frac{\gamma \nu_1}{m}\right) + O\left(\varepsilon^2\right), \qquad R_{0,2} = \frac{\beta}{m} + \varepsilon \frac{\beta}{m}\left(b_2 - \frac{\gamma \nu_2}{m}\right) + O\left(\varepsilon^2\right) \tag{S3.2.1}$$

which implies

$$R_{0,1} \geq R_{0,2} \iff b_1 - b_2 \geq \frac{\gamma}{m}(\nu_1 - \nu_2). \tag{S3.2.2}$$

Strain 1 has higher strain-specific R_0 if and only if its advantage in transmission is bigger than its relative disadvantage in clearance rate, weighted by the global reproduction number.

For any given value $R_{0,1}$, when we fix any value of ν_1, we can find unique value b_1 such that strain 1 is associated to b_1 and ν_1 and has the basic reproductive number $R_{0,1}$. The same holds true for strain 2. This is plausible because of the formulas $R_{0,i} = \dfrac{\beta_i}{m_i} = \dfrac{\beta(1+\varepsilon b_i)}{m+\varepsilon\gamma\nu_i}$.
Hence, with only values of basic reproductive number $R_{0,1}$ and $R_{0,2}$, in a coinfection model, we cannot determine the long time behavior of the dynamics. Figure S1 is actually a vertical slice in each sub figure of figure S6 when we keep $\Delta\nu$ unchanged and change the value of Δb to vary $(R_{0,1}, R_{0,2})$. It is consistent with figure S6 when the smaller $\Delta\nu$ leads to the smaller possibility of coexistence in the same range of Δb but the possibility of the exclusion of strain 2 stays the same.

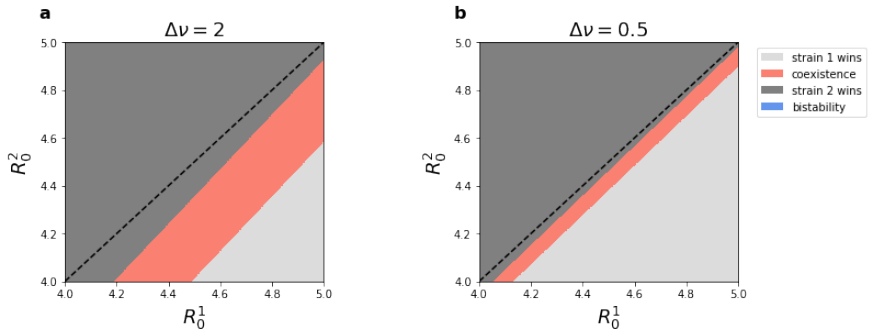


Fig. S1 **Ecological scenarios do not depend just on relative basic reproduction numbers of strains $R_{0,1}$ and $R_{0,2}$.** We compare the case of large variation in clearance rate between two strains $\nu_1 = -1$, $\nu_2 = 1$ (**a**) with lower one $\nu_1 = -0.25$, $\nu_2 = 0.25$ (**b**). The other parameters are assumed: $k = 3$, $m = 1.5$, $\beta = 6$, $\varepsilon = 0.1$. We vary Δb in each case to obtain the relative comparison between $R_{0,1}$ and $R_{0,2}$ shown in the figures. The variations are in transmission rates β_i, single clearance rates γ_i, and for coinfection clearance rate we assume γ_{ij} with $u_{12} = u_{21} = \min\{\nu_1, \nu_2\}$. Comparing **(a)** and **(b)**, for the same combination of strain-specific R_0, the scenarios can diverge depending on the actual difference in strain-specific parameters leading to the particular reproduction number. Particularly, where the $\Delta\nu$ is higher (a), the relative benefit of strain 2 from lower clearance in mixed co-colonization is larger, leading to a larger strain-2 only region, and a larger coexistence region, over the same R_0 range. (Code)

S3.3 Speed of strain dynamics depends on global parameters

Our model allows explicit quantification of the speed of strain dynamics as a function of epidemiological parameters. Next we illustrate a dynamics example for a 2-strain system tending to exclusion. For the same relative variation between two strains, the dynamics are much faster when R_0 is lower, in this case obtained by changes in β. For the dynamics in figure S2, we calculate the values of theta's and pairwise invasion fitnesses

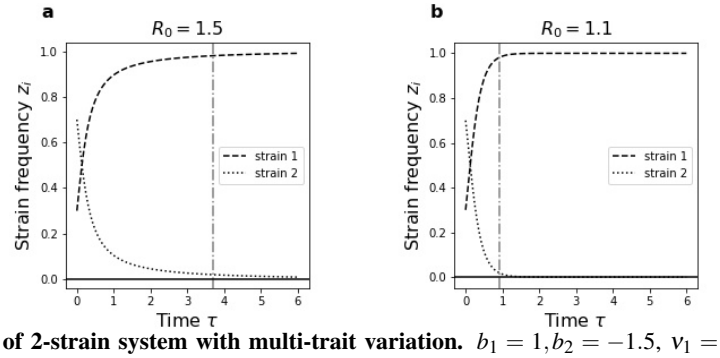


Fig. S2 **An example of 2-strain system with multi-trait variation.** $b_1 = 1, b_2 = -1.5$, $\nu_1 = -2, \nu_2 = -3$, $\left(u_{ij}\right)_{ij} = \begin{pmatrix} 1 & -2 \\ -2 & 3 \end{pmatrix}$, $\left(\omega_{ij}^i\right)_{ij} = \begin{pmatrix} -1 & -3 \\ -1 & 3 \end{pmatrix}$ and $\left(\alpha_{ij}\right)_{ij} = \begin{pmatrix} -\sqrt{2} & 0 \\ \sqrt{2} & 0 \end{pmatrix}$. We choose mean transmission rate $\beta = 4$, mortality $r = 0.2$, mean co-colonization interaction factor (altered susceptibility to co-colonization) $k = 1.5$, and the initial values of strain frequencies $(z_1, z_2) = \begin{pmatrix} 0.3 & 0.7 \end{pmatrix}$. Both figures present periods of transient coexistence before the exclusion of strain 1. The dashed-dotted grey vertical lines, denoted by d in two figures show the region where $\min(z_1, z_2) > 0.02$. Figure **(a)** plots for $R_0 = 1.5$, in which, d goes through a point between 3 and 4 (near 4). Figure **(b)** plots for $R_0 = 1.1$, in which, d goes through a point very near and less than 1. Thus, it can be seen that the the region of "effective transient coexistence" when $R_0 = 1.5$ is larger than when $R_0 = 1.1$.

in each sub figure as follows:

- **(a)**: $\begin{pmatrix} \theta_1 & \theta_2 & \theta_3 & \theta_4 & \theta_5 \end{pmatrix} \approx \begin{pmatrix} 0.48 & 0.2 & 0.09 & 0.03 & 0.2 \end{pmatrix}$, and $\left(\lambda_1^2, \lambda_2^1\right) \approx (10.37, -0.20)$.
- **(b)**: $\begin{pmatrix} \theta_1 & \theta_2 & \theta_3 & \theta_4 & \theta_5 \end{pmatrix} \approx \begin{pmatrix} 0.51 & 0.39 & 0.03 & 0.07 & 0.0 \end{pmatrix}$, and $\left(\lambda_1^2, \lambda_2^1\right) \approx (8.28, -4.42)$.

It is easy to verify that because of the values of λ_i^j, we have exclusion of strain 2 from the system and only strain 1 persists, but depending on the value of R_0 the dynamics will be quicker or slower. In this particular case, reducing R_0, hence overall endemic prevalence of the two strains, leads to faster exclusion dynamics and shorter period of transient coexistence.

S3.4 Qualitative transitions in the same system when varying μ

S3.4.1 A result for variation occurring only in two out of: i) transmission rates β_i, ii) clearance rates γ_i and ii) transmission biases from coinfection p_{ij}^i

In this subsection, we prove that, in the case of variation between 2 strains only, can be in transmission rates β_i, clearance rates γ_i and transmission probability from co-colonized hosts, there is at most one shift for the ecological outcome as a function of μ (single to coinfection ratio in the system). Indeed, we show this result in three cases.

Case 1: Transmission rates β_i and clearance rates γ_i vary between strains.

By formulae (3.4.1) with $\Delta\omega = 0$ we have that mutual invasion fitnesses are:

$$\begin{cases} \lambda_1^2 = \dfrac{\gamma}{2(\gamma+r)}\mu(2\mu+1)\Delta\nu + (\mu+1)^2\Delta b \\ \lambda_2^1 = -\lambda_1^2 \end{cases}. \tag{S3.4.1}$$

It suffices to show that the following equation can not have two positive roots, when solved for μ

$$\frac{\gamma}{2(\gamma+r)}\mu(2\mu+1)\Delta\nu + (\mu+1)^2\Delta b = 0.$$

Assume the contradiction, which means that it has two positive roots, which can be denoted by μ_1 and μ_2. By Viete's theorem, we have that

$$\mu_1+\mu_2 = -\frac{\frac{\gamma}{2(\gamma+r)}\Delta\nu + 2\Delta b}{\frac{\gamma}{\gamma+r}\Delta\nu + \Delta b} > 0, \qquad \mu_1\mu_2 = \frac{\Delta b}{\frac{\gamma}{\gamma+r}\Delta\nu + \Delta b} > 0. \tag{S3.4.2}$$

Thus, (S3.4.2) implies $\mu_1+\mu_2+\dfrac{3}{2}\mu_1\mu_2 > 0$, which is equivalent to

$$-\frac{\frac{\gamma}{2(\gamma+r)}\Delta\nu + 2\Delta b}{\frac{\gamma}{\gamma+r}\Delta\nu + \Delta b} + \frac{\frac{3}{2}\Delta b}{\frac{\gamma}{\gamma+r}\Delta\nu + \Delta b} > 0,$$

which is absurd because the left-hand side is equal to $-\dfrac{1}{2}$.

Case 2: Clearance rates γ_i and transmission probability from coinfected hosts p_{ij}^i vary.

By formulae (3.4.1) with $\Delta b = 0$ we have that mutual invasion fitnesses are given by:

$$\begin{cases} \lambda_1^2 = \dfrac{\gamma}{2(\gamma+r)}\mu(2\mu+1)\Delta\nu + (\mu+1)\Delta\omega \\ \lambda_2^1 = -\lambda_1^2 \end{cases}. \tag{S3.4.3}$$

It suffices to show that the following equation can not have two positive roots, when solved for μ

$$\frac{\gamma}{2(\gamma+r)}\mu(2\mu+1)\Delta\nu + (\mu+1)\Delta\omega = 0.$$

Assume the contradiction, which means that it has two positive roots, which can be denoted by μ_1 and μ_2. By Viete's theorem, we have that

$$\mu_1+\mu_2 = -\frac{\frac{\gamma}{2(\gamma+r)}\Delta\nu + \Delta\omega}{\frac{\gamma}{\gamma+r}\Delta\nu} > 0, \qquad \mu_1\mu_2 = \frac{\Delta\omega}{\frac{\gamma}{\gamma+r}\Delta\nu} > 0. \tag{S3.4.4}$$

Thus, (S3.4.4) implies $\mu_1+\mu_2+\mu_1\mu_2>0$, which is equivalent to

$$-\frac{\frac{\gamma}{2(\gamma+r)}\Delta\nu+\Delta\omega}{\frac{\gamma}{\gamma+r}\Delta\nu}+\frac{\Delta\omega}{\frac{\gamma}{\gamma+r}\Delta\nu}>0,$$

which is absurd because the left-hand side is equal to $-\frac{1}{2}$.

Case 3: Transmission probabilities from coinfected hosts p_{ij}^i and transmission rates β_i vary.

By formulae (3.4.1) with $\Delta\nu=0$ we have that

$$\begin{cases}\lambda_1^2=(\mu+1)^2\Delta b+(\mu+1)\Delta\omega\\ \lambda_2^1=-\lambda_1^2\end{cases}. \tag{S3.4.5}$$

The quadratic equation

$$(\mu+1)^2\Delta b+(\mu+1)\Delta\omega=0$$

has one root to be -1. Hence, it can not have two positive roots, which implies our requirement.

S3.4.2 The case of four ecological outcomes in the same system according to μ

It is possible to have four distinct ecological outcomes between which the same system with two strains can shift, as a function of global μ. Initially, we will prove that the necessary condition for the presenting of fully four survival outcomes: E1, E2, C, B is that, variations are in coinfection clearance rates γ_{ij} and co-colonization interaction k_{ij}.

Proof. • **Firstly, we prove that if variations are only in β_i, γ_i, p_{ij}^i and k_{ij} then we can not have four survival scenarios as $\mu\to\infty$.**
The formulas for pairwise invasion fitnesses are

$$\begin{aligned}\lambda_1^2=&\frac{\gamma}{2(\gamma+r)}\mu(2\mu+1)\Delta\nu+(\mu+1)\Delta\omega+\frac{(R_0-1)\mu}{2}(\mu(\alpha_{21}-\alpha_{12})+\alpha_{21}-\alpha_{22})+(\mu+1)^2\Delta b\\ \lambda_2^1=&-\frac{\gamma}{2(\gamma+r)}\mu(2\mu+1)\Delta\nu-(\mu+1)\Delta\omega+\frac{(R_0-1)\mu}{2}(-\mu(\alpha_{21}-\alpha_{12})+\alpha_{12}-\alpha_{11})-(\mu+1)^2\Delta b\end{aligned}. \tag{S3.4.6}$$

We have that two equations $\lambda_1^2=0$ and $\lambda_2^1=0$ are respectively equivalent to

$$(\mathbf{Eq1}):\quad A_1\mu^2+B_1\mu^2+C_1=0,\qquad A_1>0$$

and

$$(\mathbf{Eq2}):\quad A_2\mu^2+B_2\mu^2+C_2=0,\qquad A_2>0,$$

in which $A_1=A_2$ and $C_1=C_2$, easily deduced from (S3.4.6).

On the other side, by direct verification, to have fully four outcomes, the signs of $\left(\lambda_1^2,\lambda_2^1\right)$ must change at least three times, so one in two equations **(Eq 1)** or **(Eq 2)** must have two distinguished positive solutions and the other must have at least one positive solution. By Viete's theorem, the products of solutions in two equations **(Eq 1)** and **(Eq 2)** are equal. Hence, **(Eq 1)** and **(Eq 2)** must have both positive solutions.

Denote two positive solutions of **(Eq 1)** are $x_1 < x_2$, two positive solutions of **(Eq 2)** are $y_1 < y_2$.
By direct checking, and note that when $\mu \to +\infty$ or $\mu \to -\infty$, $\lambda_1^2 + \lambda_2^1 \to 0$, we have that, to have fully four outcomes, it must be

$$x_1 < y_1 < x_2 < y_2$$

or

$$y_1 < x_1 < y_2 < x_2.$$

Both of these inequality requires the products of solutions in **(Eq 1)** is strictly less or more than the products of solutions in **(Eq 2)**, which is a contradiction of the equality of two products mentioned before.
Thus, if variations are only in β_i, γ_i, p_{ij}^i and k_{ij} then we can not have four survival scenarios as $\mu \to \infty$.

• Secondly, we prove that if variations are only in β_i, γ_i, γ_{ij} and p_{ij}^i then we can not have four survival scenarios as $\mu \to \infty$.

The formulas for pairwise invasion fitnesses now read

$$\begin{aligned} \lambda_1^2 &= \frac{\gamma}{2(\gamma+r)}\mu(2\mu+1)\Delta\nu + \frac{\gamma}{2(\gamma+r)}(\mu+1)\Delta_2 u + (\mu+1)\Delta\omega + (\mu+1)^2\Delta b \\ \lambda_2^1 &= -\frac{\gamma}{2(\gamma+r)}\mu(2\mu+1)\Delta\nu + \frac{\gamma}{2(\gamma+r)}(\mu+1)\Delta_1 u - (\mu+1)\Delta\omega - (\mu+1)^2\Delta b \end{aligned}. \tag{S3.4.7}$$

For the sake of simplicity, we denote that $m := \frac{2(\gamma+r)}{\gamma}$, $d_1 := -u_{12} - u_{21} + 2u_{22}$ and $-d_2 := -u_{12} - u_{21} + 2u_{11}$.

We have that two equations $\lambda_1^2 = 0$ and $\lambda_2^1 = 0$ are respectively equivalent to

$$(\mathbf{Eq3}): \quad A_1'\mu^2 + B_1'\mu^2 + C_1' = 0, \qquad A_1' > 0$$

and

$$(\mathbf{Eq4}): \quad A_2'\mu^2 + B_2'\mu^2 + C_2' = 0, \qquad A_2' > 0,$$

now in which

$$A_1' = 2\Delta\nu + m\Delta b, \qquad B_1' = \Delta\nu + d_1 + 2m\Delta b, \qquad C_1' = d_1 + m\Delta b$$

and

$$A_2' = 2\Delta\nu + m\Delta b, \qquad B_2' = \Delta\nu + d_2 + 2m\Delta b, \qquad C_2' = d_2 + m\Delta b.$$

Denote two solutions of **(Eq 3)** to be $x_3 < x_4$, and two solutions of **(Eq 4)** to be $y_3 < y_4$. By the same arguments, it can be deduce that, to have fully four survival outcome, it must be

$$(*) \qquad x_3 < y_3 < x_4 < y_4$$

or

$$(**) \qquad y_3 < x_3 < y_4 < x_4$$

where at least $0 < x_3 < x_4$ or $0 < y_3 < y_4$. Without loss of generality, similarly to the previous arguments, we assume that $0 < x_3 < x_4$ and $y_4 > 0$.

According to Viete's theorem, we obtain, note that $A_1' = A_2'$

$$x_3 + x_4 = -\frac{B_1'}{A_1'}, \qquad x_3 x_4 = \frac{C_1'}{A_1'},$$

$$y_3 + y_4 = -\frac{B_2'}{A_2'}, \qquad y_3 y_4 = \frac{C_2'}{A_2'}.$$

Case 1: $A_1' = A_2' > 0$.

- If $d_1 > d_2$ then $-\frac{B_1'}{A_1'} < -\frac{B_2'}{A_2'}$ and $\frac{C_1'}{A_1'} > \frac{C_2'}{A_2'}$, which implies $x_3 + x_4 < y_3 + y_4$ and $x_3 x_4 > y_3 y_4$. From $x_3 + x_4 < y_3 + y_4$, then $0 < x_3 < y_3 < x_4 < y_4$, which contradicts $x_3 x_4 > y_3 y_4$.
- If $d_1 < d_2$ then $-\frac{B_1'}{A_1'} > -\frac{B_2'}{A_2'}$ and $\frac{C_1'}{A_1'} < \frac{C_2'}{A_2'}$, which implies $x_3 + x_4 > y_3 + y_4$ and $x_3 x_4 < y_3 y_4$. From $x_3 x_4 < y_3 y_4$, then $0 < x_3 < y_3 < x_4 < y_4$, which contradicts $x_3 + x_4 > y_3 + y_4$.

Case 2: $A_1' = A_2' < 0$.

- If $d_1 > d_2$ then $-\frac{B_1'}{A_1'} > -\frac{B_2'}{A_2'}$ and $\frac{C_1'}{A_1'} < \frac{C_2'}{A_2'}$, which implies $x_3 + x_4 > y_3 + y_4$ and $x_3 x_4 < y_3 y_4$. From $x_3 x_4 < y_3 y_4$, then $0 < x_3 < y_3 < x_4 < y_4$, which contradicts $x_3 + x_4 > y_3 + y_4$.
- If $d_1 < d_2$ then $-\frac{B_1'}{A_1'} < -\frac{B_2'}{A_2'}$ and $\frac{C_1'}{A_1'} > \frac{C_2'}{A_2'}$, which implies $x_3 + x_4 < y_3 + y_4$ and $x_3 x_4 > y_3 y_4$. From $x_3 + x_4 < y_3 + y_4$, then $0 < x_3 < y_3 < x_4 < y_4$, which contradicts $x_3 x_4 > y_3 y_4$.

Hence, if variations are only in β_i, γ_i, γ_{ij} and p_{ij}^i then we can not have four survival scenarios as $\mu \to \infty$. □

In the main text, we give an example, in which varying μ from 0 to ∞ may give us 4 outcome exclusion of either strain, coexistence and bistability.

S3.5 Examples for 3 possible global outcomes

We illustrate possible outcomes as a function of Δb and $\Delta \nu$, similarly to Figure 3.6, when the perturbations are only in transmission rates β_i, clearance rates γ_i and co-colonization interactions k_{ij}. We recall that the borders separating exclusion regions are lines representing $\lambda_1^2 = 0$ and $\lambda_2^1 = 0$.

According to the explicit formulas of $\left(\lambda_1^2, \lambda_2^1\right)$, the border lines have the same slope, thus leading to parallelism. Thus, there are at most 3 possible outcomes for each fixed value (R_0, k). Figures S3 and S4 shows that changing the matrix (α_{ij}) may generate different final outcomes.

According to (3.3.1), we can deduce that regardless of changing (α_{ij}), we can not observe coexistence and bistability for a fixed value of (α_{ij}). Indeed, according (3.3.1), two coefficients of $\Delta \nu$ in the formulae of λ_1^2 and λ_2^1 are of opposite signs, and the same holds for Δb.

Then, for $(\Delta \nu, \Delta b) \to (\infty, \infty)$ and $(-\infty, -\infty)$, we have the opposing signs of $\left(\lambda_1^2, \lambda_2^1\right)$, leading to exclusion.

S3.6 The analytical expression for coexistence prevalence

Here we compute the value of z_2^* in the case of symmetric co-colonization interaction matrix. Thus: $\alpha_{ij} = \alpha_{ji}$, $\alpha_{ii} = \alpha_{jj}$. We also assume that strain-specific transmission probability from mixed coinfected hosts is in this case $p_{ij}^i = \frac{1}{2}$ (no priority effects).

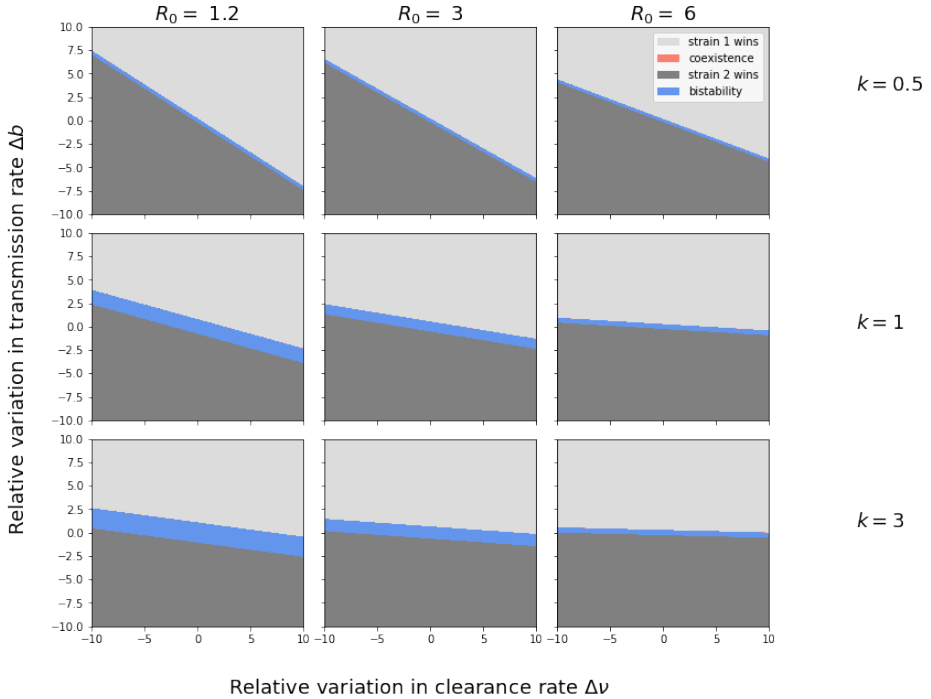


Fig. S3 **Illustration of 3 possible outcomes, as a function of $\Delta\nu$ and Δb, for different values of k and R_0.** We highlight the respective regions in different colors, according to the critical relationship between Δb, $\Delta\nu$, k and R_0 when perturbations happen only β_i, γ_i, and k_{ij} with $\gamma = 1$, $r = 0.2$ and the matrix (α_{ij}) be $\begin{pmatrix} 0 & -\sqrt{2} \\ -\sqrt{2} & 0 \end{pmatrix}$. Three possible survival outcomes include the exclusion of strain 1, bistability and the exclusion of strain 2.

- Perturbation in $\beta_i, \gamma_i, \gamma_{ij}, k_{ij}$

$$z_2^* = \frac{1}{2} \cdot \frac{-\Delta b - \dfrac{\theta_2}{\theta_1}\Delta\nu + \dfrac{\theta_3}{\theta_1}(2u_{11} - u_{12} - u_{21}) + \dfrac{\theta_5}{\theta_1}(\alpha_{12} - \alpha_{11})}{\dfrac{\theta_3}{\theta_1}(u_{11} + u_{22} - u_{12} - u_{21}) + \dfrac{\theta_5}{\theta_1}(\alpha_{12} - \alpha_{11})} \tag{S3.6.1}$$

- Perturbation in $\beta_i, \gamma_i, k_{ij}$

$$\begin{aligned} z_2^* &= \frac{1}{2} \cdot \frac{-\Delta b - \dfrac{\theta_2}{\theta_1}\Delta\nu + \dfrac{\theta_5}{\theta_1}(\alpha_{12} - \alpha_{11})}{\dfrac{\theta_5}{\theta_1}(\alpha_{12} - \alpha_{11})} = \frac{1}{2}\left(\frac{-\Delta b - \dfrac{\theta_2}{\theta_1}\Delta\nu}{\dfrac{\theta_5}{\theta_1}(\alpha_{12} - \alpha_{11})} + 1\right) \\ &= \frac{1}{2}\left(-\frac{1}{\alpha_{12} - \alpha_{11}} \cdot \frac{2(\gamma + r) + k(\beta - \gamma - r)}{k(\beta - \gamma - r)^2} \cdot \gamma\Delta\nu - \frac{1}{\alpha_{12} - \alpha_{11}} \cdot \frac{2(\gamma + r + k(\beta - \gamma - r))^2}{k(\beta - \gamma - r)^2}\Delta b + 1\right). \end{aligned} \tag{S3.6.2}$$

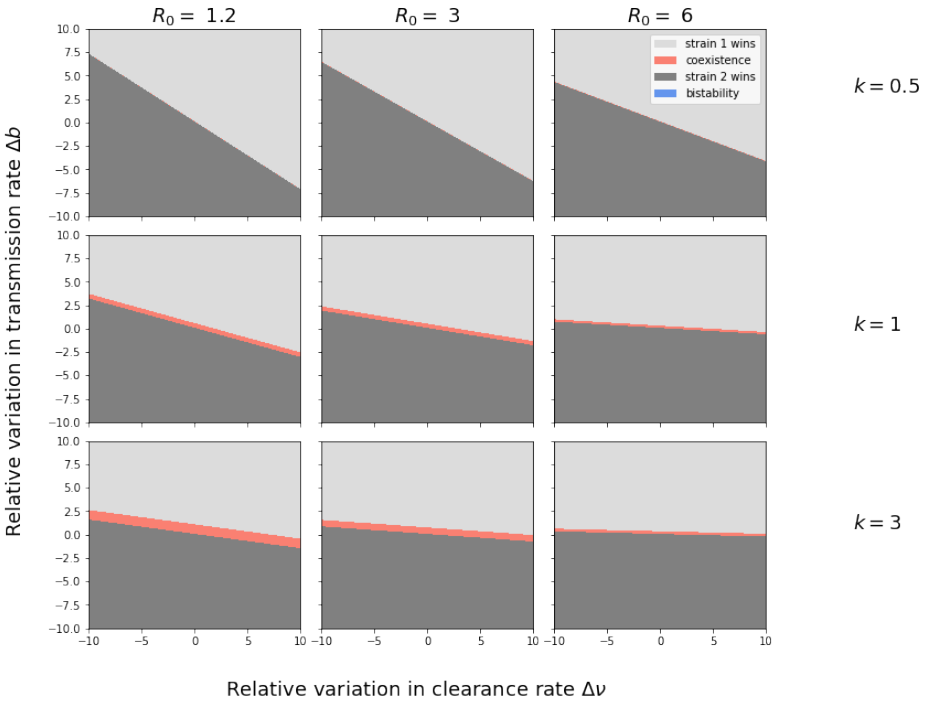


Fig. S4 **Illustration of 3 possible outcomes, as a function of difference in transmission and clearance rates between two strains, for different values of k and R_0.** We highlight the respective regions in different colors, according to the critical relationship between Δb, $\Delta \nu$, k and R_0 when perturbations happen only β_i, γ_i, and k_{ij} with $\gamma = 1$, $r = 0.2$ and the matrix (α_{ij}) be $\begin{pmatrix} -2 & 0 \\ 0 & 0 \end{pmatrix}$. This changing of matrix (α_{ij}) flips the position of two lines $\lambda_1^2 = 0$ and $\lambda_2^1 = 0$ leads to the other possible outcomes. Three possible system outcomes include: the exclusion of strain 1, coexistence and the exclusion of strain 2.

S3.6.1 Studying the monotonicity of z_2^*: increasing vs. decreasing equilibrium strain frequency

For the sake of simplicity, and for the purposes of illustration, we consider the equilibrium of resistance strain in section 4. Recall the formula of z_2^* in (3.5.9)

$$z_2^* = \frac{1}{2} - \Delta\nu \cdot \gamma \frac{2(\gamma+r)+k(\beta-\gamma-r)}{2\sqrt{2k}(\beta-\gamma-r)^2} - \Delta b \frac{[\gamma+r+k(\beta-\gamma-r)]^2}{\sqrt{2k}(\beta-\gamma-r)^2}. \tag{S3.6.3}$$

To investigate the monotonicity of strain frequency, we need to compute the first partial derivative of z_2^* with respect to β and γ, noting that $m = \gamma + r$ and investigate its sign:

$$\frac{\partial z_2^*}{\partial \beta} = -\Delta\nu \frac{1}{2\sqrt{2k}} \frac{(k-4)m-\beta k}{(\beta-m)^3} - \Delta b \frac{1}{\sqrt{2k}} \frac{2m((k-1)m-\beta k)}{(\beta-m)^3}, \tag{S3.6.4}$$

and

$$\frac{\partial z_2^*}{\partial \gamma} = -\Delta\nu \frac{1}{2\sqrt{2k}} \frac{\beta(\beta k-km+4m)-r((k+2)\beta-(k-2)m)}{(\beta-m)^3} - \Delta b \frac{1}{\sqrt{2k}} \frac{2\beta(\beta k-km+m)}{(\beta-m)^3}. \tag{S3.6.5}$$

It can be seen that the expressions are entirely explicit but they display nonlinear dependence on many parameters, including strain variation $(\Delta b, \Delta \nu)$, as well as mean parameters β, m or coinfection determinants (k, etc.). This allows to obtain a full analytic understanding of the coexistence between two strains at the epidemiological level and how their relative hierarchies depend and may shift with overall context.

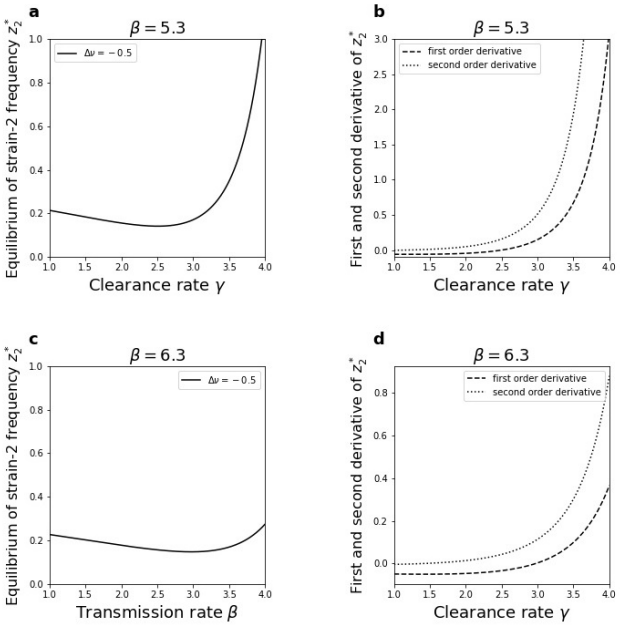


Fig. S5 **Monotonicity and convexity of z_2^* according to transmission rate β and clearance rate γ.** We choose $r = 0.2$, $k = 1$ and matrix including entries α_{ij} is $\begin{pmatrix} 0 & \sqrt{2} \\ \sqrt{2} & 0 \end{pmatrix}$. The behavior observed is by plotting for $\Delta\nu = -0.5$. In **(a)** and **(c)**, setting $\Delta b = 0.3$, we plot z_2^* according to γ with $\beta = 5.3$ and $\beta = 6.3$ respectively. Figures **(c)** and **(b)** present the curve of first and second orders derivatives of z_2^* according to γ, respective for figures **(a)** and **(b)**.

S3.6.2 Studying the convexity of z_2^*: accelerating vs. decelerating behavior

In this subsection, we consider the convexity of z_2^* with respect to β and γ, global strain-transcending parameters, that can be controlled via interventions (antibiotic treatment, vaccines, etc.). Thus, we compute the second-order derivative of z_2^* related to β as follows, noting that $m = \gamma + r$:

$$\frac{\partial^2 z_2^*}{\partial \beta^2} = -\Delta\nu \frac{1}{\sqrt{2}k} \frac{6\beta m + (-4m+k)(\beta - m)}{(\beta - m)^4} - \Delta b \frac{\sqrt{2}m(3m - 2km + 2\beta k)}{k(\beta - m)^4}, \tag{S3.6.6}$$

and the second-order derivative of z_2^* related to γ as follows:

$$\frac{\partial^2 z_2^*}{\partial \gamma^2} = -\Delta\nu \frac{\beta k - (k-6)m}{\sqrt{2}k(\beta - m)^4} - \Delta b \frac{2\beta((2k+1) - 2(k-1)m)}{k(\beta - m)^4}. \tag{S3.6.7}$$

These expressions reveal whether the behavior of the strain frequency as a function of global parameters is decelerating or accelerating. This has implications for the control of the strain composition at the population level if we intervene via changing global transmissibility of all strains β or global clearance rate γ. If the behavior is increasing and accelerating, this means that the frequency of that strain can reach fixation for

some value of the global parameter. If the behavior is increasing and decelerating (first derivative positive and second derivative negative) this means that the strain frequency saturates at a particular value < 1 indicating coexistence in the limit of high values of the control parameter. The following simulation in Figure S5 allows us to see how we can use these analytical expressions above to predict exactly the behavior the strain coexistence frequencies as a function of our parameters, and possible responses to interventions in global transmission rate β or clearance rate γ. In these figures, we can see that z_2^* is always convex related to γ in the considered period of varying γ.

S3.6.3 Effect of transmission priority effects from mixed coinfection $\Delta\omega$

To see the effect of transmission biases from mixed coinfection, as in the section 3.4 we plot the values of pairwise invasion fitnesses according to the single to co-colonization ratio μ. From 3.4.1, it is known that if variations are only in transmission rates β_i, clearance rates γ_i and transmission probability from coinfected hosts p_{ij}^i, the outcome is always the exclusion of either strain. In this figure, the perturbations are in transmission rates, clearance rates, transmission probability from co-colonized hosts as well as in co-colonization interaction factor k_{ij} to break the anti-symmetry. Except the transmission biases from mixed coinfection p_{ij}^i favor to strain 2, other trait differences in β_i, γ_i favors to strain 1.

We consider in three cases including co-colonization interaction factor k_{ij} favors to strain 2, disfavors to strain 1 and counter balance, respectively. Although the difference $\Delta\omega$ is bias to strain 2, when μ is large enough, the difference in transmission probability does not effect too much. This can be seen in the similar trending of $\left(\lambda_1^2, \lambda_2^1\right)$ in figures **(a, b, c)** in figure 3.4, whose values of Δb and Δv are kept the same.

These phenomena are plausible because the weight (written in μ) of transmission probability from coinfected hosts is $\mu + 1$ in the formulae of invasion fitnesses (3.4.5). Of course, unlike 3.4.3, when μ is small enough, $\Delta\omega$ leads to the exclusion of strain 2 in figure S6 **(a)** and the exclusion of strain 1 in figure S6 **(b)**. Meanwhile, λ_1^2 is always positive in figure 3.4 **(a)** and λ_2^1 is always negative in figure 3.4 **(b)**.

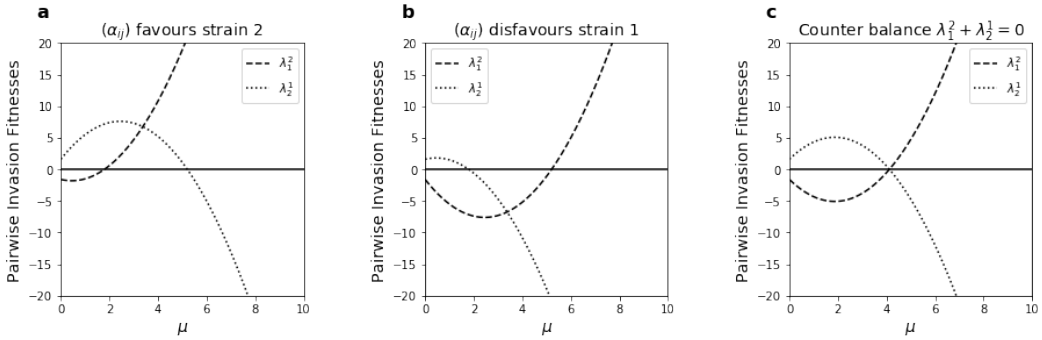


Fig. S6 **Additional effects of within-host transmission advantage from mixed coinfection $p_{ij}^i \neq \frac{1}{2}$ (for comparison with Figure 3.4 a-b-c)** Here we combine variation in co-colonization interactions k_{ij} with variations in transmission rates, infection clearance rates, and transmission probability from coinfected hosts (See Eqs (3.4.9)). We compute pairwise invasion fitnesses $\left(\lambda_1^2, \lambda_2^1\right)$according to μ in various cases of co-colonization interaction matrix $\left(\alpha_{ij}\right)$. We illustrate the cases of transmission and clearance superiority of strain 1 (parameters as in top row of Figure 4): $\Delta b = 0.4$, $\Delta v = 0.8$, $R_0 = 5$, $r = 0.5$ and $\gamma = 1.5$. We choose the value $\Delta\omega = -2$ to increase the advantage in transmission probability of strain 2 from mixed coinfection. We consider 3 structures: **(a)** $\begin{pmatrix} -2 & 0 \\ 0 & 0 \end{pmatrix}$; **(b)** $\begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$; **(c)** $\begin{pmatrix} -\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{pmatrix}$ for variation in co-colonization interactions.

Secondly, as mentioned in Sections 3.6.1 and 3.6.2, we point out that depending on how the fitness differential between strain 1 and its competitor strain 2, is manifested $(\Delta b, \Delta v)$ increasing strain-transcending

clearance or transmission rates (γ or β) in a population can have opposing effects. Besides other trait parameters which are kept unchanged, we analyze the effect of $\Delta\omega < 0$ favouring strain 2, and how this modifies the range of γ guaranteeing the coexistence, originally observed in Figure S7. We can see that the shapes of curves representing z_2^* are similar to the Figures 3.8 **(a, b, c, d)**. respectively. However, in Figures S7 for all **(a, b, c, d)**, for $\Delta\nu > 0$ (which favours strain 1), compared to figures 3.8 **(a, b, c, d)** respectively, the ranges for survival of strain 2 are larger, highlighting the positive effect of its precedence in transmission from mixed coinfection. In contrast to this, when $\Delta\nu < 0$, the range for coexistence decreases, to a larger advantage of strain 2-only persistence. In all cases, the values of equilibrium of strain 2 in Figure S7 corresponding to each parameter γ or β are higher than in Figure 3.8.

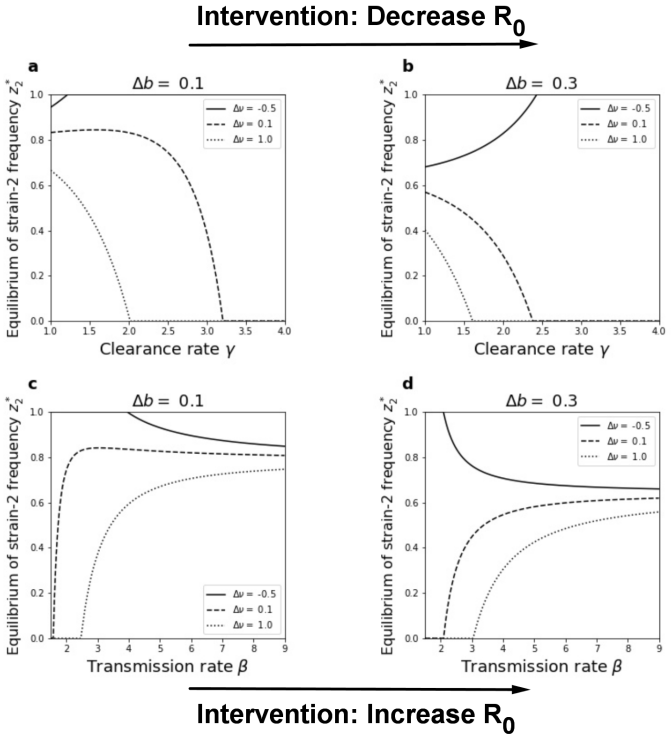


Fig. S7 **Strain frequency at the endemic equilibrium vs. γ and β depends on underlying trait variation (related to Fig.3.8 but with $\Delta\omega \neq 0$).** We plot the prevalence of strain 2, z_2^* at the coexistence equilibrium (Eq. 3.5.9), as a function of mean transmission rate and mean clearance rate. Variation among 2-strains is encoded in the transmission and clearance rate axes: β_i and γ_i, transmission probability from co-colonized hosts p_{ij}^i and co-colonization vulnerabilities k_{ij}. In this simulation, we keep the same values of r, k and the matrix of standardized interactions in Figure 3.8. We choose $\Delta\omega = -0.5$, which favors strain 2 in precedence of transmission from mixed coinfection. **a-b.** We plot the equilibrium z_2^* as a function of mean clearance rate γ (varied between 1 and 4) for 2 cases of fitness differentials in transmission Δb and 3 cases of variability in clearance $\Delta\nu$. The global transmission rate is $\beta = 4.5$ to ensure $R_0 \geq 1$. **c-d.** We plot the equilibrium frequency of strain 2, z_2^*, as a function of mean transmission rate β (varied between 1.5 and 9) for 2 cases of different variation Δb and 3 cases of $\Delta\nu$. In these plots, overall clearance rate is held fixed at $\gamma = 1$ to ensure $R_0 \geq 1$.

S3.7 Error estimates of the slow dynamics approximation

In this section, we present numerical simulations to verify the error estimates when approximating the quasi-neutral system (3.2.1) using the corresponding replicator equation when $N = 2$. We recall the following

estimate in main result of [120], using the solution of the slow-system at time τ to approximate the system of the quasi-neutral system at time $\frac{\tau}{\varepsilon}$, for all $\tau > 0$.

$$\text{Error estimate} = \left|S\left(\frac{\tau}{\varepsilon}\right) - S^*\right| + \sum_{i=1}^{N}\left|I_i\left(\frac{\tau}{\varepsilon}\right) - I^* z_i(\tau)\right| + \sum_{i,j=1}^{N}\left|I_{ij}\left(\frac{\tau}{\varepsilon}\right) - D^* z_i z_j(\tau)\right| \tag{S3.7.1}$$

Using the ODE solving processes of scipy.integrate.odeint for two models (quasi-neutral and slow dyanmics), we also can have the comparison in time complexity, see Figure S8.

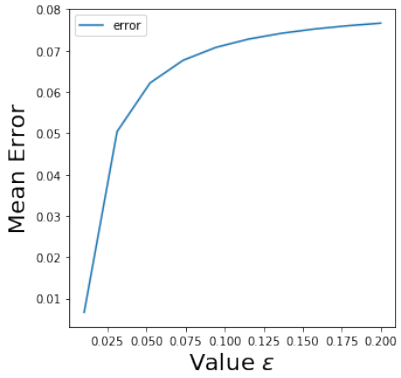


Fig. S8 **Error of the slow dynamics approximation, starting from the slow manifold. The error (mean over all model variables and all time) between the original system trajectories and the slow dynamics.** We create randomly 15 2-strain SIS quasi-neutral systems in which the dynamics start near the slow manifold. Recall that parameters include transmission rates β_i, single clearance rates γ_i, coinfection clearance rates γ_{ij}, transmission probability p^s_{ij} and relative factors of altered susceptibilities to co-colonization k_{ij}, $i,j \in \{1,2\}$. We only fix the value recruitment rate as well as the mortality rate to be 0.3 in this simulation. To ensure these 15 systems are quasi-neutral, we create their parameters using the formulas: $\beta_i = \beta\,(1+\varepsilon_0 b_i)$, $\gamma_i = \gamma(1+\varepsilon_0 \nu_i)$, $\gamma_{ij} = \gamma\left(1+\varepsilon_0 u_{ij}\right)$, $p^s_{ij} = 1/2 + \varepsilon_0 \omega^s_{ij}$ and $k_{ij} = k + \varepsilon_0 \alpha_{ij}$. Choosing $\varepsilon_0 = 0.01$, to create these systems, we generate random values for β, γ, k in their appropriate ranges, and b_i, ν_i, u_{ij}, ω^s_{ij}, α_{ij}, for $i,j \in \{1,2\}$, in their appropriate ranges as well. Then, we solve these SIS systems for time $t = 1000$. For each ε given, we have a replicator system and we solve them for $\tau = \varepsilon t$. For each of these 15 systems, we can compute the error estimate base on the formulas $\|S(\tau) - S^*\| + \sum \|I_i(\tau/\varepsilon) - I^* z_i(\tau)\| + \sum \left\|I_{ij}(\tau/\varepsilon) - D^* z_i(\tau) z_j(\tau)\right\|$, with $\tau = \varepsilon t$. This leads us to observe the average error estimates for each value ε over 15 systems and over all time. As expected from the mathematical theory, the mean error increases with ε, but remains small even for $\varepsilon = 0.2$, pointing to the wider applicability of our method. (Code)

We also note that, the initial values for the replicator system is generated randomly as well. Accordingly, the initial vales for the original quasi-neutral system is calculated using the formula

$$I_i(0) = I^* z_i(0), \qquad I_{ij}(0) = D^* z_i(0) z_j(0), \qquad S_0 = S^* = \frac{m}{\beta}.$$

We can change the range for creating neutral parameters r, β, γ and k to see the different cases. Using the neutral parameters and values S^*, T^*, I^*, D^* computed, we can input and solve the replicator system by its pairwise invasion fitness matrix.

We can note that, there is a trade-off, since the two time-scale in (S3.7.1) is $\dfrac{\tau}{\varepsilon}$ and τ. We can choose a smaller ε but the fast-time that can be used to approximated becomes larger.

Chapter 4

A Reaction-Advection-Diffusion Model for Quasi-neutral Dynamics of Coinfected Strains

This chapter is the work in the paper "A Reaction-Advection-Diffusion Model for Quasi-neutral Dynamics of Coinfected Strains", with Dr. Sten Madec, 2021, see [121].

4.1 Introduction

Heterogeneity is a common feature of real world infections. Heterogeneous susceptibilities may arise, for instance, through individuals having differing histories of prior exposure to infection or vaccination. Thus, it remains challenging to accurately describe diffusion process of bacteria/virus and investigate the transmission dynamics of free-living bacteria/virus in the contaminated environment on disease infection. There are many studies the mathematical framework on the predator-prey models within heterogeneous environment [174, 175]. In particular, many studies deeply solution for compartmental models in epidemiology with diffusion terms. For instants, [205, 204] studies the existence and non-existence of travelling wave solutions for a general class of diffusive Kermack–McKendrick SIR models with nonlocal and delayed disease transmission. However, there is a lack of a comprehensive theoretical framework for spatial models of co-infection though it frequently appears in models with migration, evolution, and heterogeneous environment. It is known that co-infection dynamics have received considerable attention [4, 6, 145], because of their importance to biology, especially in the outbreaks of infectious diseases. For instance, [143, 206] studied different co-infection models to help diagnose and treat infectious diseases.

Even without a spatial structuration, the interactions between traits and strains yield complex consequences on the population dynamics [140, 120]. However, under a quasi-neutral hypothesis, this complexity is decoded into a replicator equation. In a heterogeneous environment, the dynamics surely become a PDE system, which is more complex to studies. In this study, with diffusion terms and under appropriate conditions, the dynamical system of co-infection, now becomes a reaction-advection-diffusion system, will be coded again through a replicator equation, with or without diffusion, depending on the rate of diffusion.

In this article, we describe and study the spatial version of dynamics considered in [120], i.e. the quasi-neutral SIS model between similar strains, with diffusion and zero flux assumption on the boundaries, in two cases, including slow ($\varepsilon\Delta$) and fast diffusions ($\frac{1}{\varepsilon}\Delta$). The choice of terms presenting heterogeneity depends on a priori on the type of population considered. For the case of a large population in a bounded domain, which leads to a large density, diffusion is a good approach to model spatial movements because organisms are assumed to have random motions. In a mathematical sense, the term diffusion-presented by the Laplacian operator is a strong elliptic operator. Most of results of this paper may be extended to more general elliptic operators. For the sake of simplicity, we restrict ourselves to the Laplacian.

We focus on modeling the host-to-host transmission of different strains using the SIS (susceptible-infected-susceptible) modeling approach. Despite the assumption on compactness and smoothness of domain, the main difficulty is to take into account the impact of strain traits under propagation in space. It is useful to take the viewpoint of reaction-diffusion equations, which are studied deeply in [41]. Moreover, the assumptions of zero total flux on the boundary make our system isolated. An important point is that, in the case of slow diffusion, we assume that whole the considered domain is at high risk of infection, which means at any point in the considered domain, the transmission rate of either is larger than the sum of the clearance rate and the mortality rate. This assumption leads to the existence, the uniqueness and the stability of the endemic equilibrium. When low risk site, the set of all points in which the transmission rate is less than the clearance rate, is non empty as well, we also have a disease-free equilibrium, which is studied concretely in [9].

For each type of diffusion, we present a specific method to approximate the solutions under a quasi-neutral assumption on the parameters. In the case of slow diffusion, we first consider the reaction-diffusion model with symmetric interactions, which is the neutral system with diffusion. Similar to [120], and K_f, K_g are operators which are computed later, we find how to rewrite the original system in the form $\frac{\partial}{\partial t}f(x,t) = F(f,g,x,t,\varepsilon) + K_f f$ and $\frac{\partial}{\partial t}g(x,t) = \varepsilon G(f,g,x,t,\varepsilon) + \varepsilon K_g g$, where f describes the fast dynamics and g the slow dynamics. The Tikhonov's theorem used in [120] now is improved to a Tikhonov-like theorem applied for PDEs model with appropriate assumptions. Accordingly, at the slow time scale $\tau = \dfrac{t}{\varepsilon}$, we obtain the slow dynamics on the slow manifold.

For the case of fast diffusion, the Central Manifold Theorem [42] is applied directly on the original SIS system under an appropriate rewriting, yielding to an ordinary differential (ODE) SIS system under the mean variables. In this system, we invoke the quasi-neutral assumptions on the traits to use the main result of [120]. For a clearer view, this theorem plays the main role in [42, 43].

Analogously to the non-spatial models in [120], we obtain the diffusion replicator system at the end and they are in different types due to the distinguished kinds of diffusions. The replicator system with diffusion attracts much attention and be studied in [35]. Comparing two cases of diffusion, the Tikhonov-like approximates the slow-fast form to the replicator system, in which variables are prevalences of strains. Meanwhile, the Central Manifold Theorem leads us to equations of total masses over the domain of susceptible, infected and coinfected strains. Although, both of them claim that the original system's solution can be approached based on the solution of a simpler system in any bounded time interval as $\varepsilon \to 0$. Despite the distinction in variables of system in slow-manifold, error estimates in both cases are computed in L^2 and L^1 norms, respectively.

This article is organized as follows. Sections 4.2 and 4.3 are dedicated to the case when the coefficients of diffusion are ε. In the beginning of section 4.2, we present the model and state some general results including the existence of a unique solution and the introduction of new variables. Next, we analyzes the semi-neutral

system and the slow-fast form to prepare for application the approximation theorems. Similarly to [120], we solve the system with slow diffusion in each elementary sub-case in which only one trait depends on the strains. For this sake, a lemma showing of to combine the elementary cases is presented, starting section 4.3. With these sufficient materials, the replicator system with diffusion follows with proofs and finalizes the case of small diffusion rates. The model with fast diffusion ($\frac{1}{\varepsilon}\Delta$) is studied in section 4.4. We refer the Central Manifold Theorem in [43] and make some conventions at the beginning to apply this result. As mentioned after the application of the Central Manifold Theorem, we invoke the quasi-neutral assumptions on traits. These ingredients are combined and used to derive the replicator system, by the main result in [120], in which the variables are total masses over domain of strains. Section 4.5 is to compare the two cases of diffusion in some respects including the relations with basic reproduction ratio R_0 and three examples for different behaviors. Section 4.6 draws remarkable results and concluding. The final section Appendix A4.1 closes this article with the proofs of the theorems stated in section 4.2.

4.2 General and Semi-neutral Systems with Slow Diffusion

4.2.1 The general N-strain model

The dynamics studied in this article groups the pathogen types in N subsets, indexed by i, $1 \leq i \leq N$. With a set of ordinary differential equations, we then track the proportion of hosts in $1+N+N^2$ compartments: susceptible: $S(x,t)$, hosts colonized by strain-i: $I_i(x,t)$, hosts co-colonized by strain-i then strain-j: $I_{ij(x,t)}$. Notice that we include also same strain coinfection, as argued in [120].

We formulate the general model based on the same structure as that in [120] but here allow for strains to vary in their transmission rates $\beta_i(x)$, clearance rates of single infection $\gamma_i(x)$ (or duration of carriage $1/\gamma_i(x)$), clearance rates from mixed co-colonization $\gamma_{ij}(x)$, within-host competition reflected in relative transmissibilities from mixed coinfected hosts ($p^i_{ij}(x)$ and $p^i_{ji}(x)$), as well as co-colonization vulnerabilities $k_{ij}(x)$, already studied in [120]. In a compact domain $\Omega \subset \mathbb{R}^{\mathrm{d}}$ with smooth boundary Γ, we consider the general SIS dynamics in a coinfection system with diffusion as follows

$$\begin{cases} \dfrac{\partial S}{\partial t} &= r(1-S) + \displaystyle\sum_{i=1}^{N} \gamma_i I_i + \sum_{i,j=1}^{N} \gamma_{ij} I_{ij} - S \sum_{i=1}^{N} \beta_i J_i + \quad \varepsilon \Delta S, \\ \dfrac{\partial I_i}{\partial t} &= \beta_i J_i S - (r+\gamma_i) I_i - I_i \displaystyle\sum_{j=1}^{N} k_{ij} \beta_j J_j + \qquad \varepsilon \Delta I_i, \qquad 1 \leq i \leq N, \\ \dfrac{\partial I_{ij}}{\partial t} &= k_{ij} I_i \beta_j J_j - (r+\gamma_{ij}) I_{ij} + \qquad \varepsilon \Delta I_{ij}, \qquad 1 \leq i,j \leq N, \end{cases} \tag{4.2.1}$$

where J_i is proportion of all hosts transmitting strain i, including singly- and co-colonized hosts and has the explicit formula

$$J_i = I_i + \sum_{j=1}^{N} \left(p^i_{ij} I_{ij} + p^i_{ji} I_{ji} \right).$$

We will assume no-flux boundary conditions, i.e. the Neumann boundary conditions $\partial_n S = \partial_n I_i = \partial_n I_{ij} = 0$ on the boundary Γ of Ω and the given initial values.

Note that $\beta_i J_i$ is the infection force of strain i, for all i. In (4.2.1), for $1 \leq i,j \leq N$, parameters (that all depend on space) are interpreted as follows

Table 4.1 **Conventions and notations of parameters**

	Parameter	**Interpretation**	Under strain similarities
1.	$\beta_i(x)$	Strain-specific transmission rates	$\beta_i(x) = \beta(x)(1+\varepsilon b_i(x))$
2.	$\gamma_i(x)$	Strain-specific clearance rates of single colonization	$\gamma_i(x) = \gamma(x)(1+\varepsilon \nu_i(x))$
3.	$\gamma_{ij}(x)$	Clearance rates of co-colonization with i and j	$\gamma_{ij}(x) = \gamma(x)(1+\varepsilon u_{ij}(x))$
4.	$p^s_{ij}(x)$	Transmission capacity of the strain $s \in \{i,j\}$ by a host co-colonized by strain-i then strain-j, $\left(p^i_{ij}(x) + p^j_{ij}(x) = 1\right)$	$p^s_{ij}(x) = \frac{1}{2} + \varepsilon \omega^s_{ij}(x)$
5.	$k_{ij}(x)$	Relative factor of altered susceptibility to co-colonization between colonizing strain i and co-colonizing strain j	$k_{ij}(x) = k(x) + \varepsilon \alpha_{ij}(x)$
	$r(x)$	Susceptible recruitment rate (Equal to natural mortality)	

Assumption 4.1. We assume the regularity for the intial values and parameters as follows.

- Initial values $S(x,0)$, $I_i(x,0)$, and $I_{ij}(x,0)$ are smooth enough in $x \in \Omega$, for $1 \leq i,j \leq N$.
- All the parameters in (4.2.1), which are included in Table 4.1, are all smooth enough in $x \in \Omega$.

From Assumption 4.1, we can see that the perturbations considered on the epidemiological parameters depend on the spatial position, bring a real richness of "spatial" to the model.

It is classical that this systems conserved the positive quadrant and then we consider only positive solutions.

For the sake of simplicity, we denote the inverse duration of a carriage episode by strain i with $m_i = r + \gamma_i$, of a co-carriage episode by strains i and j with $m_{ij} = r + \gamma_{ij}$ and the corresponding inverse duration of carriage if all strains were equivalent with $m = r + \gamma$.

In this paper, we use the notation ∇u and Δu when $u(x,\cdot) = \begin{pmatrix} u_1(x,\cdot) & u_2(x,\cdot) & \dots & u_k(x,\cdot) \end{pmatrix}$, for $k \in \mathbb{N}$, with the meaning

$$\nabla u = \begin{pmatrix} \nabla u_1 & \nabla u_2 & \dots & \nabla u_k \end{pmatrix}, \qquad \text{and} \quad \Delta u = \begin{pmatrix} \Delta u_1 & \Delta u_2 & \dots & \Delta u_k \end{pmatrix}.$$

Such a very general pattern of considered system forms $\dfrac{\partial X}{\partial t} = \tilde{F}(X,x,\varepsilon) + \varepsilon \Delta X$ with Neumann boundary condition, where $X = (X_1, X_2, \dots, X_n) \in \mathbb{R}^n$ and is equivalent to $\dfrac{\partial X}{\partial t} = F(X,x) + O(\varepsilon) + \varepsilon \Delta X$ after some algebraic transformations. The part $\dfrac{dX}{dt} = F(X,x) + \varepsilon \Delta X$ is called as the *semi-neutral system*, consistently stays unaltered and be investigated in the subsection 4.3.1. It is important to note that this system is structurally unstable. Then, the part $O(\varepsilon)$ is a slow perturbation of the neutral system. To treat such an emergence by a Tikhonov-liked theorem, it's essential to rewrite $\dfrac{\partial X}{\partial t} = F(X,x) + O(\varepsilon) + \varepsilon \Delta X$ into equivalent slow-fast form

$$\begin{cases} \dfrac{dU}{dt} = & f(U,V,x) + O(\varepsilon) + \varepsilon \Delta U \\ \dfrac{dV}{dt} = \varepsilon & (g(U,V,x) + O(\varepsilon) + \Delta V) \\ \dfrac{\partial U}{\partial n} = & \dfrac{\partial V}{\partial n} = 0, \quad \text{on} \quad \partial\Omega \end{cases} \tag{4.2.2}$$

where $U \in \mathbb{R}^N$ is the slow variable and $V \in \mathbb{R}^N$ is the fast variable. This step is achieved thanks to the ansatz (4.3.2) will be yielded from the study of the semi-neutral system.
This subsection makes a change of variables then allows to rewrite the system in an equivalent structure explicitly dependent on ε. Then, we study the important *semi-neutral system* which is obtained for $\varepsilon = 0$ **except in diffusion terms**. The study on the semi-neutral system leads to the definition of the appropriate slow and fast variables (z_i, v_i). These variables together with the ansatz (4.3.2) are the key for the slow-fast study of the next section.

• **Initially,** sum up all the equations of (4.2.1), we have that

$$\begin{cases} \dfrac{\partial}{\partial t}\left(S+\sum\limits_{i=1}^{N} I_i+\sum\limits_{i,j=1}^{N} I_{ij}\right) = r(1-S)-r\left(\sum\limits_{i=1}^{N} I_i+\sum\limits_{i,j=1}^{N} I_{ij}\right)+\varepsilon\Delta\left(S+\sum\limits_{i=1}^{N} I_i+\sum\limits_{i,j=1}^{N} I_{ij}\right) \\ \partial_n\left(S+\sum\limits_{i=1}^{N} I_i+\sum\limits_{i,j=1}^{N} I_{ij}\right) = 0 \qquad \text{on} \quad \Gamma. \end{cases}$$

Denoting

$$T = \sum_{i=1}^{N} I_i + \sum_{i,j=1}^{N} I_{ij}, \tag{4.2.3}$$

we have the following equation

$$\frac{\partial}{\partial t}(S+T) = r[1-(S+T)] + \varepsilon\Delta(S+T).$$

with the Neumann boundary condition. The assumed smoothness of $\partial\Omega$ implies that $\varepsilon\Delta$ generates a C^0 semi group of contraction on $C^0\left(\bar{\Omega}\right)$, see [27].
Note that $S+T=1$ are the solution of $r[1-(S+T)]+\varepsilon\Delta(S+T)=0$ and the linearized operator becomes $\varepsilon\Delta - r$ which has spectrum lies in the left-half plane (since the Laplacian has the negative spectrum and $r(x) > 0$). By the Theorem 11.20 in [189], we deduce that $S+T=1$ is asymptotically stable, which implies that $S+\sum_{i=1}^{N} I_i + \sum_{i,j=1}^{N} I_{ij} \to 1$ as $t \to \infty$ asymptotically for all x. Therefore, we can assume that $S+\sum_{i=1}^{N} I_i + \sum_{i,j=1}^{N} I_{ij} = 1$ in this article.

From this convention, we deduce that, (4.2.1) has unique solution for every $\varepsilon > 0$. Indeed, (4.2.1) can be rewritten in the form of

$$\frac{\partial}{\partial t} u(x,t) = F(u(x,t), x, \varepsilon) + \varepsilon\Delta u(x,t), \qquad x \in \Omega \tag{4.2.4}$$

with $u = \begin{pmatrix} S & I_1 & \dots & I_N & I_{11} & \dots & I_{NN} \end{pmatrix}^T$. We state the following result on the unique existence of solution of (4.2.4). The proof is given in Appendix A4.1.1.

Theorem 4.2. *Given compact domain $\Omega \in \mathbb{R}^{\mathrm{d}}$ and $u : \Omega \times [0,+\infty) \to \mathbb{R}^n$, $(x,t) \mapsto u(x,t)$. Assume that $F : \mathbb{R}^n \times \Omega \times \mathbb{R}^+$ is continuous in $x \in \Omega$ and $F : \mathbb{R}^n \times \Omega \times \mathbb{R}^+ \to \mathbb{R}^n$ is a Lipschitz map in $u \in \mathbb{R}^n$, i.e. there is a constant L such that*

$$\|Fu - Fv\| \le L\|u-v\|, \qquad \forall u, v \in \mathbb{R}^n, \quad \forall x \in \Omega.$$

Then (4.2.4) *admits a solution in $C^2(\Omega \times \cdot, \mathbb{R}^n) \cap C^1(\cdot \times [0,\infty), \mathbb{R}^n)$, and this solution is unique.*

For the sake of clarify later, we now make conventions for the norms used in this article.

Definition 4.3. Let $v : \Omega \times \mathbb{R}^+ \to \mathbb{R}^n$ and $v \in L^2(\Omega)$ for each $t \geq 0$, we define.

- The norm $|\cdot|_1$ for $v(x,t) \in \mathbb{R}^n$ for each $x \in \Omega$ and $t \in \mathbb{R}^+$:

$$|v(x,t)|_1 := \sum_{i=1}^{n} |v_i(x,t)| \tag{4.2.5}$$

  with $v(x,t) = \begin{pmatrix} v_1(x,t) & v_2(x,t) & \dots & v_n(x,t) \end{pmatrix}$.

- The norm $|\cdot|_2$ for $v(x,t) \in \mathbb{R}^n$ for each $x \in \Omega$ and $t \in \mathbb{R}^+$:

$$|v(x,t)|_2 := \left(\sum_{i=1}^{n} |v_i(x,t)|^2 \right)^{1/2} \tag{4.2.6}$$

  with $v(x,t) = \begin{pmatrix} v_1(x,t) & v_2(x,t) & \dots & v_n(x,t) \end{pmatrix}$.
  However, for the sake of simplicity, we only write $|\cdot|$ instead of $|\cdot|_2$.

- The norm for $v(\cdot,t) \in L^2(\Omega)$ for each $t \in \mathbb{R}^n$:

$$\|v(\cdot,t)\|_2 := \left(\int_\Omega |v(x,t)|^2 dx \right)^{1/2}. \tag{4.2.7}$$

Note that in (4.2.4), $u(x,t) \in \mathbb{R}^{N^2+N+1}$ and in this finite-dimensional space, the norms $|\cdot|$ and $|\cdot|_1$ are equivalent, we recall our previous convention $S + \sum_{i=1}^{N} I_i + \sum_{i,j=1}^{N} I_{ij} = 1$. Then, thanks to the positivity of the solutions, $|u(\cdot,t)|_1 = S + \sum_{i=1}^{N} I_i + \sum_{i,j=1}^{N} I_{ij} = 1$ and satisfies the Theorem 4.2. Hence, the system (4.2.1) always has unique solution.

• **Secondly,** for the sake of simplicity, we denote $m_i = r + \gamma_i$, $m_{ij} = r + \gamma_{ij}$ and $m = r + \gamma$. Then, we define total mass of single infected I, the total mass of double infected D and and the total mass of infected T, as in [120], which reads

$$I = \sum_{i=1}^{N} I_i, \quad D = \sum_{i=1}^{N} I_{ij}, \quad T = I + D. \tag{4.2.8}$$

(4.2.8) yields $\sum_{i=1}^N J_i = T$. For later computations, remark that $\sum_{i=1}^N \beta_i J_i = \beta T + \varepsilon \sum_{i=1}^N b_i J_i$. Thanked to the new variables, the systems for (S,T), (I_i,J_i) and $(I_{ij})_{1\le i,j\le N}$ reads

$$\begin{cases}
\dfrac{\partial S}{\partial t} &= r(1-S) + \gamma T + \varepsilon\gamma\left(\displaystyle\sum_{i=1}^N \nu_i I_i + \sum_{i,j=1}^N u_{ij} I_{ij}\right) - \beta S T - \varepsilon\beta S \displaystyle\sum_{i=1}^N b_i J_i + \varepsilon\Delta S \\
\dfrac{\partial T}{\partial t} &= \beta S T - mT + \varepsilon\beta S \displaystyle\sum_{i=1}^N b_i J_i - \varepsilon\gamma\left(\sum_{i=1}^N \nu_i I_i + \sum_{i,j=1}^N u_{ij} I_{ij}\right) + \varepsilon\Delta T \\
\dfrac{\partial I_i}{\partial t} &= \beta\,(1+\varepsilon b_i) J_i S - (m + \varepsilon\gamma\nu_i) I_i - \beta I_i \displaystyle\sum_{j=1}^N (k+\varepsilon\alpha_{ij})(1+\varepsilon b_j) J_j + \varepsilon\Delta I_i \\
\dfrac{\partial J_i}{\partial t} &= \beta\,(1+\varepsilon b_i) J_i S - \beta I_i \displaystyle\sum_{j=1}^N (k+\varepsilon\alpha_{ij})(1+\varepsilon b_j) J_j - \varepsilon\gamma\left[\nu_i I_i + \sum_{j=1}^N \left(\left(\frac{1}{2} + \varepsilon\omega_{ij}^i\right) u_{ij} I_{ij} + \left(\frac{1}{2} + \varepsilon\omega_{ji}^i\right)\nu_{ji} I_{ji}\right)\right] \\
&\quad - mJ_i + \beta \displaystyle\sum_{j=1}^N \left(\left(\frac{1}{2} + \omega_{ij}^i\right)(k+\varepsilon\alpha_{ij})(1+\varepsilon b_j) I_i J_j + \left(\frac{1}{2} + \varepsilon\omega_{ji}^i\right)(k+\varepsilon\alpha_{ji})(1+\varepsilon b_i) I_j J_i\right) + \varepsilon\Delta J_i \\
\dfrac{\partial I_{ij}}{\partial t} &= \beta\,(k+\varepsilon\alpha_{ij})(1+\varepsilon b_j) I_i J_j - (m + \varepsilon\gamma u_{ij}) I_{ij} + \varepsilon\Delta I_{ij}, \qquad 1 \le i,j \le N.
\end{cases} \tag{4.2.9}$$

4.2.2 The semi-neutral system

Take $\varepsilon = 0$ in (4.2.9) except the diffusion rates, we obtain the *semi-neutral system*[1] for (S,T,I,I_i,J_i), which reads

$$\begin{cases}
\dfrac{\partial S}{\partial t} &= m(1-S) - \beta S T &+\varepsilon\Delta S \\
\dfrac{\partial T}{\partial t} &= \beta S T - mT &+\varepsilon\Delta T \\
\dfrac{\partial I}{\partial t} &= \beta T S - (m + \beta k T) I &+\varepsilon\Delta I, \\
\dfrac{\partial I_i}{\partial t} &= \beta J_i S - m I_i - \beta k I_i T &+\varepsilon\Delta I_i, \quad 1 \le i \le N \\
\dfrac{\partial J_i}{\partial t} &= (\beta S - m) J_i - \beta k I_i T + \dfrac{\beta k}{2}(I_i T + J_i I) &+\varepsilon\Delta J_i, \quad 1 \le i \le N \\
\dfrac{\partial I_{ij}}{\partial t} &= \beta k I_i J_j - m I_{ij} &+\varepsilon\Delta I_{ij}
\end{cases} \tag{4.2.10}$$

with the Neumann boundary condition and the initial condition

$$\begin{cases}
\dfrac{\partial S}{\partial n} = \dfrac{\partial T}{\partial n} = \dfrac{\partial I_i}{\partial n} = \dfrac{\partial I_{ij}}{\partial n} = 0 \quad \text{on} \quad \partial\Omega \\
S(x,0) = S_0(x), \quad T(x,0) = T_0(x), \quad I(x,0) = I_{i,0}(x), \quad I_{ij}(x,0) = I_{ij,0}(x).
\end{cases}$$

[1] The name *semi-neutral system* comes from the fact that if $\varepsilon = 0$, except the coefficients of diffusion terms, then the parameters do not depend on the strains as in the neutral theory.

• **Firstly,** we consider the semi-neutral equation for (S,T), that reads

$$\begin{cases} \dfrac{\partial S}{\partial t} & = m(1-S) - \beta ST + \varepsilon \Delta S \\ \dfrac{\partial T}{\partial t} & = -mT + \beta ST + \varepsilon \Delta T. \\ S(x,0) & = S_0(x), \quad T(x,0) = T_0(x), \\ \dfrac{\partial S}{\partial n}|_{\partial\Omega} & = \dfrac{\partial T}{\partial n}|_{\partial\Omega} = 0. \end{cases} \tag{4.2.11}$$

By the Theorem 4.2 that (4.2.11) has the unique solution.

Before analyzing, similar to [9], we say that x is a low-risk site if the local disease transmission rate $\beta(x)$ is lower than the local disease recovery rate (which is the sum of clearance rate and mortality rate) $m(x)$. A high-risk site is defined in a similar manner. Let

$$H^- = \{x \in \Omega : \beta(x) < m(x)\} \quad \text{and} \quad H^+ = \{x \in \Omega : \beta(x) > m(x)\} \tag{4.2.12}$$

denote the set of these low- and high-risk sites, respectively. Accordingly, the term $R_0(x)$ is the local reproduction number at $x \in \Omega$. Then $R_0(x) < 1$ for low-risk sites $x \in H^-$ and $R_0(x) > 1$ for high-risk sites $x \in H^+$. It is well-known that without movement, the disease can persist at high-risk sites but not at low-risk sites. We say that, a domain Ω' is a low-risk domain if $\int_{\Omega'} \beta < \int_{\Omega'} m$ and a high-risk domain if $\int_{\Omega'} \beta > \int_{\Omega'} m$.

In this case of slow diffusion, i.e. in sections 4.2 and 4.3, we make an assume that

Assumption 4.4. The domain Ω is high-risk everywhere, i.e. $\beta(x) > m(x)$ for all $x \in \Omega$.

Denoting $S^*(x) = \dfrac{m(x)}{\beta(x)} = \dfrac{1}{R_0(x)}$ and $T^*(x) = 1 - S^*(x)$, then $0 \le S^*, T^* \le 1$ for all $x \in \Omega$, which is well-defined. At each $x \in \Omega$, consider the differential equations of variables $\left(\tilde{S}(\cdot,t), \tilde{T}(\cdot,t)\right)$

$$\begin{cases} \dfrac{d\tilde{S}}{dt} = m(x)\left(1-\tilde{S}\right) - \beta(x)\tilde{S}\tilde{T} \\ \dfrac{d\tilde{T}}{dt} = -m(x)\tilde{T} + \beta(x)\tilde{S}\tilde{T} \end{cases}, \tag{4.2.13}$$

with initial condition $\left(\tilde{S}(0), \tilde{T}(0)\right) = (S(x,0), T(x,0))$. It is claimed that $\left(\tilde{S}(x,t), \tilde{T}(x,t)\right) \to (S^*(x), T^*(x))$ for each $x \in \Omega$ and $t \to \infty$, see [120].

Furthermore, noting that $\left|\tilde{S}(x,t) - S^*(x)\right|^2 \le 2$ for all $(x,t) \in \Omega \times \mathbb{R}^+$ and Ω is compact, by Dominated Convergence Theorem, for all sequence $(t_n)_n$ satisfying $t_0 < t_1 < \cdots < t_n < \ldots$, $t_n \to +\infty$, we obtain that $\left\|\tilde{S}(x,t) - S^*(x)\right\|_2 \to 0$ as $t \to \infty$.

Another important point is that, for any $t \in \mathbb{R}^+$, $\tilde{S}(x,t)$ is smooth enough with respect to x. Indeed, we have the differential equation for $\tilde{S}(x) - \tilde{S}(x')$ as follows

$$\frac{d}{dt}\left(\tilde{S}(x) - \tilde{S}(x')\right) = \left(\tilde{S}(x) - \tilde{S}(x')\right)\left[-(m+\beta) + \beta\left(\tilde{S}(x) + \tilde{S}(x')\right)\right].$$

Noting that $\tilde{S}(x) + \tilde{S}(x') \le 2$ for all $t \in \mathbb{R}^+$ and $x, x' \in \Omega$, applying the Gronwall's inequality, we deduce that $\frac{|\tilde{S}(x) - \tilde{S}(x')|}{|x-x'|}$ can be controlled at each t, since $\tilde{S}(0)$ is smooth enough with respect to x.

Analogously, we observe $x \mapsto \tilde{T}(x,t)$ is smooth enough as well.

Alternatively, we have that

$$\frac{\partial}{\partial t}\left(S-\tilde{S}\right)=\left[\mathscr{F}(S)-\mathscr{F}\left(\tilde{S}\right)\right]+\varepsilon\Delta S,\qquad\text{with}\quad\mathscr{F}(X)=\mathscr{F}(X,\cdot)=m(\cdot)(1-X(\cdot))-\beta(\cdot)X(\cdot)(1-X(\cdot))$$

then $\mathscr{F}(X)$ is Lipschitz continuous with coefficient $C>0$. We have the following transformations

$$\begin{aligned}&\left(S-\tilde{S}\right)\frac{\partial}{\partial t}\left(S-\tilde{S}\right)=\left(S-\tilde{S}\right)\left[\mathscr{F}(S)-\mathscr{F}\left(\tilde{S}\right)\right]+\varepsilon\left(S-\tilde{S}\right)\Delta S\\ \Longrightarrow&\frac{1}{2}\frac{\partial}{\partial t}\left|S-\tilde{S}\right|^2\leq C\left|S-\tilde{S}\right|^2+\varepsilon\left(S-\tilde{S}\right)\Delta\left(S-\tilde{S}\right)+\varepsilon\left(S-\tilde{S}\right)\Delta\tilde{S}\\ \Longrightarrow&\frac{1}{2}\frac{\partial}{\partial t}\int_\Omega\left|S-\tilde{S}\right|^2\leq C\int_\Omega\left|S-\tilde{S}\right|^2+\varepsilon\int_\Omega\left(S-\tilde{S}\right)\Delta\tilde{S}\end{aligned}$$

By the definition of $\tilde{S}$ and $\max\left|S-\tilde{S}\right|\leq 2$ then by the Gronwall's inequality we have that

$$\left\|S-\tilde{S}\right\|_2=O\left(\sqrt{\varepsilon}\right),\qquad\text{since}\quad S(\cdot,0)=S^*(\cdot,0),$$

which leads to $\|S-S^*\|_2\to O\left(\sqrt{\varepsilon}\right)$ as $t\to\infty$ since Ω is compact.

Recalling that $S+T=1$ then $\|T-T^*\|_2\to O\left(\sqrt{\varepsilon}\right)$ when $t\to\infty$.

• **Secondly,** we consider the semi-neutral equation for $I(x,t)$ which reads

$$\begin{cases}\dfrac{\partial I}{\partial t}=\beta TS-(m+\beta kT)I+\varepsilon\Delta I\\ \dfrac{\partial I}{\partial n}|_{\partial\Omega}=0\end{cases}$$

Similarly to the previous proof for the stability of $(S(x,t),T(x,t))$, we consider the equation for $\tilde{I}(\cdot,t)$ at each $x\in\Omega$

$$\frac{\partial\tilde{I}}{\partial t}=\beta T^*S^*-(m+\beta kT^*)\tilde{I},\qquad\tilde{I}(\cdot,0)=I(x,0),$$

which implies $\tilde{I}(x,t)\to I^*(x):=\dfrac{mT^*}{m+\beta kT^*}$ at each $x\in\Omega$ as $t\to\infty$, as proved in [120].

By the same arguments for $\left\|\tilde{S}(x,t)-S^*(x)\right\|_2\to 0$ as $t\to\infty$ previously, we also obtain that $\left\|\tilde{I}(x,t)-I^*(x)\right\|_2\to 0$ as $t\to\infty$.

Similarly to the proof of the smoothness of $x\mapsto\tilde{S}(x,t)$, we can prove that $x\mapsto\tilde{I}(x,t)$ is smooth enough as well, for all $t\in\mathbb{R}^+$.

Accordingly, we deduce the equation

$$\frac{\partial}{\partial t}\left(I-\tilde{I}\right)=\left[\mathscr{G}(I)-\mathscr{G}\left(\tilde{I}\right)\right]+\varphi(S,T,S^*,T^*)+\varepsilon\Delta I,$$

where $\mathscr{G}(X)=\beta T^*S^*-(m+\beta kT^*)X$ and

$$\varphi(S,T,S^*,T^*)=\beta\left[(T-T^*)S^*+(S-S^*)T^*+(T-T^*)(S-S^*)\right]$$

which means $\mathscr{G}$ is Lipschitz continuous and $\|\varphi\|_2 = O\left(\sqrt{\varepsilon}\right)$. We make the similar process as before, then combining the Holder's inequality, we deduce that

$$\frac{1}{2}\frac{\partial}{\partial t}\int_\Omega \left|I-\tilde{I}\right|^2 \leq (C+\varepsilon)\int_\Omega \left|I-\tilde{I}\right|^2 + \varepsilon\int_\Omega \left(I-\tilde{I}\right)\Delta\tilde{I}$$

which implies $\left\|I-\tilde{I}\right\|_2 = O\left(\sqrt{\varepsilon}\right)$ by applying the Gronwall's inequality with noting that $I(\cdot,0) = \tilde{I}(\cdot,0)$. Thus, we have that $\|I-I^*\|_2 \to O\left(\sqrt{\varepsilon}\right)$ as $t\to\infty$.

For later reference, we also write

$$\boxed{S^* = \frac{m}{\beta},\quad T^* = 1-\frac{m}{\beta},\quad I^* = \frac{mT^*}{m+\beta kT^*},\quad D^* = T^*-I^* = \frac{\beta kT^{*2}}{m+\beta kT^*}.} \tag{4.2.14}$$

Note that, since all parameters depend on x, then $S^*(x)$, $T^*(x)$, $I^*(x)$ and $D^*(x)$ depend on x.

Hence, for (S,T,I) satisfying the semi-neutral system (4.2.10), we have that

$$\|S-S^*\|_2 \to O\left(\sqrt{\varepsilon}\right),\qquad \|T-T^*\|_2 \to O\left(\sqrt{\varepsilon}\right),\qquad \|I-I^*\|_2 \to O\left(\sqrt{\varepsilon}\right). \tag{4.2.15}$$

when $t\to\infty$.

• **Thirdly,** the N^2 equations for I_{ij} in (4.2.10) yields that, for $1\leq i\leq N$,

$$\frac{\partial I_{ij}}{\partial t} = \beta kI_iJ_j - mI_{ij} + \varepsilon\Delta I_{ij}. \tag{4.2.16}$$

Whose dynamics is trivial once I_i and J_i are known. Indeed, assume that for each i, there exists $\left(\tilde{I}_i(x,t),\tilde{J}_i(x,t)\right)$ such that $\left\|I_i-\tilde{I}_i\right\|_2 = O\left(\sqrt{\varepsilon}\right)$ and $\left\|J_i-\tilde{J}_i\right\|_2 = O\left(\sqrt{\varepsilon}\right)$, then we can rewrite (4.2.16) into

$$\frac{\partial I_{ij}}{\partial t} = -mI_{ij} + \beta k\tilde{I}_i\tilde{J}_j + \beta k\left[\left(I_i-\tilde{I}_i\right)\tilde{J}_j + \left(J_j-\tilde{J}_j\right)\tilde{I}_i + \left(I_i-\tilde{I}_i\right)\left(J_j-\tilde{J}_j\right)\right] + \varepsilon\Delta I_{ij}.$$

At each $x\in\Omega$, we consider the equation of $\tilde{I}_{ij}(\cdot,t)$

$$\frac{\partial\tilde{I}_{ij}}{\partial t} = -m\tilde{I}_{ij} + \beta k\tilde{I}_i\tilde{J}_j,\qquad \tilde{I}_{ij}(\cdot,0) = I_{ij}(x,0)$$

Once again, by the same argument for the smoothness of $x\mapsto\tilde{S}(x,t)$, we obtain that $x\mapsto\tilde{I}_{ij}(x,t)$, for all $1\leq i,j\leq N$.

Then we can obtain the differential equation for $\left(I_{ij}-\tilde{I}_{ij}\right)$

$$\frac{\partial}{\partial t}\left(I_{ij}-\tilde{I}_{ij}\right) = -m\left(I_{ij}-\tilde{I}_{ij}\right) + \beta k\left[\left(I_i-\tilde{I}_i\right)\tilde{J}_j + \left(J_j-\tilde{J}_j\right)\tilde{I}_i + \left(I_i-\tilde{I}_i\right)\left(J_j-\tilde{J}_j\right)\right] + \varepsilon\Delta\left(I_{ij}-\tilde{I}_{ij}\right) + \varepsilon\Delta\tilde{I}_{ij}.$$

Denoting

$$\phi(x,t) = \beta k\left[\left(I_i-\tilde{I}_i\right)\tilde{J}_j + \left(J_j-\tilde{J}_j\right)\tilde{I}_i + \left(I_i-\tilde{I}_i\right)\left(J_j-\tilde{J}_j\right) + \varepsilon\Delta\tilde{I}_{ij}\right]$$

then $\|\phi\|_2 = O\left(\sqrt{\varepsilon}\right)$. By the same process as previous, we have that

$$\frac{1}{2}\frac{\partial}{\partial t}\int_\Omega \left|I_{ij}-\tilde{I}_{ij}\right|^2 = -m\int_\Omega \left|I_{ij}-\tilde{I}_{ij}\right|^2 + \int_\Omega \phi\left(x,t\right)\left(I_{ij}-\tilde{I}_{ij}\right) - \varepsilon\int_\Omega \left|\nabla\left(I_{ij}-\tilde{I}_{ij}\right)\right|^2.$$

Using the Holder's inequality for the term $\int_\Omega \phi\left(x,t\right)\left(I_{ij}-\tilde{I}_{ij}\right)$ then applying the Gronwall's inequality again, note that $I_{ij}\left(\cdot,0\right) = \tilde{I}_{ij}\left(\cdot,0\right)$, we have that $\left\|I_{ij}-\tilde{I}_{ij}\right\|_2 \to O\left(\sqrt{\varepsilon}\right)$ as $t\to\infty$, for all $1\le i,j\le N$.

4.2.3 The slow-fast form and approximations theorems

Next, we consider the semi-neutral system for $\begin{pmatrix} I_i \\ J_i \end{pmatrix}$ for all $1\le i\le N$

$$\begin{cases} \dfrac{\partial I_i}{\partial t} = \beta J_i S - mI_i - \beta k I_i T & +\varepsilon\Delta I_i \\ \dfrac{\partial J_i}{\partial t} = \left(\beta S - m\right)J_i - \beta k I_i T + \dfrac{\beta k}{2}\left(I_i T + J_i I\right) & +\varepsilon\Delta J_i \end{cases}. \tag{4.2.17}$$

Denoting $D^* = T^* - I^*$, we set

$$A\left(x\right) = \begin{pmatrix} -\left(m+\beta k T^*\right) & m \\ -\dfrac{\beta k T^*}{2} & \dfrac{\beta k I^*}{2} \end{pmatrix},$$

and

$$P = \begin{pmatrix} 2T^* & I^* \\ D^* & T^* \end{pmatrix}, \qquad P^{-1} = \frac{1}{|P|}\begin{pmatrix} T^* & -I^* \\ -D^* & 2T^* \end{pmatrix}. \tag{4.2.18}$$

We have

$$A\left(x\right) = P\left(x\right)\begin{pmatrix} -\xi\left(x\right) & 0 \\ 0 & 0 \end{pmatrix}P^{-1}\left(x\right)$$

where $\xi = m + \beta k T^* - \dfrac{1}{2}\beta k I^* > 0$ and $\left|P\left(x\right)\right| = 2T^{*2} - I^* D^* > 0$.

In the equations for $\left(I_i, J_i\right)$ in (4.2.17), we substitute $\left(S,T,I\right)$ by $\left(S^*,T^*,I^*\right)$ and note that

$$\left\|S-S^*\right\|_2 = O\left(\sqrt{\varepsilon}\right), \quad \left\|T-T^*\right\|_2 = O\left(\sqrt{\varepsilon}\right), \quad \left\|I-I^*\right\|_2 = O\left(\sqrt{\varepsilon}\right).$$

Now, we have the semi-neutral system of equations for $\left(I_i,J_i\right)_{1\le i\le N}$, in the sense of norm $\|\cdot\|_2$ of $L^2\left(\Omega\right)$:

$$\frac{\partial}{\partial t}\begin{pmatrix} I_i \\ J_i \end{pmatrix} = A\left(x\right)\begin{pmatrix} I_i \\ J_i \end{pmatrix} + O\left(\sqrt{\varepsilon}\right)\begin{pmatrix} I_i \\ J_i \end{pmatrix} + \varepsilon\begin{pmatrix} \Delta I_i \\ \Delta J_i \end{pmatrix}. \tag{4.2.19}$$

Applying Theorem 4.6, we have that

$$\left\|I_i - \tilde{I}_i\right\|_2 = O\left(\sqrt{\varepsilon}\right), \qquad \left\|J_i - \tilde{J}_i\right\|_2 = O\left(\sqrt{\varepsilon}\right), \tag{4.2.20}$$

where $\left(I_i,J_i\right)_{1\le i\le N}$ are solutions of the semi-neutral system (4.2.17) and $\left(\tilde{I}_i,\tilde{J}_i\right)_{1\le i\le N}$ are solutions of

$$\frac{\partial}{\partial t}\begin{pmatrix} I_i \\ J_i \end{pmatrix} = A\left(x\right)\begin{pmatrix} I_i \\ J_i \end{pmatrix} + \varepsilon\begin{pmatrix} \Delta I_i \\ \Delta J_i \end{pmatrix} \tag{4.2.21}$$

Hence, it suffices to consider the system (4.2.21). For every $1 \leq i \leq N$, set

$$\begin{pmatrix} v_i \\ z_i \end{pmatrix} = P^{-1} \begin{pmatrix} I_i \\ J_i \end{pmatrix} \tag{4.2.22}$$

From (4.2.21) we infer an equation for $\begin{pmatrix} z_i \\ v_i \end{pmatrix}$ for each $1 \leq i \leq N$:

$$\frac{\partial}{\partial t} \begin{pmatrix} v_i \\ z_i \end{pmatrix} = \begin{pmatrix} -\xi(x) & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} v_i \\ z_i \end{pmatrix} + \varepsilon \left[P^{-1}(x) \begin{pmatrix} \Delta & 0 \\ 0 & \Delta \end{pmatrix} P(x) \right] \begin{pmatrix} v_i \\ z_i \end{pmatrix} \tag{4.2.23}$$

This step of changing to (z_i, v_i) plays an important role. Since under these new variables, we can rewrite into the slow-fast form. It allows us to apply the approximation theorem introduced in the next subsection.

When $v_i = 0$ - which will be asymptotically true - then z_i is exactly $\frac{I_i}{I^*} = \frac{J_i}{T^*}$ the prevalence of strain i in the total of infected, see the proof in [140].

$$\begin{cases} \dfrac{\partial v_i}{\partial t} &= -\xi v_i + O(\varepsilon) + \varepsilon \Delta v_i \quad + \varepsilon \dfrac{1}{|P|} \left[(2T^* \nabla T^* - I^* \nabla D^*) \nabla v_i + (T^* \nabla I^* - I^* \nabla T^*) \nabla z_i \right] \\ \dfrac{\partial z_i}{\partial t} &= O(\varepsilon) + \varepsilon \Delta z_i \qquad\qquad + \varepsilon \dfrac{1}{|P|} \left[(-D^* \nabla T^* + 2T^* \nabla D^*) \nabla v_i + (-D^* \nabla I^* + 2T^* \nabla T^*) \nabla z_i \right] \end{cases} . \tag{4.2.24}$$

Next, by setting $\tau = \varepsilon t$, (4.2.24) can be read as the slow time scale

$$\begin{cases} \varepsilon \dfrac{\partial v_i}{\partial \tau} &= -\xi v_i + O(\varepsilon) + \varepsilon \dfrac{1}{|P|} \left[(2T^* \nabla T^* - I^* \nabla D^*) \nabla v_i + (T^* \nabla I^* - I^* \nabla T^*) \nabla z_i \right] + \varepsilon \Delta v_i \\ \dfrac{\partial z_i}{\partial \tau} &= O(1) + \dfrac{1}{|P|} \left[(-D^* \nabla T^* + 2T^* \nabla D^*) \nabla v_i + (-D^* \nabla I^* + 2T^* \nabla T^*) \nabla z_i \right] + \Delta z_i. \end{cases} \tag{4.2.25}$$

We need to compute explicitly the perturbation $O(1)$ in (4.2.24). This computation is quite complex especially when involving perturbation in each parameters, so it's worthwhile of dividing this progress into five sub single cases wherein only one perturbation at the time occurs.

After that, we will treat the slow-fast form by a Tikhonov-like theorem, that is presented in the Theorem 4.5. This result is for the parameter-dependent reaction-diffusion system with Neumann boundary condition as following,

$$\begin{cases} \dfrac{\partial}{\partial t} f(x,t) = F(f(x,t), g(x,t), x, t) \qquad\qquad + K_f f(x,t) \\ \varepsilon \dfrac{\partial}{\partial t} g(x,t) = G(f(x,t), g(x,t), x, t) + \varepsilon G_1(x) \cdot \nabla f(x,t) \quad + \varepsilon K_g g(x,t) \\ \dfrac{\partial}{\partial n} f(x,t) = \dfrac{\partial}{\partial n} g(x,t) = 0, \qquad x \in \partial\Omega, \\ f(x,0) = f^0(x), \quad g(x,0) = g^0(x) \end{cases} \tag{4.2.26}$$

in which,

- $f : \Omega \times \mathbb{R} \to \mathbb{R}^n$ and $g : \Omega \times \mathbb{R} \to \mathbb{R}^m$,
- $G_1 : \Omega \to \mathbb{R}^m$ is continuously differentiable and $\cdot$ denotes the scalar product,

- the operators K_f, K_g defined on $C^\infty(\Omega\times[0,t_1])$ by $K_f u := a_f(x)\nabla u + \Delta u$, $K_g u := a_g(x)\nabla u + \Delta u$, in which $a_f(x)$ is an $n\times n$ diagonal matrix and $a_g(x)$ is an $m\times m$ diagonal matrix, in which entries of each matrix depends on $x\in\Omega$. We assume that $a_f(x)$ and $a_g(x)$ are differentially continuous in x.

Theorem 4.5. *Let $f_0(x,t):\Omega\times[t_0,t_1]\mapsto\mathbb{R}^n$, $g_0(x,t):\Omega\times[t_0,t_1]\mapsto\mathbb{R}^m$ be continuous functions satisfying equations*

$$\begin{cases} \dfrac{\partial}{\partial t}f(x,t) &= F(f(x,t),g(x,t),x,t)+K_f f(x,t)\\ 0 &= G(f(x,t),g(x,t),x,t)\\ \dfrac{\partial}{\partial n}f(x,t) &= \dfrac{\partial}{\partial n}g(x,t)=0, \qquad x\in\partial\Omega \end{cases} \tag{4.2.27}$$

where $F:\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}\mapsto\mathbb{R}^n$ and $G:\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}\mapsto\mathbb{R}^m$ are continuous functions. We make an addition assumption that $g_0\in C^1(\Omega\times\mathbb{R})$.

For any $(x,t)\in\Omega\times\mathbb{R}^+$ and $f(x,t)\in\mathbb{R}^n$, we denote $A(x,t)$ is the Jacobian matrix of $G(f(x,t),\cdot,x,t)$ with respect to the second variable.

Alternatively, we assume that F,G are continuously differentiable with respect to their first two arguments in a neighborhood of the trajectory $f_0(x,t)$, $g_0(x,t)$, and that $A(x,t)$ is a Hurwitz matrix, i.e. every eigenvalue of it has strictly negative real part, for all $t\in[t_0,t_1]$ and $x\in\Omega$.

Then there exists $\varepsilon_0>0$ and $C>0$ such that inequalities

$$\begin{cases} \displaystyle\int_\Omega |f_0(x,t)-f(x,t)|^2dx\le C\varepsilon, \qquad \forall t\in[t_0,t_1]\\ \displaystyle\int_\Omega |g_0(x,t)-g(x,t)|^2dx\le C\varepsilon, \qquad \forall t\in[t_0,t_1] \end{cases} \tag{4.2.28}$$

hold for all solutions of (4.2.26) *with $\int_\Omega|f_0(x,t_0)-f(x,t_0)|^2dx\le\varepsilon$, $\int_\Omega|g_0(x,t_0)-g(x,t_0)|^2dx\le\varepsilon$ and $\varepsilon\in(0,\varepsilon_0)$.*

The conclusion of this theorem means that, for the initial values closed enough to $f_0(x,t_0)$ and $g_0(x,t_0)$ in the sense of $L^2(\Omega)$ norm, we have the approximation for the solution of (4.2.26). Explicitly, this can be rewritten as follows

$$\|f_0(x,t)-f(x,t)\|_2=O\left(\sqrt{\varepsilon}\right), \qquad \|g_0(x,t)-g(x,t)\|_2=O\left(\sqrt{\varepsilon}\right), \qquad \forall t\in[t_0,t_1]$$

for all solutions of (4.2.26) with $\|f_0(x,t_0)-f(x,t_0)\|_2=O\left(\sqrt{\varepsilon}\right)$, $\|g_0(x,t_0)-g(x,t_0)\|_2=O\left(\sqrt{\varepsilon}\right)$ and $\varepsilon\in(0,\varepsilon_0)$.

Next, we claim a result that allows us approximate the original system by the semi-neutral system. The following error estimate gives a more precise description of these limits.

Theorem 4.6. *Given $\Omega\in\mathbb{R}^n$ compact domain with smooth boundary. Let F and G be two continuously differentiable functions on $\Omega\times[0,\infty)$ and suppose that F is Lipschitz continuous. Assume there exists a*

bounded function u satisfies the reaction diffusion equation with Neumann boundary condition

$$\begin{cases} \dfrac{\partial u}{\partial t} = F(u,x) + \varepsilon G(u,x) + \varepsilon \Delta u, \\ u(x,0) = u_0(x), \quad x \in \Omega, \\ \dfrac{\partial u}{\partial n}|_{\partial\Omega} = 0. \end{cases} \tag{4.2.29}$$

Then for every fixed $T > 0$, $\forall t < T$, we have that $\int_\Omega |u(x,t) - v(x,t)|^2 dx = O(\varepsilon)$, i.e. $\|u(x,t) - v(x,t)\|_2 = O\left(\sqrt{\varepsilon}\right)$, with $v(t)$ is the solution of the problem

$$\begin{cases} \dfrac{\partial v}{\partial t} = F(v) + \varepsilon \Delta v, \\ v(x,0) = u_0(x), \quad x \in \Omega, \quad \forall x \in \Omega. \\ \dfrac{\partial v}{\partial n}|_{\partial\Omega} = 0. \end{cases} \tag{4.2.30}$$

4.3 Approximation theorems, derivations of original dynamics and main results for the case of slow diffusion

4.3.1 Lemmas and derivation of non-semi neutral dynamics

Next we develop a lemma showing allowing to linearly combine all the relevant simple cases directly into the slow equation. For this purpose, we use the following notations in system (4.2.1).

$$\begin{aligned} &\beta_i = \beta(1 + \chi_1 \varepsilon b_i); \qquad \gamma_i = \gamma(1 + \chi_2 \varepsilon \nu_i); \qquad \gamma_{ij} = \gamma(1 + \chi_3 \varepsilon u_{ij}); \\ &p_{ij}^s = \frac{1}{2} + \chi_4 \varepsilon \omega_{ij}^s, \quad s \in \{i,j\}, \quad \left(\omega_{ij}^i + \omega_{ij}^j = 0\right); \qquad k_{ij} = k + \chi_5 \varepsilon \alpha_{ij}, \end{aligned} \tag{4.3.1}$$

where $\chi_d \in \{0,1\}$ for $d = 1,2,3,4,5$.

Any combination of axes of trait variation among strains, can be captured via $\mathscr{A}$ where $\mathscr{A}$ is a subset of $\{1,2,3,4,5\}$, and for some fixed initial values given, denote $C_{\mathscr{A}}$ be the system (4.2.4) with $\chi_d = 1$ if $d \in \mathscr{A}$ and $\chi_d = 0$ if $d \notin \mathscr{A}$. For simplicity, we note also $C_{\{d\}}$ by C_d for $d \in \{1,2,3,4,5\}$ denote the absence/presence of perturbations in that parameter among strains.

Remark 4.1. If $\mathscr{A} = \emptyset$ then there is no trait perturbation and the system $C_\emptyset$ is exactly the *semi neutral* model (4.2.10).

In order to capture all the perturbations of order 1 in the equation of the z_i we need these additional changes of variables:

$$\boxed{S(x,t) = S^*(x,t) - \varepsilon X(x,t) + O\left(\varepsilon^2\right); \quad T(x,t) = T^*(x,t) + \varepsilon X(x,t) + O(\varepsilon^2); \quad I(x,t) = I^*(x,t) + \varepsilon Y(x,t) + O\left(\varepsilon^2\right)} \tag{4.3.2}$$

where S^*, T^* and I^* are defined in (4.2.14), and for $i = 1, \cdots, N$:

$$L_i(x,t) = \frac{1}{2} \sum_{j=1}^{N} \left(u_{ij} I_{ij}(x,t) + u_{ji} I_{ji}(x,t)\right). \tag{4.3.3}$$

With these notations, $C_{\mathscr{A}}$ reads

$$\begin{cases} \dfrac{\partial X}{\partial t} &= -\beta T^* X + \chi_1 \beta S^* \sum_{i=1}^{N} b_i J_i - \chi_2 \gamma \sum_{i=1}^{N} \nu_i I_i - \chi_3 \gamma \sum_{i=1}^{N} L_i + \varepsilon \Delta X + O(\varepsilon) \\ \dfrac{\partial Y}{\partial t} &= \beta(S^* - T^* - kI^*)X - (m + \beta k T^*)Y + \chi_1 \beta (S^* - kI^*) \sum_{i=1}^{N} b_i J_i - \chi_2 \gamma \sum_{i=1}^{N} \nu_i I_i - \chi_5 \beta \sum_{i,j=1}^{N} \alpha_{ij} I_i J_j + \varepsilon \Delta Y + O(\varepsilon) \\ \dfrac{\partial L_i}{\partial t} &= -mL_i + \chi_3 \frac{1}{2} \beta \gamma k I_i \sum_{j=1}^{N} u_{ij} J_j + \chi_3 \frac{1}{2} \gamma \beta k J_i \sum_{j=1}^{N} \nu_{ji} I_j + \varepsilon \Delta L_i + O(\varepsilon) \end{cases} \tag{4.3.4}$$

together with

$$\begin{aligned} \frac{\partial}{\partial t} \begin{pmatrix} I_i \\ J_i \end{pmatrix} = & \begin{pmatrix} -(m + \beta k T^*) & m \\ -\frac{\beta k T^*}{2} & \frac{\beta k I^*}{2} \end{pmatrix} \begin{pmatrix} I_i \\ J_i \end{pmatrix} + \varepsilon \begin{pmatrix} \Delta I_i \\ \Delta J_i \end{pmatrix} \\ & - \varepsilon \left[\beta \begin{pmatrix} k & 1 \\ \frac{k}{2} & 1 \end{pmatrix} \begin{pmatrix} I_i \\ J_i \end{pmatrix} X + \frac{\beta k}{2} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} I_i \\ J_i \end{pmatrix} Y + \mathscr{M}_{\mathscr{A}} \begin{pmatrix} I_i \\ J_i \end{pmatrix} - \chi_3 \gamma \begin{pmatrix} 0 \\ L_i \end{pmatrix} \right] \end{aligned} \tag{4.3.5}$$

where $\mathscr{M}_{\mathscr{A}}$ is the matrix

$$\begin{pmatrix} -\chi_1 \beta k \sum_{i=1}^{N} b_i J_i - \chi_2 \gamma \nu_i - \chi_5 \beta \sum_{j=1}^{N} \alpha_{ij} J_j & \chi_1 \beta b_i S^* \\ \beta \sum_{j=1}^{N} \left(\chi_4 k \omega_{ij}^i - \chi_5 \frac{\alpha_{ij}}{2} \right) J_j - \chi_1 \frac{\beta k}{2} \sum_{i=1}^{N} b_i J_i - \chi_2 \gamma \nu_i & \chi_1 \beta b_i \left(S^* + \frac{kI^*}{2} \right) + \beta \sum_{j=1}^{N} \left(\chi_4 k \omega_{ji}^i + \chi_5 \frac{\alpha_{ji}}{2} \right) I_j \end{pmatrix} \tag{4.3.6}$$

In order to apply the Theorem (4.5), we rewrite system $C_{\mathscr{A}}$ using the changes of variables $\begin{pmatrix} v_i \\ z_i \end{pmatrix} = P^{-1} \begin{pmatrix} I_i \\ J_i \end{pmatrix}$ with P^{-1} in (4.2.18). Let us note

$$\mathbf{L} = (L_i)_i, \quad \mathbf{v} = (v_i)_i, \quad \mathbf{z} = (z_i)_i,$$

and $-\xi = -(m + \beta k T^*) + \dfrac{\beta k I^*}{2} < 0$. The system $C_{\mathscr{A}}$ reads now as the slow-fast form

$$\begin{cases} \dfrac{\partial X}{\partial t} &= -\beta T^* X + \chi_1 F_X^1(\mathbf{v}, \mathbf{z}) + \chi_2 F_X^2(\mathbf{v}, \mathbf{z}) + \chi_3 F_X^3(\mathbf{L}) + \varepsilon \Delta X + O(\varepsilon) \\ \dfrac{\partial Y}{\partial t} &= \beta(S^* - T^* - kI^*)X - (m + \beta k T^*)Y + \chi_1 F_Y^1(\mathbf{v}, \mathbf{z}) + \chi_2 F_Y^2(\mathbf{v}, \mathbf{z}) + \chi_5 F_Y^5(\mathbf{v}, \mathbf{z}) + \varepsilon \Delta Y + O(\varepsilon) \\ \dfrac{\partial L_i}{\partial t} &= -mL_i + \chi_3 F_{L_i}(\mathbf{v}, \mathbf{z}) + O(\varepsilon) + \varepsilon \Delta L_i \\ \dfrac{\partial v_i}{\partial t} &= -\xi v_i + O(\varepsilon) + \varepsilon \Delta v_i + \varepsilon \frac{1}{|P|} \left[(2T^* \nabla T^* - I^* \nabla D^*) \nabla v_i + (T^* \nabla I^* - I^* \nabla T^*) \nabla z_i \right] \\ \dfrac{\partial z_i}{\partial t} &= \varepsilon \left(F_{z_i}(X, Y, \mathbf{L}, \mathbf{v}, \mathbf{z}) + O(\varepsilon) \right) + \varepsilon \Delta z_i + \varepsilon \frac{1}{|P|} \left[(-D^* \nabla T^* + 2T^* \nabla D^*) \nabla v_i + (-D^* \nabla I^* + 2T^* \nabla T^*) \nabla z_i \right] \end{cases} \tag{4.3.7}$$

For $i = 1, \cdots, N$, the functions F_X^i, F_Y^i, F_{L_i} are obviously deduced from the right term of (4.3.4) and are linear in theirs variables, $X, Y, \mathbf{L}$, respectively. The function F_Y^5 is quadratic in $(\mathbf{v}, \mathbf{z})$. Finally, F_{z_i} is given by the second

line of the right term of (4.3.6) after the linear change of variables (4.2.22):

$$F_{z_i}(X,Y,\mathbf{L},\mathbf{v},\mathbf{z}) = \begin{pmatrix} 0 & 1 \end{pmatrix} P^{-1} \left(\beta \begin{pmatrix} -k & -1 \\ -\frac{k}{2} & -1 \end{pmatrix} X + \frac{\beta k}{2} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} Y + \mathscr{M} \right) P \begin{pmatrix} v_i \\ z_i \end{pmatrix} + \begin{pmatrix} 0 & 1 \end{pmatrix} P^{-1} \chi_3 \gamma \begin{pmatrix} 0 \\ L_i \end{pmatrix}. \tag{4.3.8}$$

Lemma 4.7. Let $\varepsilon = 0$ in (4.3.11). Then there exist a function $\Phi(\mathbf{z}) = (X^*(\mathbf{z}), Y^*(\mathbf{z}), \chi_3 \mathbf{L}^*(\mathbf{z}), 0)$ such that the solution $(X,Y,\mathbf{L},\mathbf{v},\mathbf{z})$ of (4.3.7) with any initial condition

$$(X,Y,\mathbf{L},\mathbf{v},\mathbf{z})(0) = (X_0,Y_0,\mathbf{L}_0,\mathbf{v}_0,\mathbf{z}_0) \in \mathbb{R} \times \mathbb{R} \times (\mathbb{R}^n)^3$$

verifies $\mathbf{z}(t) = \mathbf{z}_0$ for all $t \geq 0$ and

$$\lim_{t \to +\infty} (X,Y,\mathbf{L},\mathbf{v})(t) = \Phi(\mathbf{z}_0)$$

exponentially. Moreover, X^* and Y^* are linear function of the χ_i.

Proof. First, in (4.3.11), we can write the system for $X,Y,\mathbf{L},\mathbf{v}$ when $\varepsilon = 0$ in the following form

$$0 = G(\mathbf{z}, (X,Y,\mathbf{L},\mathbf{v}))$$

with function $G(x_1,x_2) : \mathbb{R}^N \times \mathbb{R}^{2N+2} \to \mathbb{R}^{2N+2}$, $x_1 = \mathbf{z}$, $x_2 = (X,Y,\mathbf{L},\mathbf{v})$.
The Jacobian matrix of G respected to $\begin{pmatrix} X & Y & \mathbf{L} & \mathbf{v} \end{pmatrix}$ reads as as follows

$$A(x,t) = \begin{pmatrix} -\beta T^* & 0 & * & * & \dots & * & * & & & \\ \beta(S^* - T^* - kI^*) & -(m+\beta k T^*) & 0 & \dots & 0 & * & & & & \\ 0 & 0 & -m & 0 & \dots & 0 & * & & & \\ 0 & 0 & 0 & -m & \dots & 0 & * & & & \\ \ddots & & & & & & & & & \\ 0 & 0 & 0 & 0 & \dots & -m & * & & & \\ 0 & 0 & 0 & 0 & \dots & 0 & -\xi & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & -\xi & \dots & 0 \\ \ddots & & & & & & & & & \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & -\xi \end{pmatrix}. \tag{4.3.9}$$

Since $A(x,t)$ is block-diagonal matrix, it is easy to find the characteristic polynomial

$$(\lambda + \beta T^*)(\lambda + m + \beta k T^*)(\lambda + m)^n (\lambda + \xi)^n$$

which implies that all the eigenvalue of A have the negative real part.

Using the triangular structure of (4.3.11) the idea is to compute the limits when $\varepsilon \to 0$ step by step of $\mathbf{v}$, $\mathbf{L}$, X and Y in this order. Here we make a quick formal computation by simply plugging the limits obtained at one step into the equation of the next step.
Indeed, since (4.3.11) is equivalent to (4.3.7) but in the slow motion, we take $\varepsilon = 0$ in (4.3.7). We have directly $\mathbf{z}(t) = \mathbf{z}_0$ for all $t \geq 0$ and $v_i = e^{-\xi t} v_i(0) \to 0$ exponentially as $t \to +\infty$. Remark that taking $v_i = 0$ in the others equations leads to the simple change of variables : $I_i = I^* z_i$ and $J_i = T^* z_i$ that we can plug in

(4.3.4)-(4.3.5)-(4.3.6) to simplify the explicit computations.
Now we have the following exponential limits

$$L_i(t) \to \chi_3 \frac{1}{m} F_{L_i}(0,\mathbf{z}_0) = \chi_3 L_i^*(\mathbf{z}_0),$$

Denoting $\mathbf{L}^* = (L_i^*)_i$ and plugging this into the equation of X we have that exponentially:

$$X(t) \to -\frac{1}{\beta T^*}\left(\chi_1 F_X^1(0,\mathbf{z}_0) + \chi_2 F_X^2(0,\mathbf{z}_0) + \chi_3 F_X^3(\chi_3 \mathbf{L}^*(\mathbf{z}_0))\right) = X^*(\mathbf{z}_0).$$

Remark that by linearity of the F_X^i and the fact that $\chi_d^2 = \chi_d$ for each d, we have the simpler formula

$$X^*(\mathbf{z}_0) = -\frac{1}{\beta T^*}\left(\chi_1 F_X^1(0,\mathbf{z}_0) + \chi_2 F_X^2(0,\mathbf{z}_0) + \chi_3 F_X^3(\mathbf{L}^*(\mathbf{z}_0))\right). \tag{4.3.10}$$

Finally, using the same arguments we get

$$Y(t) \to Y^*(\mathbf{z}_0) \qquad \text{exponentially}$$

wherein we have note

$$Y^*(\mathbf{z}_0) = \frac{1}{m+\beta k T^*}\left(\beta(S^* - T^* - kI^*)X^*(\mathbf{z}_0) + \chi_1 F_Y^1(0,\mathbf{z}_0) + \chi_2 F_Y^2(0,\mathbf{z}_0) + \chi_4 F_Y^4(0,\mathbf{z}_0)\right).$$

□

The next step is to change the time scale. Taking $\tau = \varepsilon t$ in (4.3.7) we obtain[2] the following system which is equivalent to (4.3.7) but in the slow motion τ

$$\begin{cases}
\varepsilon \dfrac{\partial X}{\partial \tau} &= -\beta T^* X + \chi_1 F_X^1(\mathbf{v},\mathbf{z}) + \chi_2 F_X^2(\mathbf{v},\mathbf{z}) + \chi_3 F_X^3(\mathbf{L}) + O(\varepsilon) + \varepsilon \Delta X \\
\varepsilon \dfrac{\partial Y}{\partial \tau} &= \beta(S^* - T^* - kI^*)X - (m+\beta k T^*)Y + \chi_1 F_Y^1(\mathbf{v},\mathbf{z}) + \chi_2 F_Y^2(\mathbf{v},\mathbf{z}) + \chi_5 F_Y^5(\mathbf{v},\mathbf{z}) + O(\varepsilon) + \varepsilon \Delta Y \\
\varepsilon \dfrac{\partial L_i}{\partial \tau} &= -mL_i + \chi_3 F_{L_i}(\mathbf{v},\mathbf{z}) + O(\varepsilon) + \varepsilon \Delta L_i \\
\varepsilon \dfrac{\partial v_i}{\partial \tau} &= -\xi v_i + O(\varepsilon) + \varepsilon \dfrac{1}{|P|}\left[(2T^*\nabla T^* - I^*\nabla D^*)\nabla v_i + (T^*\nabla I^* - I^*\nabla T^*)\nabla z_i\right] + \varepsilon \Delta v_i \\
\dfrac{\partial z_i}{\partial \tau} &= F_{z_i}(X,Y,\mathbf{L},\mathbf{K},\mathbf{v},\mathbf{z}) + O(\varepsilon) + \dfrac{1}{|P|}\left[(-D^*\nabla T^* + 2T^*\nabla D^*)\nabla v_i + (-D^*\nabla I^* + 2T^*\nabla T^*)\nabla z_i\right] + \Delta z_i
\end{cases} \tag{4.3.11}$$

Using the notation of the Theorem 4.5, we see that the fast variables is $y(\tau) = (X,Y,\mathbf{L},\mathbf{v})$ and the slow variable is $x(\tau) = \mathbf{z}(\tau)$. The first step in applying the Theorem 4.5 is to take $\varepsilon = 0$ in (4.3.11) and to show that the fast variable converge exponentially to an attractor $\phi(\mathbf{z})$ which is parametrized by the slow variable.

Now, we take $\varepsilon = 0$ in (4.3.11) and

$$(X,Y,\mathbf{L},\mathbf{v})(\tau) = \Phi(\mathbf{z}(\tau)), \tag{4.3.12}$$

[2]We use the usual notation abuse. Rigorously speaking, we have to define $\widetilde{X}(\tau) = X\left(\frac{\tau}{\varepsilon}\right)$ and the same for each variables. Here we remove the $\widetilde{}$ for simplicity.

the $2N+2$ first equations are satisfied and the N last equations give the *slow system*

$$\frac{dz_i}{d\tau} = F_{z_i}(X^*(\mathbf{z}), Y^*(\mathbf{z}), \mathbf{L}^*(\mathbf{z}), 0, \mathbf{z}) + \frac{1}{|P|}\left(-D^*\nabla I^* + 2T^*\nabla T^*\right)\nabla z_i + \Delta z_i. \tag{4.3.13}$$

It's important to note that, since $\mathbf{v} = 0$ then (2.3.12) gives $\sum_{i=1}^{N} z_i = 1$. This is plausible because z_i reflects the frequency of strain i by the formula $I_i = I^* z_i$ for all i.
The Theorem 4.5 imply that the solutions of (2.3.12) together with (4.3.12) gives a good approximation of the original system (4.3.11) for a small enough but positive ε. Coming back to the original variables of the SIS system, we deduce the following result on error estimate, whose proof will be given in section 4.3.2.

Lemma 4.8. Let $T > 0$ be fixed. There exists $\varepsilon_0 > 0$ and $C_T > 0$ such that for any $\varepsilon \in (0, \varepsilon_0)$ we have for any solution of $(S, (I_i)_i, (I_{ij})_{ij})_{i,j}$ of (4.2.1) and $(z_i)_i$ of (2.3.12)

$$\int_\Omega \left| S\left(\frac{\tau}{\varepsilon}\right) - S^* \right|^2 + \sum_{i=1}^{N} \int_\Omega \left| I_i\left(\frac{\tau}{\varepsilon}\right) - I^* z_i(\tau) \right|^2 + \sum_{i,j=1}^{N} \int_\Omega \left| D^* z_i(\tau) z_j(\tau) - I_{ij}\left(\frac{\tau}{\varepsilon}\right) \right|^2 \leq \varepsilon C_T, \tag{4.3.14}$$

Proof. See section 4.3.2 □

It remains to compute explicitly the slow system (2.3.12). The following lemma shows that it suffices to compute independently each perturbation, that is $\mathscr{A} = \{d\}$ for $d = 1, \cdots, 5$. The case of a general $\mathscr{A}$ being just a sum of each simple case thanked to the following result.

Lemma 4.9. Let $\mathscr{A} \subset \{1, \cdots, 5\}$. Recall that $\chi_d = 1$ if $d \in \mathscr{A}$ and $\chi_d = 0$ if $d \notin \mathscr{A}$. The functions F_{z_i} for $i = 1, \cdots, N$ in (2.3.12) read

$$F_{z_i}(X^*(\mathbf{z}), Y^*(\mathbf{z}), \mathbf{L}^*(\mathbf{z}), 0, \mathbf{z}) = \sum_{d=1}^{5} \chi_d z_i f_{z_i}^d(\mathbf{z}),$$

where the functions $f_{z_i}^d$ do not depend on χ_d.
In particular, if $\mathscr{A} = \{d\}$ for some $d \in \{1,2,3,4,5\}$, then

$$F_{z_i}(X^*(\mathbf{z}), Y^*(\mathbf{z}), \mathbf{L}^*(\mathbf{z}), 0, \mathbf{z}) = z_i f_{z_i}^d(\mathbf{z}).$$

Proof. Taking $v_i = 0$ in (4.3.8) we see that there is two constant C_X and C_Y such that

$$F_{z_i}(X^*(\mathbf{z}), Y^*(\mathbf{z}), \mathbf{L}^*(\mathbf{z}), 0, \mathbf{z}) = z_i \left(C_X X^*(\mathbf{z}), + C_Y Y^*(\mathbf{z}) + \begin{pmatrix} 0 & 1 \end{pmatrix} P \mathscr{M}_{\mathscr{A}} P^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) + \chi_3 \gamma \begin{pmatrix} 0 & 1 \end{pmatrix} P \begin{pmatrix} 0 \\ L_i^*(\mathbf{z}) \end{pmatrix}.$$

Firstly, as it is show in the proof of the lemma 4.7, the expression of X^* and Y^* are both a linear combination of the χ_d.

Secondly, recalling that we have at this step $I_i = I^* z_i$, $J_i = T^* z_i$, $\mathbf{L}_i = \chi_3 \mathbf{L}^*$ and $\chi_d^2 = \chi_d$, for $d = 3$ and $d = 5$ particularly. Plugging this in (4.3.7), we see that the matrix $\mathscr{M}_{\mathscr{A}}$ is also a linear combination of the χ_d which yields for some functions $m_d(\mathbf{z})$ which do not depend on χ_d :

$$\begin{pmatrix} 0 & 1 \end{pmatrix} P \mathscr{M}_{\mathscr{A}} P^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \sum_{d \in \{1,2,3,4,5\}} \chi_d m_d(\mathbf{z}). \tag{4.3.15}$$

Thirdly, plugging $I_i = I^* z_i$ and $J_i = T^* z_i$, for all i in (4.3.4) we prove that

$$L_i^*(\mathbf{z}) = \frac{1}{2m}\beta k I^* T^* z_i \sum_{j=1}^{N} (u_{ij} + u_{ji}) z_j.$$

The result follows directly from three previous points. □

In the next section 4.3.2, these functions $f_{z_i}^d$ are explicitly compute for any d.

4.3.2 Main results and proofs

We reuse the computations in [120], in each case of $\mathscr{A} = \{d\}$, $d \in \{1,2,3,4,5\}$. We set that

$$\vec{\vartheta} = \frac{1}{|P|}\left(-D^* \nabla I^* + 2T^* \nabla T^*\right).$$

Note that $\vec{\vartheta} = 0$ if I^* and T^* do not depend on x.

In each following case of perturbation, by the similar argument, we obtain the slow system (2.3.12), respectively.

- Perturbations in transmission rates, $\mathscr{A} = \{1\}$

$$\begin{cases} \dfrac{\partial z_i}{\partial \tau} = \dfrac{2\beta S^* T^{*2}}{|P|} z_i \left(b_i - \sum\limits_{j=1}^{N} b_j z_j \right) + \vec{\vartheta} \cdot \nabla z_i + \Delta z_i, \\ \dfrac{\partial z_i}{\partial n}|_{\partial\Omega} = 0, \end{cases} \qquad 1 \le i \le N.$$

- Perturbations in clearance rates γ_i, $\mathscr{A} = \{2\}$

$$\begin{cases} \dfrac{\partial z_i}{\partial \tau} = \dfrac{\gamma I^* (I^* + T^*)}{|P|} \left(\nu_i - \sum\limits_{j=1}^{N} \nu_j z_j \right) z_i + \vec{\vartheta} \cdot \nabla z_i + \Delta z_i, \\ \dfrac{\partial z_i}{\partial n}|_{\partial\Omega} = 0, \qquad 1 \le i \le N. \end{cases}$$

- Perturbations in co-infection clearance rate γ_{ij}, $\mathscr{A} = \{3\}$

$$\begin{cases} \dfrac{\partial z_i}{\partial \tau} = \dfrac{\gamma T^* D^*}{|P|} \left[\sum\limits_{j=1}^{N} (u_{ij} + u_{ji}) z_j - \sum\limits_{j,l=1}^{N} \left(u_{jl} + u_{lj} \right) z_l z_j \right] z_i + \vec{\vartheta} \cdot \nabla z_i + \Delta z_i, \\ \dfrac{\partial z_i}{\partial n}|_{\partial\Omega} = 0, \qquad 1 \le i \le N. \end{cases}$$

- Perturbations in perturbations in transmission coefficients from mixed carriage p_{ij}^i, $\mathscr{A} = \{4\}$

$$\begin{cases} \dfrac{\partial z_i}{\partial \tau} = \dfrac{2mT^* D^*}{|P|} z_i \sum\limits_{j=1}^{N} \left(\omega_{ij}^i - \omega_{ji}^j \right) z_j + \vec{\vartheta} \cdot \nabla z_i + \Delta z_i, \\ \dfrac{\partial z_i}{\partial n}|_{\partial\Omega} = 0, \end{cases} \qquad 1 \le i \le N.$$

- Perturbations co-colonization interaction k_{ij}, $\mathscr{A}=\{5\}$

$$\begin{cases}\dfrac{\partial z_i}{\partial \tau}=\dfrac{-\beta T^*I^*D^*}{|P|}z_i\left[\displaystyle\sum_{j=1}^{N}\left(\frac{T^*}{D^*}\alpha_{ji}-\frac{I^*}{D^*}\alpha_{ij}\right)z_j-\sum_{j,l=1}^{N}\alpha_{jl}z_jz_l\right]+\vec{\vartheta}\cdot\nabla z_i+\Delta z_i, & 1\le i\le N.\\ \dfrac{\partial z_i}{\partial n}|_{\partial\Omega}=0,\end{cases}$$

Let $\mathscr{A}\subset\{1,2,3,4,5\}$. Using the notations in the previous section, (2.3.12) reads.

$$\boxed{\begin{aligned}\frac{\partial z_i}{\partial \tau}=&\Theta_1 z_i\left(b_i-\sum_{j=1}^{N}b_jz_j\right)+\Theta_2 z_i\left(-\nu_i+\sum_{j=1}^{N}\nu_jz_j\right)+\Theta_3 z_i\left[-\sum_{j=1}^{N}(u_{ij}+u_{ji})z_j+\sum_{j,l=1}^{N}(u_{jl}+u_{lj})z_lz_j\right]\\&+\Theta_4 z_i\sum_{j=1}^{N}\left(\omega_{ij}^{i}-\omega_{ji}^{j}\right)z_j+\Theta_5 z_i\left[\sum_{j=1}^{N}\left(\frac{T^*}{D^*}\alpha_{ji}-\frac{I^*}{D^*}\alpha_{ij}\right)z_j-\sum_{j,l=1}^{N}\alpha_{jl}z_jz_l\right]+\quad\vec{\vartheta}\cdot\nabla z_i+\Delta z_i\end{aligned}}\tag{4.3.16}$$

where Θ_i, $i=1,2,3,4,5$, are given by

$$\Theta_1(x)=\chi_1\frac{2\beta S^*T^{*2}}{|P|},\quad \Theta_2(x)=\chi_2\frac{\gamma I^*(I^*+T^*)}{|P|},\quad \Theta_3(x)=\chi_3\frac{\gamma T^*D^*}{|P|},\quad \Theta_4(x)=\chi_4\frac{2mT^*D^*}{|P|},\quad \Theta_5(x)=\chi_5\frac{\beta T^*I^*D^*}{|P|}.\tag{4.3.17}$$

It is useful to rewrite (4.3.16) using the pairwise invasion fitness between strains. Define

$$\Theta(x)=\Theta_1(x)+\Theta_2(x)+\Theta_3(x)+\Theta_4(x)+\Theta_5(x)\quad\text{ and }\theta_i(x)=\frac{\Theta_i(x)}{\Theta(x)}.\tag{4.3.18}$$

we see that $\theta_i(x)>0$ for each $i=1,2,3,4,5$ and $\theta_1+\theta_2+\theta_3+\theta_4+\theta_5=1$ for all x. For completeness, if $\mathscr{A}=\emptyset$ then we set $\Theta=1$. Using these notations, we obtain our main result.

Theorem 4.10. *Consider the system of equations*

$$\boxed{\begin{cases}\dfrac{\partial z_i}{\partial \tau}=\Theta z_i\left((\Lambda(x)\mathbf{z})_i-\mathbf{z}^T\Lambda(x)\mathbf{z}\right)+\vec{\vartheta}\cdot\nabla z_i+\Delta z_i,\ i=1,\cdots,N,\\ z_1+z_2+\cdots+z_N=1.\end{cases}}\tag{4.3.19}$$

where $\Lambda(x)$ is the square matrix of size $N\times N$ whose coefficient $(i;j)$ are the pairwise fitness $\lambda_i^j(x)$ which satisfy

$$\boxed{\lambda_i^j(x)=\theta_1(b_i-b_j)+\theta_2(-\nu_i+\nu_j)+\theta_3(-u_{ij}-u_{ji}+2u_{jj})+\theta_4\left(\omega_{ij}^{i}-\omega_{ji}^{j}\right)+\theta_5\left(\mu(\alpha_{ji}-\alpha_{ij})+\alpha_{ji}-\alpha_{jj}\right).}\tag{4.3.20}$$

with $\mu=\dfrac{I^*}{D^*}$.

Then, for any initial values of (4.2.1)*, for each $\tau_0>0$, $T>\tau_0$ arbitrarily and independent on ε, there is $\varepsilon_0>0$, $C>0$ and a vector of positive coefficients $\mathbf{z}_0\in\mathbb{R}^N$ verifying $\sum_{i=1}^{N}\mathbf{z}_{0,i}=1$, such that $\forall\varepsilon<\varepsilon_0$*

$$\int_\Omega\left|S^*(x)-S\left(x,\frac{\tau}{\varepsilon}\right)\right|^2+\sum_{i=1}^{N}\int_\Omega\left|I^*z_i(x,\tau)-I_i\left(x,\frac{\tau}{\varepsilon}\right)\right|^2+\sum_{i,j=1}^{N}\int_\Omega\left|D^*z_i(x,\tau)z_j(x,\tau)-I_{ij}\left(x,\frac{\tau}{\varepsilon}\right)\right|^2\le\varepsilon C,\quad\forall\tau\in(\tau_0,T).$$

where S, $(I_1, I_2, \ldots, I_N)$, $(I_{ij})_{i,j\in\{1,\ldots,N\}}$ *is the solution of* (4.2.1) *and* $(z_1, z_2, \ldots, z_N)$ *is the solution of reduced system* (4.3.19) *together with* $\mathbf{z}(0) = \mathbf{z}_0$.

This system (4.3.19) is a general replicator system with diffusion, which is studied in [35]. We back to the proof of Theorem 4.8.

Proof. We separate this proof into three steps, in which, we respectively show the approximation for S, I_i, $i = 1, \ldots, N$ using the theorems 4.5, 4.6, then prove the approximation holds for I_{ij}, $i, j = 1, \ldots, N$.

• **Firstly,** use the Theorem 4.6,we have that

$$\|S(x,t) - S^*(x)\|_2 = O\left(\sqrt{\varepsilon}\right). \tag{4.3.21}$$

On the other side, we note that the algebraic linear transformations to the new variables $(z_i, v_i)_{1\le i\le N}$; and $v_i \to 0$ when $\varepsilon \to 0$ (by the Theorem 4.5), which deduces that

$$\|S(x,t) - S^*(x)\|_2 + \sum_{i=1}^{N} \|I_i(x,t) - I^* z_i^s(x,t)\|_2 = O\left(\sqrt{\varepsilon}\right), \tag{4.3.22}$$

where $(z_1^s, z_2^s, \ldots, z_N^s)$ are solution of slow-fast system (4.3.7), noting that and changing time scale yielding the equivalent system .

• **Secondly,** by the lemma 4.7 and the same arguments in [120], we can verify the exponential stability condition of the Theorem 4.5. Hence, the solution of system (4.3.11) after changing time scale $\tau = \varepsilon t$ tends to the solution of (4.3.19) as $\varepsilon \to 0$ on $[\tau_0, T]$, with $\tau_0 > 0$, $T > \tau_0$. arbitrary and independent on ε.
Combine with the previous claim (4.3.22), we obtain that

$$\left\|S\left(x, \frac{\tau}{\varepsilon}\right) - S^*(x)\right\|_2 + \sum_{i=1}^{N} \left\|I^* z_i(x,\tau) - I_i\left(x, \frac{\tau}{\varepsilon}\right)\right\|_2 = O\left(\sqrt{\varepsilon}\right). \tag{4.3.23}$$

• **Thirdly,** we make a result for solutions $I_{ij}(x,t)$, $1 \le i, j \le N$. For the sake of shortness, we remark that each partial differential equation in this proof associates with Neumann boundary condition and we will not remark it in each equation. Assume $\left(I_{ij}^r\right)_{1\le i,j\le N}$ to be the solution of

$$\frac{\partial I_{ij}}{\partial t} = -mI_{ij} + \beta k I^*(x) T^*(x) z_i(x,\tau) z_j(x,\tau) + \varepsilon \Delta I_{ij}, \qquad 1 \le i, j \le N \tag{4.3.24}$$

Then, for each $\tau_0 > 0$ and $T > \tau_0$, we claim that $\sum_{i,j=1}^{N} \left\|I_{ij}\left(x, \frac{\tau}{\varepsilon}\right) - I_{ij}^r\left(x, \frac{\tau}{\varepsilon}\right)\right\|_2 = O(\sqrt{\varepsilon})$ for any $\tau \in [\tau_0, T]$.
Indeed, by the property of solutions of (4.2.16) and (4.3.24), we have that (we omit the term (x)-dependence on x, of parameters, for the sake of convenience)

$$\begin{aligned}
\frac{\partial I_{ij}}{\partial t}\left(x, \frac{\tau}{\varepsilon}\right) &= -m_{ij} I_{ij}\left(x, \frac{\tau}{\varepsilon}\right) + \beta_j k_{ij} I_i\left(x, \frac{\tau}{\varepsilon}\right) J_j\left(x, \frac{\tau}{\varepsilon}\right) &&+ \varepsilon \Delta I_{ij}\left(x, \frac{\tau}{\varepsilon}\right) \\
\frac{\partial I_{ij}^r}{\partial t}\left(x, \frac{\tau}{\varepsilon}\right) &= -m I_{ij}^r\left(x, \frac{\tau}{\varepsilon}\right) + \beta k I^*(x) T^*(x) z_i(x,\tau) z_j(x,\tau) &&+ \varepsilon \Delta I_{ij}^r\left(x, \frac{\tau}{\varepsilon}\right)
\end{aligned} \tag{4.3.25}$$

which implies

$$\begin{aligned}\frac{\partial}{\partial t}\left(I_{ij}\left(x,\frac{\tau}{\varepsilon}\right)-I^r_{ij}\left(x,\frac{\tau}{\varepsilon}\right)\right)=&\varepsilon\Delta\left(I_{ij}\left(x,\frac{\tau}{\varepsilon}\right)-I^r_{ij}\left(x,\frac{\tau}{\varepsilon}\right)\right)-m\left(I_{ij}\left(x,\frac{\tau}{\varepsilon}\right)-I^r_{ij}\left(x,\frac{\tau}{\varepsilon}\right)\right)\\&+\left(\beta_j k_{ij} I_i\left(x,\frac{\tau}{\varepsilon}\right)J_j\left(x,\frac{\tau}{\varepsilon}\right)-\beta k I^*(x)T^*(x)z_i(x,\tau)z_j(x,\tau)\right)-\varepsilon\gamma u_{ij}I_{ij}\left(x,\frac{\tau}{\varepsilon}\right).\end{aligned} \tag{4.3.26}$$

Then for all $1\le i,j\le N$, using the Theorem 4.6, we observe that

$$\left\|I_{ij}\left(x,\frac{\tau}{\varepsilon}\right)-I^r_{ij}\left(x,\frac{\tau}{\varepsilon}\right)\right\|_2=O(\sqrt{\varepsilon}). \tag{4.3.27}$$

Note that, $k(x)\frac{I^*(x)T^*(x)}{S^*(x)}=D^*(x)$, it suffices to compute the solution I^r_{ij}, which satisfies $\left\|I^r_{ij}-D^*(x)z_i(\varepsilon t)z_j(\varepsilon t)\right\|_2=O\left(\sqrt{\varepsilon}\right)$ for all $1\le i,j\le N$, by the Theorem 11.19 in [189] again. Combining with (4.3.27) implies that $\left\|I_{ij}-D^*(x)z_i(\varepsilon t)z_j(\varepsilon t)\right\|_1=O\left(\sqrt{\varepsilon}\right)$.

Combine the results in three above steps, we get the conclusion of the lemma 4.8. □

4.4 Models with fast diffusion

4.4.1 The general model and the Central Manifold Theorem

Keeping the same notations of the previous sections, we now study the following system, where the rates of diffusion are large.

$$\begin{cases}\dfrac{\partial S}{\partial t} &=r(x)(1-S)+\displaystyle\sum_{i=1}^N\gamma_i(x)I_i+\sum_{i,j=1}^N\gamma_{ij}(x)I_{ij}-S\sum_{i=1}^N\beta_iJ_i \quad+\dfrac{d}{\varepsilon}\Delta S & ,\\ \dfrac{\partial I_i}{\partial t} &=\beta_iJ_iS-(r(x)+\gamma_i(x))I_i-I_i\displaystyle\sum_{j=1}^N k_{ij}\beta_jJ_j \quad+\dfrac{d}{\varepsilon}\Delta I_i, & 1\le i\le N,\\ \dfrac{\partial I_{ij}}{\partial t} &=k_{ij}(x)\beta_jI_iJ_j-(r(x)+\gamma_{ij}(x))I_{ij} \quad+\dfrac{d}{\varepsilon}\Delta I_{ij}, & 1\le i,j\le N,\end{cases} \tag{4.4.1}$$

with the Neumann boundary conditions $\dfrac{\partial S}{\partial n}=\dfrac{\partial I_i}{\partial n}=\dfrac{\partial I_{ij}}{\partial n}=0$ for all $1\le i,j\le N$ on the boundary of Ω and given initial conditions.
Accordingly, this system (4.4.1) can be shortly written as

$$\begin{cases}\dfrac{\partial}{\partial t}\mathbf{W}(x,t)=\mathscr{F}(x,\mathbf{W}(x,t))+\dfrac{1}{\varepsilon}K\mathbf{W}(x,t),\\ \dfrac{\partial}{\partial n}\mathbf{W}(x,t)=0, \qquad x\in\partial\Omega\end{cases} \tag{4.4.2}$$

where

$$\mathbf{W}(x,t)=(S,I_1,\dots,I_N,I_{11},\dots,I_{NN})$$

and K is the operator $\begin{pmatrix}d\Delta\cdot & \dots & 0\\ \vdots & \ddots & \vdots\\ 0 & \dots & d\Delta\end{pmatrix}$.

We set the notation $A = d\Delta$. When seen as an operator on $L^2(\Omega)$, the operator A^2 with the formula of A, accompanied with homogeneous Neumann boundary conditions, is defined as follows, see [43].

$$\begin{aligned} &D\left(A^2\right) = \left\{U \in H^1(\Omega) : \exists V \in L^2(\Omega), \forall \phi \in H^1(\Omega), \int \nabla U(x) \nabla \phi(x) dx = -d \int V(x) \phi(x) dx\right\}, \\ &A^2 U := V, \quad U \in D\left(A^2\right). \end{aligned} \tag{4.4.3}$$

In order to obtain uniform estimates, we prefer to focus on the operator $A^\infty := A$ acting on $C^0(\bar{\Omega})$ with sup norm. Hence, we define

$$\begin{aligned} &D(A^\infty) := \left\{U \in D\left(A^2\right) \cap C\left(\bar{\Omega}\right), A^2 U \in C\left(\bar{\Omega}\right)\right\}, \\ &A^\infty U = A^2 U, \quad U \in D(A^\infty). \end{aligned} \tag{4.4.4}$$

Then we have that

$$E_0 := \ker(A^\infty) = \operatorname{span}(1) = \mathbb{R} \text{ and } \operatorname{Im}(A^\infty) \subset \left\{U \in C^0(\bar{\Omega}), \int_\Omega U = 0\right\} = F_0. \tag{4.4.5}$$

One gets $C^0\left(\bar{\Omega}\right) = \ker A^\infty \oplus \operatorname{Im} A^\infty$. Now we define the Banach space $\left(C^0\left(\bar{\Omega}\right)\right)^{N^2+N+1}$ together with the norm

$$\left\|\left(U_1, \ldots, U_{N^2+N+1}\right)\right\|_\infty = \|U_1\|_\infty + \cdots + \left\|U_{N^2+N+1}\right\|_\infty \tag{4.4.6}$$

and the operator $\left(A^\infty\right)^{N^2+N+1}$ acting on each coordinate of $\left(C^0\left(\bar{\Omega}\right)\right)^{N^2+N+1}$. The kernel and the range of this operator are respectively

$$E := \ker\left(\left(A^\infty\right)^{N^2+N+1}\right) = \mathbb{R}^{N^2+N+1} \qquad \text{and} \qquad F := \left(F_0\right)^{N^2+N+1}. \tag{4.4.7}$$

Hence we have $\left(C^0\left(\bar{\Omega}\right)\right)^{N^2+N+1} = \bar{E} \oplus \bar{F}$. The projection of $\left(C^0\left(\bar{\Omega}\right)\right)^{N^2+N+1}$ on E and F, denoted by Π_E and Π_F respectively, given explicit by

$$\Pi_E\left(V_1, \ldots, V_{N^2+N+1}\right) = \frac{1}{|\Omega|}\left(\int_\Omega V_1, \ldots, \int_\Omega V_{N^2+N+1}\right); \qquad \Pi_F = \mathrm{Id} - \Pi_E. \tag{4.4.8}$$

For all $u \in \left(C^0\left(\bar{\Omega}\right)\right)^{N^2+N+1}$ we rewrite it into $u = \mathbf{X} + \mathbf{Y}$ with $\mathbf{X} \in E$ and $\mathbf{Y} \in F$. We change the system (4.4.1) on an equivalent slow-fast form by projecting (4.4.1) on E and F respectively. The slow variable $\mathbf{X} := \Pi_E(\mathbf{W}) \in E$ is the vector

$$\mathbf{X} = \left(\frac{1}{|\Omega|}\int_\Omega S, \frac{1}{|\Omega|}\int_\Omega I_1, \ldots, \frac{1}{|\Omega|}\int_\Omega I_N, \frac{1}{|\Omega|}\int_\Omega I_{11}, \ldots, \frac{1}{|\Omega|}\int_\Omega I_{NN}\right) \in \mathbb{R}^{N^2+N+1}$$

and the fast variable is $\mathbf{Y} := \Pi_F \mathbf{W} = \mathbf{W} - \mathbf{X} \in F$. Projecting the system (4.4.1) on E and F yields to the equivalent system

$$(S_\varepsilon) : \begin{cases} \dfrac{d}{dt}\mathbf{X}(t) = f(\mathbf{X},\mathbf{Y}) \\ \dfrac{d}{dt}\mathbf{Y}(t) = g(\mathbf{X},\mathbf{Y}) + \dfrac{1}{\varepsilon}K\mathbf{Y} \\ \dfrac{\partial}{\partial n}\mathbf{X} = 0 \\ \dfrac{\partial}{\partial n}\mathbf{Y} = 0 \\ \mathbf{X}(0) = \Pi_E(\mathbf{W}(0)) \\ \mathbf{Y}(0) = \Pi_F(\mathbf{W}(0)) \end{cases} \tag{4.4.9}$$

For the end of this section, we state the Central Manifold Theorem 4.11 and the Theorem of convergence towards the central manifold. These theorems may be proved in [42, 43]. Let us begin by a version of the central manifold Theorem for an elliptic operator K. This Theorem claims the existence of an invariant manifold for the slow-fast system which allows to defined several reduced systems.

Theorem 4.11. ***(Central Manifold Theorem)*** *Let E and F be two Banach spaces. Define $f(X,Y) \in C^1(E \times F; E)$ and $g(X,Y) \in C^1(E \times F; F)$. Assume that f and g are uniformly bounded as well than their first derivatives. Let K be an operator with domain $D(K) \subset F$. Assume that K generates an analytical semi-group $\exp(tK)$ of linearly operators on F and that there exists $\mu > 0$ such that*

$$\forall t \geq 0; \qquad \forall \varepsilon \in (0,1], \qquad \left\| \exp\left(\frac{t}{\varepsilon}K\right) Y \right\|_F \leq C \|Y\|_F \exp\left(-\mu \frac{t}{\varepsilon}\right). \tag{4.4.10}$$

For all initial condition $(x_0, y_0) \in E \times F$ and for all $\varepsilon \in (0,1]$, one defines $X^\varepsilon(t, x_0, y_0) \equiv X^\varepsilon(t)$ and $Y^\varepsilon(t, x_0, y_0) \equiv Y^\varepsilon(t)$ the solution, for $t \geq 0$, of the differential system

$$(S_\varepsilon) : \qquad \begin{cases} \dfrac{d}{dt}X^\varepsilon(t) = f(X^\varepsilon(t), Y^\varepsilon(t), \varepsilon), \\ \dfrac{d}{dt}Y^\varepsilon(t) = g(X^\varepsilon(t), Y^\varepsilon(t), \varepsilon) + \dfrac{1}{\varepsilon}KY^\varepsilon(t), \\ X^\varepsilon(0) = x_0, Y^\varepsilon(0) = y_0. \end{cases} \tag{4.4.11}$$

Then there exists $\varepsilon_0 > 0$ such that, for all $\varepsilon \in (0, \varepsilon_0)$, the system (S_ε) admits a central manifold C_ε in the following sense.

1. *There exists a function $h(X,\varepsilon) \in C^1(E \times [0,\varepsilon_0]; F)$ such that, for all $\varepsilon \in (0,\varepsilon_0]$, $C_\varepsilon = \{(X, h(X,\varepsilon)); X \in E\}$ is invariant under the semi flow generated by S_ε for $t \geq 0$. Moreover, we have that $\|h(\cdot,\varepsilon)\|_{L^\infty(E,F)} = O(\varepsilon)$ as $\varepsilon \to 0$.*

2. *The function $h(x,\varepsilon)$ satisfies the partial differential equation*

$$D_x h(x,\varepsilon) f(x, h(x,\varepsilon), \varepsilon) = \frac{K}{\varepsilon} h(x,\varepsilon) + g(x, h(x,\varepsilon), \varepsilon), \tag{4.4.12}$$

where $D_x h$ stands for $\dfrac{Dh}{Dx}$. On top of that, any bounded function $\tilde{h}$ such that $\left\|\tilde{h}\right\|_{L^\infty}$, $\left\|D_x\tilde{h}\right\|_{L^\infty} \leq 1$, and such that we have

$$D_x\tilde{h}(x,\varepsilon) f\left(x,\tilde{h}(x,\varepsilon),\varepsilon\right) = \frac{K}{\varepsilon}\tilde{h}(x,\varepsilon) + g\left(x,\tilde{h}(x,\varepsilon),\varepsilon\right) + O(\varepsilon) \tag{4.4.13}$$

in L^∞, also necessarily satisfies

$$\left\|h-\tilde{h}\right\|_{L^\infty} = O(\varepsilon). \tag{4.4.14}$$

This Theorem provides the existence of a manifold C_ε which is invariant for the system (4.4.11) and parametrized by the slow variable $X^\varepsilon \in E$ In our application, E is finite dimensional so that the system on C^ε is a finite dimensional system. After showing that the solutions are close to the central manifold, up to an exponentially small error term, we can reduce the study to a system on the invariant manifold C^ε. This finite dimensional system approach the original problem in a sense that is specified below.
More precisely, let us define the following reduced system. We do not precise the initial data at this step.

$$\left(S_\varepsilon^{[\infty]}\right): \quad \frac{d}{dt}X^{\varepsilon,[\infty]}(t) = f\left(X^{\varepsilon,[\infty]}(t), h\left(X^{e,[\infty]}(t),\varepsilon\right),\varepsilon\right), \quad Y^{\varepsilon,[\infty]}(t) = h\left(X^{\varepsilon,[\infty]}(t),\varepsilon\right). \tag{4.4.15}$$

When the original data belongs to this manifold, that is if $Y^\varepsilon(0) = h(X^\varepsilon(0),0)$, (4.4.15) describes the exact dynamics of (4.4.11). In general, if $Y^\varepsilon(0) \neq h(X^\varepsilon(0),\varepsilon)$ and the solutions do not belong to C_ε. However, the initial data can be slightly modified so that the solution of (4.4.11) are exponentially close to the solution of (4.4.15).

Note that, $h(X,\varepsilon)$ admits an asymptotic expansion of the form $h(X,\varepsilon) = \sum_{k=1}^{r-1}\varepsilon^k h_k(X) + O(\varepsilon^r)$, which is explicitly calculable provided the functions f and f have C^r smoothness. The approximate $h(X,\varepsilon) \approx \sum_{k=1}^{r}\varepsilon^k h_k(X)$ leads to the writing of reduced systems of order r (see [42]). This paper focus only on the case $r=1$. By this assumption, we obtain the following reduced system

$$\left(S_\varepsilon^{[0]}\right): \quad \frac{d}{dt}X^{\varepsilon,[0]}(t) = f\left(X^{\varepsilon,[0]}(t),0,\varepsilon\right), \quad Y^{\varepsilon,[0]}(t) = h\left(X^{\varepsilon,[0]}(t),\varepsilon\right). \tag{4.4.16}$$

An important fact in the sequel is that the dynamic of $S_\varepsilon^{[\infty]}$ is completely determined by its first equation: the following O.D.E system

$$(S_\varepsilon^c): \quad \frac{d}{dt}X^{\varepsilon,[\infty]}(t) = f\left(X^{\varepsilon,[\infty]}(t), h\left(X^{\varepsilon,[\infty]}(t),\varepsilon\right)\right)$$

and S_ε^c can be seen as a regular perturbation of the first equation of $S_\varepsilon^{[0]}$, that is

$$(S_0^c), \quad \frac{d}{dt}X^{[0]}(t) = f\left(X^{[0]}(t),0\right).$$

4.4.2 Application of the Central Manifold Theorem and main results

In order to apply the Central Manifold Theorem and related results, we need that the operator K define a C^0 semi-group of contraction on F. Note that, the assumed smoothness of $\partial\Omega$ implies that the operator A^∞ generates a C^0 semi group of contraction on $\left(C\left(\bar{\Omega}\right)\right)^{N^2+N+1}$, see [27]. Denoting $\exp\left(tA_2^\infty\right)$ this semi-group, we

deduce that

$$\forall t \geq 0, \qquad \|\exp(tT^{\infty})v\|_{\infty} \leq \|v\|_{\infty}. \tag{4.4.17}$$

Lemma 4.12. The restriction of $\tilde{A}$ of A^{∞} to the subspace $F_0 = \left\{u \in C^0\left(\bar{\Omega}\right) : \int_{\Omega} u = 0\right\}$ is the generator of a C^0 semi-group of strict contraction $\exp\left(t\tilde{A}\right)$ on F_0 verifying for some $\mu > 0$

$$\forall v \in F_0, \qquad \left\|\exp\left(t\tilde{A}\right)v\right\|_{\infty} \leq e^{-\mu t}\|v\|_{\infty}. \tag{4.4.18}$$

Proof. F_0 is closed in $C^0\left(\bar{\Omega}\right)$ and is clearly invariant under $\exp(tA^{\infty})$ by its definition. It follows (from [169] p. 123) that $\tilde{A}$ is the generator of a C^0 semi-group of contraction on F_0.

On the other side, it is well known that the the Laplacian operator on $C^0\left(\bar{\Omega}\right)$ has the discrete spectrum $\sigma(A)$ which totally lies in the negative half line. Since $\sigma\left(\tilde{A}\right) \subset \sigma(A^{\infty})$ and $0 \notin \sigma\left(\tilde{A}\right)$, one has that $\sigma\left(\tilde{A}_2\right) \subset (-\infty, -\lambda_1]$ (for some $\lambda_1 > 0$). Apply the Theorem 4.3 (p.118) in [169], we have the conclusion of the lemma. □

We have the following result.

Proposition 4.13. K is the generator of a C^0 semi group $\exp(tK)$ on F verifying

$$\|\exp(tK)v\|_F \leq e^{-\mu t}\|v\|_F. \tag{4.4.19}$$

Now, we need to show that the function $f = \Pi_E \mathscr{F}$ and $g = \Pi_F \mathscr{F}$ are smooth enough. By the same arguments of Lemma 4.3 in [43] and note that $\mathscr{F}$ is the vector-valued function whose each component is a multi-variable polynomial. This result can be stated as follows.

Lemma 4.14. The function f and g have C^1 smoothness when acting on $E \times F$.

By the Central Manifold Theorem, there exists a manifold $\mathscr{M}^{\varepsilon} = \{(x, h(x,\varepsilon)), x \in E\} \in E \times F$ which is invariant for (S^{ε}). It verifies moreover $h(x^{\varepsilon}, \varepsilon) = O(\varepsilon)$ and $\mathscr{M}^{\varepsilon}$ attracts any trajectory exponentially fast in time.

Recalling E_0 defined in (4.4.5) and denoting $\Pi_{E_0}(U) = \dfrac{1}{|\Omega|}\int_{\Omega} U$ for all $U \in C^0\left(\bar{\Omega}\right)$.

Setting that $\bar{S} = \Pi_{E_0}(S)$, $\bar{I}_i = \Pi_{E_0}(I_i)$ and $\bar{I}_{ij} = \Pi_{E_0}(I_{ij})$, for all $1 \leq i, j \leq N$.

Since $h(x^{\varepsilon}, \varepsilon) = O(\varepsilon)$ as $\varepsilon \to 0$, one obtains the approximation of the slow manifold to be $\dfrac{\partial \mathbf{X}}{\partial t} = f(\mathbf{X}, 0)$ as follows

$$\begin{cases} \dfrac{d}{dt}\bar{S} &= \Pi_{E_0}(r)\left(1-\bar{S}\right) + \Pi_{E_0}(\gamma_i)\bar{I}_i + \Pi_{E_0}(\gamma_{ij})\bar{I}_{ij} - \bar{S}\displaystyle\sum_{i=1}^{N}\Pi_{E_0}(\beta_i J_i) \\ \dfrac{d}{dt}\bar{I}_i &= \Pi_{E_0}(\beta_i J_i)\bar{S} - \left(\Pi_{E_0}(r) + \Pi_{E_0}(\gamma_i)\right)\bar{I}_i - \bar{I}_i\displaystyle\sum_{j=1}^{N}\left(\Pi_{E_0}(\beta_j k_{ij})\bar{I}_i + \sum_{j=1}^{N}\left(\Pi_{E_0}\left(\beta_j p_{ij}^i k_{ij}\right)\bar{I}_{ij} + \Pi_{E_0}\left(\beta_j p_{ji}^i k_{ij}\right)\bar{I}_{ji}\right)\right) \\ \dfrac{d}{dt}\bar{I}_{ij} &= \bar{I}_i\left(\Pi_{E_0}(\beta_j k_{ij})\bar{I}_j + \displaystyle\sum_{l=1}^{N}\left(\Pi_{E_0}\left(\beta_j p_{jl}^i k_{ij}\right)\bar{I}_{jl} + \Pi_{E_0}\left(\beta_j p_{lj}^i k_{ij}\right)\bar{I}_{lj}\right)\right) - \left(\Pi_{E_0}(r) + \Pi_{E_0}(\gamma_{ij})\right)\bar{I}_{ij} \end{cases} \tag{4.4.20}$$

Now, we make a *quasi neutral* assumption as in Table 4.1 and wish to transform (4.4.20) to apply result in [120]. It suffices to write the parameters $\Pi_{E_0}(\beta_j k_{ij})$, $\Pi_{E_0}\left(\beta_j p_{ij}^i k_{ij}\right)$, etc, in (4.4.20) as the forms in [120].

Indeed, we first denote

$$\begin{aligned} &\bar{r} = \Pi_{E_0}(r), \quad \bar{\beta} = \Pi_{E_0}(\beta), \quad \bar{\gamma} = \Pi_{E_0}(\gamma), \\ &\bar{b}_i = \frac{\Pi_{E_0}(\beta b_i)}{\Pi_{E_0}(\beta)}, \quad \bar{\nu}_i = \frac{\Pi_{E_0}(\gamma \nu_i)}{\Pi_{E_0}(\gamma)}, \quad \bar{u}_{ij} = \frac{\Pi_{E_0}(\gamma u_{ij})}{\Pi_{E_0}(\gamma)}, \quad 1 \le i,j \le N \end{aligned} \tag{4.4.21}$$

then $\Pi_{E_0}(\beta_i) = \bar{\beta}\left(1+\varepsilon \bar{b}_i\right) := \bar{\beta}_i$, $\Pi_{E_0}(\gamma_i) = \bar{\gamma}(1+\varepsilon\bar{\nu}_i) := \bar{\gamma}_i$ and $\Pi_{E_0}(\gamma_{ij}) = \bar{\gamma}(1+\varepsilon\bar{u}_{ij}) := \bar{\gamma}_{ij}$. Next, we set that

$$\bar{p}^s_{ij} = \frac{1}{2} + \varepsilon \bar{\omega}^s_{ij}, \qquad \text{with} \quad \bar{\omega}^s_{ij} = \frac{\Pi_{E_0}\left(\beta_i \omega^s_{ij}\right)}{\Pi_{E_0}(\beta_i)};$$

and

$$\bar{k}_{ij} = \bar{k} + \varepsilon \bar{\alpha}_{ij}, \qquad \text{with} \quad \bar{k} = \frac{\Pi_{E_0}(\beta k)}{\Pi_{E_0}(\beta)} \quad \text{and} \quad \bar{\alpha}_{ij} = \frac{1}{\varepsilon}\left[\frac{\Pi_{E_0}(\beta_j k_{ij})}{\Pi_{E_0}(\beta_j)} - \bar{k}\right]. \tag{4.4.22}$$

It is necessary to note that $\bar{p}^s_{ij} \neq \Pi_{E_0}\left(p^s_{ij}\right)$ and $\bar{k}_{ij} \neq \Pi_{E_0}(k_{ij})$ to not make mistakes.

Then, we have that,

$$\Pi_{E_0}\left(\beta_i p^s_{ij}\right) = \bar{\beta}_i \bar{p}^s_{ij}, \qquad \Pi_{E_0}(\beta_j k_{ij}) = \bar{\beta}_j \bar{k}_{ij}, \quad \forall 1 \le i,j \le N. \tag{4.4.23}$$

We will show that $\bar{\alpha}_{ij} = O(1)$, indeed,

$$\begin{aligned} \bar{\alpha}_{ij} &= \frac{1}{\varepsilon}\left[\frac{\int_\Omega (\beta+\varepsilon b_j)(k+\varepsilon\alpha_{ij})}{\int_\Omega \beta + \varepsilon\int_\Omega b_j} - \frac{\int_\Omega \beta k}{\int_\Omega \beta}\right] \\ &= \frac{1}{\varepsilon}\left[\left(\frac{\int_\Omega \beta k}{\int_\Omega \beta + \varepsilon\int_\Omega b_j} - \frac{\int_\Omega \beta k}{\int_\Omega \beta}\right) + \varepsilon\frac{\int_\Omega (k b_j + \beta\alpha_{ij})}{\int_\Omega \beta + \varepsilon\int_\Omega b_j} + \varepsilon^2 \frac{\int_\Omega b_j \alpha_{ij}}{\int_\Omega \beta + \varepsilon\int_\Omega b_j}\right] \\ &= \frac{\int_\Omega \beta k}{\int_\Omega \beta} \cdot \frac{-\frac{\int_\Omega b_j}{\int_\Omega \beta}}{1+\varepsilon\frac{\int_\Omega b_j}{\int_\Omega \beta}} + \frac{\int_\Omega (k b_j + \beta\alpha_{ij})}{\int_\Omega \beta + \varepsilon\int_\Omega b_j} + \varepsilon\frac{\int_\Omega b_j\alpha_{ij}}{\int_\Omega \beta + \varepsilon\int_\Omega b_j}. \end{aligned}$$

Combining this with direct calculations, we have that

$$\left\|\Pi_{E_0}(\beta_j k_{ij} p^s_{mn}) - \bar{\beta}_j \bar{k}_{ij} \bar{p}^s_{mn}\right\| = O(\varepsilon), \quad s \in \{m,n\}, \qquad \forall 1 \le i,j,m,n \le N.$$

Indeed, for $s \in \{m,n\}$, for all $1 \le i,j,m,n \le N$, denote that $\varepsilon\psi^{s,ij}_{mn}$ to be $\Pi_{E_0}(\beta_j k_{ij} p^s_{mn}) - \bar{\beta}_j\bar{k}_{ij}\bar{p}^s_{mn}$ then

$$\begin{aligned} \varepsilon\psi^{s,ij}_{mn} &= \varepsilon\left[\frac{1}{2}\left(\Pi_{E_0}(k b_j) - \bar{k}\bar{b_j}\right) + \left(\Pi_{E_0}(\beta k \omega^s_{mn}) - \bar{\beta}\bar{k}\bar{\omega}^s_{mn}\right)\right] \\ &+\varepsilon^2\left[\frac{1}{2}\left(\Pi_{E_0}(\alpha_{ij} b_j) - \bar{\alpha}_{ij}\bar{b}_j\right) + \frac{1}{2}\left(\Pi_{E_0}(\beta\alpha_{ij}\omega^s_{mn}) - \bar{\beta}\bar{\alpha}_{ij}\bar{\omega}^s_{mn}\right) + \left(\Pi_{E_0}(k b_j \omega^s_{mn}) - \bar{k}\bar{b}_j\bar{\omega}^s_{mn}\right)\right] \\ &+\varepsilon^3\left[\Pi_{E_0}(\alpha_{ij} b_j \omega^s_{mn}) - \bar{\alpha}_{ij}\bar{b}_j\bar{\omega}^s_{mn}\right]. \end{aligned}$$

For the sake of applying the result in [120], we make an assumption that

Assumption 4.15. p^s_{ij} does not depend on x for all $1 \le i,j \le N$ and $s \in \{i,j\}$.

Hence, $\Pi_{E_0}\left(\beta_j k_{ij} p^s_{mn}\right) = \bar{\beta}_j \bar{k}_{ij} \bar{p}^s_{mn}$ and the system (4.4.20) becomes

$$\begin{cases} \dfrac{d\bar{S}}{dt} &= \bar{r}\left(1-\bar{S}\right) + \displaystyle\sum_{i=1}^{N} \bar{\gamma}_i \bar{I}_i + \sum_{i,j=1}^{N} \bar{\gamma}_{ij} \bar{I}_{ij} - \bar{S} \sum_{i=1}^{N} \bar{\beta}_i \bar{J}_i \\ \dfrac{d\bar{I}_i}{dt} &= \bar{\beta}_i \bar{J}_i \bar{S} - \left(\bar{r} + \bar{\gamma}_i\right) \bar{I}_i - \bar{I}_i \displaystyle\sum_{j=1}^{N} \bar{k}_{ij} \bar{\beta}_j \bar{J}_j, \qquad 1 \le i \le N, \\ \dfrac{d\bar{I}_{ij}}{dt} &= \bar{k}_{ij} \bar{\beta}_j \bar{I}_i \bar{J}_j - \left(\bar{r} + \bar{\gamma}_{ij}\right) \bar{I}_{ij}, \qquad 1 \le i, j \le N, \end{cases} \tag{4.4.24}$$

where

$$\bar{\beta}_i = \bar{\beta}\left(1 + \varepsilon \bar{b}_i\right), \quad \bar{\gamma}_i = \bar{\gamma}\left(1 + \varepsilon \bar{\nu}_i\right), \quad \bar{\gamma}_{ij} = \bar{\gamma}\left(1 + \bar{u}_{ij}\right), \quad \bar{p}^s_{ij} = \frac{1}{2} + \varepsilon \bar{\omega}^s_{ij}, \quad \bar{k}_{ij} = \bar{k} + \varepsilon \bar{\alpha}_{ij}$$

and

$$\bar{J}_i = \bar{I}_i + \sum_{j=1}^{N} \left(\bar{p}^i_{ij} \bar{I}_{ij} + \bar{p}^i_{ji} \bar{I}_{ji}\right), \qquad \forall 1 \le i \le N$$

Before applying the result in [120], we make the following assumption on the basic reproduction ratio.

Assumption 4.16. Assume that $\int_\Omega \beta\left(x\right) dx > \int_\Omega m\left(x\right) dx$, which means $\dfrac{\int_\Omega \beta\left(x\right) dx}{\int_\Omega m\left(x\right) dx} > 1$.

Applying the result in [120] for (4.4.24), we have the following theorem. Initially, we define that

$$\boxed{S^* = \frac{\bar{m}}{\bar{\beta}}, \qquad T^* = 1 - S^*, \qquad I^* = \frac{\bar{m} T^*}{\bar{m} + \bar{\beta} \bar{k} T^*}, \qquad D^* = T^* - I^*} \tag{4.4.25}$$

and

$$\Theta = \Theta_1 + \Theta_2 + \Theta_3 + \Theta_4 + \Theta_5 \quad \text{ and } \theta_i = \frac{\Theta_i}{\Theta} \tag{4.4.26}$$

where

$$\Theta_1 = \chi_1 \frac{2\bar{\beta} S^* T^{*2}}{|P|}, \quad \Theta_2 = \chi_2 \frac{\bar{\gamma} I^* \left(I^* + T^*\right)}{|P|}, \quad \Theta_3 = \chi_3 \frac{\bar{\gamma} T^* D^*}{|P|}, \quad \Theta_4 = \chi_4 \frac{2\bar{m} T^* D^*}{|P|}, \quad \Theta_5 = \chi_5 \frac{\bar{\beta} T^* I^* D^*}{|P|}. \tag{4.4.27}$$

We see that $\theta_i > 0$ for each $i = 1,2,3,4,5$ and $\theta_1 + \theta_2 + \theta_3 + \theta_4 + \theta_5 = 1$. Using these notations, we obtain our main result.

Theorem 4.17. *Consider the system of equations*

$$\boxed{\begin{cases} \dot{z}_i = \Theta z_i \left(\left(\bar{\Lambda}\mathbf{z}\right)_i - \mathbf{z}^T \bar{\Lambda}\mathbf{z}\right), \; i = 1, \cdots, N, \\ z_1 + z_2 + \cdots + z_N = 1. \end{cases}} \tag{4.4.28}$$

where $\bar{\Lambda}$ is the square matrix of size $N \times N$ whose coefficient $(i;j)$ are the pairwise fitness $\bar{\lambda}_i^{\,j}$ which satisfy

$$\boxed{\bar{\lambda}_i^{\,j} = \theta_1\left(\bar{b}_i - \bar{b}_j\right) + \theta_2\left(-\bar{\nu}_i + \bar{\nu}_j\right) + \theta_3\left(-\bar{u}_{ij} - \bar{u}_{ji} + 2\bar{u}_{jj}\right) + \theta_4\left(\bar{\omega}^i_{ij} - \bar{\omega}^j_{ji}\right) + \theta_5\left(\mu\left(\bar{\alpha}_{ji} - \bar{\alpha}_{ij}\right) + \bar{\alpha}_{ji} - \bar{\alpha}_{jj}\right).} \tag{4.4.29}$$

with $\mu = \dfrac{I^*}{D^*}$.

Then, for any initial values of (4.4.1)*, for each $\tau_0 > 0$, $T > \tau_0$ arbitrarily and independent on ε, there is $\varepsilon_0 > 0$,*

$C > 0$ and a vector of positive coefficients $\mathbf{z}_0 \in \mathbb{R}^N$ verifying $\sum_{i=1}^N \mathbf{z}_{0,i} = 1$, such that $\forall \varepsilon < \varepsilon_0$

$$\left|\bar{S}\left(\frac{\tau}{\varepsilon}\right) - S^*\right| + \sum_{i=1}^{N}\left|I^* z_i(\tau) - \bar{I}_i\left(\frac{\tau}{\varepsilon}\right)\right| + \sum_{i,j=1}^{N}\left|D^* z_i(\tau) z_j(\tau) - \bar{I}_{ij}\left(\frac{\tau}{\varepsilon}\right)\right| \leq \varepsilon C, \quad \forall \tau \in (\tau_0, T). \tag{4.4.30}$$

where $\bar{S}$, $(\bar{I}_1, \bar{I}_2, \ldots, \bar{I}_N)$, $(\bar{I}_{ij})_{i,j\in\{1,\ldots,N\}}$ are the mean values over Ω of the solution for (4.4.1) *and $(z_1, z_2, \ldots, z_N)$ is the solution of reduced system* (4.4.28) *together with $\mathbf{z}(0) = \mathbf{z}_0$.*

4.5 Comparison between two cases of slow and fast diffusions

Initially, we recall the two replicator system used to approximate in both cases

Case 1. Slow diffusion $\varepsilon\Delta$:

$$\begin{cases} \dot{z}_i = \Theta z_i \left((\Lambda(x)\mathbf{z})_i - \mathbf{z}^T \Lambda(x) \mathbf{z} \right) + \vec{\vartheta} \cdot \nabla z_i + \Delta z_i, \ i = 1, \cdots, N, \\ z_1 + z_2 + \cdots + z_N = 1. \end{cases} \tag{4.5.1}$$

where $\vec{\vartheta}(x) = \dfrac{1}{|P|}\left(-D^* \nabla I^* + 2T^* \nabla T^*\right)$ and $\Lambda(x)$ is the square matrix of size $N \times N$ whose coefficient $(i; j)$ are the pairwise fitness λ_i^j which satisfy

$$\lambda_i^j(x) = \theta_1(b_i - b_j) + \theta_2(-\nu_i + \nu_j) + \theta_3(-u_{ij} - u_{ji} + 2u_{jj}) + \theta_4\left(\omega_{ij}^i - \omega_{ji}^j\right) + \theta_5\left(\mu(\alpha_{ji} - \alpha_{ij}) + \alpha_{ji} - \alpha_{jj}\right). \tag{4.5.2}$$

Case 2. Fast diffusion $\dfrac{1}{\varepsilon}\Delta$:

$$\begin{cases} \dot{z}_i = \Theta z_i \left(\left(\bar{\Lambda}\mathbf{z}\right)_i - \mathbf{z}^T \bar{\Lambda} \mathbf{z} \right), \ i = 1, \cdots, N, \\ z_1 + z_2 + \cdots + z_N = 1. \end{cases} \tag{4.5.3}$$

where $\bar{\Lambda}$ is the square matrix of size $N \times N$ whose coefficient $(i; j)$ are the pairwise fitness $\bar{\lambda}_i^j$ which satisfy

$$\bar{\lambda}_i^j = \theta_1\left(\bar{b}_i - \bar{b}_j\right) + \theta_2(-\bar{\nu}_i + \bar{\nu}_j) + \theta_3(-\bar{u}_{ij} - \bar{u}_{ji} + 2\bar{u}_{jj}) + \theta_4\left(\bar{\omega}_{ij}^i - \bar{\omega}_{ji}^j\right) + \theta_5\left(\mu(\bar{\alpha}_{ji} - \bar{\alpha}_{ij}) + \bar{\alpha}_{ji} - \bar{\alpha}_{jj}\right). \tag{4.5.4}$$

We first note that, in Case 1, the replicator system is partial differential equations, in which, its variables are prevalences of strains depending in space $x \in \Omega$ and time (in slow time scale) $\tau \in \mathbb{R}^+$. Moreover, it is not actually the same type of replicator equations with diffusion studied in [35] since there is a term of gradient in each equation, which is interesting. The parameters in the replicator system of this case, including the pairwise invasion fitness matrix $\left(\lambda_i^j\right)_{1\leq i,j\leq N}$ and $\vec{\vartheta} = \frac{1}{|P|}\left(-D^* \nabla I^* + 2T^* \nabla T^*\right)$, are taken from the parameters of the neutral equations then depends on space. In Case 2, meanwhile, the replicator system is ordinary differential equations, in which, its variables are total masses over the domain of strain frequencies. Thus, the system's parameters- the pairwise invasion fitness matrix, can be taken directly from original model's ones, but their mean values over domain Ω.

One point need to note is the basic reproductive ratio R_0. In Case 1, we assume in Assumption 4.4 that all domain Ω is high-risk site, i.e. $\beta(x) > m(x)$ for all $x \in \Omega$. Hence, the equilibrium of susceptible $S^* = \dfrac{m(x)}{\beta(x)}$ is

well-defined and proved to be stable as in section 4.2. In this case, we denote spatial basic reproductive ratio $R_0(x) = \dfrac{\beta(x)}{m(x)}$, which exceeds 1, leading to the equilibrium of endemic mentioned in the Introduction.

However, in Case 2, we make a slighter assumption that Ω is a high-risk domain, i.e. $\int_\Omega \beta(x) > \int_\Omega m(x)$. Hence, there can exist non empty low risk site, i.e. the set H^- in (4.2.12) is non empty. Next, we make some analyzing on the basic reproductive ratio for the quasi-neutral SIS system with fast diffusion (i.e. Case 2). First, we assume that

Assumption 4.18.
1. $\int_\Omega T(x,0)\,dx > 0$, i.e. at the beginning, the total mass of infected and coinfected individuals is positive.
2. H^+ and H^- are nonempty, with H^+, H^- are in (4.2.12).

Thank the singular perturbation in transmission rates $\beta_i = \beta(1+\varepsilon b_i)$ and clearance rates $\gamma_i = \gamma(1+\varepsilon v_i)$, $\gamma_{ij} = \gamma(1+u_{ij})$, we now define a basic reproductive ratio R_0 for (4.4.1), recalling $m = \gamma + r$.

Theorem 4.19. *Similarly in [9], for each $\varepsilon > 0$, let*

$$R_0 = \sup_{\substack{\phi \in H^1(\Omega),\\ \phi \neq 0}} \left\{ \frac{\int_\Omega \beta \phi^2}{\int_\Omega \left(\frac{1}{\varepsilon} |\nabla \phi|^2 + m\phi^2 \right)} \right\}. \tag{4.5.5}$$

Then, we have that

$$R_0 \to \frac{\int_\Omega \beta}{\int_\Omega m} \qquad as \quad \varepsilon \to 0.$$

Note that, our variational characterization of the basic reproduction number R_0 is in keeping with the next generation approach for heterogeneous populations [65] which occupy a continuous spatial habitat. It is interesting that $\frac{\int_\Omega \beta}{\int_\Omega m}$ is the basic reproductive ratio R_0 of (4.4.24).

Proof. Firstly, we recall the semi-neutral system for (S,T) in Case 2

$$\begin{cases} \dfrac{\partial S}{\partial t} = mT - \beta TS + \dfrac{1}{\varepsilon}\Delta S \\ \dfrac{\partial T}{\partial t} = -mT + \beta ST + \dfrac{1}{\varepsilon}\Delta T \end{cases}. \tag{4.5.6}$$

with the same initial value condition of (4.4.1) and Neumann boundary condition.

By similar proof for Theorem 4.6, we have that the solution (S,T) of (4.4.1) can be approximated by the solution (S,T) of (4.5.6) with error $O(\varepsilon)$.

Apply the Theorem 2 in [9], we have that $R_0 \to \dfrac{\int_\Omega \beta}{\int_\Omega m}$ as $\varepsilon \to 0$. □

Next, we come to three following examples, to see more detailed comparison between two cases.

Example 4.20. Firstly, we consider the simplest example of an N-strain system and compact domain Ω, when all the parameters in Table 4.1 do not depend on x. In addition in this example, we consider the perturbations are only in the transmission rates β_i, i.e. v_i, u_{ij}, ω_{ij}^s and α_{ij} are all zeros, for all i, j and $s \in \{i, j\}$. Without loss of generality, we assume that $b_1 > b_2 \geq b_3 \geq \cdots \geq b_N$.
In the Case 2, when diffusion is fast $\frac{1}{\varepsilon}\Delta$, apply the result in [120], the strain with biggest transmission rate, in this case is strain 1, becomes the unique survivor.

Meanwhile, in the Case 2, when diffusion rates are singular $\varepsilon\Delta$, we have the replicator equation system as follows

$$\begin{cases} \dot{z}_i = \Theta_1 z_i \sum_{j=1}^{N} (b_i - b_j) z_j + \Delta z_i, \quad i = 1, \dots, N \\ z_1 + z_2 + \cdots + z_N = 1 \end{cases} \tag{4.5.7}$$

with $\Theta_1 = \dfrac{2\beta S^* T^{*2}}{|P|}$ which can be regarded as $\dot{\mathbf{z}} = f(\mathbf{z}) + \Delta\mathbf{z}$. We can compute the linearized operator $df|_{\bar{\mathbf{z}}} + \Delta$ with stable state $\bar{\mathbf{z}} = (1, 0, \dots, 0)$ as follows

$$df|_{\bar{\mathbf{z}}} + \Delta = \begin{pmatrix} \Delta & \Theta_1(b_1 - b_2) & \Theta_1(b_1 - b_3) & \dots & \Theta_1(b_1 - b_N) \\ 0 & \Delta & 0 & \dots & 0 \\ 0 & 0 & \Delta & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \Delta \end{pmatrix}$$

which has the negative spectrum, since $b_1 - b_j > 0$ for all $j \neq 1$ and the Laplacian has negative spectrum. Apply Theorem 11.20 in [189], the state $(1, 0, \dots, 0)$ is linearly stable, implying the unique survival of strain 1.

In this example, the survival outcomes in two strain are the same.

When perturbations are only in single-infection clearance rates γ_{ij} or transmission capacity of the strain s by a co-colonized host by strain-i then strain-j p_{ij}^s, we can have the similar results by applying the same arguments.

Roughly speaking, it can happen that, in both cases: slow diffusion and fast diffusion, the unique survivors are the same.

To close this section, we consider two other examples, in which, the longtime behaviors of strains distinguish in two cases of diffusions.

Example 4.21. We consider in two cases the systems of two strains $N = 2$ and $\Omega = [0, 1]$ when the neutral values of parameters as follows

$$\beta = 3, \qquad k = 0.1, \qquad m = \frac{3(\psi - 0.36) + 3\sqrt{\psi(\psi - 0.8)}}{2(1.62 + \psi)}, \quad \text{with} \quad \psi(x) = \frac{1}{-\frac{1}{3}x + \frac{1}{2}}, \forall x \in [0, 1]. \tag{4.5.8}$$

It can be verified directly that $m < \beta$ for all x, which satisfies our assumption 4.4.

From (4.3.17), we recall that $\mu = \dfrac{1}{k(R_0 - 1)}$ and

$$\frac{\Theta_1}{\Theta_5 \mu} = 2k^2 (\mu + 1)^2 (R_0 - 1). \tag{4.5.9}$$

Substituting (4.5.8) into (4.5.9), by direct calculation, we can verify that

$$\frac{\Theta_5 \mu}{\Theta_1} = -\frac{1}{3}x + \frac{1}{2}. \tag{4.5.10}$$

In this case, we consider perturbations in transmission rates β_i and co-colonization interaction k_{ij}, which are given as follows

$$\begin{aligned} &b_1(x) = \frac{x}{3}, \qquad b_2(x) = \frac{1-x}{3}, \\ &\alpha_{12} = x, \qquad \alpha_{21} = 1-x, \qquad \alpha_{11} = \alpha_{12}, \quad \alpha_{22} = \alpha_{21}, \end{aligned} \tag{4.5.11}$$

for all $x \in [0,1]$.

In the case of fast diffusion $\frac{1}{\varepsilon}\Delta$, using (4.5.3) and (4.5.4), we only need to compute the pairwise-invasion fitness for the slow-system to determine the unique survivor. From (4.5.11), we have that $\int_0^1 b_1 dx = \int_0^1 b_2 dx$, leading to $\bar{b}_1 = \bar{b}_2$ and $\bar{\beta}_1 = \bar{\beta}_2$. From the definition of $\bar{\alpha_{ij}}$ in (4.4.22), we deduce that $\bar{\lambda}_1^2 + \bar{\lambda}_2^1 < 0$, indeed, we recall the formula (4.5.4) in this case

$$\begin{cases} \bar{\lambda}_1^2 = \theta_5 \left(\mu \left(\bar{\alpha}_{21} - \bar{\alpha}_{12}\right) + \bar{\alpha}_{21} - \bar{\alpha}_{22}\right) \\ \bar{\lambda}_2^1 = \theta_5 \left(\mu \left(\bar{\alpha}_{12} - \bar{\alpha}_{21}\right) + \bar{\alpha}_{12} - \bar{\alpha}_{11}\right) \end{cases}. \tag{4.5.12}$$

Then we have that

$$\begin{aligned} \frac{1}{\theta_5}\left(\bar{\lambda}_1^2 + \bar{\lambda}_2^1\right) &= \left(\bar{\alpha}_{12} - \bar{\alpha}_{11}\right) + \left(\bar{\alpha}_{21} - \bar{\alpha}_{22}\right) = \frac{1}{\varepsilon \int_0^1 \beta_1 dx} \int_0^1 (k_{21} - k_{12})(\beta_1 - \beta_2)\, dx \\ &= \frac{\beta}{\varepsilon \int_0^1 \beta_1 dx} \int_0^1 (\varepsilon\alpha_{21} - \varepsilon\alpha_{12})(\varepsilon b_1 - \varepsilon b_2)\, dx = -\frac{2\varepsilon}{3\int_0^1 (1+\varepsilon b_1)} \int_0^1 \left(x - \frac{1}{2}\right)^2 dx \\ &= -\frac{\beta\varepsilon}{18\int_0^1 (1+\varepsilon b_1)} < 0. \end{aligned}$$

Moreover, we observe that

$$\bar{\alpha}_{12} - \bar{\alpha}_{21} = \frac{\int_0^1 \beta_2 k_{12} dx}{\varepsilon \int_0^1 \beta_2 dx} - \frac{\int_0^1 \beta_1 k_{21} dx}{\varepsilon \int_0^1 \beta_1 dx} = \frac{1}{\varepsilon \int_0^1 \beta_1 dx} \left[\int_0^1 (3 + \varepsilon(1-x))(0.1 + \varepsilon x)\, dx - \int_0^1 (3+\varepsilon x)(0.1 + \varepsilon(1-x))\, dx \right],$$

which implies $\bar{\alpha}_{12} = \bar{\alpha}_{21}$. From (4.5.12), we have that

$$\bar{\lambda}_2^1 - \bar{\lambda}_1^2 = \bar{\alpha}_{22} - \bar{\alpha}_{11} = \frac{\varepsilon}{\int_0^1 (1+\varepsilon b_1)\, dx} \int_0^1 (b_2\alpha_{21} - b_1\alpha_{12})\, dx = 0.$$

Then $\bar{\lambda}_1^2 = \bar{\lambda}_2^1$. Combining with $\bar{\lambda}_1^2 + \bar{\lambda}_2^1 < 0$ then $\bar{\lambda}_1^2 = \bar{\lambda}_2^1 < 0$, which leads to the bistability.

When the diffusion is slow $\varepsilon\Delta$, we compute the pairwise invasion fitnesses of both strains at each $x \in [0,1]$. From (4.5.2), we have the explicit formulas for pairwise invasion fitness in this case as follows

$$\begin{cases} \lambda_1^2(x) = \theta_1 (b_1 - b_2) + \theta_5 \mu (\alpha_{21} - \alpha_{12}) \\ \lambda_2^1(x) = \theta_1 (b_2 - b_1) + \theta_5 \mu (\alpha_{12} - \alpha_{21}) \end{cases} \tag{4.5.13}$$

It is easy to see that $\lambda_1^2(x) + \lambda_2^1(x) = 0$ for all $x \in [0,1]$. We claim that $\lambda_1^2 > 0$ for all $x \in [0,1]$. Indeed, we will show that

$$b_1 - b_2 \geq \frac{\Theta_5 \mu}{\Theta_1} (\alpha_{12} - \alpha_{21}). \tag{4.5.14}$$

Because $3(b_1-b_2)=\alpha_{12}-\alpha_{21}=2x-1$ and $\frac{\Theta_5\mu}{\Theta_1}=-\frac{1}{3}x+\frac{1}{2}$, we have that

$$(b_1-b_2)-\frac{\Theta_5\mu}{\Theta_1}(\alpha_{12}-\alpha_{21})=\frac{1}{3}(2x-1)^2\geq 0,$$

implies the inequality (4.5.14). According to the formulas for pairwise invasion fitnesses (4.5.13), this means that, at every point $x\in\Omega$, strain 1 excludes strain 2 in the case of asymptotically small diffusion.

Roughly speaking, it can happen that, when the diffusion rates are singular, a strain is the unique survivor at each point of domain; meanwhile, in the case of large rates of diffusion, the longtime behavior is bistability.

The following example is similar to the Example 4.21. In which, strain 1 is the unique survivor at each point of domain in the case of slow diffusion, but strain 2 excludes strain 1 when the diffusion is asymptotically fast.

Example 4.22. We consider in two cases the systems of two strains $N=2$ and $\Omega=[0,1]$ when the neutral values of parameters as follows

$$\beta=2,\qquad k=0.2,\qquad m=\frac{\psi-0.64+\sqrt{\psi(\psi-1.6)}}{1.28+\psi},\quad \text{with}\quad \psi(x)=\frac{1}{-\frac{1}{3}x+\frac{1}{2}},\forall x\in[0,1]. \tag{4.5.15}$$

It can be verified directly that $m<\beta$ for all x, which satisfies our assumption 4.4.

Analogously to the previous Example 4.21, by direct calculation, we can verify that

$$\frac{\Theta_5\mu}{\Theta_1}=-\frac{1}{3}x+\frac{1}{2}. \tag{4.5.16}$$

In this case, we consider perturbations in transmission rates β_i and co-colonization interaction k_{ij}, which are given as follows

$$\begin{aligned}&b_1(x)=\frac{x}{2},\qquad b_2(x)=\frac{1-x}{2},\\&\alpha_{12}=x(x+1),\qquad \alpha_{21}=(1-x)(x+1),\qquad \alpha_{11}=\alpha_{12},\quad \alpha_{22}=\alpha_{21},\end{aligned} \tag{4.5.17}$$

for all $x\in[0,1]$.

When the diffusion rates are singular $\varepsilon\Delta$, we compute the pairwise invasion fitnesses of both strains at each $x\in[0,1]$. From (4.5.2), we have the explicit formula for pairwise invasion fitnesses in this case as follows

$$\begin{cases}\lambda_1^2(x)=\theta_1(b_1-b_2)+\theta_5\mu(\alpha_{21}-\alpha_{12})\\ \lambda_2^1(x)=\theta_1(b_2-b_1)+\theta_5\mu(\alpha_{12}-\alpha_{21})\end{cases} \tag{4.5.18}$$

It is easy to see that $\lambda_1^2(x)+\lambda_2^1(x)=0$ for all $x\in[0,1]$. We claim that $\lambda_1^2>0$ for all $x\in[0,1]$. Indeed, we will show that

$$b_1-b_2\geq\frac{\Theta_5\mu}{\Theta_1}(\alpha_{12}-\alpha_{21}). \tag{4.5.19}$$

Because $2(b_1-b_2)=2x-1$, $\alpha_{12}-\alpha_{21}=(2x-1)(x+1)$ and $\frac{\Theta_5\mu}{\Theta_1}=-\frac{1}{3}x+\frac{1}{2}$, we have that

$$(b_1-b_2)-\frac{\Theta_5\mu}{\Theta_1}(\alpha_{12}-\alpha_{21})=\frac{1}{6}x(2x-1)^2\geq 0,$$

implies the inequality (4.5.19). According to the formulas for pairwise invasion fitnesses (4.5.18), this means that, at every point $x \in \Omega$, strain 1 excludes strain 2 in the case of asymptotically slow diffusion.

Meanwhile, in the case of fast diffusion $\frac{1}{\varepsilon}\Delta$, using (4.5.3) and (4.5.4), we only need to compute the pairwise-invasion fitness for the slow-system to determine the unique survivor. From (4.5.17), we have that $\int_0^1 b_1 dx = \int_0^1 b_2 dx$, leading to $\bar{b}_1 = \bar{b}_2$ and $\bar{\beta}_1 = \bar{\beta}_2$. From the definition of $\bar{\alpha_{ij}}$ in (4.4.22), we deduce that $\bar{\lambda}_1^2 + \bar{\lambda}_2^1 \to 0$ as $\varepsilon \to 0$, indeed, we recall the formula (4.5.4) in this case

$$\begin{cases} \bar{\lambda}_1^2 = \theta_5 \left(\mu \left(\bar{\alpha}_{21} - \bar{\alpha}_{12}\right) + \bar{\alpha}_{21} - \bar{\alpha}_{22}\right) \\ \bar{\lambda}_2^1 = \theta_5 \left(\mu \left(\bar{\alpha}_{12} - \bar{\alpha}_{21}\right) + \bar{\alpha}_{12} - \bar{\alpha}_{11}\right) \end{cases}. \tag{4.5.20}$$

Then we have that

$$\begin{aligned} \frac{1}{\theta_5}\left(\bar{\lambda}_1^2 + \bar{\lambda}_2^1\right) &= \left(\bar{\alpha}_{12} - \bar{\alpha}_{11}\right) + \left(\bar{\alpha}_{21} - \bar{\alpha}_{22}\right) = \frac{1}{\varepsilon \int_0^1 \beta_1 dx} \int_0^1 \left(k_{21} - k_{12}\right)\left(\beta_1 - \beta_2\right) dx \\ &= \frac{\beta}{\varepsilon \int_0^1 \beta_1 dx} \int_0^1 \left(\varepsilon\alpha_{21} - \varepsilon\alpha_{12}\right)\left(\varepsilon b_1 - \varepsilon b_2\right) dx = -\frac{2\varepsilon}{\int_0^1 \left(1 + \varepsilon b_1\right)} \int_0^1 \left(x - \frac{1}{2}\right)^2 (x+1)\, dx \\ &= -\frac{\beta\varepsilon}{4\int_0^1 \left(1 + \varepsilon b_1\right)} \to 0 \quad \text{when} \quad \varepsilon \to 0. \end{aligned}$$

Moreover, we observe that

$$\begin{aligned} \bar{\alpha}_{12} - \bar{\alpha}_{21} &= \frac{\int_0^1 \beta_2 k_{12} dx}{\varepsilon \int_0^1 \beta_2 dx} - \frac{\int_0^1 \beta_1 k_{21} dx}{\varepsilon \int_0^1 \beta_1 dx} \\ &= \frac{1}{\varepsilon \int_0^1 \beta_1 dx}\left[\int_0^1 \left(2 + \varepsilon\left(1 - x\right)\right)\left(0.2 + \varepsilon x\left(x+1\right)\right) dx - \int_0^1 \left(2 + \varepsilon x\right)\left(0.2 + \varepsilon\left(1 - x\right)\left(x+1\right)\right) dx\right] \\ &= \frac{1}{\int_0^1 \beta_1 dx} \int_0^1 \left(2x - 1\right)\left(2x + 1.8\right) dx = \frac{1}{3\int_0^1 \beta_1 dx} > 0. \end{aligned}$$

From (4.5.20), we have that

$$\bar{\lambda}_2^1 - \bar{\lambda}_1^2 = \left(2\mu + 1\right)\left(\bar{\alpha}_{12} - \bar{\alpha}_{21}\right) + \left(\bar{\alpha}_{22} - \bar{\alpha}_{11}\right) = \frac{1}{\int_0^1 \left(1 + \varepsilon b_1\right) dx}\left[\left(\frac{2\mu + 1}{6}\right) - \frac{1}{6} - \frac{\varepsilon}{12}\right] = \frac{1}{\int_0^1 \left(1 + \varepsilon b_1\right) dx}\left(\frac{\mu}{3} - \frac{\varepsilon}{12}\right) > 0$$

for ε small enough. Then $\bar{\lambda}_1^2 < 0 < \bar{\lambda}_2^1$ for ε small enough, since $\bar{\lambda}_2^1 - \bar{\lambda}_1^2 = O(1)$. Therefore, when the diffusion is fast, strain 2 excludes strain 1 in long time.

Roughly speaking, it can happen that, when the diffusion rates are singular, a strain, denoted by strain 1, is the unique survivor at each point of domain; meanwhile, in the case of large rates of diffusion, the other strain, denoted strain 2, will exclude strain 1 over the domain.

4.6 Conclusion

Epidemiology for homogeneous environment receives many intention so far [14] because invasion of disease is now an international public health problem. In reality, populations tend not to be homogeneous and there are nonlocal interactions. Hence, people investigate more theory on the geographical spread of infectious diseases. The mechanisms of invasion of disease to new territories may take many different forms and

there are several ways to model such problems [67, 116, 130, 155], in which, the equilibrium behavior has been studied. This mathematical study provides a fundamental advance in understanding analytically quasi-neutral dynamics between multiple strains in a co-infection diffusion system. Until now, explicit and general derivations of coinfection dynamics among N strains are very rare in the literature, especially models with diffusion. Nevertheless, many models have been proposed to investigate effect of diffusion of disease infection [167, 79, 78, 183].

Motivated by the dynamics without diffusion in [120], we formulate an SIS-type reaction diffusion equations among similar strains, in both cases of slow and fast diffusions. Naturally in this present model, infectious strains compete for susceptible and singlycolonized hosts, which are the only resources that can favor their growth and propagation. The different traits provide each strain with variable fitness advantages or disadvantages in exploiting such dynamic resources in the system, and interact together to shape multi-strain selection. We aim to simplify the dynamics when small perturbations arise in the clearance rates, transmission rates, within-host competitiveness coefficients, as well as co-colonization susceptibility interaction factors between strains. However, with spatial structure, it requires us to add some appropriate assumptions, especially, the assumption of high-risk site Ω with slow diffusion and the assumption of

When diffusion rates are singular ($\varepsilon\Delta$), we base on the framework in [120] and adapt for our current system, including proving a Tikhonov-like Theorem. The details of this framework are not mentioned again here. We derive the corresponding slow-fast form for the global dynamics, with the system of strain frequencies completely explicit, and provide the formal approximation for solutions of all epidemiological variables by quantifying error estimates. We reduce the complexity of N^2+N+1 equations of the original SIS compartmental model to the N-equations of replicator dynamics with diffusion, which reduces substantially time for computation. Meanwhile, for the case of fast diffusion ($\frac{1}{\varepsilon}\Delta$), we apply the Central Manifold Theorem to obtain an SIS system for total masses of susceptible, infected and coinfected individuals, which allows us to use the main result in [120]. Accordingly, the reduced system in this case is the replicator equation, which is studied widely [102]. A similar point in both approaches is that, the error in approximation is estimated for total masses of susceptible, infected and coinfected strains.

When the diffusion is fast, we can use the result about survival outcome of strains in [120] to study the longtime behavior of total mass of each strain. However, there is not much study on the replicator equations with diffusion and gradient, so there is no general theory for the long time phenomena of individuals in the case of slow diffusion. Though, it is exciting to envision how this approach could be extended to other epidemiological models of multi- strain dynamics with diffusion or even more with general spatial structure. Like the non-spatial model, an essential requirement is that their embedded neutral system admits a central manifold which is globally stable. The challenge would then be to identify the equations governing slow motion on this manifold in each specific model. It is essential to note that we use strong assumption of high-risk site Ω in Case 1 and high-risk domain Ω in Case 2, which lead to the endemic equilibrium. In general, without these assumption, people are interested in the theory of disease-free equilibrium and endemic equilibrium, [9, 208].

In Case 1, when diffusion rates are singular, without the assumption of high risk site Ω, i.e. $H^- \neq \emptyset$, there are points x's at which $S^* = 1$, that may not allow the smoothness of S^* in x. Then, our approach may not work because $\|S - S^*\|_2 \to O\left(\sqrt{\varepsilon}\right)$ may not hold anymore.

One more thing, until now we have not considered a spatial component of intermediate diffusion ($d\Delta$, $d > 0$) to the multi-strain dynamics. A further perspective is considering the application of the Central Manifold Theorem to this model.

A4.1 Appendix: Proof for theorems

A4.1.1 Proof for theorem 4.2

In this proof, we will show that equation, recalling Ω compact,

$$\begin{cases} \dfrac{\partial u}{\partial t} = F\left(u\left(x,t\right),x\right) + d\Delta u, \\ \dfrac{\partial u}{\partial n} = 0 \quad \text{on} \quad \partial\Omega, \qquad u\left(x,0\right) = u_0\left(x\right) \end{cases} \tag{A4.1.1}$$

has unique solution $u : \Omega \times [0,\infty) \to \mathbb{R}^n$, satisfying $u \in C^2\left(\Omega \times \cdot, \mathbb{R}^n\right) \cap C^1\left(\cdot \times [0,\infty), \mathbb{R}^n\right)$ when $F : \mathbb{R}^n \times \cdot \to \mathbb{R}^n$ is a Lipschitz map, i.e. there exists a constant L such that

$$\|F\tilde{u} - F\tilde{v}\| \leq L\|\tilde{u} - \tilde{v}\|, \qquad \forall \tilde{u}, \tilde{v} \in \mathbb{R}^n, \quad \forall x \in \Omega. \tag{A4.1.2}$$

First, we denote that $Q_{\mathscr{T}} = \Omega \times [0,\infty)$ and $\bar{Q}_{\mathscr{T}} = \Omega \times [0,\infty)$ and $u\left(x,t\right) \in \mathbb{R}^n$ for $(x,t) \in \bar{Q}_{\mathscr{T}}$. When seen the Laplacian as an operator on $L^2\left(\Omega\right)$, the operator A^2 with homogeneous Neumann boundary conditions is defined as

$$\begin{aligned} &D\left(A^2\right) = \left\{ U \in H^1\left(\Omega\right) : \exists V \in L^2\left(\Omega\right), \forall \phi \in H^1\left(\Omega\right), \int \nabla U\left(x\right) \nabla \phi\left(x\right) dx = -d \int V\left(x\right) \phi\left(x\right) dx \right\}, \\ &A^2 U := V, \quad U \in D\left(A^2\right). \end{aligned} \tag{A4.1.3}$$

In order to obtain uniform estimates, we prefer to focus on the operator $A^\infty := A$ acting on $C^2(\Omega)$. Denoting by operator A to be the Laplacian Δ acting on $\left(C^2\left(\Omega \times \cdot\right)\right)^n$. Hence, we define

$$\begin{aligned} &D\left(A^\infty\right) := \left\{ U \in D\left(A^2\right) \cap C\left(\bar{\Omega}\right), A^2 U \in C\left(\bar{\Omega}\right) \right\}, \\ &A^\infty U = A^2 U, \quad U \in D\left(A^\infty\right). \end{aligned} \tag{A4.1.4}$$

Firstly, by Duhamel's formula and [169], (A4.1.1) implies that

$$u\left(x,t\right) = e^{At} u_0 + \int_0^t e^{A(t-s)} F\left(u\left(x,s\right),x\right) ds, \qquad \forall (x,t) \in Q_{\mathscr{T}} \tag{A4.1.5}$$

where $\exp\left(At\right)$ is the semi-group generated by the operator Δ with the Neumann boundary condition. We consider the operator T defined by

$$Tu\left(x,t\right) := e^{At} u_0 + \int_0^t e^{A(t-s)} F\left(u\left(x,s\right),x\right) ds, \qquad \forall (x,t) \in Q_{\mathscr{T}}.$$

Given $k > 0$, to be fixed later, set

$$X = \left\{ u \in C^1\left(\Omega \times [0,+\infty), \mathbb{R}^n\right); \sup_{\substack{t \geq 0 \\ x \in \Omega}} e^{-kt} \left\|u\left(x,t\right)\right\| \leq +\infty \right\}$$

We can check that X is a Banach space for the norm

$$\|u\|_X = \sup_{\substack{t\geq 0\\ x\in\Omega}} e^{-kt}\|u(x,t)\|.$$

For every $u\in X$, the Tu also belongs to X. To prove this, using the argument in the beginning of subsection 4.4.2,we first recall that A is the generator of a C^0 semi-group $\exp(tA)$ on $C^2(\Omega\times\cdot)$ verifying $\|\exp(tA)v\|\leq \exp(-\mu t)\|v\|$, for $\mu\geq 0$.

Then we observe that

$$\begin{aligned}
e^{-kt}\|Tu\| &\leq e^{(A-kI)t}u_0 + e^{-kt}\int_0^t \left\|e^{A(t-s)}F(u(x,s),x)\right\| ds\\
&\leq e^{(A-kI)t}u_0 + e^{-kt}\int_0^t e^{-\mu(t-s)}\|F(u(x,s),x)\| ds\\
&\leq e^{(A-kI)t}u_0 + e^{-kt}\int_0^t e^{-\mu(t-s)}\left(L\|u(x,s)-u_0\| + \|Fu_0\|\right) ds
\end{aligned}$$

according to (A4.1.2). Hence, we deduce that

$$e^{-kt}\|Tu\| \leq e^{(A-kI)t}u_0 + \frac{e^{-kt}}{\mu}\left(L\|u_0\| + \|Fu_0\|\right)\left(1-e^{-\mu t}\right) + e^{-kt}L\int_0^t \|u(x,s)\| ds.$$

Alternatively, we have that

$$e^{-kt}L\int_0^t \|u(x,s)\| ds = e^{-kt}L\int_0^t e^{-ks}\|u(x,s)\|\cdot e^{ks} ds \leq e^{-kt}L\|u\|_X\int_0^t e^{ks} ds$$

which implies

$$e^{-kt}\|Tu\| \leq e^{(A-kI)t}u_0 + \frac{e^{-kt}}{\mu}\left(L\|u_0\| + \|Fu_0\|\right)\left(1-e^{-\mu t}\right) + \frac{1}{k}L\|u\|_X\left(1-e^{-kt}\right),$$

leading to $Tu\in X$ whenever $u\in X$. Moreover, for all $u,v\in X$, we have that

$$\begin{aligned}
\|Tu-TV\|_X &\leq e^{-kt}\int_0^t \left\|e^{A(t-s)}\left[F(u(x,s),x)-F(v(x,s),x)\right]\right\| ds\\
&\leq e^{-kt}\int_0^t e^{-\mu(t-s)}\|F(u(x,s),x)-F(v(x,s),x)\| ds\\
&\leq Le^{-(k+\mu)t}\int_0^t e^{(\mu+s)s}\cdot e^{-ks}\|u(x,s)-v(x,s)\| ds \leq \frac{L}{\mu+k}\left(1-e^{-(\mu+k)t}\right)\|u-v\|_X.
\end{aligned}$$

Fixing $k>0$ such that $k+\mu>L$ then applying the Banach fixed point theorem, we obtain that (A4.1.1) has at least one solution.

For the uniqueness, assume there exists functions u and v, which satisfy for (A4.1.5). For any given $\mathscr{T}>0$, we have that

$$\begin{aligned}
\|u(x,t)-v(x,t)\| &\leq \left\|e^{At}(u_0-v_0)\right\| + \int_0^t \left\|e^{A(t-s)}\left[F(u(x,s),x)-F(v(x,s),x)\right]\right\| ds\\
&\leq ML\int_0^t \|u(x,s)-v(x,s)\| ds, \qquad \forall 0\leq t\leq \mathscr{T}
\end{aligned}$$

By the Gronwall's inequality and the same initial value of u and v, we have that $u(,t) = v(x,t)$, for all $x \in \Omega$ and $0 \leq t \leq \mathscr{T}$. This holds for all $\mathscr{T} \geq 0$, which yields the uniqueness of solution.

Therefore, the equation (4.2.4) has the unique solution.

A4.1.2 Proof for theorem 4.5

The idea of our proof bases on the technique mentioned in [148].

Proof. Firstly, we make a convention for the norm using in this proof. For each $t \in \mathbb{R}^+$, for every $f_1, f_2 \in L^2(\Omega \times \mathbb{R}, \mathbb{R}^n)$ we denote

$$\langle f_1, f_2 \rangle = \int_\Omega f_1(x,t) \cdot f_2(x,t)\, dx,$$

where the $f_1 \cdot f_2$ representing for the usual scalar product $\sum_{i=1}^n f_1^i f_2^i$ in $\mathbb{R}^n$. This scalar product $\langle \cdot, \cdot \rangle$ induces the norm

$$\|f(\cdot,t)\|_2 = \left(\int_\Omega f(x,t) \cdot f(x,t)\, dx \right)^{1/2}$$

For the sake of convenience in this proof, we only write $\|\cdot\|$ instead of $\|\cdot\|_2$.

We do the same convention for $\langle g_1, g_2 \rangle$ and $\|g(\cdot,t)\|$ for all $g_1, g_2, g \in C^1(\Omega \times \mathbb{R}, \mathbb{R}^m)$.

Because in the finite dimensional space, all norms are equivalent, we then denote $|\cdot|$ to be the usual 2-Euclidean norm. Moreover, we recall the notation $A \prec 0$ for a symmetric matrix A if A is definitely negative, and $A \succ 0$ for definitely positive symmetric matrix.

First, let us show that the interval $[t_0, t_1]$ can be subdivided into subinterval $\Delta_k = [\tau_{k-1}, \tau_k]$, where $k \in \{1, 2, \ldots, N\}$ and $t_0 = \tau_0 < \tau_1 < \cdots < \tau_N = t_1$ in such a way that for every k, there exists a symmetric matrix $P_k = P_k^T \succ 0$ for which

$$P_k A(x,t) + A^T(x,t) P_k \prec -I. \tag{A4.1.6}$$

Indeed, since $A(x,t)$ is a Hurwitz matrix for every $t \in [t_0, t_1]$, according to [66], there exists $P(x,t) = P^T(x,t) \succ 0$ such that

$$P(x,t)A(x,t) + A^T(x,t)P(x,t) \prec -I.$$

Since A depends continuously on t, there exists an open interval $\Delta(t)$ such that $t \in \Delta(t)$ and

$$P(x,t)A(x,\tau) + A^T(x,\tau)P(x,t) \prec -I, \quad \forall \tau \in \Delta(t).$$

Now the open intervals $\Delta(t)$ with $t \in [t_0, t_1]$ cover the whole closed bounded interval $[t_0, t_1]$ and taking a finite number of τ_k, $k = 1, \ldots, N$ such that $[t_0, t_1]$ is completely covered by $\Delta(\tau_k)$ yields the desired partition subdivision.

We can note that a strictly negative upper bound is not required on the real parts eigenvalues uniformly in space, because the spatial domain is supposed to be compact.
Note that, from (A4.1.6), for all $y \in \mathbb{R}^m$ we have that

$$y^T \left(P_k A + A^T P_k \right) y \prec -y^T y. \tag{A4.1.7}$$

Second, because F, G are continuously differential in x and t, then for every $\mu > 0$ there exists $C, r > 0$ such that

$$\left\| F\left(f_0(x,t) + \bar{\delta}_f(x,t), g_0(x,t) + \bar{\delta}_g(x,t), x, t \right) - F(f_0, g_0, x, t) \right\| \leq C \left(\left\| \bar{\delta}_f(x,t) \right\| + \left\| \bar{\delta}_g(x,t) \right\| \right) \tag{A4.1.8}$$

for all $t \in \mathbb{R}$, $\bar{\delta}_f(x,t) \in \mathbb{R}^n$, $\bar{\delta}_g(x,t) \in \mathbb{R}^m$ satisfying

$$\forall t \in [t_0,t_1], \forall x \in \Omega, \qquad |\bar{\delta}_f(x,t)| \le r, \qquad |\bar{\delta}_g(x,t)| \le r.$$

For the sake of simplicity, we write $\bar{\delta}_f$ and $\bar{\delta}_g$ instead of $\bar{\delta}_f(x,t)$ and $\bar{\delta}_g(x,t)$. We now have the Taylor expansion as follows, noting that $G(f_0(x,t), g_0(x,t), x, t) = 0$,

$$G\left(f_0(x,t) + \bar{\delta}_f, g_0(x,t) + \bar{\delta}_g, x, t\right) = A(x,t)\,\bar{\delta}_g + B(x,t)\bar{\delta}_f + o\left(|\bar{\delta}_g|\right) + o\left(|\bar{\delta}_g|\right), \tag{A4.1.9}$$

with $B(x,t)$ is the Jacobian matrix of $G(\cdot,\cdot,t)$ with respect to the first variable.

For each $k = 1,\dots,N$, and $u \in \mathbb{R}^m$, set $|u|_k = \left(u^T P_k u\right)^{1/2}$, then $|\cdot|_k$ is a norm in $\mathbb{R}^m$. Indeed, because $P_k \succ 0$ then $|\cdot|_k$ is well-defined, it suffices to check the condition $|u+v|_k \le |u|_k + |v|_k$, which is equivalent to

$$\left(u^T P_k v\right)^2 \le \left(u^T P_k u\right)\left(v^T P_k v\right).$$

It now becomes

$$\left(\left(L^T u\right)^T \left(L^T v\right)\right)^2 \le \left(\left(L^T u\right)^T \left(L^T u\right)\right)\left(\left(L^T v\right)^T \left(L^T v\right)\right), \tag{A4.1.10}$$

thanks to the Cholesky's factorization, which states that, if $P_k \succ 0$, there exist a square matrix such that $P_k = L_k^T L_k$. Note that, (A4.1.10) holds because of the inequality Cauchy-Schwarz. Hence, $|\cdot|_k$ is a norm in $\mathbb{R}^m$ and it is equivalent to an arbitrary norm in $\mathbb{R}^m$.

Then, for $\delta_f(x,t) = f(x,t) - f_0(x,t)$, $\delta_g(x,t) = g(x,t) - g_0(x,t)$, we have that

$$\begin{cases} \dfrac{d}{dt}\left\|\delta_f\right\|^2 & \le C_1\left(\left\|\delta_f\right\| + \left\|\delta_g\right\|\right)\left\|\delta_f\right\|, \\ \varepsilon\dfrac{d}{dt}\left\|\delta_g\right\|_k^2 & \le -q\left\|\delta_g\right\|_k^2 dt + C_1\left(\left\|\delta_f\right\|^2 + \varepsilon\right)dt \end{cases} \tag{A4.1.11}$$

as long as δ_f, δ_g are sufficiently small, where C_1, q are positive constants which do not depend on k.
Initially, for the sake of simplicity, in the following arguments, we write f, g instead of $f(x,t)$ and $g(x,t)$, respectively. Then, we have the equation for $\delta_f(x,t)$ as follows

$$\frac{\partial}{\partial t}\delta_f = F\left(f_0 + \delta_f, g_0 + \delta_g, x, t\right) - F\left(f_0, g_0, x, t\right) + K\delta_f \quad .$$

By the convention of $\|\cdot\|$, we have that

$$\begin{aligned} \frac{d}{dt}\left\|\delta_f\right\|^2 &= \frac{d}{dt}\langle\delta_f,\delta_f\rangle = 2\langle\frac{\partial}{\partial t}\delta_f,\delta_f\rangle \\ &= \langle F\left(f_0+\delta_f, g_0+\delta_g, x, t\right) - F\left(f_0, g_0, x, t\right) + K\delta_f, \delta_f\rangle \\ &= \langle F\left(f_0+\delta_f, g_0+\delta_g, x, t\right) - F\left(f_0, g_0, x, t\right), \delta_f\rangle + \langle K\delta_f, \delta_f\rangle \\ &\le \left\|F\left(f_0+\delta_f, g_0+\delta_g, x, t\right) - F\left(f_0, g_0, x, t\right)\right\|\left\|\delta_f\right\| + \langle K\delta_f, \delta_f\rangle \\ &\le C\left(\left\|\delta_f\right\| + \left\|\delta_g\right\|\right)\left\|\delta_f\right\| + \langle K\delta_f, \delta_f\rangle. \end{aligned}$$

On the other hand, recalling that $K_f = a_f(x)\nabla + \Delta$ implies

$$\langle K_f\delta_f, \delta_f\rangle = \langle \Delta\delta_f, \delta_f\rangle + \langle a_f(x)\nabla\delta_f, \delta_f\rangle = -\int_\Omega \left|\nabla\delta_f\right|^2 dx + \int_\Omega a_f(x)\nabla\delta_f\cdot\delta_f dx$$

which leads to, when we apply the Young inequality for the term $\int_\Omega a(x)\nabla\delta_f\cdot\delta_f dx$,

$$\langle K_f\delta_f, \delta_f\rangle \le -\int_\Omega \left|\nabla\delta_f\right|^2 dx + \max_{x\in\Omega}\left(\left|a_f(x)\right|\right)\left[\frac{1}{\max\limits_{x\in\Omega}\left(\left|a_f(x)\right|\right)}\int_\Omega \left|\nabla\delta_f\right|^2 dx + C\left(\max_{x\in\Omega}\left(\left|a_f(x)\right|\right)\right)\int_\Omega \left|\delta_f\right|^2 dx\right],$$

where $\left|a_f(x)\right|$ is the matrix in which entries are absolute values of corresponding coordinates of $a_f(x)$.

Accordingly, we have the estimation for $\frac{d}{dt}\left\|\delta_f\right\|^2$ as follows

$$\frac{d}{dt}\left\|\delta_f\right\|^2 \le C_1\left(\left\|\delta_f\right\| + \left\|\delta_g\right\|\right)\left\|\delta_f\right\| \tag{A4.1.12}$$

Next, we come to control the growth of $\left\|\delta_g\right\|_k$. We first observe that

$$\varepsilon\frac{\partial}{\partial t}\delta_g = G\left(f_0(x,t)+\delta_f, g_0(x,t)+\delta_g, x, t\right) + \varepsilon K_g\delta_g + \varepsilon\left[K_g g_0(x,t) + \frac{\partial}{\partial t}g_0(x,t) + G_1(x)\cdot\nabla f_0\right] + \varepsilon G_1(x)\cdot\nabla\delta_f.$$

We denote $\varepsilon\left[K_g g_0(x,t) + \frac{\partial}{\partial t}g_0(x,t) + G_1(x)\cdot\nabla f_0\right]$ as $O(\varepsilon)$, then

$$\varepsilon\frac{\partial}{\partial t}\delta_g = G\left(f_0(x,t)+\delta_f, g_0(x,t)+\delta_g, x, t\right) + \varepsilon K_g\delta_g + O(\varepsilon) + \varepsilon G_1(x)\cdot\nabla\delta_f. \tag{A4.1.13}$$

Using the Taylor expansion for G in (A4.1.9) and the equation (A4.1.13), we obtain the following computations

$$\begin{aligned}
\varepsilon\frac{d}{dt}\left\|\delta_g\right\|_k^2 &= \varepsilon\frac{d}{dt}\langle\delta_g, P_k\delta_g\rangle = \varepsilon\langle\frac{\partial}{\partial t}\delta_g, P_k\delta_g\rangle + \varepsilon\langle\delta_g, P_k\frac{\partial}{\partial t}\delta_g\rangle \\
&= \left(\langle A\delta_g, P_k\delta_g\rangle + \langle\delta_g, P_k A\delta_g\rangle\right) + 2B(x,t)\langle\delta_f,\delta_g\rangle + \langle o\left(\left|\delta_g\right|\right) + o\left(\left|\delta_g\right|\right) + O(\varepsilon), P_k\delta_g + \delta_g\rangle \\
&+ 2\varepsilon\langle G_1(x)\cdot\nabla\delta_f, P_k\delta_g + \delta_g\rangle + \varepsilon\langle K_g\delta_g, P_k\delta_g + \delta_g\rangle \\
&= \langle\delta_g, \left(A^T P_k + P_k A\right)\delta_g\rangle + 2B(x,t)\langle\delta_f,\delta_g\rangle + \langle o\left(\left|\delta_f\right|\right) + o\left(\left|\delta_g\right|\right) + O(\varepsilon), P_k\delta_g + \delta_g\rangle \\
&+ 2\varepsilon\langle G_1(x)\cdot\nabla\delta_f, P_k\delta_g + \delta_g\rangle + \varepsilon\langle K_g\delta_g, P_k\delta_g + \delta_g\rangle \\
&\le -\left\|\delta_g\right\|^2 + 2C_1\left\|\delta_f\right\|\left\|\delta_g\right\| + \langle o\left(\left|\delta_f\right|\right) + o\left(\left|\delta_g\right|\right) + O(\varepsilon), P\delta_g + \delta_g\rangle + 2\varepsilon\langle G_1(x)\cdot\nabla, P_k\delta_g + \delta_g\rangle + \varepsilon\langle K_g\delta_g, P\delta_g + \delta_g\rangle.
\end{aligned} \tag{A4.1.14}$$

Using the Young inequality, we have the estimation for $\langle o\left(\left|\delta_f\right|\right) + o\left(\left|\delta_g\right|\right) + O(\varepsilon), P\delta_g + \delta_g\rangle$ as follows

$$\langle o\left(\left|\delta_f\right|\right) + o\left(\left|\delta_g\right|\right) + O(\varepsilon), P\delta_g + \delta_g\rangle \le O(\varepsilon) + \tilde{C}\left\|\delta_g\right\|^2, \qquad \text{with} \quad \tilde{C} \ll 1. \tag{A4.1.15}$$

Alternatively, applying the Young inequality, we have that

$$\langle G_1(x)\cdot\nabla\delta_f, P_k\delta_g + \delta_g\rangle \le C(G_1)\left\|\nabla\delta_f\right\|^2 + \left\|\delta_g\right\|^2 \le C(G_1) + \left\|\delta_g\right\|^2 \tag{A4.1.16}$$

since $\nabla\delta_f$ is bounded in Ω.

For the term $\langle K_g\delta_g, P\delta_g+\delta_g\rangle$, we get that

$$\begin{aligned}\langle K_g\delta_g, P\delta_g+\delta_g\rangle &= \langle \Delta\delta_g, P\delta_g\rangle + \langle \Delta\delta_g, \delta_g\rangle + \langle a_g(x)\nabla\delta_g, P\delta_g\rangle + \langle a_g(x)\nabla\delta_g, \delta_g\rangle \\ &= \int_\Omega \Delta\delta_g\cdot P\delta_g dx + \int_\Omega \Delta\delta_g\cdot\delta_g dx + \int_\Omega a_g(x)\nabla\delta_g\cdot P\delta_g dx + \int_\Omega a_g(x)\nabla\delta_g\cdot\delta_g dx \\ &= -\int_\Omega \nabla\delta_g\cdot\nabla(P\delta_g)dx - \int_\Omega |\nabla\delta_g|^2 dx + \int_\Omega a_g(x)\nabla\delta_g\cdot P\delta_g dx + \int_\Omega a_g(x)\nabla\delta_g\cdot\delta_g dx \\ &= -\int_\Omega \nabla\delta_g P\nabla\delta_g dx - \int_\Omega \nabla\delta_g\cdot(\nabla P)\delta_g dx - \int_\Omega |\nabla\delta_g|^2 dx + \int_\Omega a_g(x)\nabla\delta_g\cdot P\delta_g dx + \int_\Omega a_g(x)\nabla\delta_g\cdot\delta_g dx.\end{aligned}$$

Note that $P\succ 0$ then $\int_\Omega \nabla\delta_g P\nabla\delta_g dx \geq \lambda\left\|\nabla\delta_g\right\|^2$. Applying the Young inequality once more for the terms

$$\int_\Omega \nabla\delta_g\cdot(\nabla P)\delta_g dx, \qquad \int_\Omega a_g(x)\nabla\delta_g\cdot P\delta_g dx, \qquad \int_\Omega a_g(x)\nabla\delta_g\cdot\delta_g dx,$$

we have that

$$\langle K_g\delta_g, P\delta_g+\delta_g\rangle \leq -(1+\lambda)\left\|\nabla\delta_g\right\|^2 + (1+\lambda)\left\|\nabla\delta_g\right\|^2 + C(1+\lambda)\left\|\delta_g\right\|^2$$

which implies

$$\langle K_g\delta_g, P\delta_g+\delta_g\rangle \leq C(1+\lambda)\left\|\delta_g\right\|^2, \tag{A4.1.17}$$

with $C(1+\lambda)$ denoting a constant depending on $1+\lambda$.

Combining these equations (A4.1.14), (A4.1.15), (A4.1.16), and (A4.1.17), and noting that two norms $\|\cdot\|_k$ and $\|\cdot\|_k$ are equivalent, we observe that

$$\varepsilon\frac{\partial}{\partial t}\left\|\delta_g\right\|_k^2 \leq (2\varepsilon+\varepsilon C(1+\lambda)-1)\left\|\delta_g\right\|_k^2 + C_1\left\|\delta_f\right\|\left\|\delta_g\right\| + C_1\varepsilon$$

which implies when ε small enough

$$\varepsilon\frac{\partial}{\partial t}\left\|\delta_g\right\|_k^2 \leq -q\left\|\delta_g\right\|_k^2 + C_1\left\|\delta_f\right\|\left\|\delta_g\right\| + C_1\varepsilon \tag{A4.1.18}$$

Thus, combine (A4.1.12) and (A4.1.18) and we obtain that

$$\frac{d}{dt}\left(\left\|\delta_f\right\|^2 + \varepsilon\frac{C_1}{q}\left\|\delta_g\right\|^2\right) \leq C_1\left\|\delta_f\right\|^2 - \left\|\delta_g\right\|_k^2 + C_1\varepsilon. \tag{A4.1.19}$$

for some constant C_1 independent of k.

By the Gronwall's inequality for $\left(\left\|\delta_f\right\|^2 + \dfrac{\varepsilon C_1}{q}\left\|\delta_g\right\|_k^2\right)dx$, for each $k\geq 1$, we can regard τ_{k-1} as the initial value, and then deduce that

$$\left\|\delta_f(\tau_{k-1}+\tau)\right\|^2 \leq e^{C_3\tau}\left(\left\|\delta_f(x,\tau_{k-1})\right\|^2 + \varepsilon\frac{C_1}{q}\left\|\delta_g(x,\tau_{k-1})\right\|_k^2\right)dx + C_1\varepsilon$$

for $\tau \in [0, \tau_k - \tau_{k-1}]$. With the aid of this bound for the growth of $|\delta_f|$, the second inequality of (A4.1.11) implies a bound for $\|\delta_g\|_k$ as following

$$\|\delta_g(\tau_{k-1}+\tau)\|_k^2 dx \leq e^{-q\tau/\varepsilon}\|\delta_g(\tau_{k-1})\|_k^2 + C_4\left(\|\delta_f(x,\tau_{k-1})\|^2 dx + \varepsilon\frac{C_1}{q}\|\delta_g(x,\tau_{k-1})\|_k^2\right) + C_4\varepsilon.$$

We already have that $\delta_f(x,t_0) = \delta_f(x,\tau_0) \leq \varepsilon$ and $\delta_g(x,t_0) = \delta_g(x,\tau_0) \leq \varepsilon_0$ for ε_0 small enough. Then, by the compactness of Ω, for $\tau \in [0,\tau_1 - \tau_0]$, $\|\delta_f(\tau)\|^2 \leq O(\varepsilon)$, for all $x \in \Omega$. Make a process similarly and successively for $k = 1,2,\ldots$, we have that $\|\delta_f\|^2 \leq O(\varepsilon)$ for all $x \in \Omega$. Analogously, we can also prove that $\|\delta_g\|^2 \leq O(\varepsilon)$.

Therefore, $\int_\Omega |f(x,t) - f_0(x,t)|^2 dx \leq C\varepsilon$ and $\int_\Omega |g(x,t) - g_0(x,t)|^2 dx \leq C\varepsilon$, and we have the conclusion of the theorem. □

A4.1.3 Proof for theorem 4.6

Proof. Note that $\|F(u_1,x) - F(u_2,x)\| \leq C\|u_1 - u_2\|, \forall u_1, u_2 \in D(F)$ and $|G(u,x)v|$ is bounded, $\forall u, v$ bounded due to the continuous differentiability of G in a bounded domain. Consider

$$\begin{aligned}\frac{1}{2}\frac{\partial}{\partial t}|u-v|^2 = (u-v)\frac{\partial}{\partial t}(u-v) &= (u-v)[F(u,x) - F(v,x)] + \varepsilon(u-v)G(u,x) + \varepsilon(u-v)\Delta(u-v) \\ &\leq C|u-v|^2 + O(\varepsilon) + \varepsilon(u-v)\Delta(u-v).\end{aligned} \tag{A4.1.20}$$

Taking the integral of (A4.1.20) over Ω and using the Neumann boundary condition implies that

$$\frac{1}{2}\frac{\partial}{\partial t}\int_\Omega |u-v|^2 dx \leq C\int_\Omega |u-v|^2 dx + O(\varepsilon) - \varepsilon\int_\Omega \|\nabla(u-v)\|^2 dx,$$

which leads to

$$\frac{\partial}{\partial t}\int_\Omega |u-v|^2 dx \leq C\int_\Omega |u-v|^2 dx + O(\varepsilon).$$

Apply the Gronwall's in equality, we have that

$$\int_\Omega |u-v|^2 dx \leq O(\varepsilon) + O(\varepsilon)e^{Ct},$$

which implies $\int_\Omega |u-v|^2 dx = O(\varepsilon)$ for all $t < T$ with given $T > 0$, by the compactness of Ω. □

References

[1] Abdullahi, O., Karani, A., Tigoi, C. C., Mugo, D., Kungu, S., Wanjiru, E., Jomo, J., Musyimi, R., Lipsitch, M., and Scott, J. A. G. (2012). Rates of acquisition and clearance of pneumococcal serotypes in the nasopharynges of children in kilifi district, kenya. *Journal of Infectious Diseases*, 206(7):1020–1029.

[2] Abrams, P. A. and Ginzburg, L. R. (2000). The nature of predation: prey dependent, ratio dependent or neither? *Trends in Ecology and Evolution*, 15(8):337–341.

[3] Abu-Raddad, L. J. and Ferguson, N. M. (2004). The impact of cross-immunity, mutation and stochastic extinction on pathogen diversity. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 271(1556):2431–2438.

[4] Adler, F. R. and Brunet, R. C. (1991). The dynamics of simultaneous infections with altered susceptibilities. *Theoretical population biology*, 40(3):369–410.

[5] Alizon, S. (2008). Decreased overall virulence in coinfected hosts leads to the persistence of virulent parasites. *The American Naturalist*, 172(2):E67–E79.

[6] Alizon, S. (2013). Co-infection and super-infection models in evolutionary epidemiology. *Interface focus*, 3(6):20130031.

[7] Alizon, S., de Roode, J. C., and Michalakis, Y. (2013). Multiple infections and the evolution of virulence. *Ecology letters*, 16(4):556–567.

[8] Allen, L. J. (2008). An introduction to stochastic epidemic models. In *Mathematical epidemiology*, pages 81–130. Springer.

[9] Allen, L. J. S., Bolker, B. M., Lou, Y., and Nevai, A. L. (2008). Asymptotic profiles of the steady states for an sis epidemic reaction-diffusion model. *Discrete and Continuous Dynamical Systems*, 21(1):1–20.

[10] Allesina, S. and Levine, J. M. (2011). A competitive network theory of species diversity. *Proceedings of the National Academy of Sciences*, 108(14):5638–5642.

[11] Alon, U. (2006). *An Introduction to Systems Biology: Design Principles of Biological Circuits (1st ed.)*. Chapman and Hall/CRC.

[12] Alonso D, McKane AJ, P. M. (2007). Stochastic amplification in epidemics. *J R Soc Interface*, 4.

[13] Anderson, R. and May, R. (1988). Epidemiological parameters of hiv transmission. *Nature*, 333.

[14] Anderson, R. M. and May, R. (1991). *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press.

[15] Aratchige, P., McIntyre, P., Quinn, H., and Gilbert, G. (2008). Recent increases in mumps incidence in australia: the "forgotten" age group in the 1998 australian measles control campaign. *Med J Aust.*, 189.

[16] Arditi, R. and Ginzburg, L. R. (1989). Coupling in predator-prey dynamics: Ratio-dependence. *Journal of Theoretical Biology*, 139(3):311–326.

[17] Arnold, V. I. (1963). Proof of a theorem of a. n. kolmogorov on the invariance of quasi-periodic motions under small perturbations of the hamiltonian. *Russian Mathematical Surveys*, 18:9–36.

[18] Arnold, V. I. (1978). *Mathematical Methods of Classical Mechanics*. Springer.

[19] Arnold, V. I. (1994). *Bifurcation theory and catastrophe theory*. Springer.

[20] Ball, F. (1999). Stochastic and deterministic models for sis epidemics among a population partitioned into households. *Mathematical biosciences*, 156(1-2):41–67.

[21] Balmer, O. and Tanner, M. (2011). Prevalence and implications of multiple-strain infections. *Lancet Infect Dis*, 11.

[22] Bardina, X., Cuadrado, S., and Rovira, C. (2019). Coinfection in a stochastic model for bacteriophage systems. *Discrete and Continuous Dynamical Systems-Series B*, 24:6607–6620.

[23] Barrow-Green, J. (1996). *Poincare and the Three Body Problem*, volume 11. American Mathematical Society and London Mathematical Society.

[24] Bartlett, M. (1957). Measles periodicity and community size. *J R Stat Soc A*, 120.

[25] Bartlett, M. S. (1960). *Stochastic population models; in ecology and epidemiology*. New York and London:John Wiley.

[26] Berryman, A. A. (1992). The orgins and evolution of predator-prey theory. *Ecology*, 73(5):1530–1535.

[27] Biegert, M. (2003). The neumann laplacian on spaces of continuous functions. *Note di Matematica*, 22(1):65–74.

[28] Billah, M., Miah, M., and Khan, M. (2020). Reproductive number of coronavirus: A systematic review and meta-analysis based on global level evidence. *PLoS One*, 15.

[29] Birger, R., Kouyos, R., Dushoff, J., and Grenfell, B. (2015a). Modeling the effect of hiv coinfection on clearance and sustained virologic response during treatment for hepatitis c virus. *Epidemics*, 12:1–10. Papers arising from Epidemics 4.

[30] Birger, R. B., Kouyos, R. D., Cohen, T., Griffiths, E. C., Huijben, S., Mina, M. J., Volkova, V., Grenfell, B., and Metcalf, C. J. E. (2015b). The potential impact of coinfection on antimicrobial chemotherapy and drug resistance. *Trends in microbiology*, 23(9):537–544.

[31] Birkhoff, G. D. (1942). What is the ergodic theorem. *American Mathematical Monthly*, 49:222–226.

[32] Bjørnstad, O. N., Finkenstädt, B., and Grenfell, B. T. (2002). Dynamics of measles epidemics: Estimating scaling of transmission rates using a time series sir model. *Ecological Monographs*, 72:169–184.

[33] Boden, L. A. and McKendrick, I. J. (2017). Model-based policymaking: A framework to promote ethical "good practice" in mathematical modeling for public health policymaking. *Frontiers in Public Health*, 5:68.

[34] Borrelli, R. L. and Coleman, C. S. (1987). *Differential Equations: A Modeling Approach*. Prentice-Hall, Englewood Cliffs, NJ.

[35] Bratus, A. S., Posvyanskii, V. P., and Novozhilov, A. S. (2014). Replicator equations and space. *Mathematical Modelling of Natural Phenomena*, 9(3):47–67.

[36] Brauer, F. (2017). Mathematical epidemiology: Past, present, and future. *Infectious Disease Modelling*, 2(2):113–127.

[37] Brauer, F. and Castillo-Chavez, C. (2001). *Mathematical Models in Population Biology and Epidemiology*. Springer.

[38] Brauer, F., Driessche, P. v. d., and Wu, J. (2008). *Mathematical Epidemiology*. Springer.

[39] Brezis, H. (2010). *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer.

[40] Buck, J. (1988). Synchronous rhythmic flashing of fireflies. ii. *The Quarterly Review of Biology*, 63(3):265–289.

[41] Cantrell, R. S. and Cosner, C. (2004). *Spatial Ecology via Reaction-Diffusion Equations*. John Wiley and Sons, Ltd.

[42] Castella, F., Hoffbeck, J.-P., and Lagadeuc, Y. (2009). A reduced model for spatially structured predator-prey systems with fast spatial migrations and slow demographic evolutions. *Asymptotic Analysis*, 61(3-4):125–175.

[43] Castella, F., Madec, S., and Lagadeuc, Y. (2016). Global behavior of N competing species with strong diffusion: diffusion leads to exclusion. *Applicable Analysis*, 95(2):341–372.

[44] Cesari, L. (1963). *Asymptotic Behavior and Stability Problems in Ordinary Differential Equations*. Academic, New York.

[45] Chakerian, D. and Mandelbrot, B. B. (1982). The fractal geometry of nature. *College Mathematics Journal*, 15:175.

[46] Champneys, A. R., Hunt, G. W., and Thompson, J. M. T. (1999). *Localization and Solitary Waves in Solid Mechanics*. World Scientific.

[47] Chawanya, T. and Tokita, K. (2002). Large-dimensional replicator equations with antisymmetric random interactions. *Journal of the Physical Society of Japan*, 71(2):429–431.

[48] Chen, C. and Kang, Y. (2017). Dynamics of a stochastic multi-strain sis epidemic model driven by lévy noise. *Communications in Nonlinear Science and Numerical Simulation*, 42:379–395.

[49] Chen, L., Ghanbarnejad, F., and Brockmann, D. (2017). Fundamental properties of cooperative contagion processes. *New J. of Phys.*, 19(10).

[50] Chowell, G., Miller, M. A., and Viboud, C. (2008). Seasonal influenza in the united states, france, and australia: transmission and prospects for control. *Epidemiology and Infection*, 136(6):852–864.

[51] Cobey, S. and Lipsitch, M. (2012). Niche and neutral effects of acquired immunity permit coexistence of pneumococcal serotypes. *Science (New York, NY)*, 335:1376–1380.

[52] Cobey, S. and Lipsitch, M. (2013). Pathogen diversity and hidden regimes of apparent competition. *The American Naturalist*, 181(1):12–24.

[53] Cohen, T., van Helden, P. D., Wilson, D., Colijn, C., McLaughlin, M. M., Abubakar, I., and Warren, R. M. (2012). Mixed-strain mycobacterium tuberculosis infections and the implications for tuberculosis treatment and control. *Clinical microbiology reviews*, 25(4):708–719.

[54] Cole, J. D. and Kevorkian, J. (1963). Uniformly valid asymptotic approximations for certain non-linear differential equations. *Nonlinear Differential Equations and Nonlinear Mechanics*.

[55] Constable, G. W., Rogers, T., McKane, A. J., and Tarnita, C. E. (2016). Demographic noise can reverse the direction of deterministic selection. *Proceedings of the National Academy of Sciences*, 113(32):E4745–E4754.

[56] Coulson, T., Rohani, P., and Pascual, M. (2004). Skeletons, noise and population growth: the end of an old debate? *Trends in Ecology & Evolution*, 19(7):359–364.

[57] Courchamp, F., Berec, L., and Gascoigne, J. C. (2008). *Allee Effects in Ecology and Conservation*. Oxford University Press.

[58] Cox, F. (2001). Concomitant infections, parasites and immune responses. *Parasitology*, 122:S23 – S38.

[59] Cressman, R. and Tao, Y. (2014). The replicator equation and other game dynamics. *Proceedings of the National Academy of Sciences*, 111(Supplement 3):10810–10817.

[60] David, A. (2020). Special report: The simulations driving the world's response to covid-19. *Nature*, 580(7803):316–318.

[61] Davies, N. G., Flasche, S., Jit, M., and Atkins, K. E. (2019). Within-host dynamics shape antibiotic resistance in commensal bacteria. *Nature ecology and evolution*, 3(3):440.

[62] Dawes, J. H. P. and Gog, J. R. (2002). The onset of oscillatory dynamics in models of multiple disease strains. *Journal of mathematical biology*, 45.

[63] De Roode, J. C., Helinski, M. E., Anwar, M. A., and Read, A. F. (2005). Dynamics of multiple infection and within-host competition in genetically diverse malaria infections. *The American Naturalist*, 166(5):531–542.

[64] Delamater, P., Street, E., Leslie, T., Yang, Y., and Jacobsen, K. (2019). Complexity of the basic reproduction number (r0). *Emerg Infect Dis.*, 25.

[65] Diekmann, O., Heesterbeek, J., and Metz, J. A. (1990). On the definition and the computation of the basic reproduction ratio r_0 in models for infectious diseases in heterogeneous populations. *Journal of mathematical biology*, 28(4):365–382.

[66] Duan, G.-R. and Patton, R. J. (1998). A note on hurwitz stability of matrices. *Automatica*, 34.

[67] Dushoff, J. and Levin, S. (1995). The effects of population heterogeneity on disease invasion. *Mathematical biosciences*, 128 1-2:25–40.

[68] Epstein, I. R., Kustin, K., De Kepper, P., and Orbán, M. (1983). Oscillating chemical reactions. *Scientific American*, 248(3):112–123.

[69] Eshel, I. (1977). On the founder effect and the evolution of altruistic traits: an ecogenetical approach. *Theoretical population biology*, 11 3:410–24.

[70] Faust, K. and Raes, J. (2012). Microbial interactions: from networks to models. *Nat Rev Microbiol*, 10.

[71] Fenichel, N. (1977). Asymptotic stability with rate conditions, ii. *Indiana University Mathematics Journal*, 26(1):81–93.

[72] Fenichel, N. (1979). Geometric singular perturbation theory for ordinary differential equations. *Journal of Differential Equations*, 31(1):53–98.

[73] Ferguson, N. and Viggo, A. (2002). *The influence of different forms of cross-protective immunity on the population dynamics of antigenically diverse pathogens." Mathematical Approaches for Emerging and Reemerging Infectious Diseases: Models, Methods, and Theory*. Springer, New York, NY.

[74] Ferguson, N. M., Galvani, A. P., and Bush, R. M. (2003). Ecological and immunological determinants of influenza evolution. *Nature*, 422(6930):428–433.

[75] Fife, P. C. (1979). *Mathematical Aspects of Reacting and Diffusing Systems*. Springer.

[76] Fine, P., Eames, K., and Heymann, D. L. (2011). "Herd Immunity": A Rough Guide. *Clinical Infectious Diseases*, 52(7):911–916.

[77] Fisher, R. A. (1937). The wave of advance of advantageous genes. *Annals of Human Genetics*, 7:355–369.

[78] Fitzgibbon, W. and Langlais, M. (2008). Simple models for the transmission of microparasites between host populations living on noncoincident spatial domains. In *Structured Population Models in Biology and Epidemiology*.

[79] Fitzgibbon, W., Langlais, M., and Morgan, J. (2004). A reaction-diffusion system modeling direct and indirect transmission of diseases. *Discrete and Continuous Dynamical Systems-series B*, 4:893–910.

[80] Fraser, C., Donnelly, C. A., Cauchemez, S., Hanage, W. P., Van Kerkhove, M. D., Hollingsworth, T. D., Griffin, J., Baggaley, R. F., Jenkins, H. E., Lyons, E. J., Jombart, T., Hinsley, W. R., Grassly, N. C., Balloux, F., Ghani, A. C., Ferguson, N. M., Rambaut, A., Pybus, O. G., Lopez-Gatell, H., Alpuche-Aranda, C. M., Chapela, I. B., Zavala, E. P., Guevara, D. M. E., Checchi, F., Garcia, E., Hugonnet, S., and Roth, C. (2009). Pandemic potential of a strain of influenza a (h1n1): Early findings. *Science*, 324(5934):1557–1561.

[81] Furuya-Kanamori, L., Liang, S., Milinovich, G., Magalhaes, R. J. S., Clements, A. C., Hu, W., Brasil, P., Frentiu, F. D., Dunning, R., and Yakob, L. (2016). Co-distribution and co-infection of chikungunya and dengue viruses. *BMC infectious diseases*, 16(1):1–11.

[82] Gaivão, M., Dionisio, F., and Gjini, E. (2017). Transmission fitness in co-colonization and the persistence of bacterial pathogens. *Bulletin of mathematical biology*, 79(9):2068–2087.

[83] García-Rodríguez, J. Á. and Martinez, M. F. (2002). Dynamics of nasopharyngeal colonization by potential respiratory pathogens. *Journal of Antimicrobial Chemotherapy*, 50(suppl 3):59–74.

[84] Garnier, J., Roques, L., and Hamel, F. (2012). Success rate of a biological invasion in terms of the spatial distribution of the founding population. *Bulletin of Mathematical Biology*, 74:453–473.

[85] Geritz, S. A., Mesze, G., Metz, J. A., et al. (1998). Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evolutionary ecology*, 12(1):35–57.

[86] Gjini, E. and Madec, S. (2017). A slow-fast dynamic decomposition links neutral and non-neutral coexistence in interacting multi-strain pathogens. *Theoretical Ecology*, 10(1):129–141.

[87] Gjini, E. and Madec, S. (2019). Predicting multi-strain coexistence from co-colonization interactions. *in prep.*

[88] Gjini, E. and Madec, S. (2021a). The ratio of single to co-colonization is key to complexity in interacting systems with multiple strains. *Ecology and Evolution (in press) doi:10.1002/ece3.7259*.

[89] Gjini, E. and Madec, S. (2021b). Towards a mathematical understanding of colonization resistance. *bioRxiv*.

[90] Gjini, E., Valente, C., Sá-Leão, R., and Gomes, M. G. M. (2016). How direct competition shapes coexistence and vaccine effects in multi-strain pathogen systems. *Journal of Theoretical Biology*, 388:50–60.

[91] Goel, N. S., Maitra, S. C., and Montroll, E. W. (1971). On the volterra and other nonlinear models of interacting populations. *Rev. Mod. Phys.*, 43:231–276.

[92] Gog, J. and Grenfell, B. (2002). Dynamics and selection of many-strain pathogens. *Proc Natl Acad Sci USA.*, 99.

[93] Golubitsky, M. and Stewart, I. (2002). *The Symmetry Perspective: From Equilibrium to Chaos in Phase Space and Physical Space*. Birkhauser.

[94] Griffiths, E. C., Pedersen, A. B., Fenton, A., and Petchey, O. L. (2011). The nature and consequences of coinfection in humans. *Journal of Infection*, 63(3):200–206.

[95] Grilli, J., Adorisio, M., and Suweis, S. e. a. (2017). Feasibility and coexistence of large ecological communities. *Nat Commun*, 8.

[96] Guerra, F. M., Bolotin, S., Lim, G., Heffernan, J., Deeks, S. L., Li, Y., and Crowcroft, N. S. (2017). The basic reproduction number (r0) of measles: a systematic review. *The Lancet Infectious Diseases*, 17(12):e420–e428.

[97] Halliday, F. W., Penczykowski, R. M., Barrès, B., Eck, J. L., Numminen, E., and Laine, A.-L. (2020). Facilitative priority effects drive parasite assembly under coinfection. *Nature Ecology and Evolution*, 4(11):1510–1521.

[98] Hansen, J. and Day, T. (2014). Coinfection and the evolution of drug resistance. *Journal of evolutionary biology*, 27(12):2595–2604.

[99] Harko, T., Lobo, F. S., and Mak, M. (2014). Exact analytical solutions of the susceptible-infected-recovered (sir) epidemic model and of the sir model with equal death and birth rates. *Applied Mathematics and Computation*, 236:184–194.

[100] Hébert-Dufresne, L. and Althouse, B. M. (2015). Complex dynamics of synergistic coinfections on realistically clustered networks. *PNAS*, 112(33):10551–10556.

[101] Helleman, R. H. G. (1980). Self-generated chaotic behavior in nonlinear mechanics. in e. g. d. cohen, ed. *Fundamental Problems in Statistical Mechanics*, 5.

[102] Hofbauer, J. and Sigmund, K. (1998). *Evolutionary Games and Population Dynamics*. Cambridge University Press.

[103] Hofbauer, J. and Sigmund, K. (2003). Evolutionary game dynamics. *Bulletin of the American Mathematical Society*, 40(4):479–519.

[104] Holmes, M. H. (1995). *Introduction to Perturbation Methods*. Springer.

[105] Jordan, D. and Smith, P. (1987). *Nonlinear Ordinary Differential Equations*. Oxford University Press, Oxford, England.

[106] Kermack, W. O. and McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society A, London*, 115.

[107] Keshet, E. L. (1988). *Mathematical Models in Biology*. Random House, New York.

[108] King, A. A., Nguyen, D., Ionides, E. L., et al. (2016). Statistical inference for partially observed markov processes via the r package pomp. *Journal of Statistical Software*, 69(i12).

[109] Kirupaharan, N. and Allen, L. J. (2004). Coexistence of multiple pathogen strains in stochastic epidemic models with density-dependent mortality. *Bulletin of Mathematical Biology*, 66(4):841–864.

[110] Knight, G. M., Dharan, N. J., Fox, G. J., Stennis, N., Zwerling, A., Khurana, R., and Dowdy, D. W. (2016). Bridging the gap between evidence and policy for infectious diseases: How models can aid public health decision-making. *International Journal of Infectious Diseases*, 42:17–23.

[111] Kolmogorov, A. N. (1954). On the conservation of conditionally periodic motions under small perturbation of the hamiltonian. *Dokl. Akad. Nauk SSR*, 98.

[112] Kretzschmar, M., Teunis, P., and Pebody, R. (2010). Incidence and reproduction numbers of pertussis: estimates from serological and social contact data in five european countries. *PLoS Med.*, 7.

[113] Kucharski, A. J., Viggo, A., and Julia, R. G. (2016). Capturing the dynamics of pathogens with many strains. *Journal of mathematical biology*, 72.

[114] Kuehn, C. (2010). *Multiple Time Scale Dynamics*. Applied Mathematical Sciences book series (AMS, volume 191). Springer.

[115] Kuzmak, G. E. (1959). Asymptotic solutions of nonlinear second order differential equations with variable coefficients. *Journal of Applied Mathematics and Mechanics*, 23:730–744.

[116] Lajmanovich, A. and Yorke, J. (1976). A deterministic model for gonorrhea in a nonhomogeneous population. *Bellman Prize in Mathematical Biosciences*, 28:221–236.

[117] Lambrechts, L., Paaijmans, K. P., Fansiri, T., Carrington, L. B., Kramer, L. D., Thomas, M. B., and Scott, T. W. (2011). Impact of daily temperature fluctuations on dengue virus transmission by aedes aegypti. *Proceedings of the National Academy of Sciences*, 108(18):7460–7465.

[118] Lasalle, J. P. (1960). Some extensions of liapunov's second method. *IEEE Transactions on Circuits and Systems I-regular Papers*, 7:520–527.

[119] Le, T.-M.-T., Gjini, E., and Madec, S. (2021a). Disentangling how multiple traits drive 2 strain frequencies in sis dynamics with coinfection. *biorXiv, DOI: 10.1101/2021.04.29.442023*.

[120] Le, T. M. T., Gjini, E., and Madec, S. (2021b). Quasi-neutral dynamics in a coinfection system with n strains and multiple trait variation. *arXiv:2104.07289*.

[121] Le, T. M. T. and Madec, S. (2021). A reaction-advection-diffusion model for quasi-neutral dynamics of coinfected strains. *hal-0346720*.

[122] Levin, S. and Pimentel, D. (1981). Selection of intermediate rates of increase in parasite-host systems. *The American Naturalist*, 117(3):308–315.

[123] Li, Y., Liu, X., and Wang, L. (2017). Modelling the transmission dynamics and control of mumps in mainland china. *International Journal of Environmental Research and Public Health*, 15.

[124] Lin, Y. T., Kim, H., and Doering, C. R. (2012). Features of fast living: on the weak selection for longevity in degenerate birth-death processes. *Journal of Statistical Physics*, 148(4):647–663.

[125] Lindstedt, A. (1882). über die integration einer für die störungstheorie wichtigen differentialgleichung. *Astron Nachr*, 103.

[126] Lion, S., Boots, M., and Sasaki, A. (2021). Multi-morph eco-evolutionary dynamics in structured populations. *bioRxiv*.

[127] Lipsitch, M. (1997). Vaccination against colonizing bacteria with multiple serotypes. *PNAS*, 94(12):6571–6576.

[128] Lipsitch, M., Colijn, C., Cohen, T., Hanage, W. P., and Fraser, C. (2009). No coexistence for free: neutral null models for multistrain pathogens. *Epidemics*, 1(1):2–13.

[129] Liu, Y., Lillepold, K., Semenza, J. C., Tozan, Y., Quam, M. B., and Rocklöv, J. (2020). Reviewing estimates of the basic reproduction number for dengue, zika and chikungunya across global climate zones. *Environmental research*, 182:109114.

[130] Lloyd, A. and May, R. (1996). Spatial heterogeneity in epidemic models. *Journal of theoretical biology*, 179 1:1–11.

[131] Locatelli, I., Trächsel, B., and Rousson, V. (2021). Estimating the basic reproduction number for covid-19 in western europe. *PLoS ONE*, 16.

[132] Longo, G. and Soto, A. M. (2016). Why do we need theories? *Progress in Biophysics and Molecular Biology*, 122(1):4–10. From the Century of the Genome to the Century of the Organism: New Theoretical Approaches.

[133] Lorenz, E. (1972). Predictability: Does the flap of a butterfly's wings in brazil set off a tornado in texas? *American Association for the Advancement of Science*.

[134] Lorenz, E. N. (1963). Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, 20:130–141.

[135] Lotka, A. J. (1910). Contribution to the theory of periodic reactions. *The Journal of Physical Chemistry*, 14(3):271–274.

[136] Lotka, A. J. (1926). Elements of physical biology. *Science Progress in the Twentieth Century (1919-1933)*, 21(82):341–343.

[137] Lourenco, J. and Recker, M. (2013). Natural, persistent oscillations in a spatial multi-strain disease system with application to dengue. *PLoS Comput Biol.*, 9.

[138] Lyapunov, A. (1892). The general problem of the stability of motion (in russian). *Doctoral dissertation*.

[139] Ma, J. and Ma, Z. (2006). Epidemic threshold conditions for seasonally forced seir models. *Mathematical Biosciences and Engineering*, 3:161.

[140] Madec, S. and Gjini, E. (2020). Predicting n-strain coexistence from co-colonization interactions: epidemiology meets ecology and the replicator equation. *Bulletin of mathematical biology*, 82(11):1–26.

[141] Majumder, M. and Mandl, K. (2020). Early in the epidemic: impact of preprints on global discourse about covid-19 transmissibility. *Lancet Glob Health*, 8.

[142] Mandelbrot, B. B. and Aizenman, M. (1977). *Fractals: Form, Chance and Dimension*. W.H.Freeman and Co Ltd.

[143] Marchaim, D., Perez, F., Lee, J., Bheemreddy, S., Hujer, A. M., Rudin, S., Hayakawa, K., Lephart, P. R., Blunden, C., Shango, M., Campbell, M. L., Varkey, J., Manickam, P., Patel, D., Pogue, J. M., Chopra, T., Martin, E. T., Dhar, S., Bonomo, R. A., and Kaye, K. S. (2012). "swimming in resistance": Co-colonization with carbapenem-resistant enterobacteriaceae and acinetobacter baumannii or pseudomonas aeruginosa. *American journal of infection control*, 40(9):830—835.

[144] Margolis, E., Yates, A., and Levin, B. R. (2010). The ecology of nasal colonization of Streptococcus pneumoniae, Haemophilus influenzae and Staphylococcus aureus: the role of competition and interactions with host's immune response. *BMC Microbiology*, 10(1):59.

[145] Martcheva, M. (2009). A non-autonomous multi-strain sis epidemic model. *Journal of biological dynamics*, 3(2-3):235–251.

[146] May, R. M. and Nowak, M. A. (1995). Coinfection and the evolution of parasite virulence. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 261(1361):209–215.

[147] McKane, A. and Newman, T. (2005). Predator-prey cycles from resonant amplification of demographic stochasticity. *Physical Review Letters*, 94.

[148] Megretski, A. (2003). Singular perturbations and averaging. Lecture Notes, Massachusetts Institute of Technology.

[149] Meszéna, G., Gyllenberg, M., Jacobs, F. J., and Metz, J. A. (2005). Link between population dynamics and dynamics of darwinian evolution. *Physical review letters*, 95(7):078105.

[150] Metcalf, C., Edmunds, W., and Lessler, J. (2015). Six challenges in modelling for public health policy. *Epidemics*, 10:93–96. Challenges in Modelling Infectious Disease Dynamics.

[151] Miller, J. C. (2013). Cocirculation of infectious diseases on networks. *Physical Review E*, 87(6):060801.

[152] Minayev, P. and Ferguson, N. (2009). Incorporating demographic stochasticity into multi-strain epidemic models: application to influenza a. *Journal of the Royal Society Interface*, 6(40):989–996.

[153] Moser, J. K. (1962). On invariant curves of area-preserving mappings of an anulus. *Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II*.

[154] Mosquera, J. and Adler, F. R. (1998). Evolution of virulence: a unified framework for coinfection and superinfection. *Journal of Theoretical Biology*, 195(3):293–313.

[155] Mottoni, P., Orlandi, E., and Tesei, A. (1979). Asymptotic behavior for a system describing epidemics with migration and spatial spread of infection. *Nonlinear Analysis-theory Methods & Applications*, 3:663–675.

[156] Mulberry, N., Rutherford, A., and Colijn, C. (2019). Systematic comparison of coexistence in models of drug-sensitive and drug-resistant pathogen strains. *Theoretical Population Biology*.

[157] Murray, J. D. (2002). *Mathematical Biology, I. An Introduction*. Springer.

[158] Murray, J. D. (2003). *Mathematical Biology, II Spatial Models and Biomedical Applications*. Springer.

[159] Nakamura, G. and Martinez, A. (2019). Hamiltonian dynamics of the sis epidemic model with stochastic fluctuations. *Scientific Reports*, 9.

[160] Nåsell, I. (2002). Stochastic models of some endemic infections. *Mathematical biosciences*, 179(1):1–19.

[161] Neumann, J. v. (1932a). Physical applications of the ergodic hypothesis. *Proceedings of the National Academy of Sciences of the United States of America*, 18 3:263–6.

[162] Neumann, J. v. (1932b). Proof of the quasi-ergodic hypothesis. *Proceedings of the National Academy of Sciences of the United States of America*, 18 1:70–82.

[163] Nicoli, E. J., Ayabina, D., Trotter, C. L., Turner, K. M., and Colijn, C. (2015). Competition, coinfection and strain replacement in models of bordetella pertussis. *Theoretical population biology*, 103:84–92.

[164] Nowak, M. A. (2006). *Evolutionary Dynamics: Exploring the Equations of Life*. Harvard University Press.

[165] Nowak, M. A. and May, R. M. (1994). Superinfection and the evolution of parasite virulence. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 255(1342):81–89.

[166] Otsuka, J. (2019). *The Role of Mathematics in Evolutionary Theory*. Elements in the Philosophy of Biology. Cambridge University Press.

[167] Pang, D. and Xiao, Y. (2019). The sis model with diffusion of virus in the environment. *Mathematical Biosciences and Engineering*, 16.

[168] Parsons, T. L., Quince, C., and Plotkin, J. B. (2008). Absorption and fixation times for neutral and quasi-neutral populations with density dependence. *Theoretical population biology*, 74(4):302–310.

[169] Pazy, A. (1983). *Semigroups of linear Operators and Applications to Partial Differential Equations*, volume 44 of *Applied Mathematical Sciences*. Springer.

[170] Pearl, R. (1927). The growth of populations. *Quart. Rev. Biol.*, 2.

[171] Petney, T. N. and Andrews, R. H. (1998). Multiparasite communities in animals and humans: frequency, structure and pathogenic significance. *International Journal for Parasitology*, 28(3):377–393.

[172] Pinky, L., Gonzalez-Parra, G., and Dobrovolny, H. M. (2019). Effect of stochasticity on coinfection dynamics of respiratory viruses. *BMC bioinformatics*, 20(1):1–12.

[173] Pinotti, F., Ghanbarnejad, F., Hövel, P., and Poletto, C. (2019). Interplay between competitive and cooperative interactions in a three-player pathogen system. *arXiv preprint arXiv:1912.07289*.

[174] Poggiale, J. (1998a). Lotka-volterra's model and migrations: Breaking of the well-known center. *Mathematical and Computer Modelling*, 27(4):51–61.

[175] Poggiale, J. (1998b). Predator-prey models in heterogeneous environment: Emergence of functional response. *Mathematical and Computer Modelling*, 27(4):63–71.

[176] Poincaré, H. (1886). Sur les intégrales irrégulières des équations linéaires. *Acta Math*, 8.

[177] Poincaré, H. (1900). Sur l'équilibre d'une masse fluide animée d'un mouvement de rotation. *Acta Mathematica*, 7:259–380.

[178] Richards, F. J. (1959). A flexible growth function for empirical use. *Journal of Experimental Botany*, 10(29):290–300.

[179] Rohani, P., Keeling, M. J., and Grenfell, B. T. (2002). The interplay between determinism and stochasticity in childhood diseases. *The American Naturalist*, 159(5):469–481.

[180] Ross, R. (1916). An application of the theory of probabilities to the study of a priori pathometry.—part i. *Proc. R. Soc. Lond.*, A(92).

[181] Ross, R. and Hudson, H. P. (1917). An application of the theory of probabilities to the study of a priori pathometry.—part i. *Proceedings of The Royal Society A: Mathematical, Physical and Engineering Sciences*, 93:204–230.

[182] Rozhnova, G., Metcalf, C. J. E., and Grenfell, B. T. (2013). Characterizing the dynamics of rubella relative to measles: the role of stochasticity. *Journal of the Royal Society Interface*, 10.

[183] Ruan, S. (2007). Spatial-temporal dynamics in nonlocal epidemiological models. In *Mathematics for Life Science and Medicine*.

[184] Sasaki, A. and Dieckmann, U. (2011). Oligomorphic dynamics for analyzing the quantitative genetics of adaptive speciation. *Journal of mathematical biology*, 63(4):601–635.

[185] Schnakenberg, J. (1979). Simple chemical reaction systems with limit cycle behaviour. *Journal of Theoretical Biology*, 81(3):389–400.

[186] Shi, J. and Shivaji, R. (2006). Persistence in reaction diffusion models with weak allee effect. *Journal of Mathematical Biology*, 52:807–829.

[187] Sigmund, K. (2010). *The Calculus of Selfishness*. Princeton Series in Theoretical and Computational Biology. Princeton University Press.

[188] Smith, H. L. and Waltman, P. (1995). *The Theory of the Chemostat: Dynamics of Microbial Competition*. Cambridge Studies in Mathematical Biology. Cambridge University Press.

[189] Smoller, J. (1994). *Shock Waves and Reaction—Diffusion Equations*, volume 258 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag New York.

[190] Stokes, G. G. (1843). On some cases of fluid motion. *Trans Cambridge Philos Soc*, 8.

[191] Strogatz, S. (2018). *Nonlinear dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering, Second Edition*. CRC Press.

[192] Susi, H., Barrès, B., Vale, P. F., and Laine, A.-L. (2015). Co-infection alters population dynamics of infectious disease. *Nature communications*, 6(1):1–8.

[193] Takeuchi, Y. (1996). *Global Dynamical Properties of Lotka-Volterra Systems*. World Scientific Publishing Company.

[194] Teschl, G. (2010). *Ordinary differential equations and Dynamical systems*, volume 140 of *Graduate Studies in Mathematics*. American Mathematical Society.

[195] Thieme, H. R. (2007). Pathogen competition and coexistence and the evolution of virulence. In *Mathematics for Life Science and Medicine*, pages 123–153. Springer.

[196] Tikhonov, A. N. (1952). Systems of differential equations containing a small parameter multiplying the derivative. *Mat. SB. (NS)*, 31.

[197] Turchin, P. (2002). Does population ecology have general laws? *Zhurnal obshchei biologii*, 63 1:3–14.

[198] Turing, A. M. (1952). The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 237:37 – 72.

[199] van Baalen, M. and Sabelis, M. W. (1995). The dynamics of multiple infection and the evolution of virulence. *The American Naturalist*, 146(6):881–910.

[200] Verhulst, F. (1996). *Nonlinear Differential Equations and Dynamical Systems (Second Edition)*. Springer.

[201] Verhulst, P. F. (1838). Notice sur la loi que la population suit dans son accroissement. *Correspondance Mathematique et Physique Publiee par A*.

[202] Volpert, A. I., Volpert, V., and Volpert, V. A. (1994). *Traveling Wave Solutions of Parabolic Systems*. American Mathematical Society.

[203] Volterra, V. (1926). Fluctuations in the abundance of a species considered mathematically.

[204] Wang, X.-S., Wang, H., and Wu, J. (2012). Traveling waves of diffusive predator-prey systems: Disease outbreak propagation. *Discrete and Continuous Dynamical Systems*, 32:3303–3324.

[205] Wang, Z.-C. and Wu, J. (2009). Travelling waves of a diffusive kermack–mckendrick epidemic model with non-local delayed transmission. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 466:237 – 261.

[206] Warren, D. K., Nitin, A., Hill, C., Fraser, V. J., and Kollef, M. H. (2004). Occurrence of co-colonization or co-infection with vancomycin-resistant enterococci and methicillin-resistant staphylococcus aureus in a medical intensive care unit. *Infection Control and Hospital Epidemiology*, 25(2):99–104.

[207] World-Health-Organization (2003). Consensus document on the epidemiology of severe acute respiratory syndrome (sars).

[208] Wu, Y. and Zou, X. (2016). Asymptotic profiles of steady states for a diffusive sis epidemic model with mass action infection mechanism. *Journal of Differential Equations*, 261:4424–4447.

[209] Zhang, Y., Auranen, K., and Eichner, M. (2004). The influence of competition and vaccination on the coexistence of two pneumococcal serotypes. *Epidemiology and Infection*, 132(06):1073–1081.

Appendix A

Lyapunov Functions

This section of appendix is to remind the Lyapunov functions and the LaSalle's invariance principle, which we use in chapter 2. The presented theory can be referred in [188, 191].

In many cases, even for systems that have nothing to do with mechanics, it is occasionally possible to construct an energy-like function that decreases along trajectories. Such a function is called a Lyapunov function, which is defined as follows.

Definition A.1. (Lyapunov function) Given the a system $\dot{x} = f(x)$ with a fixed point at x^*, i.e. $f(x^*) = 0$. A function $V(x)$ is called Lyapunov function if it satisfies:

- $V(x)$ is real-valued and continuously differentiable.
- $V(x) > 0$ for all $x \neq x^*$, and $V(x^*) = 0$. (We say that V is *positive definite*.)
- $\dot{V}(x) < 0$ for all $x \neq x^*$ with $\dot{V}(x) = \nabla V(x) \cdot \dot{x} < 0$ (All trajectories flow "downhill" toward x^*.)

The following result is of fundamental importance in system theory [138]. It asserts the possibility of establishing stability or asymptotic stability of equilibrium points without explicitly computing trajectories.

Theorem A.2. ***(Lyapunov)*** *Given the system*

$$\dot{x}(t) = f(x(t)), \quad x(0) = x_0, \qquad f(0) = 0 \tag{A4.0.1}$$

and let $\phi(t;0,x_0)$ denote the unique solution $x(t)$ to (A4.0.1).

Let $x_e = 0$ be an equilibrium point for (A4.0.1). *Let $V : \mathbb{R}^n \to \mathbb{R}$ radially unbounded, be a positive definite continuously differentiable function.*

1. *If $\dot{V} : \mathbb{R}^n \to \mathbb{R}$ is negative semi-definite, then x_e is stable.*
2. *If $\dot{V}$ is negative definite, then x_e is asymptotically stable.*

Proof. Suppose that $\dot{V} : \mathbb{R}^n \to R$ is negative semi-definite. Given $\varepsilon > 0$, consider the closed ball $\bar{B}(0,\varepsilon)$. Since its boundary $S(0,\varepsilon)$ is compact (closed and bounded) and V is continuous, V admits a minimum m on $S(0,\varepsilon)$ by Weierstrass's theorem. Such minimum is positive because V is positive definite:

$$\min_{\{x:\|x\|=\varepsilon\}} V(x) = m > 0$$

Since V is continuous, in particular at the origin, there exists a $\delta > 0$ such that

$$x_0 \in B(0,\delta) \Rightarrow |V(x) - V(0)| = V(x) < m.$$

We claim that this δ is the δ required in the definition of stability, so that any trajectory starting from $B(0,\delta)$ never exits $B(0,\varepsilon)$.

Choose indeed $x_0 \in B(0,\delta)$ as the initial condition for (A4.0.1), and for the sake of contradiction suppose that the trajectory $\phi(t;0,x_0)$ is not entirely contained in the ball $B(0,\varepsilon)$. Then there exists a time T in which the trajectory intersects the boundary of $\bar{B}(0,\varepsilon)$, i.e. $V(\phi(T;0;x_0)) \geq m$. But the derivative of V with respect to time, that is $\dot{V}$, is negative semi-definite, hence V is non-increasing along the corresponding trajectory (that is, $V(\phi(T;0,x_0)) \leq V(x_0)$). Therefore,

$$m \leq V(\phi(T;0,x_0)) \leq V(x_0) < m$$

which is a contradiction. Hence, the trajectory is contained in $B(0,\varepsilon)$. Given $\varepsilon > 0$, we have constructed a $\delta > 0$ such that if $x_0 \in B(0,\delta)$ then $\phi(t;0,x_0) \in B(0;\varepsilon)$ for all $t \geq 0$. Hence, 0 is a stable equilibrium point.

Suppose now that $\dot{V}$ is negative definite, which implies that $\dot{V}$ is also negative semi-definite. Hence the first property in the definition of asymptotic stability is trivially satisfied. This means that, given $\varepsilon > 0$, there exists $\delta > 0$ such that if $x_0 \in B(0;\delta)$ then $\phi(t;0;x_0) \in B(0;\varepsilon)$ for all $t \geq 0$.

We claim that $\lim_{t\to\infty} \phi(t;0,x_0) = 0$ or, more explicitly that for all ε' such that $0 < \varepsilon' < \varepsilon$, there exists a certain time T such that $\phi(t;0,x_0) \in B(0,\varepsilon')$ for all $t \geq T$. Indeed, in view of stability and time invariance, for all $\varepsilon' > 0$ there exists a $\delta' > 0$ such that, if $x(T) \in B(0;\delta')$, then $\phi(t;T;x(T)) \in B(0,\varepsilon')$ for all $t \geq T$. Hence, we just need to prove that there exists T such that $x(T) \in B(0,\delta')$.

For the sake of contradiction, suppose that this is not the case. Then, for all $t \geq 0$ we have

$$\phi(t;0,x_0) \in \bar{B}(0,\varepsilon) \setminus B(0,\delta')$$

Since $\bar{B}(0,\varepsilon) \setminus B(0,\delta')$ is compact and $\dot{V}$ is continuous and negative definite, $\dot{V}$ attains a negative maximum $-\mu$ there. Hence, $\dot{V}(x) < -\mu$ if $x \in \bar{B}(0,\varepsilon) \setminus B(0,\delta')$ and then

$$V(\phi(t;0,x_0)) = V(x_0) + \int_0^t \dot{V}(\phi(\tau;0,x_0))\, d\tau \leq V(x_0) - \mu t.$$

Letting $t \to \infty$, we obtain the contradiction because $V(x) > 0$ for all $x \in \bar{B}(0,\varepsilon)$ but the right-hand side tends to $-\infty$. Therefore there must exist T such that $x(T) \in B(0,\delta')$. This proves the theorem. □

The LaSalle's invariance principle [118] gives a criterion for asymptotic stability in the case when $\dot{V}(x)$ is only negative semi-definite.

Theorem A.3. ***(LaSalle's invariance principle)*** *Let Ω be a closed set with the property that every solution of*

$$\dot{x} = f(x), \qquad \textit{where } x \textit{ is the vector of variables} \tag{A4.0.2}$$

which begins in Ω remains for all future time in Ω. Suppose that we can find a C^1 function $V(x)$ such that $\dot{V}(x) \leq 0$ and $V(x) > 0$ for all $x \neq 0$. Let E be the set of all points in Ω where $\dot{V}(x) = 0$. Let M be the largest invariant set in E. Then every solution starting in Ω approaches M as $t \to +\infty$.

By this, we can deduce that, if we can find a Lyapunov function $V(x)$, x^* is globally asymptotically stable: for all initial conditions, $x(t) \to x^*$ as $t \to \infty$. For the proof of this, see [105]. We note that, using the LaSalle's invariance principle, we can prove that a system has no closed orbits.

The intuition is that all trajectories move monotonically down the graph of $V(x)$ toward x^*, see figure A.1 The solutions can not get stuck anywhere because if they did, $V(x)$ would stop changing, but by assumption,

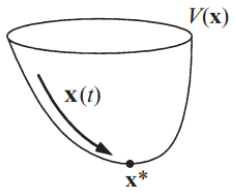


Fig. A.1 Intuition of the Lyapunov function

$\dot{V} < 0$ everywhere except at x^*. Unfortunately, there is not systematic way to construct Lyapunov functions. One of ideas, which occasionally works, is sum of squares, as we do in section 2.

Proof. Let $x(t)$ be a solution of (A4.0.2) starting in Ω. Since $\dot{V}(x) < 0$ in Ω, $V(x(t))$ is a non-increasing function of t. Recalling that $V(x(t)) > 0$, then $V(x(t))$ has a limit c as $t \to +\infty$.

Note also that the positive limiting set Γ^+ (the set of all limit points when $t \to +\infty$) is a subset of Ω (because Ω is closed), and since V is continuous on Ω, $V(x) = c$ on Γ^+. Since Γ^+ is invariant under (A4.0.2), we have that $\dot{V}(x) = 0$ on Γ^+, which implies $\Gamma^+ \subset M$ i.e. $x(t) \to M$ as $t \to +\infty$. Hence, all solutions starting in Ω approach M as t approaches infinity. □

We close this section by applying Lyapunov's theorem to prove an algebraic result.

Theorem A.4. *Let $A \in \mathbb{R}^{n\times n}$. The following statements are equivalent:*

1. *all the eigenvalues of A have negative real part;*
2. *for all matrices $Q = Q^T \succ 0$ there exists an unique solution $P = P^T \succ 0$ to the following (Lyapunov) equation:*

$$A^T P + PA + Q = 0. \tag{A4.0.3}$$

Proof. $(2) \Rightarrow 1)$) Consider the equation $\dot{x} = Ax$. To prove 1), it suffices to show that $x^* = 0$ is an asymptotically stable state. We consider

$$V(x) = x^T (Px), \qquad \text{with} \quad A^T P + PA = -I.$$

V is a Lyapunov function, indeed,

- Since $P = P^T \succ 0$ then $V(x) = x^T(Px) > 0$ for all $x \neq 0$,
- $\dot{V}(x) = (\dot{x})^T (Px) + x^T P\dot{x} = x^T A^T Px + x^T PAx = x^T \left(A^T + PA\right) x < 0$, for all $x \neq 0$.

Hence, $x^* = 0$ is asymptotically stable.
$(1) \Rightarrow 2)$) Suppose that all the eigenvalues of A have negative real parts. Define

$$P := \int_0^\infty e^{A^T t} Q e^{At} dt.$$

Since the elements of the integrand matrix are all linear combinations of functions of the form $t^k e^{\alpha t}$ where α has negative real part (this can be verified by considering the Jordan form of A), the integral exists and finite.

We claim that P satisfies $A^T P + PA + Q = 0$. Indeed,

$$A^T P + PA = \int_0^\infty \left(A^T e^{A^T t} Q e^{At} + e^{A^T t} Q e^{At} A \right) dt = \int_0^\infty \frac{d}{dt} \left(e^{A^T t} Q e^{At} \right) dt = e^{A^T t} Q e^{At} |_0^\infty$$

Alternatively, since all of eigenvalues of A have negative real parts, we have that $x^* = 0$ is an asymptotically stable equilibrium of $\dot{x} = Ax$, leading $e^{At} \to 0$ and $e^{A^T t} \to 0$ as $t \to +\infty$. Hence, we observe that $\lim_{t \to +\infty} e^{A^T t} Q e^{At} = 0$, which leads to $A^T P + PA = Q$.

Moreover, the solution is unique; indeed, let P_1 and P_2 be any two solutions of (A4.0.3), then $A^T (P_1 - P_2) + (P_1 - P_2) A = 0$. Hence

$$0 = e^{A^T t} \left(A^T (P_1 - P_2) + (P_1 - P_2) A \right) e^{At} = \frac{d}{dt} \left(e^{A^T t} (P_1 - P_2) e^{At} \right)$$

leading to $e^{A^T t} (P_1 - P_2) e^{At}$ is a constant for all t. Recalling that $\lim_{t \to +\infty} e^{A^T t} (P_1 - P_2) e^{At} = 0$, letting $t = 0$, we have that $P_1 = P_2$. □

Appendix B

Proof for Tikhonov's theorem 1952

Theorem B.1 (The Tikhonov's theorem)**.** *Consider the initial value problem*

$$\begin{cases} \dfrac{dx}{dt} &= f(x,z,t), \quad x(t_0) = x_0, \quad x \in D \subset \mathbb{R}^n, \\ \mu \dfrac{dz}{dt} &= F(x,z,t), \quad z(t_0) = z_0, \quad z \in G \subset \mathbb{R}^m. \end{cases} \tag{A4.0.1}$$

For f and F, we take sufficiently smooth vector functions in x, z and t.

a. *We assume that a unique solution of the initial value problem exists and suppose this holds also for the reduced problem*

$$\begin{cases} \dfrac{dx}{dt} &= f(x,z,t), \qquad x(t_0) = x_0, \\ 0 &= F(x,z,t), \end{cases} \tag{A4.0.2}$$

   with solutions $\bar{x}(t)$, $\bar{z}(t)$.

b. *Suppose that $0 = F(x,z,t)$ is solved by $\bar{z} = \phi(x,t)$, where $\phi(x,t)$ is a continuous function and an isolated root. Also, suppose that $\bar{z} = \phi(x,t)$ is an asymptotically stable solution of the equation $\dfrac{dz}{d\tau} = F(x,z,t)$ that is uniform in the parameters $x \in D$ and $t \in \mathbb{R}^+$.*

c. *$z(0)$ is contained in an interir subset of the domain of attraction of $\bar{z} = \phi(x,t)$ in the case of the parameter values $x = x(0)$, $t = t_0$.*

Then we have

$$\begin{cases} \lim\limits_{\varepsilon \to 0} x_\varepsilon(t) = \bar{x}(t), \qquad t_0 \le t \le L, \\ \lim\limits_{\varepsilon \to 0} z_\varepsilon(t) = \bar{z}(t), \quad t_0 < d \le t \le L \end{cases} \tag{A4.0.3}$$

with d and L are constants independent of ε.

We present the proof in the original publication of Tikhnov's. First, we recall the definition of asymptotic stability and state a lemma. For the purpose of consistence, denote by $B(x_0, r)$ the ball with center x_0 and radius r, $r > 0$.

Definition B.2. Assume that the equation

$$\frac{dz}{d\tau} = F(x,z,t) \tag{A4.0.4}$$

has a solution $\bar{z} = \phi(x,t)$ stable at any (x,t) in some closed set $\Omega \subset D \times [0,+\infty)$. Given any $\varepsilon > 0$, then there exists $\delta(\varepsilon)$ independent on (x,t) such that, the trajectory of any point in $B(\bar{z}(\tau_0), \delta(\varepsilon))$ converges to $\bar{z}(\tau)$ when $\tau \to \infty$, i.e.

$$\|z(\tau_0) - \bar{z}(\tau_0)\| < \delta(\varepsilon) \qquad \Rightarrow \qquad \lim_{\tau\to\infty} z(\tau) = \bar{z}(\tau). \tag{A4.0.5}$$

Remark: From this definition we can deduce that trajectories starting at points in $B(\bar{z}(\tau_0), \delta(\varepsilon))$ do not come out $B(\bar{z}(\tau), \varepsilon)$, i.e. for any $\tau > \tau_0$

$$\|z(\tau_0) - \bar{z}(\tau_0)\| < \delta(\varepsilon) \qquad \Rightarrow \qquad \|z(\tau) - \bar{z}(\tau)\| < \varepsilon. \tag{A4.0.6}$$

Lemma B.3. Given $\varepsilon > 0$, we can find $\delta(\varepsilon)$ and $\mu(\varepsilon)$ such that: if some trajectory $(x(t,\mu), z(t,\mu))$ which satisfies the hypothesis of Theorem B.1 at $\mu < \mu(\varepsilon)$, starts in $B(\bar{z}(0), \delta(\varepsilon))$, with $\bar{z} = \phi(x,t)$ is the asymptotically stable solution of $\dfrac{dz}{d\tau} = F(x,z,t)$ for any points $(x,t) \in \Omega \subset D \times [0,+\infty)$, then

i. the trajectory $(x(t,\mu), z(t,\mu))$ stays forever in $B(\bar{z}, \varepsilon)$;

ii. the projection of $(x(t,\mu), z(t,\mu))$ onto the space (x,t) never leaves Ω.

Proof. Take $\delta(\varepsilon) = \bar{\delta}\left(\dfrac{\varepsilon}{2}\right)$, where $\bar{\delta}$ defined in the Definition B.2. Let us prove that there exists $\mu(\varepsilon)$ satisfies Lemma B.3.
Assume the contrast that for arbitrary $\delta_n(\varepsilon) \to 0$ there exists $L_n(t)$ be trajectory of $(x(t,\mu_n), z(t,\mu_n))$ of (A4.0.1) with $\mu_n \to 0$ enters $B(\bar{z}, \delta(\varepsilon))$ but does not stay forever in $B(\bar{z}, \varepsilon)$. For each given values $\tilde{t}_n \to \infty$, let $\left(x\left(\bar{t'}_n, \mu_n\right), z\left(\bar{t'}_n, \mu_n\right)\right)$ be first point of the trajectory leaves $B(\bar{z}, \varepsilon)$ at $t > \tilde{t}_n$.

Denote by $\bar{t_n}$ last value of time corresponding intersection L_n with boundary of $B(\bar{z}, \delta(\varepsilon))$ and before $\bar{t'_n}$. Thus

$$\begin{aligned} \delta(\varepsilon) < \|z(t,\mu_n) - \phi\left(x(t,\mu_n), t\right)\| < \varepsilon, \qquad \bar{t_n} < t < \bar{t'_n}, \\ \|z\left(\bar{t'_n}, \mu_n\right) - \phi\left(x(\bar{t'_n}, \mu_n), \bar{t'_n}\right)\| = \varepsilon, \qquad t = \bar{t'_n}. \end{aligned} \tag{A4.0.7}$$

By our assumption at first, $L_n(\bar{t_n})$ converges to (x_0, z_0, t_0). Consider the trajectory $L_0(\tau)$ defined by the equations

$$\frac{dz}{d\tau} = F(x_0, z, t_0), \qquad z(0) = z_0 \tag{A4.0.8}$$

and then transform the original system to

$$\begin{cases} \dfrac{dx}{d\tau} &= \mu f(x, z, \bar{t_n} + \mu_n \tau), \quad x(0) = x(\bar{t_n}, \mu_n), \\ \dfrac{dz}{d\tau} &= F(x, z, \bar{t_n} + \mu_n \tau), \quad z(0) = z, (\bar{t_n}, \mu_n). \end{cases} \tag{A4.0.9}$$

where $\tau = \dfrac{t - \bar{t_n}}{\mu_n}$. Obviously, the solution of (A4.0.9) is $L_n(\bar{t_n} + \mu_n \tau)$. Since continuous change of solutions depends on the continuous change of the right hand side of first equation of (A4.0.9) and the initial condition $L_n(\bar{t_n}) \to (x_0, z_0, t_0) = L_0(0)$, we have that

$$L_n(\bar{t_n} + \mu_n \tau) \to L_0(\tau) \qquad \text{as} \qquad n \to \infty. \tag{A4.0.10}$$

On the other side, since $(x_0, z_0, t_0) \in B(\bar{z}(0), \delta(\varepsilon))$, it is possible to find τ_0 be strictly positive so that $L_0(\tau_0) \in B(\bar{z}, \delta(\varepsilon))$ then whole the trajectory $L_0(\tau) \in B\left(\bar{z}, \dfrac{\varepsilon}{2}\right)$ for $0 \le \tau \le \tau_0$.

For $0 \leq \tau \leq \tau_0$, $L_n(\bar{t_n} + \mu_n \tau)$ approaches $L_0(\tau)$ as $n \to \infty$, and therefore, starting from some n_0, the trajectories $L_n(\bar{t_n} + \mu_n \tau)$ for $n > n_0$, $0 \leq \tau \leq \tau_0$ lie entirely in $B\left(\bar{z}, \frac{3}{4}\varepsilon\right)$ and for $\tau = \tau_0$ are in $B(\bar{z}, \delta(\varepsilon))$. Note that $(x_0, z_0, t_0) \in B(\bar{z}(0), \delta(\varepsilon))$, consequently, $L_0(\tau)$ always stays in $B\left(\bar{z}, \frac{\varepsilon}{2}\right)$ by the definition $\delta(\varepsilon) = \bar{\delta}\left(\frac{\varepsilon}{2}\right)$. Hence, $L_n(\bar{t_n} + \mu_n \tau)$ cannot escape from $B(\bar{z}, \varepsilon)$ for $\tau > \tau_0$ when $\mu_n \to 0$.

This, however, contrasts to the behavior of $L_n(t)$ assumed at first. □

Now, we can back to the theorem.

Proof. Take an arbitrary ε and choose $\delta(\varepsilon)$ defined by Lemma B.3. Consider the associated system with solution $\bar{\bar{z}}$

$$\frac{d\bar{\bar{z}}}{d\tau} = F(x_0, \bar{\bar{z}}, t_0), \qquad \bar{\bar{z}}(0) = z_0 \tag{A4.0.11}$$

where (x_0, z_0, t_0) are initial values of (A4.0.1). Since z_0 is in the domain of attraction of the solution $z = \phi(x,t)$ by the condition 2, the trajectory $\bar{\bar{z}}(\tau)$ approaches a stable point $\phi(x_0, t_0)$ as $\tau \to \infty$. Hence, from some τ_0, $\bar{\bar{z}}(\tau)$ stays forever in $B\left(\phi(x_0, t_0), \frac{\delta(\varepsilon)}{2}\right)$.

Rewrite the original system into a new form

$$\begin{cases} \frac{dx}{d\tau} &= \mu f(x, z, t_0 + \mu\tau), \quad x(\tau)|_{\tau=0} = x_0, \\ \frac{dz}{d\tau} &= F(x, z, t_0 + \mu\tau), \quad z(\tau)|_{\tau=0} = z_0. \end{cases} \tag{A4.0.12}$$

By Lemma B.3, as $\mu \to 0$, $z(\tau, \mu)$ approaches $\bar{\bar{z}}(\tau)$ (of (A4.0.11)), therefore, there exists μ_0 such that if $\mu < \mu_0$ then $z(t, \mu) \in B\left(\phi(x_0, t_0), \frac{\delta(\varepsilon)}{2}\right)$ from $t_1 = t_0 + \mu\tau_0$. On the other side, we can also choose this μ_0 small enough such that for $\mu < \mu_0$,

$$\|\phi(x(t,\mu), t) - \phi(x_0, t_0)\| < \frac{\delta(\varepsilon)}{2}, \quad \forall t < t_0 + \mu_0\tau_0. \tag{A4.0.13}$$

Hence, we deduce that

$$\|z(t,\mu) - \phi(x(t,\mu), t)\| < \delta(\varepsilon), \qquad t = t_1 = t_0 + \mu\tau_0, \tag{A4.0.14}$$

which implies that the trajectory $(x(t,\mu), z(t,\mu)) \subset B(\phi(x,t), \delta(\varepsilon))$ by the asymptotic stability of $\phi(x,t)$.

By Definition B.1, the trajectory $(x(t,\mu), z(t,\mu))$ never exits out of $B(\bar{z}, \varepsilon)$, while $(x(t,\mu), t) \subset \Omega$ for all t. Thus

$$z(t,\mu) = \phi(x(t,\mu), t) + \varepsilon(t,\mu), \qquad t \geq t_1 = t_0 + \mu\tau_0. \tag{A4.0.15}$$

Moreover, $\varepsilon(t,\mu)$ can be small arbitrarily by take μ small enough. Consequently,

$$\frac{dx}{dt} = f(x(t,\mu), \phi(x(t,\mu), t) + \varepsilon(t,\mu), t), \qquad t \geq t_1 \tag{A4.0.16}$$

with initial condition $x|_{t=t_1} = x_0 + \bar{\delta}(\mu)$ with $\bar{\delta}(\mu) \to 0$ as $\mu \to 0$.

By Poincare expansion theorem, there exists a limit

$$\bar{x}(t) = \lim_{\mu \to 0} x(t, \mu) \tag{A4.0.17}$$

satisfies the equation

$$\begin{cases} \dfrac{d\bar{x}}{dt} = f\left(\bar{x}, \phi(\bar{x}, t), t\right), \\ \bar{x}(t_0) = x_0. \end{cases} \tag{A4.0.18}$$

Combine this derivation with (A4.0.15), which concludes the theorem. □

Appendix C

Simulations

C.1 Error estimates and time-complexity comparison between quasi-neutral system and the replicator equation

In this section, we present numerical simulations to verify the error estimates when approximating the quasi-neutral system in Chapter 2 using the corresponding replicator equation. We recall the following estimate which is the equation (2.3.21) in Theorem 2.7, using the solution of the slow-system at time τ to approximate the system of the quasi-neutral system at time $\frac{\tau}{\varepsilon}$, for all $\tau > 0$.

$$\text{Error estimate} = \left|S\left(\frac{\tau}{\varepsilon}\right) - S^*\right| + \sum_{i=1}^{N}\left|I_i\left(\frac{\tau}{\varepsilon}\right) - I^* z_i(\tau)\right| + \sum_{i,j=1}^{N}\left|I_{ij}\left(\frac{\tau}{\varepsilon}\right) - D^* z_i z_j(\tau)\right| \tag{C.1.1}$$

Using the ODE solving processes of scipy.integrate.odeint for two models (quasi-neutral and slow dyanmics), we also can have the comparison in time complexity, see C.1. Since we choose $\varepsilon = 0.1$, the time T is taken to be large enough $T = 200$ and solve the replicator equation for $\tau \in (0, 20)$.

We also note that, the initial values for the replicator system is generated randomly as well. Accordingly, the initial vales for the original quasi-neutral system is calculated using the formula

$$I_i(0) = I^* z_i(0), \qquad I_{ij}(0) = D^* z_i(0) z_j(0), \qquad S_0 = S^* = \frac{m}{\beta}.$$

We can change the range for creating neutral parameters r, β, γ and k to see the different cases. Using the neutral parameters and values S^*, T^*, I^*, D^* computed, we can input and solve the replicator system by its pairwise invasion fitness matrix.

The figure C.1 (a) illustrates our results in error estimate in Chapter 2, in which the Error Estimate in (C.1.1) is equal to $O(\varepsilon)$. However, we can note that, there is a trade-off, since the two time-scale in (C.1.1) is $\frac{\tau}{\varepsilon}$ and τ. We can choose a smaller ε but the fast-time that can be used to approximated becomes larger.

The figure C.1 (b) illustrates the comparison of time-complexity of two solving progresses. This is plausible because the dimension of the original system is $N^2 + N + 1$ and the dimension of the replicator equation is N. For example, this fact is displayed clearly when $N = 20$, the quasi-neutral system now contains 401 equations whereas the corresponding replicator model has only 20 equations.

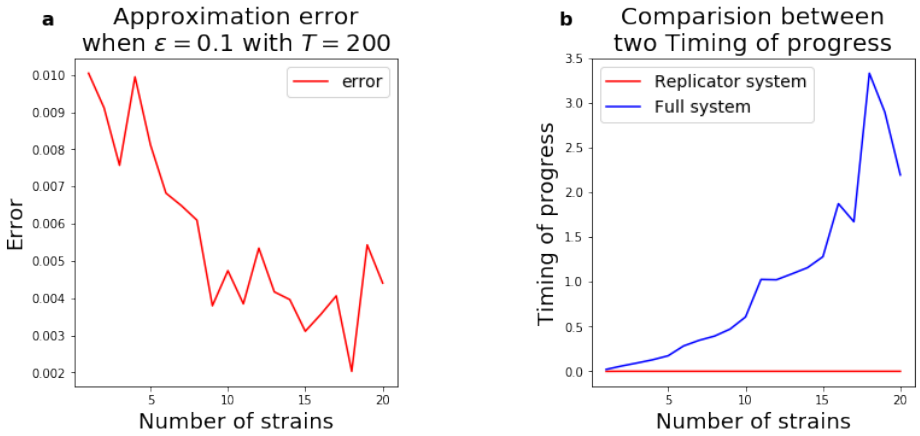


Fig. C.1 **Error estimates and time-complexity comparison between quasi-neutral system and its corresponding replicator equations.** We make the simulations for multiple cases of number of strains, varying from 2 to 20. In each case, we compute the mean error estimates at the last time (when $\tau = 20$ and $t = 200$) and mean time of processes over 10 pairs of SIS system and its corresponding replicator system, randomly created. We choose randomly these parameters, including transmission rate β, the clearance rate γ, and the co-colonization interaction coefficient $k > 1$ meaning the cooperation. Meanwhile the mortality as well as the recruitment rate are kept $r = 0.3$. The value $\varepsilon = 0.1$ and all the variations in dimension traits are randomly created. We solve the original SIS systems for $t = 200$ which implies $\tau = 20$ for the replicator systems. **(a)** The error estimates using the formulas (C.1.1). **(b)** The time-complexity in each case solving the original quasi-neutral system and replicator equation. (Data & Code)

C.2 Two examples of ten-strain systems and their space phases

We make an illustration for a replicator system of ten species. In these example, we consider two cases when the the difference only occurs in co-colonization interaction coefficient k_{ij} or in the transmission capacity p^s_{ij}, with $s \in \{i, j\}$, of a host co-colonized by strain i then strain j. The code can be found in ccode. Each vertex i of the decagon in each figure represent for the unique survival of strain i.

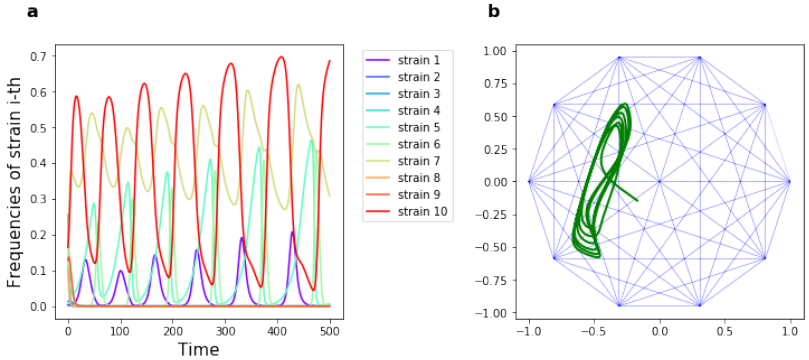


Fig. C.2 **An example of replicator system for 10 strains when variation in only k_{ij}.** The figure **(a)**, plots the frequencies of all strains. The figure **(b)** plots the space phase. The variation α_{ij}'s are chosen randomly whereas the transmission rate β is equal to 2, the mortality as well as the recruitment rate is 0.3 and the clearance rate γ is chosen to be 1.2. The co-colonization interaction k is chosen randomly from 0 to 10.

For this first case, when the perturbation is only in k_{ij}, we recall that the pairwise fitness invasion now becomes

$$\lambda_i^j = \Theta_5 \left(\mu \left(\alpha_{ji} - \alpha_{ij} \right) + \alpha_{ji} - \alpha_{jj} \right)$$

and note that $\mu = \dfrac{1}{k(R_0-1)}$ now becomes large because $R_0 = \dfrac{4}{3}$ - quite near 1, which yield μ turns large. In this figure C.2, the matrix of pairwise invasion fitness is nearly anti-symmetric, which makes the space-phase tends to go around in a submanifold of the invariant set $\{z_1 + \cdots + z_N = 1\}$ (actually, it is not a cycle).

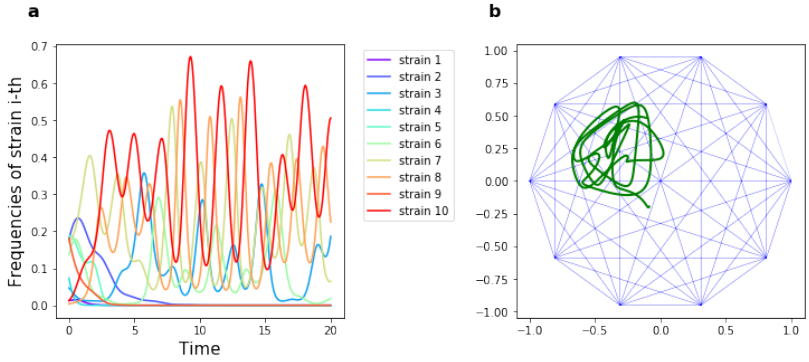


Fig. C.3 **An example of replicator system for 10 strains when variation in only p^s_{ij}.** The figure **(a)**, plots the frequencies of all strains. The figure **(b)** plots the space phase. The variation ω_{ij}'s are chosen randomly whereas the transmission rate β is equal to 2, the mortality as well as the recruitment rate is 0.3 and the clearance rate γ is chosen to be 1.2. The co-colonization interaction k is chosen randomly from 0 to 10.

In this case, as we mention in preceding chapter, the matrix of pairwise invasion fitness is anti-symmetric since

$$\lambda_i^j = \Theta_4\left(\omega_{ij}^i - \omega_{ji}^j\right).$$

We can note that there is 5 strain surviving in long time and the space phase tends to go around in a submanifold $\{z_{i1} + \cdots + z_{i5} = 1\}$. The trajectory is not cycle but tends to go around and have no equilibrium because the frequency of strain j surviving is

$$z_j = \sum_{k=1}^{n} A_k e^{-\mathbf{i}\xi_k \tau}.$$

C.3 Probability of final outcome in two-strain system

In chapter 3, we state that in a two-strain dynamics, there are only four possibility of final outcome including exclusion of strain 1 either strain 2, bistable state and coexistence. In this section, we will compute the probability of exclusion of either strain when the variations in traits are given by uniform distribution. We can make similar proofs for the case normal distribution.

C.3.1 Variations in only co-infection clearance rates γ_{ij}

First, we compute for the case the variation in only coinfection clearance rates $\gamma_{ij} = \gamma(1+\varepsilon u_{ij})$ and the variation u_{ij}'s are given by the uniform distribution, in C.4. The case of normal distribution is similar. The code for this type of simulation can be found at [code].

We consider the case u_{ij}'s are independent variables and given by the uniform distribution in $\left[-\frac{1}{2}, \frac{1}{2}\right]$. The case for normal distribution can be solved similarly. We already have that the outcome is exclusion of either strains if and only if

$$u_{11} \leq \frac{u_{12}+u_{21}}{2} \leq u_{22} \qquad \text{or} \qquad u_{22} \leq \frac{u_{12}+u_{21}}{2} \leq u_{11}.$$

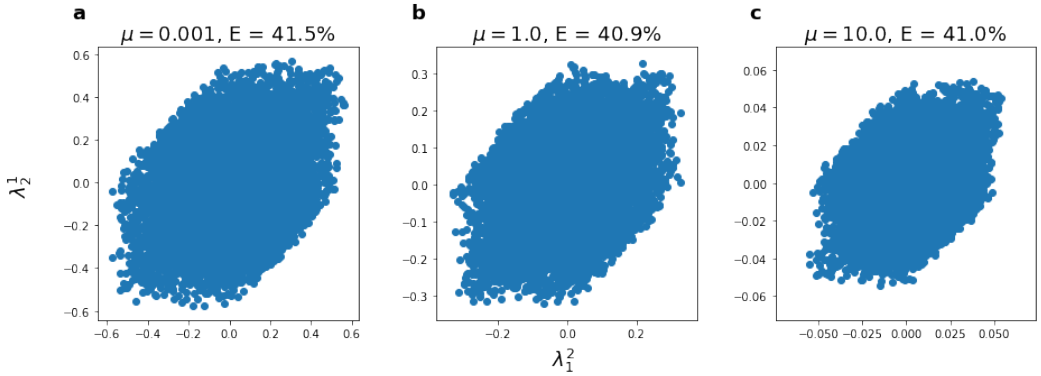


Fig. C.4 **Probability of exclusion of either strains when variation in only coinfection clearance rates γ_{ij} and the variation u_{ij}'s are given by the uniform distribution.** We simulate in three cases of μ varies from small to large. The sample space is taken by 10000 and the neutral parameters are chose to be $\gamma = 1.2$, $r = 0.3$, $k = 3$. The transmission rate depends on value of μ. However, in the chapter 2, we already know that, actually, the final outcome does not depend on the neutral parameters.

For the sake of simplicity, set $u_{11} = a$, $u_{12} = b$, $u_{21} = c$, $u_{22} = d$, we need to compute the probability of the event: $\frac{b+c}{2}$ lies between a and d with a, b, c, d are given by uniform distribution in $\left[-\frac{1}{2}, \frac{1}{2}\right]$. This probability is equal to

$$\mathscr{P} = 1 - P\left(\frac{b+c}{2} < a, \frac{b+c}{2} < d\right) - P\left(\frac{b+c}{2} > a, \frac{b+c}{2} > d\right). = 1 - 2P\left(\frac{b+c}{2} > a, \frac{b+c}{2} > d\right) \tag{C.3.1}$$

Note that $\frac{b+c}{2}$ now is given by Irwin-Hall distribution with $n = 2$. The CDF of $\frac{b+c}{2}$ now is given by:

$$F_{\frac{b+c}{2}}(x) = \begin{cases} 0, & x \leq -\frac{1}{2} \\ 2\left(x + \frac{1}{2}\right)^2, & -\frac{1}{2} < x \leq 0 \\ 1 - 2\left(\frac{1}{2} - x\right)^2, & 0 < x \leq \frac{1}{2} \\ 0, & x > \frac{1}{2} \end{cases} \tag{C.3.2}$$

Denote the CDF of random variables a and d to be F_a and F_d respectively. Then $F_a(x) = F_d(x) = x + \frac{1}{2}$.

We now compute $\mathscr{Q} = P\left(\frac{b+c}{2} > a, \frac{b+c}{2} > d\right)$ which is equal to

$$\begin{aligned} \mathscr{Q} &= \int_{-\infty}^{\infty} F_a(x) F_d(x) F'_{\frac{b+c}{2}}(x)\, dx = \int_{-1/2}^{1/2} F_a(x) F_d(x) F'_{\frac{b+c}{2}}(x)\, dx \\ &= \int_{-1/2}^{0} F_a(x) F_d(x) F'_{\frac{b+c}{2}}(x)\, dx + \int_{0}^{1/2} F_a(x) F_d(x) F'_{\frac{b+c}{2}}(x)\, dx \\ &= \int_{-1/2}^{0} 4\left(x + \frac{1}{2}\right)^3 dx + \int_{0}^{1/2} \left(x + \frac{1}{2}\right)^2 \cdot 4\left(\frac{1}{2} - x\right) dx = \frac{1}{16} + \frac{11}{48} = \frac{7}{24}. \end{aligned} \tag{C.3.3}$$

Then $\mathscr{P} = 1 - 2 \times \dfrac{7}{24} = \dfrac{5}{12}$.

C.3.2 Variations in only co-colonization interaction factor k_{ij}

We compute the probability of exclusion of either strains in a two-strain system, in which there is only variation in co-colonization interaction factors $k_{ij} = k + \varepsilon\alpha_{ij}$, $i, j \in \{1,2\}$ and α_{ij}'s are given by the uniform distribution in $\left[-\frac{1}{2}, \frac{1}{2}\right]$. Recall the infection-to-coinfection factor $\mu = \frac{I^*}{D^*}$, by the formulas of pairwise invasion fitnesses we have that

$$\lambda_1^2 = \Theta_5 \left(\mu \left(\alpha_{21} - \alpha_{12}\right) + \alpha_{21} - \alpha_{22}\right)$$
$$\lambda_1^2 = \Theta_5 \left(\mu \left(\alpha_{12} - \alpha_{21}\right) + \alpha_{12} - \alpha_{11}\right)$$

The exclusion of eithers strains occurs when $\left(\lambda_1^2, \alpha_2^1\right)$ have opposite signs. Thus, it suffices to compute the probability of the event, in which

$$\begin{array}{l} (\mu+1)\,\alpha_{21} - \mu\alpha_{12} < \alpha_{22} \\ (\mu+1)\,\alpha_{12} - \mu\alpha_{21} > \alpha_{11} \end{array}.$$

For the sake of simplicity, this problem can be rewritten as follows.

The conditions are equivalent to

$$\alpha_{12} > \max\left\{\frac{(\mu+1)\alpha_{21}}{\mu} - \frac{\alpha_{22}}{\mu}, \frac{\mu\alpha_{21}}{\mu+1} + \frac{\alpha_{11}}{\mu+1}\right\}. \tag{C.3.4}$$

so the problem is simply to determine the volume of a region in 4-space bounded by hyperplanes. If we make the change of variables

$$W = \alpha_{11} + \frac{1}{2}, \qquad X = \alpha_{21} + \frac{1}{2}, \qquad Y = \alpha_{12} + \frac{1}{2}, \qquad Z = \alpha_{22} + \frac{1}{2}.$$

Then (C.3.4) is equivalent to

$$Y > \mathscr{Y}_{\max} := \max\left\{\frac{(\mu+1)X}{\mu} - \frac{Z}{\mu}, \frac{\mu X}{\mu+1} + \frac{W}{\mu+1}\right\}, \tag{C.3.5}$$

in which W, X, Y, Z are independent random variables, uniformly distributed on $[0,1]$.

By direct computations, we find that

$$\frac{(\mu+1)X}{\mu} - \frac{Z}{\mu} \geq \frac{\mu X}{\mu+1} + \frac{W}{\mu+1} \iff X \geq \frac{\mu}{2\mu+1}W + \frac{\mu+1}{2\mu+1}Z.$$

Since the term $\frac{\mu}{2\mu+1}W + \frac{\mu+1}{2\mu+1}Z$ is a convex combination of W and Z, it always lies in $[0,1]$.

Whichever function gives the maximum, we must distinguish several cases. First, if $\mathscr{Y}_{\max} < 0$, then Y can take any value in $[0,1]$. Second, if $0 \leq \mathscr{Y}_{\max} \leq 1$, then Y can take values between $\mathscr{Y}_{\max}$ and 1. Finally, if $\mathscr{Y}_{\max} > 1$, Y cannot possibly exceed it, so we can ignore this case.

We see that the problem breaks up into several cases. First, we suppose that $X \geq \frac{\mu}{2\mu+1}W + \frac{\mu+1}{2\mu+1}Z$, so $\mathscr{Y}_{\max} = \frac{(\mu+1)X}{\mu} - \frac{Z}{\mu}$. It is straightforward to see that

$$0 \leq \frac{(\mu+1)X}{\mu} - \frac{Z}{\mu} \leq 1 \iff \frac{Z}{\mu+1} \leq X \leq \frac{Z+\mu}{\mu+1}$$

Both the extreme values are non negative numbers not exceeding 1, so that's not a problem, but in order for this case to obtain, we must have $X \geq \frac{\mu}{2\mu+1}W + \frac{\mu+1}{2\mu+1}Z$. In the first place, we need

$$\frac{Z+\mu}{\mu+1} \geq X \geq \frac{\mu}{2\mu+1}W + \frac{\mu+1}{2\mu+1}Z. \tag{C.3.6}$$

For this to hold, we need the left-hand side to be greater than or equal to the right-hand side, which is true if and only if

$$(2\mu+1)(Z+\mu) \geq \mu(\mu+1)W + (\mu+1)^2 Z$$

or

$$2\mu^2 + \mu \geq (\mu^2+\mu)W + \mu^2 Z$$

which is true since $0 \leq W, Z \leq 1$.

So (C.3.6) doesn't need adjustment, but we must have

$$X \geq \max\left\{\frac{Z}{\mu+1}, \frac{\mu}{2\mu+1}W + \frac{\mu+1}{2\mu+1}Z\right\}.$$

However, it's easy to see that $\frac{\mu+1}{2\mu+1} - \frac{1}{\mu+1} = \frac{\mu^2}{(2\mu+1)(\mu+1)}$, so $\frac{Z}{\mu+1} \leq \frac{\mu}{2\mu+1}W + \frac{\mu+1}{2\mu+1}Z$ and (C.3.6) gives the conditions for this case to hold.

We may write the probability associated with this case as

$$\int_0^1\int_0^1\int_{x_0}^{x_1}\int_{y_0}^1 dy\cdot dx\cdot dz\cdot dw \quad \text{where} \quad x_0 = \frac{\mu w}{2\mu+1} + \frac{(\mu+1)z}{2\mu+1}, \qquad x_1 = \frac{z+\mu}{\mu+1}, \qquad y_0 = \frac{(\mu+1)x}{\mu} - \frac{z}{\mu}$$

which becomes

$$\int_0^1\int_0^1\int_{x_0}^{x_1}\int_{y_0}^1 dy\cdot dx\cdot dz\cdot dw = \frac{\mu\left(7\mu^2+7\mu+2\right)}{12\left(4\mu^3+8\mu^2+5\mu+1\right)}.$$

Moreover, $\mathscr{Y}_{\max} < 0$ when $X \leq \frac{Z}{\mu+1}$ but we have just seen that this is impossible in this case, so we can move on to the case where $\mathscr{Y}_{\max} = \frac{\mu X}{\mu+1} + \frac{W}{\mu+1}$ in (C.3.6).

In this case, $\mathscr{Y}_{\max}$ is a convex combination of W and Z, so it is always between 0 and 1, and there are no adjustments needed. We may write the probability associated with this case as

$$\int_0^1\int_0^1\int_0^{x_0}\int_{y_0}^1 dy\cdot dx\cdot dz\cdot dw \quad \text{where} \quad x_0 = \frac{\mu w}{2\mu+1} + \frac{(\mu+1)z}{2\mu+1}, \qquad y_0 = \frac{\mu x}{\mu+1} + \frac{w}{\mu+1}$$

which leads to

$$\int_0^1\int_0^1\int_0^{x_0}\int_{y_0}^1 dy\cdot dx\cdot dz\cdot dw = \frac{17\mu^3+27\mu^2+15\mu+3}{12\left(4\mu^3+8\mu^2+5\mu+1\right)}.$$

The sum of two integral, which is the probability of (C.3.4), which is

$$\frac{12\mu^2+11\mu+3}{12\left(2\mu^2+3\mu+1\right)}.$$

Hence, the probability of the exclusion of either strains is

$$\frac{12\mu^2+11\mu+3}{6\left(2\mu^2+3\mu+1\right)}$$

which makes sense with our simulation as follows.

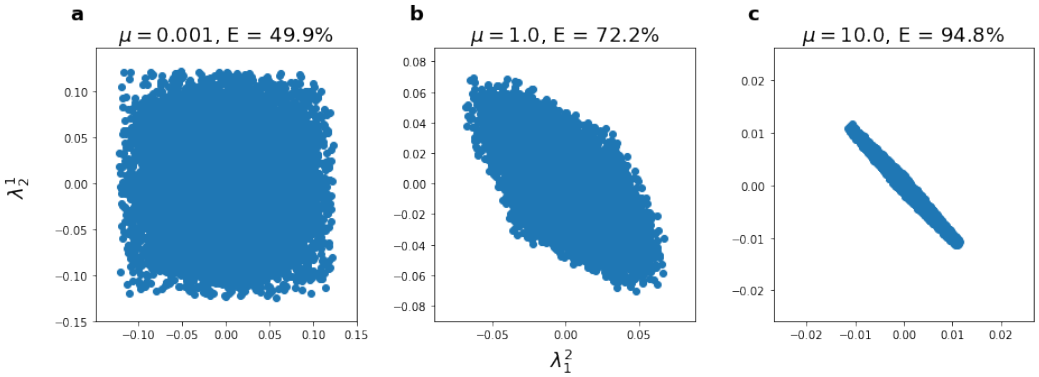


Fig. C.5 **Probability of exclusion of either strains when variation in only co-colonization interaction factors k_{ij} and the variation α_{ij}'s are given by the uniform distribution.** We simulate in three cases of μ varies from small to large. The sample space is taken by 10000 and the neutral parameters are chose to be $\gamma = 1.2$, $r = 0.3$, $k = 3$. The transmission rate depends on value of μ. However, in the chapter 2, we already know that, actually, the final outcome depends on the neutral parameters only through μ, as in [87]. When $\mu \ll 1$, the probability is near 0.5, and it rises as μ increases. When $\mu = 10$ which is large, the probability of exclusion of either strains becomes near 94.8%.

C.4 Dependence of system on infection-to-coinfection ratio μ

We investigate the relation between the infection-to-coinfection ratio $\mu = \frac{1}{k(R_0-1)}$ and each pair of strains in a whole system of $N = 150$ strains. In the following simulation, we consider each pair of strains (i, j), for $1 \leq i, j \leq N$, which has $\frac{150\times149}{2}$ such pairs. Recall that the notations b_i, ν_i, u_{ij}, ω_{ij}^s, α_{ij} in Table 2.1 are the perturbation parts of transmission rates β_i, infection clearance rates γ_i, coinfection clearance rates γ_{ij}, transmission capacity of a strain from a co-colonized host p_{ij}^s and relative factor of altered susceptibility to co-colonization k_{ij}, respectively. We consider the case when variation only in co-colonization interaction $k_{ij} = k + \varepsilon\alpha_{ij}$. α_{ij}'s are given by the uniform distribution in $\left[-\frac{1}{2}, \frac{1}{2}\right]$.

We can see that when $\mu = 0.001$ is small enough, the probability of the exclusion of either strains is near 50%. When $\mu = 1$, this probability increases to 72.4%. When $\mu = 10$ is large enough, all the points in sample seem fit to the line $\lambda_i^j + \lambda_j^i = 0$.

We can see that, clouds in three cases of Figure C.6 look similar when we simulate independent pairs of pairwise invasion fitness in Figure C.5. However, in general, the pairs $\left(\lambda_j^i, \lambda_i^j\right)$ in two figures actually are not in the same distribution. Indeed, we show as following an example for this claim.

We consider a system of 3 strains with perturbation in only co-colonization $k_{ij} = k + \varepsilon\alpha_{ij}$. We choose the variations α_{ij} satisfying $\alpha_{ij} = 0$ for all $i \neq j$. Recalling the formula for pairwise invasion fitness in chapter 2, now we have $\lambda_i^j = -\Theta_5\alpha_{jj}$, which leads to the independence between $\left(\lambda_j^i, \lambda_i^j\right)$. Without loss of generality, by normalization, we assume that the pairwise invasion fitness λ_i^j's are calculated to be belong to distribution A in $[-1, 1]$. We have two cases, corresponding two types of simulation in Figures C.5 and C.6:

1. We create three pairs of pairwise invasion fitnesses $\left(\lambda_i^j, \lambda_i^i\right)$ independently, in which $\left(\lambda_i^j, \lambda_i^i\right)$'s are given by the distribution A.

2. We create a vector (X_1, X_2, X_3) and three pairs $\left(\lambda_i^j, \lambda_i^i\right)$ are created by $\lambda_i^j \in \{X_1, X_2, X_3\}$, $i, j \in \{1, 2, 3\}$, $i < j$.

In the second case, there always exists two points in the cloud have the same first coordinates, i.e. the probability of this event is 1.

However, in the first case, the probability of the event "there are exactly two points having the same first coordinates" is less than 1. This is trivial when A is a continuous distribution. When A is a concrete distribution, this holds as well. Indeed, assume that A is a distribution in which the probability of k is p_k, $k \in \mathbb{N}$, $1 \leq k \leq n$ and $p_1 + p_2 + \cdots + p_n = 1$. Then, we have that

$$P(X_1 = X_2 \neq X_3) = \sum_{i \neq j} p_i^2 p_j = \sum_{i=1}^{n} p_i^2 (1 - p_i).$$

Alternatively, we have that $p_i \left(p_i - \frac{1}{2}\right)^2 \geq 0$ which implies $p_i^2 (1 - p_i) \leq \frac{1}{4} p_i$. Hence, $\sum_{i=1}^{n} p_i^2 (1 - p_i) \leq \frac{1}{4}$. Accordingly, the probability of the event "there are exactly two points have the same first coordinates" in this first case, which is $3 \times P(X_1 = X_2 \neq X_3)$, does not exceed $\frac{3}{4}$.

Therefore, the pairs $\left(\lambda_j^i, \lambda_i^j\right)$ in two figures C.5 and C.6 are not in the same distribution in general.

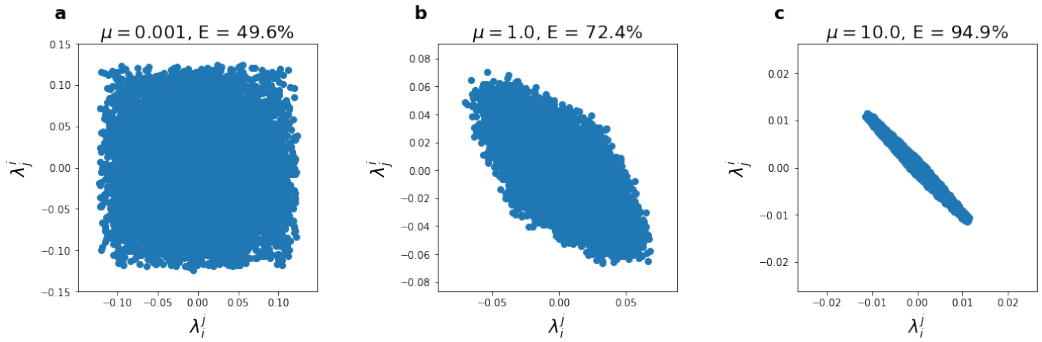


Fig. C.6 **Invasion fitness of each pairs in a 150-strain system.** Each dot in figures represents the value of $\left(\lambda_i^j, \lambda_j^i\right)$ for $1 \leq i < j \leq N$. Perturbations are in co-colonization interaction k_{ij}. We plot for three cases when infection-to-coinfection ratio μ equals to 0.001, 1 and 10. We choose the clearance rate $\gamma = 1.2$, the mortality as well as the recruitment rate $r = 0.3$ and the co-colonization interaction $k = 3$. When μ small enough, it can be seen that four scenarios have similar probability. The exclusion of either strains is the major phenomenon even with small $\mu = 1$. When μ becomes larger, this behavior appears apparently. When $\mu = 10$, all the points seem lying on the line $\lambda_i^j + \lambda_j^i = 0$. (Code).

As mentioned before, we can see there is similarity between Figures C.6 and C.5. Indeed, for $N = 150$, we can divide these 150 strains into 75 pairs of "two strains", which leads to 75 distinct pairs of invasion fitnesses $\left(\lambda_i^j, \lambda_j^i\right)$, i.e. each strain occurs uniquely in some pair. Hence, in each way of dividing, all $\left(\lambda_i^j, \lambda_j^i\right)$ are independent and taken from the same distribution, which is also the distribution for $\left(\lambda_i^j, \lambda_j^i\right)$ in previous subsection C.3.2. On the other side, there are $\binom{150}{2}$ dots, implying that there are 149 sets of "75 pairs". This leads to the clouds in Figure C.6 seem to look like in Figure C.5.

Appendix D

A special case of intermediate diffusion

D.1 The general, quasi-neutral and neutral system

D.1.1 The general and neutral systems

The dynamics studied in this article groups the pathogen types in N subsets, indexed by i, $1 \leq i \leq N$. With a set of ordinary differential equations, we then track the proportion of hosts in $1+N+N^2$ compartments: susceptible: $S(x,t)$, hosts colonized by strain-i: $I_i(x,t)$, hosts co-colonized by strain-i then strain-j: $I_{ij}(x,t)$. Notice that we include also same strain coinfection, as argued in [120].

We formulate the general model based on the same structure as in [120] but here allow for strains to vary in their transmission rates $\beta_i(x)$, clearance rates of single infection $\gamma_i(x)$ (or duration of carriage $1/\gamma_i(x)$), clearance rates from mixed co-colonization $\gamma_{ij}(x)$, within-host competition reflected in relative transmissibilities from mixed coinfected hosts ($p^i_{ij}(x)$ and $p^i_{ji}(x)$), as well as co-colonization vulnerabilities $k_{ij}(x)$, already studied in [120]. In a compact domain $\Omega \subset \mathbb{R}^n$ with smooth boundary Γ, we consider the general SIS dynamics in a coinfected system with diffusion as follows

$$\begin{cases} \dfrac{\partial S}{\partial t} &= r(1-S) + \displaystyle\sum_{i=1}^{N} \gamma_i I_i + \sum_{i,j=1}^{N} \gamma_{ij} I_{ij} - S \sum_{i=1}^{N} \beta_i J_i + \quad d\Delta S, \\ \dfrac{\partial I_i}{\partial t} &= \beta_i J_i S - (r+\gamma_i) I_i - I_i \displaystyle\sum_{j=1}^{N} k_{ij} \beta_j J_j + \quad d\Delta I_i, \qquad 1 \leq i \leq N, \\ \dfrac{\partial I_{ij}}{\partial t} &= k_{ij} I_i \beta_j J_j - (r+\gamma_{ij}) I_{ij} + \quad d\Delta I_{ij}, \qquad 1 \leq i,j \leq N, \end{cases} \tag{D.1.1}$$

where J_i is proportion of all hosts transmitting strain i, including singly- and co-colonized hosts and has the explicit formula

$$J_i = I_i + \sum_{j=1}^{N} \left(p^i_{ij} I_{ij} + p^i_{ji} I_{ji} \right)$$

with the Neumann boundary conditions $\partial_n S = \partial_n I_i = \partial_n I_{ij} = 0$ on the boundary Γ of Ω and the given initial values $S(0)$, $I_i(0)$ and $I_{ij}(0)$ for all $1 \leq i,j \leq N$.

Note that $\beta_i J_i$ is the infection force of strain i, for all i. Through out this Chapter, in (D.1.1), for $1 \leq i,j \leq N$, we assume that all parameters are given in the following Table D.1.

Table D.1 Conventions and notations of parameters for system with intermediate diffusion

	Parameter	Interpretation	Under strain similarities
1.	$\beta_i(x)$	Strain-specific transmission rates	$\beta_i = \beta(1+\varepsilon b_i(x))$
2.	$\gamma_i(x)$	Strain-specific clearance rates of single colonization	$\gamma_i = \gamma(1+\varepsilon \nu_i(x))$
3.	$\gamma_{ij}(x)$	Clearance rates of co-colonization with i and j	$\gamma_{ij} = \gamma(1+\varepsilon u_{ij}(x))$
4.	$p^s_{ij}(x)$	Transmission capacity of the strain $s \in \{i,j\}$ by a host co-colonized by strain-i then strain-j, $\left(p^i_{ij}(x)+p^j_{ij}(x)=1\right)$	$p^s_{ij}(x) = \frac{1}{2}+\varepsilon\omega^s_{ij}(x)$
5.	$k_{ij}(x)$	Relative factor of altered susceptibility to co-colonization between colonizing strain i and co-colonizing strain j	$k_{ij}(x) = k+\varepsilon\alpha_{ij}(x)$
	r	Susceptible recruitment rate (Equal to natural mortality)	

For the sake of simplicity, we denote the inverse duration of a carriage episode by strain i with $m_i = r+\gamma_i$, of a co-carriage episode by strains i and j with $m_{ij} = r+\gamma_{ij}$ and the corresponding inverse duration of carriage if all strains were equivalent with $m = r+\gamma$.

For the sake of applying the same frame work of [120], we make the following assumptions.

Assumption D.1. (Assumptions on the parameters) All the values of neutral parameters does not depend on space. In another word, mean transmission rate β, mean clearance rate γ, mortality as well as recruitment rate r, and mean co-colonization interaction factor k are all constants.

Note that, the system (D.1.1) has unique positive solution. By this, we wish to analyze the asymptotic behavior of the dynamic of (D.1.1) when $\varepsilon \to 0$.

In this section, we consider the operator Δ with Neumann condition, acting on $C^0\left(\bar{\Omega}\right)$.

• **Firstly**, we consider the total mass $S+T$. Sum up all of the equations of (D.1.1) we have that

$$\begin{cases} \dfrac{\partial}{\partial t}\left(S+\sum\limits_{i=1}^{N} I_i+\sum\limits_{i,j=1}^{N} I_{ij}\right) = r(x)-r(x)\left(S+\sum\limits_{i=1}^{N} I_i+\sum\limits_{i,j=1}^{N} I_{ij}\right)+d\Delta\left(S+\sum\limits_{i=1}^{N} I_i+\sum\limits_{i,j=1}^{N} I_{ij}\right), \\ \dfrac{\partial}{\partial n}\left(S+\sum\limits_{i=1}^{N} I_i+\sum\limits_{i,j=1}^{N} I_{ij}\right) = 0 \quad \text{on} \quad \partial\Omega. \end{cases} \tag{D.1.2}$$

which can be rewritten into the form

$$\frac{\partial}{\partial t}(S+T) = r(x)[1-(S+T)]+d\Delta(S+T). \tag{D.1.3}$$

with the Neumann boundary condition. The assumed smoothness of $\partial\Omega$ implies that $d\Delta$ generates a C^0 semi group of contraction on $C^0\left(\bar{\Omega}\right)$, see [27].
Note that $S+T=1$ are the solution of $r[1-(S+T)]+d\Delta(S+T)=0$ and the linearized operator becomes $d\Delta-r$ which has spectrum lies in the left-half plane (since the Laplacian has the strictly negative spectrum and $r>0$). By the Theorem 11.20 in [189], we deduce that $S+T=1$ is asymptotically stable.
Hence, we assume the total mass $S+T=1$ for the later.

• **Secondly,** take $\varepsilon = 0$ in (D.1.1), we have the neutral system for S,T,I as follows

$$\begin{cases} \dfrac{\partial S}{\partial t} &= m(1-S) - \beta ST &+d\Delta S, \\ \dfrac{\partial T}{\partial t} &= \beta ST - mT &+d\Delta T, \\ \dfrac{\partial I}{\partial t} &= \beta TS - (m+k\beta T)I &+d\Delta I. \\ \dfrac{\partial S}{\partial n} &= \dfrac{\partial T}{\partial n} = \dfrac{\partial I}{\partial n} = 0, \quad \text{on} \quad \partial\Omega. \end{cases} \tag{D.1.4}$$

For later reference, we denote

$$\boxed{S^* = \frac{m}{\beta}, \quad T^* = 1 - S^*, \quad I^* = \frac{mT^*}{m+\beta kT^*}, \quad D^* = T^* - I^*} \tag{D.1.5}$$

Note that (S^*,T^*,I^*) satisfy the following equations

$$\begin{cases} r(1-S) - \beta ST &+d\Delta S = 0 \\ -mT + \beta ST &+d\Delta T = 0 \\ \beta TS - (m+\beta kT)I &+d\Delta I = 0. \end{cases} \tag{D.1.6}$$

and the linearized operator in the right-hand side of (D.1.4) at (S^*,T^*,I^*) becomes

$$\begin{pmatrix} -m - \beta T^* + d\Delta & -\beta S^* & 0 \\ \beta T^* & \beta S^* - m + d\Delta & 0 \\ \beta T^* & \beta S^* - \beta kI^* & -(m+k\beta T^*) + d\Delta \end{pmatrix}$$

which has spectrum lies in the left-half plane since the Laplacian has the strictly negative spectrum. According to Theorem 11.20 again in [189], (S^*,T^*,I^*) is asymptotically stable.

• **Thirdly,** Setting the operator

$$A = \begin{pmatrix} -(m+\beta kT^*) & m \\ -\frac{\beta kT^*}{2} & \frac{\beta kI^*}{2} \end{pmatrix} \tag{D.1.7}$$

then

$$A = P\begin{pmatrix} -\xi & 0 \\ 0 & 0 \end{pmatrix}P^{-1}, \qquad \xi = m + \frac{1}{2}\beta kT^* - \frac{1}{2}\beta kI^* > 0, \tag{D.1.8}$$

with

$$P = \begin{pmatrix} 2T^* & I^* \\ D^* & T^* \end{pmatrix}, \qquad \text{and} \quad P^{-1} = \frac{1}{|P|}\begin{pmatrix} T^* & -I^* \\ -D^* & 2T^* \end{pmatrix}, \qquad |P| = 2T^{*2} - I^*D^* > 0. \tag{D.1.9}$$

Now we have the neutral system for $\begin{pmatrix} I_i & J_i \end{pmatrix}$

$$\frac{\partial}{\partial t}\begin{pmatrix} I_i \\ J_i \end{pmatrix} = A\begin{pmatrix} I_i \\ J_i \end{pmatrix} + d\begin{pmatrix} \Delta I_i \\ \Delta J_i \end{pmatrix}$$

which gives

$$\begin{cases} \dfrac{\partial v_i}{\partial t} = -\xi v_i + O(\varepsilon) + d\Delta v_i \\ \dfrac{\partial z_i}{\partial t} = O(\varepsilon) + d\Delta z_i \end{cases} \tag{D.1.10}$$

by the variable change $\begin{pmatrix} v_i \\ z_i \end{pmatrix} = P^{-1} \begin{pmatrix} I_i \\ J_i \end{pmatrix}$. Using the Theorem 11.20 in [189] again, we have that $v_i \to 0$ as $t \to \infty$ for all $x \in \Omega$, $1 \leq i \leq N$.

We define the ansatz as follows

$$\boxed{S(x,t) = S^*(x) + \varepsilon X(x,t) + O\left(\varepsilon^2\right), \quad T(x,t) = T^*(x) - \varepsilon X(x,t) + O\left(\varepsilon^2\right) \quad I(x,t) = I^*(x) + \varepsilon Y(x,t) + O\left(\varepsilon^2\right)}. \tag{D.1.11}$$

and we need to write the system for (X,Y) in later sections.

D.1.2 The slow-fast form

Initially, we define that for $i = 1, \cdots, N$:

$$L_i(x,t) = \frac{1}{2} \sum_{j=1}^{N} \left(u_{ij} I_{ij}(x,t) + u_{ji} I_{ji}(x,t) \right). \tag{D.1.12}$$

Similarly to [120], we use the following notations in system (D.1.1).

$$\begin{aligned} &\beta_i = \beta\left(1 + \chi_1 \varepsilon b_i\right); \qquad \gamma_i = \gamma(1 + \chi_2 \varepsilon \nu_i); \qquad \gamma_{ij} = \gamma(1 + \chi_3 \varepsilon u_{ij}); \\ &p_{ij}^s = \frac{1}{2} + \chi_4 \varepsilon \omega_{ij}^s \quad s \in \{i, j\} \quad \left(\omega_{ij}^i + \omega_{ij}^j = 0\right); \qquad k_{ij} = k + \chi_5 \varepsilon \alpha_{ij}; \end{aligned} \tag{D.1.13}$$

where $\chi_d \in \{0,1\}$ for $d = 1,2,3,4,5$.

Any combination of trait variation among strains, can be captured via $\mathscr{A}$ where $\mathscr{A}$ is a subset of $\{1,2,3,4,5\}$ denoting the absence/presence of perturbations in that parameter among strains: for some fixed initial values given, let $C_{\mathscr{A}}$ be the system (D.1.1) with $\chi_d = 1$ if $d \in \mathscr{A}$ and $\chi_d = 0$ if $d \notin \mathscr{A}$. For simplicity, we note also $C_{\{d\}}$ by C_d for $d \in \{1,2,3,4,5\}$.

If $\mathscr{A} = \emptyset$ then there is no perturbation and the system $C_\emptyset$ is exactly the *neutral* model deduced from (D.1.1) when $\varepsilon = 0$. With these notations, $C_{\mathscr{A}}$ reads

$$\begin{cases} \dfrac{\partial X}{\partial t} = -\beta T^* X + \chi_1 \beta S^* \displaystyle\sum_{i=1}^{N} b_i J_i - \chi_2 \gamma \sum_{i=1}^{N} \nu_i I_i - \chi_3 \gamma \sum_{i=1}^{N} L_i + d\Delta X + O(\varepsilon) \\ \dfrac{\partial Y}{\partial t} = \beta(S^* - T^* - kI^*)X - (m + \beta k T^*)Y + \chi_1 \beta (S^* - kI^*) \displaystyle\sum_{i=1}^{N} b_i J_i - \chi_2 \gamma \sum_{i=1}^{N} \nu_i I_i - \chi_5 \beta \sum_{i,j=1}^{N} \alpha_{ij} I_i J_j + d\Delta Y + O(\varepsilon) \\ \dfrac{\partial L_i}{\partial t} = -mL_i + \chi_3 \dfrac{1}{2} \beta \gamma k I_i \displaystyle\sum_{j=1}^{N} u_{ij} J_j + \chi_3 \frac{1}{2} \gamma \beta k J_i \sum_{j=1}^{N} u_{ji} I_j + d\Delta L_i + O(\varepsilon) \end{cases} \tag{D.1.14}$$

together with (we omit terms of $O\left(\varepsilon^2\right)$)

$$\frac{\partial}{\partial t}\begin{pmatrix} I_i \\ J_i \end{pmatrix} = \begin{pmatrix} -(m+\beta kT^*) & m \\ -\frac{\beta kT^*}{2} & \frac{\beta kI^*}{2} \end{pmatrix}\begin{pmatrix} I_i \\ J_i \end{pmatrix} + d\begin{pmatrix} \Delta I_i \\ \Delta J_i \end{pmatrix} - \varepsilon\left[\beta\begin{pmatrix} k & 1 \\ \frac{k}{2} & 1 \end{pmatrix}\begin{pmatrix} I_i \\ J_i \end{pmatrix} X + \frac{\beta k}{2}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} I_i \\ J_i \end{pmatrix} Y + \mathscr{M}_{\mathscr{A}}\begin{pmatrix} I_i \\ J_i \end{pmatrix} - \chi_3\gamma\begin{pmatrix} 0 \\ L_i \end{pmatrix}\right] \tag{D.1.15}$$

where $\mathscr{M}_{\mathscr{A}}$ is the matrix

$$\begin{pmatrix} -\chi_1\beta k\sum\limits_{i=1}^{N} b_iJ_i - \chi_2\gamma\nu_i - \chi_5\beta\sum\limits_{j=1}^{N}\alpha_{ij}J_j & \chi_1\beta b_iS^* \\ \beta\sum\limits_{j=1}^{N}\left(\chi_4 k\omega_{ij}^i - \chi_5\frac{\alpha_{ij}}{2}\right)J_j - \chi_1\frac{\beta k}{2}\sum\limits_{i=1}^{N} b_iJ_i - \chi_2\gamma\nu_i & \chi_1\beta b_i\left(S^* + \frac{kI^*}{2}\right) + \beta\sum\limits_{j=1}^{N}\left(\chi_4 k\omega_{ji}^i + \chi_5\frac{\alpha_{ji}}{2}\right)I_j \end{pmatrix} \tag{D.1.16}$$

Now, we rewrite system $C_{\mathscr{A}}$ using the variables $\begin{pmatrix} v_i \\ z_i \end{pmatrix}$. Let us note

$$\mathbf{L} = (L_i)_i, \qquad \mathbf{v} = (v_i)_i, \qquad \mathbf{z} = (z_i)_i.$$

The system $C_{\mathscr{A}}$ reads now as the slow-fast form

$$\begin{cases} \dfrac{\partial X}{\partial t} &= -\beta T^*X + \chi_1F_X^1(\mathbf{v},\mathbf{z}) + \chi_2F_X^2(\mathbf{v},\mathbf{z}) + \chi_3F_X^3(\mathbf{L}) + d\Delta X + O(\varepsilon) \\ \dfrac{\partial Y}{\partial t} &= \beta(S^* - T^* - kI^*)X - (m+\beta kT^*)Y + \chi_1F_Y^1(\mathbf{v},\mathbf{z}) + \chi_2F_Y^2(\mathbf{v},\mathbf{z}) + \chi_5F_Y^5(\mathbf{v},\mathbf{z}) + d\Delta Y + O(\varepsilon) \\ \dfrac{\partial L_i}{\partial t} &= -mL_i + \chi_3F_{L_i}(\mathbf{v},\mathbf{z}) + O(\varepsilon) + d\Delta L_i \\ \dfrac{\partial v_i}{\partial t} &= = -\xi v_i + O(\varepsilon) + d\Delta v_i \\ \dfrac{\partial z_i}{\partial t} &= \varepsilon\left(F_{z_i}(X,Y,\mathbf{L},\mathbf{v},\mathbf{z}) + O(\varepsilon)\right) + d\Delta z_i \end{cases} \tag{D.1.17}$$

which leads to $v_i \to 0$ as $t \to \infty$ and $\varepsilon \to 0$.

Setting $\tau = \varepsilon t$ and change (D.1.17) to the new time-scale, see [120]

$$\begin{cases} \dfrac{\partial X}{\partial \tau} &= \dfrac{1}{\varepsilon}\left[-\beta T^*X + \chi_1F_X^1(\mathbf{v},\mathbf{z}) + \chi_2F_X^2(\mathbf{v},\mathbf{z}) + \chi_3F_X^3(\mathbf{L}) + d\Delta X\right] + O(1) \\ \dfrac{\partial Y}{\partial \tau} &= \dfrac{1}{\varepsilon}\left[\beta(S^* - T^* - kI^*)X - (m+\beta kT^*)Y + \chi_1F_Y^1(\mathbf{v},\mathbf{z}) + \chi_2F_Y^2(\mathbf{v},\mathbf{z}) + \chi_5F_Y^5(\mathbf{v},\mathbf{z}) + d\Delta Y\right] + O(1) \\ \dfrac{\partial L_i}{\partial \tau} &= \dfrac{1}{\varepsilon}\left[-mL_i + \chi_3F_{L_i}(\mathbf{v},\mathbf{z}) + O(\varepsilon) + d\Delta L_i\right] \\ \dfrac{\partial v_i}{\partial \tau} &= -\dfrac{\xi}{\varepsilon}v_i + O(1) + \dfrac{d}{\varepsilon}\Delta v_i \\ \dfrac{\partial z_i}{\partial \tau} &= F_{z_i}(X,Y,\mathbf{L},\mathbf{v},\mathbf{z}) + O(\varepsilon) + \dfrac{d}{\varepsilon}\Delta z_i \end{cases} \tag{D.1.18}$$

Set $U = \begin{pmatrix} X & Y & \mathbf{L} & \mathbf{v} & \mathbf{z} \end{pmatrix}^T$ be the solution of (D.1.18), then we rewrite (D.1.18) in the form as follows:

$$\begin{cases} \dfrac{\partial}{\partial t}\mathbf{W}(x,t) = \mathscr{F}\left(\mathbf{W}(x,t)\right) + \dfrac{1}{\varepsilon}K\mathbf{W}(x,t), & t>0 \quad \& \quad x \in \Omega \\ \partial_n \mathbf{W}(x,t) = 0, & t>0 \quad \& \quad x \in \partial\Omega \\ \mathbf{W}(x,0) = \left(X^0(x), Y^0(x), z_1^0(x), \cdots, z_N^0(x)\right), & x \in \Omega. \end{cases} \tag{D.1.19}$$

with $K = diag\left(\Delta\right)$ acting on $\left(C^0\left(\bar{\Omega}\right)\right)^{3N+2}$. We need to find the null space $\ker\left(K\right)$. First, as in [43], define

$$T = d\Delta. \tag{D.1.20}$$

When seen as an operator on $L^2\left(\Omega\right)$, the operator T^2 with the formula in (D.1.20) with homogeneous Neumann boundary conditions is defined as, see [43]

$$\begin{aligned} &D\left(T^2\right) = \left\{ U \in H^1\left(\Omega\right) : \exists V \in L^2\left(\Omega\right), \forall \phi \in H^1\left(\Omega\right), \int \nabla U\left(x\right)\nabla\phi\left(x\right)dx = -d\int V\left(x\right)\phi\left(x\right)dx \right\}, \\ &T^2 U := V, \quad U \in D\left(A^2\right). \end{aligned} \tag{D.1.21}$$

In order to obtain uniform estimates, we prefer to focus on the operator $T^\infty := T$ acting on $C^0(\bar{\Omega})$ with sup norm. Hence, we define

$$\begin{aligned} &D\left(T^\infty\right) := \left\{ U \in D\left(T^2\right) \cap C\left(\bar{\Omega}\right), T^2 U \in C\left(\bar{\Omega}\right) \right\}, \\ &T^\infty U = T^2 U, \quad U \in D\left(T^\infty\right). \end{aligned} \tag{D.1.22}$$

Then we have that

$$E_0 := \ker\left(T^\infty\right) = \text{span}(1) = \mathbb{R} \text{ and } \text{Im}\left(T^\infty\right) \subset \left\{ U \in C^0(\bar{\Omega}), \int_\Omega U = 0 \right\} = F_0. \tag{D.1.23}$$

One gets $C^0\left(\bar{\Omega}\right) = \ker T^\infty \oplus \text{Im}\, T^\infty$. Now we define the Banach space $\left(C^0\left(\bar{\Omega}\right)\right)^{3N+2}$ together with the norm

$$\left\|(U_1, U_2)\right\|_\infty = \left\|U_1\right\|_\infty + \left\|U_2\right\|_\infty + \cdots + \left\|U_{3N+2}\right\|_\infty \tag{D.1.24}$$

and the operator $(T^\infty)^{3N+2}$ acting on each coordinate of $\left(C^0\left(\bar{\Omega}\right)\right)^{3N+2}$. The kernel and the range of this operator are respectively

$$E := \ker\left((T^\infty)^{3N+2}\right) = \mathbb{R}^{3N+2} \text{ and } F := \text{Im}\left((T^\infty)^{3N+2}\right). \tag{D.1.25}$$

Hence we had $\left(C^0\left(\bar{\Omega}\right)\right)^{3N+2} = E \oplus F$. The projection of $\left(C^0\left(\bar{\Omega}\right)\right)^{3N+2}$ on E and F, denoted by Π_E and Π_F respectively, given explicit by

$$\Pi_E\left(V_1, V_2, \ldots, V_{3N+2}\right) = \frac{1}{|\Omega|}\left(\int_\Omega V_1, \int_\Omega V_2, \ldots, \int_\Omega V_{3N+2}\right); \qquad \Pi_F = \text{Id} - \Pi_E. \tag{D.1.26}$$

We have that $E = \ker K$ and $F = \operatorname{Im} K$. Hence, we can find $\mathbf{W} = \mathbf{W_1} + \mathbf{W_2}$, with

$$\begin{aligned} \mathbf{W_1} &= \begin{pmatrix} \bar{X} & \bar{Y} & \bar{z_1} & \dots & \bar{z_N} & \bar{v_1} & \dots & \bar{v_N} \end{pmatrix}^T \in E, \\ \mathbf{W_2} &= \begin{pmatrix} \tilde{X} & \tilde{Y} & \tilde{z_1} & \dots & \tilde{z_N} & \tilde{v_1} & \dots & \tilde{v_N} \end{pmatrix}^T \in F. \end{aligned} \tag{D.1.27}$$

in which, the slow variable is $\mathbf{W_1} := \Pi_E(\mathbf{W}) \in E$ and the fast variable is $\mathbf{W_2} := \Pi_F \mathbf{W} = \mathbf{W} - \mathbf{W_1} \in F$. Projecting the system (D.1.18) on E and F yields to the equivalent system

$$\begin{cases} \dfrac{d}{dt}\mathbf{W_1}(t) = f(\mathbf{W_1}, \mathbf{W_2}) \\ \dfrac{d}{dt}\mathbf{W_2}(t) = g(\mathbf{W_1}, \mathbf{W_2}) + \dfrac{1}{\varepsilon} K \mathbf{W_2} \\ \dfrac{\partial}{\partial n}\mathbf{W_1} = 0 \\ \dfrac{\partial}{\partial n}\mathbf{W_2} = 0 \\ \mathbf{W_1}(0) = \Pi_E(\mathbf{W}(0)) \\ \mathbf{W_2}(0) = \Pi_F(\mathbf{W}(0)) \end{cases} \tag{D.1.28}$$

where $f(\mathbf{W_1}, \mathbf{W_2}) = \Pi_E \mathscr{F}(\mathbf{W_1} + \mathbf{W_2})$ and $g(\mathbf{W_1}, \mathbf{W_2}) = \Pi_E \mathscr{F}(\mathbf{W_1} + \mathbf{W_2}) - f(\mathbf{W_1}, \mathbf{W_2})$.

D.2 Applications of Central Manifold Theorem and main results

To apply the Central Manifold Theorem and related results, we need that the operator K define a C^0 semi-group of contraction on F. Note that, the assumed smoothness of $\partial\Omega$ implies that the operator T^∞ generates a C^0 semi group of contraction on $\left(C\left(\bar{\Omega}\right)\right)^2$, see [27]. Denoting $\exp(tT_2^\infty)$ this semi-group, we deduce that

$$\forall t \geq 0, \qquad \|\exp(tT^\infty) v\|_\infty \leq \|v\|_\infty. \tag{D.2.1}$$

Lemma D.2. The restriction of $\tilde{T}$ of T^∞ to the subspace $F_0 = \left\{u \in C^0\left(\bar{\Omega}\right) : \int_\Omega u = 0\right\}$ is the generator of a C^0 semi-group of strict contraction $\exp\left(t\tilde{T}\right)$ on F_0 verifying for some $\mu > 0$

$$\forall v \in F_0, \qquad \left\|\exp\left(t\tilde{T}\right) v\right\|_\infty \leq e^{-\mu t} \|v\|_\infty. \tag{D.2.2}$$

Proof. F_0 is closed in $C^0\left(\bar{\Omega}\right)$ and is clearly invariant under $\exp(tT^\infty)$ by its definition. It follows (from [169] p. 123) that $\tilde{T}$ is the generator of a C^0 semi-group of contraction on F_0.

On the other side, it is well known that the the Laplacian operator on $C^0\left(\bar{\Omega}\right)$ has the discrete spectrum $\sigma(T)$ which totally lies in the negative half line. Since $\sigma\left(\tilde{T}\right) \subset \sigma(T^\infty)$ and $0 \notin \sigma\left(\tilde{T}\right)$, one has that $\sigma\left(\tilde{T}_2\right) \subset (-\infty, -\lambda_1]$ (for some $\lambda_1 > 0$). Apply the Theorem 4.3 (p.118) in [169], we have the conclusion of the lemma. □

We have the following result.

Proposition D.3. K is the generator of a C^0 semi group $\exp(tK)$ on F verifying

$$\|\exp(tK) v\|_F \leq e^{-\mu t} \|v\|_F. \tag{D.2.3}$$

Now, we need to show that the function $f = \Pi_E \mathscr{F}$ and $g = \Pi_F \mathscr{F}$ are smooth enough. By the same arguments of Lemma 4.3 in [43] and note that $\mathscr{F}$ is the vector-valued function whose each component is a multi-variable polynomial. This result can be stated as follows.

Lemma D.4. The function f and g have C^1 smoothness when acting on $E \times F$.

Apply the Central Manifold Theorem in [43], we have that, there is an $\varepsilon_0 > 0$ such that for any $\varepsilon < \varepsilon_0$, a central manifold $C_\varepsilon = \{(x, h(x,\varepsilon)); x \in \Omega, t \in [0,T)\}$ is invariant under the flow generated by the system (D.1.18) and $\|h(\mathbf{W_1},\varepsilon)\|_\infty = O(\varepsilon)$ as $\varepsilon \to 0$. The function $h(x,\varepsilon)$ satisfies the partial differential equation

$$D_{\mathbf{W_1}} h(\mathbf{W_1},\varepsilon) f(\mathbf{W_1}, h(\mathbf{W_1},\varepsilon),\varepsilon) = \frac{1}{\varepsilon} \mathrm{K} h(\mathbf{W_1},\varepsilon) + g(\mathbf{W_1}, h(\mathbf{W_1},\varepsilon),\varepsilon). \tag{D.2.4}$$

For all $u \in C^0(\bar{\Omega})$, denote $\bar{u} := \Pi_{E_0}(u) = \frac{1}{|\Omega|}\int_\Omega u$ with E_0 defined in (D.1.23).

By [43], the solution of (D.1.18) can be approximated by the solution of $\frac{\partial}{\partial t}\mathbf{W_1} = f(\mathbf{W_1}, 0)$, which is

$$\begin{cases} \varepsilon \dfrac{d\bar{X}}{d\tau} &= -\beta T^* \bar{X} + \chi_1 \Pi_{E_0}\left(F_X^1(\mathbf{v},\mathbf{z})\right) + \chi_2 \Pi_{E_0}\left(F_X^2(\mathbf{v},\mathbf{z})\right) + \chi_3 \Pi_{E_0}\left(F_X^3(\mathbf{L})\right) + O(\varepsilon) \\ \varepsilon \dfrac{d\bar{Y}}{d\tau} &= \beta(S^* - T^* - kI^*)\bar{X} - (m + \beta k T^*)\bar{Y} + \chi_1 \Pi_{E_0}\left(F_Y^1(\mathbf{v},\mathbf{z})\right) + \chi_2 \Pi_{E_0}\left(F_Y^2(\mathbf{v},\mathbf{z})\right) + \chi_5 \Pi_{E_0}\left(F_Y^5(\mathbf{v},\mathbf{z})\right) + O(\varepsilon) \\ \varepsilon \dfrac{d\bar{L}_i}{d\tau} &= -m\bar{L}_i + \chi_3 \Pi_{E_0}(F_{L_i}(\mathbf{v},\mathbf{z})) + O(\varepsilon) \\ \varepsilon \dfrac{d\bar{v}_i}{d\tau} &= -\xi \bar{v}_i + O(\varepsilon) \\ \dfrac{\partial \bar{z}_i}{\partial \tau} &= \Pi_{E_0}(F_{z_i}(X,Y,\mathbf{L},\mathbf{v},\mathbf{z})) + O(\varepsilon) \end{cases}. \tag{D.2.5}$$

Apply the Tikhonov's Theorem for (D.2.5), we have the slow system

$$\frac{d\bar{z}_i}{d\tau} = \Pi_{E_0}(F_{z_i}(X^*, Y^*, \mathbf{L}^*, 0, \mathbf{z}^*)). \tag{D.2.6}$$

Use the result in [120], we easily deduce the slow equations as follows

$$\boxed{\begin{aligned} \frac{d\bar{z}_i}{d\tau} = \Theta_1 \bar{z}_i \left(\bar{b}_i - \sum_{j=1}^N \bar{b}_j \bar{z}_j\right) + \Theta_2 \bar{z}_i \left(-\bar{v}_i + \sum_{j=1}^N \bar{v}_j \bar{z}_j\right) + \Theta_3 \bar{z}_i \left[-\sum_{j=1}^N (\bar{u}_{ij} + \bar{u}_{ji})\bar{z}_j + \sum_{j,l=1}^N \left(\bar{u}_{jl} + \bar{u}_{lj}\right)\bar{z}_l \bar{z}_j\right] \\ + \Theta_4 \bar{z}_i \sum_{j=1}^N \left(\bar{\omega}_{ij}^i - \bar{\omega}_{ji}^j\right)\bar{z}_j + \Theta_5 \bar{z}_i \left[\sum_{j=1}^N \left(\frac{T^*}{D^*}\bar{\alpha}_{ji} - \frac{I^*}{D^*}\bar{\alpha}_{ij}\right)\bar{z}_j - \sum_{j,l=1}^N \bar{\alpha}_{jl}\bar{z}_j\bar{z}_l\right] \end{aligned}} \tag{D.2.7}$$

where

$$\Theta_1 = \chi_1 \frac{2\beta S^* T^{*2}}{|P|}, \quad \Theta_2 = \chi_2 \frac{\gamma I^*(I^* + T^*)}{|P|}, \quad \Theta_3 = \chi_3 \frac{\gamma T^* D^*}{|P|}, \quad \Theta_4 = \chi_4 \frac{2mT^*D^*}{|P|}, \quad \Theta_5 = \chi_5 \frac{\beta T^* I^* D^*}{|P|}. \tag{D.2.8}$$

and

$$\bar{b}_i = \Pi_{E_0}(b_i), \quad \bar{v}_i = \Pi_{E_0}(v_i), \quad \bar{u}_{ij} = \Pi_{E_0}(u_{ij}), \quad \bar{\omega}_{ij}^s = \Pi_{E_0}\left(\omega_{ij}^s\right), \quad \bar{\alpha_{ij}} = \Pi_{E_0}(\alpha_{ij}). \tag{D.2.9}$$

For the sake of convenience, we remove the bars over z_i and recall that $\bar{S} = \Pi_{E_0}(S)$, $\bar{I}_i = \Pi_{E_0}(I_i)$, $\bar{I}_{ij} = \Pi_{E_0}(I_{ij})$, $1 \leq i, j \leq N$. By the same argument used in [120], we have the following theorem.

Theorem D.5. *Consider the system of equations*

$$\boxed{\begin{cases} \dot{z}_i = \Theta z_i \left((\Lambda \mathbf{z})_i - \mathbf{z}^T \Lambda \mathbf{z} \right), & i = 1, \cdots, N, \\ z_1 + z_2 + \cdots + z_N = 1. \end{cases}} \tag{D.2.10}$$

where Λ is the square matrix of size $N \times N$ whose coefficient $(i; j)$ are the pairwise fitness λ_i^j which satisfy

$$\boxed{\lambda_i^j = \theta_1 \left(\bar{b}_i - \bar{b}_j \right) + \theta_2 \left(-\bar{\nu}_i + \bar{\nu}_j \right) + \theta_3 \left(-\bar{u}_{ij} - \bar{u}_{ji} + 2\bar{u}_{jj} \right) + \theta_4 \left(\bar{\omega}_{ij}^i - \bar{\omega}_{ji}^j \right) + \theta_5 \left(\mu \left(\bar{\alpha}_{ji} - \bar{\alpha}_{ij} \right) + \bar{\alpha}_{ji} - \bar{\alpha}_{jj} \right).} \tag{D.2.11}$$

with $\mu = \dfrac{I^*}{D^*}$ *and*

$$\Theta = \Theta_1 + \Theta_2 + \Theta_3 + \Theta_4 + \Theta_5, \qquad \theta_i = \frac{\Theta_i}{\Theta}, \qquad i = 1, 2, 3, 4, 5.$$

Then, for any initial values of (D.1.1)*, for each $\tau_0 > 0$, $T > \tau_0$ arbitrarily and independent on ε, there is $\varepsilon_0 > 0$, $C > 0$ and a vector of positive coefficients $\mathbf{z}_0 \in \mathbb{R}^N$ verifying $\sum_{i=1}^N \mathbf{z}_{0,i} = 1$, such that $\forall \varepsilon < \varepsilon_0$*

$$\left\| S^* - S\left(x, \frac{\tau}{\varepsilon}\right) \right\| + \sum_{i=1}^N \left\| I^* z_i(\tau) - \bar{I}_i\left(\frac{\tau}{\varepsilon}\right) \right\| + \sum_{i,j=1}^N \left\| D^* z_i(\tau) z_j(\tau) - \bar{I}_{ij}\left(\frac{\tau}{\varepsilon}\right) \right\| \leq \varepsilon C, \quad \forall \tau \in (\tau_0, T), \quad \forall x \in \Omega.$$

where S, $(I_1, I_2, \ldots, I_N)$, $(I_{ij})_{i,j \in \{1,\ldots,N\}}$ is the solution of (D.1.1) *and $(z_1, z_2, \ldots, z_N)$ is the solution of reduced system* (D.2.10) *together with $\mathbf{z}(0) = \mathbf{z}_0$.*

To close this section, we remark several difficulties when the parameters' mean values depend on space, if we want to keep the same approach.

1. The neutral system for (S, T, I) (D.1.4) can not be solved explicitly. However, we can work around it by denoting (S^*, T^*, I^*) to be the unique solution of (D.1.4).

2. We can not diagonalize the operator A in (D.1.7) explicitly and the new variables $\begin{pmatrix} v_i & z_i \end{pmatrix}$, $1 \leq i \leq N$ can not be defined using A. There may be another way to deal with this, by using the operator

$$\tilde{A} = \begin{pmatrix} -(m + \beta k T^*) + d\Delta & m \\ -\frac{\beta k T^*}{2} & \frac{\beta k I^*}{2} + d\Delta \end{pmatrix}.$$

   We have a claim that, $\phi_0 = \begin{pmatrix} I^* \\ T^* \end{pmatrix}$ is an eigenvector of $\tilde{A}$ respected to eigenvalue 0. The difficulty is that, due to the absence of symmetry, the spectrum of $\tilde{A}$ may be very complex and hard to study, and so may the semi-group $e^{\tilde{A}t}$.

3. The operator K is not simple as before, now it depends on the operator transforming $\begin{pmatrix} I_i \\ J_i \end{pmatrix}$ to $\begin{pmatrix} v_i \\ z_i \end{pmatrix}$.

Résumé

Cette thèse a pour objet l'étude théorique et numérique d'un modèle épidémiologique de co-infection multi-souche. Selon la situation considéré, le modèle s'écrit sous la forme d'un système d'équations différentielles ordinaires ou d'équations de réaction-advection-diffusion. Dans tous les cas, le modèle s'écrit à l'échelle de la population hôte sur la base d'un classique système susceptibles-infectés-susceptibles (SIS). L'agent infectueux est structuré en N souches, qui différent selon 5 traits : la transmissibilité, le taux de clairance des infections simples, le taux de clairance des infections doubles, la probabilité de transmission des souches et les taux de co-infection. Le système obtenu est un systèame de grande taille (N^2+N+1 équations) dont l'étude théorique complète est inaccessible en général. Cette thèse se fonde donc sur une hypothèse simplificatrice de similarité des traits - que l'on nomme hypothèse de quasi-neutralité. Dans ce cadre, il est alors possible de mettre en oeuvre des méthodes de séparations des échelles de temps de type Tikhonov. Le système est ainsi décomposé en deux sous-systèmes plus simples. Le premier est un système dit neutre - c'est-à-dire dans lequel la valeur des traits de toutes les souches sont égales - qui supporte une analyse mathématique détaillée et dont la dynamique s'avère assez simple. Le second se trouve être un système de type "équations de réplication" qui décrit la dynamique en fréquence des souches et contient toute la complexité des interactions entre souche qu'induit les petites variations dans les valeurs des traits. Le premier travail consiste à déterminer explicitement le système lent dans un cadre aspatial pour N souches faisant intervenir un système d'équations différentielles ordinaires et à justifier, que ce système décrit bien le système complet. Ce système est un système de réplication qui peut être décrit à l'aide des $N(N-1)$ fitnesses d'interaction entre les paires de souche. Il est montré que ces fitnesses sont une moyenne pondéré des perturbations de chaque traits. Le second travail consiste a utiliser les expression explicite de ces fitnesses pour décrire exhaustivement la dynamique des paires. Cette partie est illustré à l'aide de beaucoup de simulations des application sur la vaccination sont discutées. Le dernier travail consiste à reprendre cette approche dans un cadre spatialisé. Le modèle SIS est alors un système de réaction-diffusion dans lequel les coefficients sot spatialement hétérogènes. Deux cas limites sont considérés: Le cas d'un coefficient de diffusion asymptotiquement petit et celui d'un coefficient de diffusion asymptotiquement grande. Dans le cas de la diffusion lente on montre que le système lent est un système de type "équations de réplication", décrivant à nouveau l'évolution temporelles mais également spatiale des fréquences des souches. Ce système est de type réaction-advection-diffusion, le terme d'advection additionnel faisant intervenir explicitement l'hétérogénéité du système neutre associé. Dans le cas de la diffusion rapide, l'utilisation de méthodes classiques d'aggrrégation des variables permet de ramener le problème SIS spatialisé à un système SIS homogénéisé sur lequel les résultats précédents peuvent directement s'appliquer.

Mots-clé : échelles de temps multiples, microbiennes, interactions, replication, EDO, EDP

Abstract

The purpose of this thesis is the theoretical and numerical study of an epidemiological model of multi-strain co-infection. Depending on the situation, the model is written as ordinary differential equations or reaction-advection-diffusion equations. In all cases, the model is written at the host population level on the basis of a classical susceptible-infected-susceptible system (SIS). The infecting agent is structured into N strains, which differ according to 5 traits: transmissibility, clearance rate of single infections, clearance rate of double infections, probability of transmission of strains, and co-infection rates. The resulting system is a large system (N^2+N+1 equations) whose complete theoretical study is generally inaccessible. This thesis is therefore based on a simplifying assumption of trait similarity - the so-called quasi-neutrality assumption. In this framework, it is then possible to implement Tikhonov-type time scale separation methods. The system is thus decomposed into two simpler subsystems. The first one is a so-called neutral system - i.e., in which the value of the traits of all the strains are equal - which supports a detailed mathematical analysis and whose dynamics turn out to be quite simple. The second one is a "replication equation" type system that describes the frequency dynamics of the strains and contains all the complexity of the interactions between strains induced by the small variations in the trait values. The first work explicitly determines the slow system in an aspatial framework for N strains using a system of ordinary differential equations and justifies that this system describes the complete system well. This system is a replication system that can be described using the $N(N-1)$ fitnesses of interaction between the pairs of strains. It is shown that these fitnesses are a weighted average of the perturbations of each trait. The second work consists in using explicit expressions of these fitnesses to describe exhaustively the dynamics of the pairs. This part is illustrated with many simulations and applications on vaccination are discussed. The last work consists in using this approach in a spatialized framework. The SIS model is then a reaction-diffusion system in which the coefficients are spatially heterogeneous. Two limiting cases are considered: The case of an asymptotically small diffusion coefficient and the case of an asymptotically large diffusion coefficient. In the case of slow diffusion we show that the slow system is a system of type "replication equations", describing again the temporal but also spatial evolution of the frequencies of the strains. This system is of the reaction-advection-diffusion type, the additional advection term explicitly involving the heterogeneity of the associated neutral system. In the case of fast diffusion, the use of classical methods of aggregation of variables are used to reduce the spatialized SIS problem to a homogenized SIS system on which the previous results can be directly applied.

Keywords : multiple time scales, microbial, interactions, replicator, ODE, PDE