Mechanisms and Opportunities for the Design of Robust and Flexible Collective Behavior in Dynamic Multi-Agent Systems with Interactions

Renato Pagliara Vásquez

A Dissertation
Presented to the Faculty
of Princeton University
in Candidacy for the Degree
of Doctor of Philosophy

Recommended for Acceptance
by the Department of
Mechanical and Aerospace Engineering
Adviser: Naomi Ehrich Leonard

September 2019

Abstract

This dissertation examines the role of feedback and agent level adaptations in the emergence of collective behavior in multi-agent systems that is robust to perturbations and flexible in response to different environmental conditions. In particular, we consider systems with agents that stochastically interact with one another and with their environment, and study how adaptations at the agent level in response to these interactions lead to the emergence of complex group behavior that evolves in time.

Motivated by the remarkable collective behaviors of animal groups, we study the regulation of foraging in harvester ants. We use field experiments and mathematical modeling to examine how interactions between incoming foragers carrying food and available foragers inside the nest yield foraging rates that are robust to uncertainty and responsive to temperature and humidity across minute-to-hour timescales. We show that feedback from outgoing foragers returning to the nest generates stable colony foraging rates and propose that foragers modify their susceptibility to interactions after they become exposed to the environment to explain how the foraging rates adjust to temperature and humidity.

We then examine the role of reinfection and changes in susceptibility in contagion processes that propagate through local interactions. We study a model for reinfection called Susceptible-Infected-Recovered-Infected (SIRI) in which the susceptibility of individuals changes irreversibly after a first exposure to the infection. We show that in both well-mixed and network topologies, the transient and steady-state group dynamics are characterized by two critical system parameters that capture the sensitivity of the group to stimulus before and after the adjustments in susceptibility. We obtain analytical results that yield exact predictions on the effect of network structure and individual behavior on group outcomes.

Our results suggest that agent level adaptations in susceptibility, in combination with feedback across different timescales, are a general principle for robust and flex-

ible collective behavior. This presents mechanisms and opportunities for the study, design, and control of multi-agent systems with cohesive group behaviors that adapt to different conditions.

Acknowledgements

I am incredibly grateful for the opportunities that I have had and for the people in my life. The past five years have been both challenging and fulfilling, and I owe many thanks to the people who supported and helped me along the way.

First of all, I would like to thank my advisor Naomi Leonard for her continuous support, thoughtful advice, and mentorship. It has been a privilege to collaborate with someone who is as dedicated and passionate about producing high-quality work as Naomi is. Her contagious enthusiasm and vigorous creativity is only rivaled by her unending belief in her students. I am thankful for the intellectual freedom Naomi has given me, and for her devotion to clear and concise writing. I am a better researcher and professional because of her.

Much of the work in this dissertation would not have been possible without the 30 years of research on harvester ants by our collaborator Deborah Gordon. I am thankful to Deborah for her mentorship and guidance during my first few years of graduate school. I especially thank her for inviting me to assist in field experiments in the deserts of Arizona and New Mexico and to visit her lab at Stanford. I truly enjoyed my time in the desert, the local wildlife, and the star filled sky. I am thankful to Daniel Friedman for our exhilarating conversations about ants, life, and the universe. Collaborating with Deborah and her group has been a very humbling experience, and has taught me to appreciate the hard work and dedication of field biologists. As an engineer, I wanted to drive down to the desert, put down multiple cameras on top of ant nests, and return after several weeks to collect the juicy data. In reality, collecting data in the field is not that simple; months of planning, preparation and manpower are required.

I want to thank faculty members Phil Holmes, Corina Tarnita, and Andrej Košmrlj for serving in my PhD Committee. Their insightful and thoughtful comments vastly improved my research. I would like to thank Andrej and Vaibhav Srivastava for

serving as dissertation readers and providing valuable feedback and perspective. As well, I would like to thank Simon Levin and Anirudha Majumdar for serving as my dissertation examiners.

I am thankful to have shared my time in the Leonard lab with former and current members of the group: Katie Fitch, Will Scott, Peter Landgren, Liz Davidson, Bec Gray, Desmond Zhong, Anthony Savas, Anastasia Bizyaeva, Udari Madhushani, and Mari Kawakatsu. They provided friendship and a positive environment for collaboration and research. Having intellectually stimulating discussions with smart and passionate people on a daily basis is a privilege that I am thankful for. I would like to specially thank my co-years Liz and Bec for their friendship and support throughout our shared journey through graduate school. In addition, I owe a great deal of gratitude to post-docs in the group: Vaibhav Srivastava, Karla Kvaternik, Zahra Aminzare, Biswa Dey, and Kayhan Ozcimder. They served as role models and kindly shared advice about their own experiences as a graduate student. I would like to specially thank Biswa for his support, advice, and friendship.

My time in Princeton would have been more challenging without the group of talented, hard-working researchers in my cohort who provided a collegial and supportive environment. I have seen us all become better researchers and I wish you the best on your next endeavor. I would also like to thank the wonderful people in SACNAS and AI4All for their friendship and for all the fun. Attending the SACNAS conferences and participating in AI4All provided an invaluable break from research that made my time in graduate school more fun and meaningful.

The MAE department has worked very hard to create a positive and friendly community for which I am very grateful. I thank Jill Ray for all her help throughout the years and for fostering a welcoming and supportive environment in the department. She always made me feel welcome in her office and was always willing to provide a helping hand, no matter the issue.

Graduate school was not always in my plans. I owe much gratitude to my undergraduate advisor Max Donelan for his kind and thoughtful mentorship and for the opportunities he provided that ultimately lead to my enrollment at Princeton. In addition, I would like to thank Max for his commitment to science, good writing, and beautiful figures. His influence can be best observed in the time and dedication that I put into the figures in this dissertation.

Finally, I would not be at this point in my life without the love, unending encouragement, and support of my wife Katja Hyvarinen, and my family. Katja's parents, Terri and Janne, have kindly and lovingly made me part of their family. I am very thankful to them, Zak, Jari, Lucy, Matrix, and Ash for their continuous support. Most of all, I want to thank my parents, Denise and Italo, and siblings, Giancarlo and Bianca, for always believing in me and for their unconditional love and support. Katja, I am incredibly lucky and grateful to have you by my side. Thank you for your friendship and patience; your love and support these past five years have nurtured my soul.

This dissertation carries the number T#3385 in the records of the Department of Mechanical and Aerospace Engineering.

To my parents.

Contents

List of Tables

List of Figures

Part I

Robust and Flexible Collective Behavior in Dynamic Multi-Agent Systems with Interactions

Chapter 1

Introduction

Systems comprising large numbers of individuals that operate without central control, rely on complex patterns of interactions to coherently perform as a group. Individuals within the group interact with each other and their environment, generating group outcomes that feed back to the individuals and shape future interactions. The collective behavior that emerges from these dynamics allows the group to perform tasks beyond the physical and cognitive capabilities of the individuals [1–9].

A great number of biological systems consist of groups of relatively simple individuals that achieve remarkable collective behavior. Flocks of birds and schools of fish display mesmerizing and robust collective motion that increases the likelihood of survival for individuals in the presence of predators [1, 10–12]. Bacteria regulate gene expression in response to fluctuations in cell-population density, leading to virulence, antibiotic production, motility, and biofilm formation [13]. Ant, termite, and bee colonies effortlessly divide labor and make decisions as a collective in order to structure and organize their societies, allowing them to thrive in diverse and extreme environments [4, 14–17]. In all of these cases, individuals take actions based on limited local information regarding the state of the group and the environment. These

actions result in fast and accurate group behaviors that are robust to perturbations and flexible to changes in environmental conditions [15, 18].

Engineered multi-agent systems, such as computer networks, transportation systems, robotic groups, financial systems, and power grids, also exhibit collective behavior. However, unlike the case of biological systems, engineered systems often manage an inherent trade-off between robustness and flexibility. System parameters are optimized for performance according to narrow design specifications. When environmental conditions, system structure, or system inputs change in unforeseen ways, large sensitivities in system parameters can lead to considerable, and sometimes, catastrophic deviations in behavior, such as in the case of cascading failures in power grids [19] and financial markets [20].

Engineering collective behavior that is as robust and flexible as neuronal networks that operate throughout learning-related changes in synapse number and strength [21, 22], or fire ants that form living rafts to survive during floods [23], remains an elusive goal. Understanding the mechanisms that allow biological systems to perform complex collective behavior is a crucial first step in the development of new design and control methodologies for engineered multi-agent systems that exhibit some of the remarkable properties of their natural counterparts [24]. Translating knowledge of robust and flexible biological collective behavior into engineered multi-agent systems requires abstracting away from the complexities of natural swarms and focusing on the development of general principles that allow groups to respond to changes in their environment in a cohesive manner [24, 25].

A general characteristic of both biological and engineered multi-agent systems is the spread of information via local interactions. In most cases, individuals receive information about the state of their neighbors and use it to change their own state [26–29]. Commonly, it is assumed that the *susceptibility* of individuals to their neighbors is fixed, i.e., information from a particular neighbor always has the same effect on the

receiving individual. However, in many realistic scenarios, individuals modify their susceptibility in response to the full history of interactions and the state of their local environment. These adaptive changes in susceptibility at the level of the individual introduce a slow feedback on the group dynamics, leading to smooth changes in group behavior over time.

In this dissertation, we examine how adaptations in susceptibility at the level of the individual, in response to interactions with other individuals or the environment, can lead to the emergence of complex collective behaviors that adjust over time in response to different environmental conditions. Through a combination of field observations and mathematical analysis, we provide general principles for robust and flexible collective behavior. To this end, we derive analytically tractable dynamic models with a small number of critical parameters that capture the underlying mechanisms that allow animal groups to adjust and regulate collective behavior in response to environmental changes without central control.

Our models are nonlinear. The nonlinearities allows us to capture the complex relationship between individual and group level dynamics. This presents both challenges and opportunities for the design of engineered multi-agent systems. Nonlinearities complicate design and analysis as high-dimensional nonlinear systems are generally more difficult to analyze than high-dimensional linear systems [30]. However, nonlinearities can be exploited to achieve rich collective behaviors in groups of relatively simple individuals with limited actuation, sensory, and processing capabilities [24,31].

Our main source of inspiration is the robust and flexible collective behavior of ant colonies. In particular, we are motivated by field observations of foraging activity in red harvester ants.

1.1 Regulation of Foraging in Harvester Ants

Red harvester ants live in the deserts of the Southwestern United States, where they collect seeds that have been scattered by wind and flooding. Colonies obtain water by metabolizing the fat in seeds they collect. Foragers outside the nest lose water through desiccation due to the hot and dry environmental conditions [32, 33]. The colony therefore faces a trade-off in which it must spend water in order to obtain water. To balance this trade-off, colonies must regulate the number of active foragers outside the nest on a minute-to-minute and hour-to-hour basis in response to both food availability and environmental conditions. If the colony is to survive, it must be able to display collective foraging behavior that is both robust to noise and flexible in response to changing environments.

Harvester ant colonies are highly decentralized systems with limited communication capabilities. Foragers communicate through chemical cues in the form of antennal interactions that provide limited information on the state of the group and the environment. To balance the foraging trade-off, the colony must implement simple rules for individuals that lead to high-performing decentralized feedback control.

Research has shown that the rate at which foragers leave the nest depends on the rate of antennal interactions between foragers in the entrance chamber and incoming foragers carrying food [34–38]. The higher the rate of interactions, the higher the likelihood that the ant will be activated to leave the nest to forage [37]. Since foragers remain outside the nest until they find a seed, the rate of successful incoming foragers is positively correlated with food availability [39]. Available foragers waiting inside the nest can therefore acquire noisy information on the foraging conditions through the rate of interactions and use it to inform their decision to remain inside the nest or leave the nest to forage [40]. However, it is not understood how, and if, these

interactions help to regulate colony foraging rates across long timescales or in different environmental conditions.

How colonies respond to day-to-day changes in environmental conditions affects their reproductive success [41]. Colonies that forage less in hot and dry conditions tend to have higher reproductive success than colonies that display less variation in foraging rates across different environmental conditions [41]. Foraging behavior appears to be a heritable trait from parent to offspring colonies. [42]. This suggests that understanding the mechanisms that allow colonies to adjust foraging behavior in different environmental conditions could provide insight into how differences in foraging behavior across colonies at the level of the individual lead to differences in success between colonies.

In Chapter 2 we provide background material. In Chapter 3 we study the regulation of foraging in red harvester ants through field observations and mathematical modeling. We summarize three years of field observations of foraging activity collected for different colonies on different days over hour-long timescales. We derive a model with a small number of parameters that provides an explanation for the range of observed behaviors. Our model suggests that adaptations in forager susceptibility to interactions, in response to a first exposure to the environmental conditions outside the nest, can account for the observed foraging behavior under different environmental conditions. In addition, the model suggests that differences in susceptibility among colonies, in response to temperature and humidity, can produce the observed variation among colonies in the regulation of foraging. The study summarized in Chapter 3 has been published and appears as Pagliara, Gordon, and Leonard [43]. The paper can be found in Chapter 8 of Part II of this dissertation. The study was done in collaboration with Stanford University biologist Deborah Gordon and my advisor Naomi Leonard.

1.2 Contagion Processes

A *contagion process*[1] is any process in which a signal diffuses through a population. The signal may correspond to a physical event, such as a chemical reaction spreading through a batch of reactants [44, 45] or an infectious disease spreading through a community [46, 47]. Or, the signal may correspond to information, such as memes spreading on a social network [48, 49] or social behaviors spreading through the public [50–53].

Models for contagion processes commonly consider individuals with a discrete state that can take on a fixed number of values (e.g., active/inactive, susceptible/infected/recovered) [29, 44, 54]. One or more of these values correspond to a "contagious" or "infected" state. Individuals follow mutual-interaction rules that dictate how they transition between states with the defining rule being that individuals in the contagious state can "infect" their neighbors.

In general, contagion processes fall into two classes depending on the number of exposures required for individuals to become infected: *simple contagion* and *complex contagion* (see [55] and references therein). In a simple contagion process, pairwise interactions between individuals are independent of each other, and a *susceptible* individual becomes infected (with a given probability) after a single interaction with an *infected* individual. In a complex contagion process, susceptible individuals require multiple interactions with infected individuals in order to become infected. The spread of infectious diseases is a common example of a simple contagion [46, 47]. The process through which the rate of interactions activates red harvester ants to forage (see Section 1.1) is an example of a complex contagion.

[1] Contagion processes are also called spreading, epidemic or infectious processes. The latter two alternatives have a strong connotation to epidemiology, and we therefore prefer to use contagion process (we avoid using the term spreading process for consistency). We use the word epidemic in later chapters to describe the collective behavior in which there is a rapid increase in the number of infected individuals in a population.

Complex contagion is usually considered a better model for the spread of ideas and memes across populations as it takes into account how individual behavior is generally affected by the behaviors of all of its neighbors. However, experimental observations of social networks show that there is a large variation between individuals in the number of interactions required before infection [56], suggesting that both simple and complex contagion play a role in the spread of information and social behaviors [55].

Motivated by the ant foraging model presented in Chapter 3, in Chapters 4 and 5, we investigate the role of adaptations in susceptibility in the spread of contagion process across populations. To do this, we study an epidemiological model for reinfection in which the susceptibility of individuals change after a first exposure to the infection.

The work summarized in Chapter 4, which focuses on the well-mixed population setting, has been published and appears as Pagliara, Dey, and Leonard [57]. The paper can be found in Chapter 9 of Part II of this dissertation. The investigation was done in collaboration with Biswadip Dey and my advisor Naomi Leonard. The work summarized in Chapter 5, which focuses on the networked setting, is in preparation for submission and appears as [58]. The paper can be found in Chapter 10 of Part II of this dissertation. The investigation was done in collaboration with my advisor Naomi Leonard.

1.2.1 Epidemiological Models in Well Mixed Settings

Epidemiological models[2] are simple contagion models that characterize the spread of infectious diseases across populations and have been a fundamental tool in the detection, prevention, therapy, and control of the spread of infectious diseases across populations [46, 47, 59–62].[3]

[2]Epidemiological models are also called epidemic models, but we do not use this name to avoid confusion with the collective behavior in which there is a rapid increase in the number of infected individuals in a population.

[3]The first reported epidemiological model corresponds to the work of Bernoulli, published in 1760 [59], but it was not until the work of Kermack and McKendrick in 1927 [60] that the field of mathematical epidemiology as we know it truly started.

At a fundamental level, the dynamics of infectious diseases are described by stochastic reaction-diffusion processes in which individuals can belong to one of a number of discrete states called *compartments* (e.g., susceptible, infected, recovered). Two standard assumptions are made at this point. First, that the number of individuals in the system is very large and, second, that interactions between any two individuals are equally likely (this is where the name *well-mixed* comes from). The dynamics of the probability distribution of the numbers of individuals in each component are then described by the master equation [63] for the resulting Markov process [64]. The master equation is generally analytically intractable. Well-mixed models are obtained by keeping the leading order term in the van Kampen expansion [63] which provides mean-field equations that describe how the numbers of individuals in each compartment change with time [44] (see [65] for an example).

The set of rules dictating how individuals move between compartments is represented in well-mixed models by *reaction rates* in the form of scalar system parameters that usually multiply quadratic terms in the state variables.[4]

Well-mixed epidemiological models are special cases of *compartmental models* [44, 61]. The compartmental modeling framework blends high analytic tractability with flexibility, making it a common choice for modeling a wide range of contagion processes beyond epidemiological ones, including chemical reaction networks [45, 66], collective foraging [67], and collective decision-making [65].

Much of the work on epidemiological models has been dedicated to three particular models [46, 47]:[5]

1. Susceptible-Infected (SI)

$$\mathrm{S}+\mathrm{I} \xrightarrow{\beta} \mathrm{I}+\mathrm{I}$$

[4]This is analogous to using the law of mass action in Chemical Reaction Network Theory, which states that the rate of a chemical reaction is directly proportional to the product of the activities or concentrations of the reactants [66].

[5]Here we use diagrams usually used in Chemical Reaction Network Theory to describe the set of rules describing how individuals move between compartments

2. Susceptible-Infected-Susceptible (SIS)

$$\text{S+I} \xrightarrow{\beta} \text{I+I} \qquad \text{I} \xrightarrow{\delta} \text{S}$$

3. Susceptible-Infected-Recovered (SIR)

$$\text{S+I} \xrightarrow{\beta} \text{I+I} \qquad \text{I} \xrightarrow{\delta} \text{R}$$

where β is the global infection rate and δ is the global recovery rate.

In the SI model, susceptible individuals become infected through contact with infected individuals. The SIS model allows infected individuals to become susceptible again after recovering from the infection. In the SIR model, recovered individuals acquire immunity to the infection and cannot become infected again.

In addition to these models, there exist a few important variations that add or modify compartments to include relevant aspects of the infection that are not considered in the SI, SIS, and SIR models [46, 47]. The SEIR and SEIS models contain a temporary *exposed* compartment representing the incubation, or latent, period of the infection. The MSEIR model contains an additional temporary *immune* compartment, while the MSEIRS allows individuals to temporarily recover before becoming newly susceptible. The immune M and exposed E states are usually not included as they do not directly affect the susceptible-infected interaction [47].

For many of these models, there exists a critical threshold called the *basic reproduction number* R_0, defined as the average number of susceptible individuals that a single infected individual infects [47, 62, 68, 69]. In the SIS model, the basic reproduction number determines whether the fraction of infected individuals monotonically decays to zero with time ($R_0 = \beta/\delta \leq 1$), or reaches an *endemic* state ($R_0 > 1$) in which there are always infected individuals. In the SIR model with a small initial fraction of infected individuals, the basic reproduction number determines whether the fraction of infected individuals monotonically decays to zero with time ($R_0 = \beta/\delta \leq 1$), or

exhibits an epidemic in which the fraction of infected individuals monotonically grows to a maximum value before monotonically decaying to zero with time ($R_0 > 1$).

In Chapter 4 we study the Susceptible-Infected-Recovered-Infected model (SIRI):

$$\begin{array}{ccccc} \text{S+I} & \xrightarrow{\beta} & \text{I+I} & \xleftarrow{\hat{\beta}} & \text{R+I} \\ & & \text{I} & \xrightarrow{\delta} & \text{R} \end{array}$$

in which recovered individuals can become reinfected through interactions with infected individuals at the reinfection rate $\hat{\beta}$. The SIRI model contains the SI, SIS, and SIR models as special cases and can also be specialized to models in which the susceptibility of all individuals decreases to a nonzero rate after a first exposure to the infection (i.e., partial immunity) or increases after a first exposure to the infection (i.e., compromised immunity). We show that the transient and steady-state behavior of solutions in the SIRI model can be characterized by two reproduction numbers, $R_0 = \beta/\delta$ and $R_1 = \hat{\beta}/\delta$.

1.2.2 Epidemiological Models in Networks

Well-mixed epidemiological models provide group-level predictions of the number of individuals expected to belong to each compartment at any time. These predictions rely on assumptions that ignore important properties of the systems they model, such as the finite numbers of individuals in the group and the topology in which individuals interact. The topological structure is particularly important as it describes the set of allowed interactions within the group, affecting how processes and information flow through the population and, therefore, how individual-level behavior connects to group behavior [29, 70, 71].

Networks provide a natural framework for the description of contagion processes in structured populations [29]. Nodes in a network are represented by a random variable

$X_j(t) \in C$, where C is the set of q compartments in the model, and transitions between compartments are assumed to be Poisson processes with fixed rates. Nodes can represent individuals or subpopulations with edges representing connections between individuals or connections between subpopulations, respectively [72, 73].[6]

Contagion processes in network topologies are naturally described by Markov processes [72,74]. Although the Markov approach is exact, these models suffer from three main drawbacks [72]. First, the number of states in the system grows exponentially with the number of nodes in the network. For example, in an epidemiological model with q compartments and a network with N nodes in which each node represents an individual, there are q^N possible states. This limits the analysis to very small groups. Second, the structure of the associated infinitesimal generator is not particularly amenable to analysis. Third, every Markov process has an absorbing state, corresponding to an infection-free state. Thus, the Markov model is not useful in understanding stationary endemic behaviors in which the infection might take an infinite time to die out [75]. For these reasons, contagion processes in networks are usually modeled through mean-field reductions of the Markov process.

There are two main mean-field approaches commonly used to model contagion processes in networks [72]. The first approach, called the *individual-based mean-field approach* (IBMF), assumes that the state of every node is statistically independent from the state of its nearest neighbors (i.e., $\mathbb{E}[X_j X_k] = \mathbb{E}[X_j]\mathbb{E}[X_k]$) yielding dynamic equations for every node that describe the probability p_j^Ω of node j belonging to compartment $\Omega \in C$ [69,75–79]. The IBMF approach significantly reduces the number of states: from q^N to qN. Moreover, IBMF retains the exact network structure of the system, allowing for a close examination of the role of network structure in the spread or eradication of the contagion process, making it the preferred approach

[6]Models that consider subpopulations are usually referred to as metapopulation models. These models commonly assume that the number of individuals in each subpopulation remains fixed for all time but that subpopulations may affect each other.

for researchers attempting to develop control methodologies [73, 74, 80]. In many cases, the IBMF approach yields solutions that critically depend on the spectral properties of the network, similar to the basic reproduction number in well-mixed models [75, 78, 79, 81]. This connection becomes evident through the *next generation matrix* method [82], a general approach to computing basic reproduction numbers in epidemiological models. If solutions in an IBMF model converge to an infection-free equilibrium, then the stochastic Markov model reaches the infection-free absorbing state in sublinear time with respect to the size of the network in expectation [74].

The second approach, called the *degree-based mean-field approach* (DBMF), assumes that all nodes of degree d are statistically equivalent, yielding dynamic equations for every degree class that describe the probability p_d^{Ω} of a node with degree d belonging to $\Omega \in C$ [83, 84]. Unlike the IBMF, the DBMF does not retain exact information on the network topology, but contains a statistical description of the network in a mean-field sense. Because of this, DBMF is particularly useful when working with complex networks in which the exact structure of the network is unknown but in which statistical properties of the network are known [72]. The DMBF approach appears to be preferred by physicists and complex networks researchers. Solutions in DBMF models usually depend on the statistical properties of the network, making it a stronger framework for the study of time-varying networks that retain certain statistical properties through time.

All well-mixed epidemiological models, e.g., SI, SIS, SIR, have analogue network versions. Models may assume global infection and recovery rates that apply to all individuals in the network, or different infection and recovery rates for each individual.

In Chapter 5 we study the SIRI model in network topologies through the IBMF approach. We consider heterogeneities in the recovery rate across individuals and heterogeneities in the infection and reinfection rates across directed pairs of individuals.

We show that the transient and steady-state behaviors of solutions are characterized by two scalar parameters that depend on the spectral properties of the network.

1.3 Outline and Contributions

The work presented in this dissertation is motivated by our desire to understand resilient collective behavior in biological systems and to develop simply parameterized dynamic models for multi-agent systems that exhibit robust and flexible collective behaviors. To this end, we performed three years of field observations of foraging desert harvester ants, and used the data in the derivation of a dynamic model for foraging that demonstrates how simple adaptations in susceptibility in response to interactions with other individuals and the environment can lead to complex collective behaviors that are robust to uncertainty and responsive to environmental conditions across different timescales. We then generalized these principles and applied them to epidemiological models, obtaining rigorous results on how models with simple adaptations in susceptibility display complex transient and steady-state group behaviors that depend on initial conditions and communication topology. These models present opportunities for the design of multi-agent systems that exhibit robust and flexible collective behaviors which are controlled by a small number of critical parameters.

This dissertation consists of two parts, with Part II (Chapters 7 to 10) containing work that has been published or that is in preparation to be submitted.

Part I is divided into six chapters. In Chapter 2 we review some of the mathematical concepts and theories used throughout this dissertation. In Chapter 3 we describe the regulation of the foraging problem in desert harvester ants and summarize the results from Chapter 8, in which we construct a model that explains within colony variation in day-to-day rates and between colonies variation in rates on the same day. We highlight the role of a critical scalar parameter, which we call *volatility*,

that controls transient and stationary characteristics of the colony foraging rate. The volatility parameter is analogous to the average forager susceptibility to interactions. We present a simple mechanism based on changes in volatility that adjusts group behavior in response to environmental conditions. In Chapter 4 we apply general principles for robust and flexible collective behavior, obtained in Chapters 3 and 8, to the study of contagion processes in well-mixed settings. We study the SIRI epidemiological model for reinfection and summarize results from Chapter 9. We highlight the role of critical system parameters that are analogous to the volatility parameter in the ant foraging model. These parameters control transient and steady-state dynamics. We show that for certain values of these critical parameters, the dynamics exhibit bistability in initial conditions and a resurgent epidemic in which solutions rapidly decay at first and remain close to zero for an arbitrarily long length of time before rapidly increasing. In Chapter 5 we study the SIRI model on network topologies and summarize the results from Chapter 10. We extend results obtained in Chapters 4 and 9 to networks and investigate the effects of network structure and heterogeneity (between individuals in how they adapt in response to interactions) on group dynamics. Finally, in Chapter 6 we summarize the main contributions from each chapter and discuss possible future directions of inquiry.

Chapter 2

Background

In this chapter we give a brief introduction to some of the mathematical tools used throughout this dissertation. Each of the areas we touch on is the subject of a vast number of books and research papers. We focus on a selected number of key ideas and tools that are used in later chapters, and refer readers interested in a more comprehensive introduction to the referenced citations within each section.

In Section 2.1 we provide a short overview of basic concepts in dynamical systems theory. In Section 2.2 we discuss local bifurcations in continuous dynamical systems, and in Section 2.3 we introduce the concept of a bifurcation without parameters, with an emphasis on the transcritical bifurcation without parameters which appears in Chapter 4. In Section 2.4 we discuss excitability dynamics and examine the dynamics of the FitzHugh-Nagumo model which is prominently used in Chapter 3. Section 2.5 focuses on key concepts and results from Queueing Theory which are also used in Chapter 3. Finally, in Sections 2.6 and 2.7 we introduce notation, concepts, and results on graph theory and Metzler matrices, both of which are prominently used in Chapter 5.

Basic Notation: Throughout this dissertation we use the symbol $\mathbb{R}$ to denote the set of all real numbers and $\mathbb{N}$ to denote the set of natural numbers. We denote

matrices in capital letters (e.g., $M \in \mathbb{R}^{N\times N}$) and vectors in boldface lowercase letters (e.g., $\mathbf{v} \in \mathbb{R}^N$). We denote the j-th element of $\mathbf{v} \in \mathbb{R}^N$ as v_j, and the (j,k)-th element of $M \in \mathbb{R}^{N\times N}$ as m_{jk}. We define $\mathbf{0} \in \mathbb{R}^N$ as the zero vector, $\mathbf{1} \in \mathbb{R}^N$ as the vector with every element 1, and $\mathbf{e}^j$, $j = 1, \ldots, N$ as the standard basis vectors ($e^j_k = 1$ for $k = j$ and $e^j_k = 0$ for $k \neq j$). We define $\bar{\mathbf{0}} \in \mathbb{R}^{N\times N}$ as the zero square matrix, and $\mathbb{I} \in \mathbb{R}^{N\times N}$ as the identity matrix. We let $\mathrm{diag}(\mathbf{v}) \in \mathbb{R}^{N\times N}$ be the diagonal matrix with entries given by the elements of $\mathbf{v} \in \mathbb{R}^N$. For any vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^N$, we write $\mathbf{x} \gg \mathbf{y}$ if $x_j > y_j$ for all j, $\mathbf{x} \succ \mathbf{y}$ if $x_j \geq y_j$ for all j, but $\mathbf{x} \neq \mathbf{y}$, and $\mathbf{x} \succeq \mathbf{y}$ if $x_j \geq y_j$ for all j. Similarly, for any two matrices $M, Q \in \mathbb{R}^{N\times N}$ we write $M \gg Q$ if $m_{jk} > q_{jk}$, $M \succ Q$ if $m_{jk} \geq q_{jk}$, for any j, k, but $M \neq Q$, and $M \succeq Q$ if $m_{jk} \geq q_{jk}$ for any j, k. A square matrix M is Hurwitz (stable) if all its eigenvalues have negative real part, and it is unstable if at least one of its eigenvalues has positive real part. We denote the spectrum of a square matrix M as $\lambda(M) = \{\lambda_1, \lambda_2, \ldots, \lambda_N\}$ and its spectral radius by $\rho(M) = \max\{|\lambda_j| \,\big|\, \lambda_j \in \lambda(M)\}$.

2.1 Dynamical Systems Theory

A dynamical system is an abstract mathematical object that defines how a set of variables, which define the state of a system, change over time. Dynamical systems are used to study both physical and non-physical systems and are usually represented via sets of coupled ordinary differential equations (ODEs). The theory and set of tools available to study dynamical systems is vast (see for example [30, 85, 86]). Here, we focus on introducing key concepts related to the stability of stationary solutions of ODE systems which will be used throughout the remaining chapters.

2.1.1 Equilibria and Stability

Consider an autonomous system of ODEs

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}); \qquad \mathbf{x} \in \mathbb{R}^N, \tag{2.1}$$

where the vector field $\mathbf{f} : \mathbb{R}^N \to \mathbb{R}^N$ is smooth.

A *nullcline* of (2.1) is a curve in $\mathbb{R}^N$ described by setting $f_i = 0$ for $i \in \{1, \ldots, N\}$. An *equilibrium* of (2.1) is a solution $\mathbf{x}^*$ that satisfies $\mathbf{f}(\mathbf{x}^*) = \mathbf{0}$. That is, an equilibrium corresponds to the point at which all nullclines intersect.

Definition 2.1.1 (Stability [87]). *The equilibrium point* $\mathbf{x}^*$ *of* (2.1) *is*

- *stable if, for each* $\epsilon > 0$*, there is* $\delta = \delta(\epsilon) > 0$ *such that*

$$\|\mathbf{x}(0) - \mathbf{x}^*\| < \delta \implies \|\mathbf{x}(t) - \mathbf{x}^*\| < \epsilon, \quad \forall\ t \geq 0,$$

- *unstable if it is not stable,*

- *asymptotically stable if it is stable and* δ *can be chosen such that*

$$\|\mathbf{x}(0) - \mathbf{x}^*\| < \delta \implies \lim_{t \to \infty} \mathbf{x}(t) = 0.$$

2.1.2 Linear Systems

If (2.1) is a linear system $\mathbf{f}(\mathbf{x}) = A\mathbf{x}$ where $A \in \mathbb{R}^{N \times N}$ is a matrix with constant coefficients, then, if A is non-singular, (2.1) has a unique isolated equilibrium at the origin $\mathbf{x}^* = \mathbf{0}$. If A is singular, (2.1) has an infinite number of non-isolated equilibria. The eigenspaces of A are *invariant* to the dynamics, meaning that any solution that starts on one of these spaces remains in the space for all time. The eigenspace of the linear system is usually divided into stable E^s, unstable E^u, and

center E^c subspaces, where each space corresponds to the union of the eigenspaces associated with eigenvalues with negative, positive, and zero real part, respectively. Solutions that lie on E^s decay at an exponential rate to the origin as $t \to \infty$, while solutions that lie on E^u decay at an exponential rate to the origin as $t \to -\infty$. The behavior of solutions that lie on E^c is unknown without additional details. However, if E^c has dimension N, then solutions either remain constant (zero eigenvalues) or oscillate at a constant amplitude (purely imaginary eigenvalues). It follows that if all eigenvalues of A have negative real part, solutions converge to the origin at an exponential rate and we say the origin is *exponentially asymptotically stable*. If A has at least one eigenvalue with positive real part, the origin is unstable. For a linear system, asymptotic stability implies exponential stability. Throughout this dissertation, we use the term "stable" to refer to an asymptotically stable point in situations where the context is clear.

2.1.3 Nonlinear Systems

Unlike linear systems, nonlinear systems can have multiple isolated equilibria. If (2.1) is nonlinear, its *linearization* at an equilibrium $\mathbf{x} = \mathbf{x}^*$ is given by

$$\dot{\bar{\mathbf{x}}} = D_x\mathbf{f}(\mathbf{x}^*)\bar{\mathbf{x}}, \qquad \bar{\mathbf{x}} \in \mathbb{R}^N, \tag{2.2}$$

where $\bar{\mathbf{x}} = \mathbf{x} - \mathbf{x}^*$ and $D_x\mathbf{f} = \{\partial f_j/\partial x_k\} \in \mathbb{R}^{N\times N}$ is the Jacobian derivative of $\mathbf{f}(\mathbf{x})$ with respect to $\mathbf{x}$. If every eigenvalue of $D_x\mathbf{f}(\mathbf{x}^*)$ has positive or negative real part, we say the equilibrium $\mathbf{x}^*$ is *hyperbolic*.

The Hartman-Grobman Theorem [30] states that, near a hyperbolic equilibrium, the nonlinear system (2.1) is locally topologically equivalent to its linearization (2.2). This implies that a hyperbolic equilibrium $\mathbf{x}^*$ of (2.1) is exponentially asymptotically stable if all eigenvalues of $D_x\mathbf{f}(\mathbf{x}^*)$ are negative and unstable if $D_x\mathbf{f}(\mathbf{x}^*)$ has at

least one eigenvalues with positive real part. If $D_x\mathbf{f}(\mathbf{x}^*)$ has a zero or purely imaginary eigenvalue, the stability of the origin in the linearization (2.2) does not provide information on the stability of $\mathbf{x}^*$ in the nonlinear system (2.1).

For any hyperbolic equilibrium $\mathbf{x}^*$ of (2.1), the Stable Manifold Theorem [30] states that there exist stable and unstable manifolds $\mathcal{W}^s$ and $\mathcal{W}^u$ of the same dimension as the subspaces E^s and E^u of the linearized system (2.2), which are invariant to the dynamics and tangent to E^s and E^u at $\mathbf{x}^*$. Solutions that lie on $\mathcal{W}^s$ go to $\mathbf{x}^*$ as $t \to \infty$, while solutions that lie on $\mathcal{W}^u$ go to $\mathbf{x}^*$ as $t \to -\infty$.

For any non-hyperbolic equilibrium of (2.1) where (2.2) has n_+, n_-, and n_0 eigenvalues with positive, negative, and zero real part, respectively, the Center Manifold Theorem [30] states that, in addition to unique stable and unstable manifolds, there exists a non-unique center manifold $\mathcal{W}^c$ with dimension n_0, which is tangent to E^c at $\mathbf{x}^*$. Moreover, by the Reduction Principle [88], the nonlinear system (2.1) near the non-hyperbolic equilibrium $\mathbf{x}^*$ is locally topologically equivalent to the system

$$\dot{\mathbf{x}}_c = \mathbf{f}_c(\mathbf{x}_c); \qquad \mathbf{x}_c \in \mathbb{R}^{n_0}, \tag{2.3}$$

$$\dot{\mathbf{x}}_s = -\mathbf{x}_s; \qquad \mathbf{x}_s \in \mathbb{R}^{n_-}, \tag{2.4}$$

$$\dot{\mathbf{x}}_u = \mathbf{x}_u; \qquad \mathbf{x}_u \in \mathbb{R}^{n_+} \tag{2.5}$$

where (2.3) is the *restriction* of (2.1) to its center manifold. Note that (2.3), (2.4), and (2.5) are uncoupled, and that the dynamics of (2.4) and (2.5) look like the equations of a standard saddle. Therefore, the Reduction Principle can be restated as follows [88]:

> Near a non-hyperbolic equilibrium, the system (2.1) is locally topologically equivalent to the suspension of its restriction to the center manifold by the standard saddle.

If $n_+ = 0$, then all solutions that start at a point not in the center manifold, converge to the center manifold as $t \to \infty$. In this case, we say the center manifold $\mathcal{W}^c$ is *attractive*.

2.1.4 Non-isolated Equilibria

Consider systems in $\mathbb{R}^N$ with an n-dimensional ($n \leq N$) manifold of equilibria $\mathcal{M}$, such as systems with a line ($n = 1$) or a plane ($n = 2$) of equilibria. Let n_c be the number of eigenvalues with zero real part in the linearization at an arbitrary point in $\mathcal{M}$ (2.2). Then, $n_c \geq n$ with n eigenvalues corresponding to the n-dimensional space tangent to $\mathcal{M}$. The remaining $N - n$ eigenvalues are the *transverse* eigenvalues of the system corresponding to the $N - n$ dimensional space transverse to $\mathcal{M}$.

If all transverse eigenvalues of (2.2) at $\mathbf{x}^* \in \mathcal{M}$ have negative or real part, i.e., $n = n_c$, we say $\mathcal{M}$ is *normally hyperbolic at* $\mathbf{x}^*$. If $\mathcal{M}$ is normally hyperbolic at $\mathbf{x}^*$, then $\mathcal{M}$ is a center manifold of $\mathbf{x}^*$ with trivial dynamics (i.e., (2.3) is given by $\dot{\mathbf{x}}_c = \mathbf{0}$). By the Reduction Principle, $\mathbf{x}^*$ is locally stable if all transverse eigenvalues of (2.2) at $\mathbf{x}^*$ have negative real part and unstable if one or more of the transverse eigenvalues have positive real part. If $\mathcal{M}$ is normally hyperbolic at every point, then we say $\mathcal{M}$ is *normally hyperbolic*.[1]

If $\mathcal{M}$ is not normally hyperbolic at a point $\mathbf{x}^*$, i.e, $n_c > n$, then $\mathbf{x}^*$ is unstable if one or more transverse eigenvalues of (2.2) at $\mathbf{x}^*$ have positive real part. If all transverse eigenvalues have negative real part, then the stability of $\mathbf{x}^*$ depends on the dynamics on the n_c-dimensional center manifold (2.3).

[1]Normally hyperbolic manifolds of equilibria are a special case of normally hyperbolic invariant manifolds (NHIM). The topological equivalence of the linearization extends to compact NHIM, such as limit cycles (see for example Theorem 4.1 in [89]).

2.1.5 Invariant Sets Besides Equilibria

In addition to isolated and non-isolated equilibria, nonlinear systems may have other invariant sets, including *periodic orbit* solutions $\mathbf{x}(t)$ where

$$\mathbf{x}(t) = \mathbf{x}(t+T), \qquad 0 < T < \infty \tag{2.6}$$

for all t. These periodic orbits also have stable and unstable manifolds. A *limit cycle* is an isolated periodic orbit. Limit cycles can be stable, unstable or half-stable (attractive on either the inside or outside of the orbit and repelling on the other side).

A nonlinear system may also have solutions that connect one or more equilibria. A solution $\mathbf{x}(t)$ connecting two equilibria such that $\mathbf{x}(t)$ reaches one of the equilibria as $t \to \infty$ and reaches the other as $t \to -\infty$ is called a *heteroclinic orbit*. A solution $\mathbf{x}(t)$ connecting an equilibrium to itself such that $\mathbf{x}(t)$ reaches the equilibrium as $t \to \infty$ and $t \to -\infty$ is called a *homoclinic orbit*.

2.2 Bifurcations

Consider an autonomous system of ODEs of the form

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mu); \qquad \mathbf{x} \in \mathbb{R}^N, \quad \mu \in \mathbb{R}^K \tag{2.7}$$

where the vector field $\mathbf{f} : \mathbb{R}^N \times \mathbb{R}^K \to \mathbb{R}^N$ is smooth and depends on the K-dimensional parameter μ. We say that a *bifurcation* with *bifurcation value* μ_0 occurs if the dynamics of (2.7) with $\mu = \mu_0$ are not topologically equivalent to the dynamics of (2.7) with $\mu = \mu_1$, where μ_1 is arbitrarily close to μ_0. In other words, a bifurcation takes place when any arbitrarily small perturbation to the parameter at the bifurcation value results in a system with qualitatively different dynamics. Examples of qualitative

changes include changes to the number or stability of equilibria, periodic orbits, or other invariant sets.

Bifurcations are usually divided into two principal classes: local bifurcations and global bifurcations. Here we limit our scope to local bifurcations, which can be detected and analysed through changes in the local stability properties of equilibria, periodic orbits or other invariant sets.

By the Implicit Function Theorem, the equilibria of (2.7) are described by smooth functions of μ. This holds for any point that is sufficiently far away from those points at which $D_x\mathbf{f}(\mathbf{x}^*, \mu)$ has a zero eigenvalue. The graph of each of these functions is a *branch of equilibria* of (2.7). A *bifurcation point* $(\mathbf{x}^*, \boldsymbol{\mu_0})$ corresponds to an equilibrium where $D_x\mathbf{f}(\mathbf{x}^*, \boldsymbol{\mu_0})$ has a zero eigenvalue. This is the point in the $(\mathbf{x}, \mu)$ product space at which branches of equilibria meet.

The *codimension* of a bifurcation is the smallest dimension of a parameter space which contains the bifurcation in a persistent way [30]. For example, if $K = 1$, $\mu \in \mathbb{R}$ and any resulting bifurcation that happens as μ varies is a codimension-one bifurcation.

The study of bifurcations in higher-dimensional systems relies on model reductions, such as the Center Manifold Theorem (described in Section 2.1), and change of coordinates that map the bifurcation problem to the study of lower-dimensional normal forms (see [30, 88] for additional details on center manifolds and for an introduction to normal forms).

Bifurcation can be studied via the extended system

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}, \mu), \\ \dot{\mu} &= \mathbf{0}, \end{aligned} \tag{2.8}$$

where every equilibrium $\mathbf{x}^*$ of (2.7) is now a K-dimensional manifold of equilibria $\mathcal{M}_{x^*} = \{(\mathbf{x}^*, \mu) \in (\mathbb{R}^N, \mathbb{R}^K) | \mathbf{f}(\mathbf{x}^*, \mu) = \mathbf{0}\}$. The local stability of points in $\mathcal{M}_{x^*}$ can be investigated in a similar manner to the stability of points in a manifold of equilibria (see Section 2.1).

The Shoshitaishvili Reduction Principle [90] (see also [88]) extends the Reduction Principle and states that the dynamics of (2.8) near a non-hyperbolic equilibrium are topologically equivalent to the suspension of its restriction to the center manifold by the standard saddle, and that the homeomorphism which realizes the equivalence does not change the parameters in the system. In other words, the Shoshitaishvili Reduction Principle extends the Reduction Principle to systems with parameters by stating that the events near the bifurcation parameter value, that change the qualitative behavior of the nonlinear system (2.8), occur on the center manifold $\mathcal{W}^c$ and are captured by the n_c-dimensional system (2.3).

Bifurcations are usually illustrated through the use of *bifurcation diagrams* which show the different branches, along with their stability in a subset of the $(\mathbf{x}, \mu)$ product space (i.e., the phase space of the extended system (2.8)). Since the μ dynamics in (2.8) are trivial, each solution in this space with μ constant forms an invariant manifold of (2.8) that crosses the μ-axes orthogonally. We say that the manifold $\mathbb{R}^{N+K}$ is *foliated* by the *leaves* $\mu = c \in \mathbb{R}^K$, where each leaf corresponds to an invariant manifold with μ constant.

We now briefly describe the four main codimension-one bifurcations (i.e., $\mu \in \mathbb{R}$), their normal forms, and the transversality conditions that describe them (see Theorems 3.4.1 and 3.4.2 in [30]):

- A *saddle-node* bifurcation has normal form $x = \mu - x^2$. For $\mu < 0$ the system has no equilibria, while for $\mu > 0$ the system has a stable equilibrium $x = \sqrt{\mu}$ and an unstable equilibrium $x = -\sqrt{\mu}$. At $\mu = 0$, the stable and unstable

equilibria collide and disappear. A system with a saddle-node bifurcation at $(\mathbf{x}^*, \boldsymbol{\mu_0})$ satisfies the following three conditions:

(SN1) $D_x\mathbf{f}(\mathbf{x}^*, \boldsymbol{\mu_0})$ has a zero eigenvalue with algebraic multiplicity one and with left and right eigenvectors $\mathbf{w}^T$ and $\mathbf{v}$, respectively, and all other eigenvalues of $D_x\mathbf{f}(\mathbf{x}^*, \boldsymbol{\mu_0})$ have positive or negative real part.

(SN2) $\mathbf{w}^T\big(\partial\mathbf{f}/\partial\mu\big)(\mathbf{x}^*, \boldsymbol{\mu_0}) \neq 0$.

(SN3) $\mathbf{w}^T\big(D_x^2\mathbf{f}(\mathbf{x}^*, \boldsymbol{\mu_0})\big)(\mathbf{v}, \mathbf{v}) \neq 0$.

- A *transcritical bifurcation* has normal form $x = \mu x - x^2$. For $\mu < 0$, $x = 0$ is a stable equilibrium and $x = \mu$ is an unstable equilibrium. For $\mu > 0$, $x = 0$ is unstable and $x = \mu$ is stable. The branches intersect and change stability at $x = 0$. A system with a transcritical bifurcation at $(\mathbf{x}^*, \boldsymbol{\mu_0})$ satisfies conditions (SN1), (SN3), and

  (T2) $\mathbf{w}^T\big(\partial^2\mathbf{f}/\partial\mu\partial\mathbf{x}\big)(\mathbf{v})(\mathbf{x}^*, \boldsymbol{\mu_0}) \neq 0$.

- A *pitchfork bifurcation* has normal form $x = \mu x - x^3$. For $\mu < 0$, $x = 0$ is a stable equilibrium. For $\mu > 0$, $x = 0$ is unstable and the system has two additional stable equilibria at $x = \pm\mu$. All branches intersect at $x = 0$. A one-dimensional system $\dot{x} = f(x, \mu)$, has a pitchfork bifurcation at $x = x^*$,$\mu = \mu_0$ if it satisfies conditions[2]:

  (PF1) $(\partial f/\partial x)(x^*, \mu_0) = 0$.

  (PF2) $(\partial f/\partial\mu)(x^*, \mu_0) \neq 0$.

  (PF3) $\big(\partial^3 f/\partial x^3\big)(x^*, \mu_0) \neq 0$.

Condition (SN1) describes the bifurcation point and the requirement that, in the extended system (2.8), the branch described by $\mathbf{x}^*$ loses normal hyperbolicity at

[2]Conditions for the pitchfork bifurcation in general N-dimensional systems can be described using singularity theory, which is beyond the scope of this dissertation. We refer the reader to [91] for details.

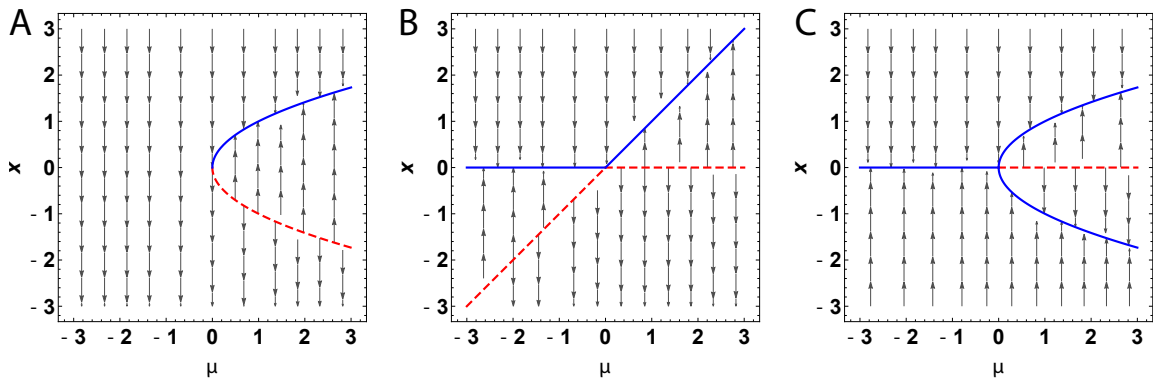

Figure 2.1: Bifurcation diagrams. A. Supercritical saddle-node. B. Transcritical. C. Supercritical pitchfork.

the bifurcation point $(\mathbf{x}^*, \mu_0)$. Condition (SN2) controls the nondegeneracy of the behavior with respect to μ. Condition (SN3) controls the dominant effect of the quadratic nonlinear term in the normal form. Condition (T2) sets the eigenvalue to cross zero with non-vanishing speed as μ increases. The pitchfork transversality conditions (PF1) and (PF2) are equal to conditions (SN1) and (SN2) for a one-dimensional system. Condition (PF3) controls the dominant effect of the cubic nonlinear term in the normal form.

Figure (2.1) shows the saddle-node, transcritical, and pitchfork bifurcation diagrams. We denote the stable branches as solid blue lines and the unstable branches as dashed red lines. The dark gray curves represent solutions of the extended system. Note that the extended phase space (x, μ) is foliated by the leaves $\mu = c \in \mathbb{R}$, which are orthogonal to the μ-axis.

In addition to changes in the number and stability of equilibria, a codimension-one bifurcation can lead to changes in the number and stability of periodic orbits.

- A *Hopf* bifurcation is a codimension-one bifurcation with normal form

$$\begin{aligned} \dot{x}_1 &= \mu x_1 - x_2 + \sigma x_1(x_1^2 + x_2^2) \\ \dot{x}_2 &= x_1 + \mu x_2 + \sigma x_2(x_1^2 + x_2^2). \end{aligned} \tag{2.9}$$

where $\sigma = \pm 1$. The only equilibrium of (2.7) is the origin with eigenvalues $\lambda_{1,2} = \mu \pm i$; the origin is an asymptotically stable spiral for $\mu < 0$ and unstable for $\mu > 0$. In polar coordinates (2.9) becomes

$$\begin{aligned} \dot{r} &= r(\mu + \sigma r^2) \\ \dot{\theta} &= 1. \end{aligned} \tag{2.10}$$

From (2.10) we see that, besides the equilibrium at the origin ($r = 0$), there is a limit cycle (an isolated periodic solution) with radius $r = \sqrt{-\mu/\sigma}$ and period 2π. For $\sigma = -1$, (2.9) has a stable limit cycle that exists for $\mu > 0$ and has radius $\sqrt{\mu}$. This is a supercritical Hopf bifurcation. If $\sigma = +1$, the origin of (2.9) has an unstable limit cycle that exists for $\mu < 0$. This is a subcritical Hopf bifurcation. A system exhibits a Hopf bifurcation at $(\mathbf{x}^*, \mu_0)$ if it satisfies the following two conditions:

(H1) $D_x\mathbf{f}(\mathbf{x}^*, \mu_0)$ has a pair of purely imaginary eigenvalues $\lambda(\mu_0)$ and $\bar{\lambda}(\mu_0)$, each with algebraic multiplicity one, and all other eigenvalues of $D_x\mathbf{f}(\mathbf{x}^*, \mu_0)$ have positive or negative real part.

(H2) $\frac{d}{d\mu}\big(\mathrm{Re}\lambda(\mu)\big)|_{\mu=\mu_0} \neq 0$.

Similarly to the other bifurcations, condition (H1) describes the bifurcation point and the requirement that, in the extended system (2.8), the branch described by $\mathbf{x}^*$ loses normal hyperbolicity at the bifurcation point $(\mathbf{x}^*, \mu_0)$, and condition (H2) describes the condition for the eigenvalues crossing zero transversally as μ increases.

Figure 2.2 shows the bifurcation diagram for the supercritical and subcritical Hopf bifurcations. In each panel, we denote the stable branch with solid blue lines, the unstable branch with red dashed lines, and the limit cycle with violet circles. The gray translucent planes orthogonal to the μ-axis correspond to the leaves $\mu = \pm 2$. The dark gray curves show solutions on these invariant planes.

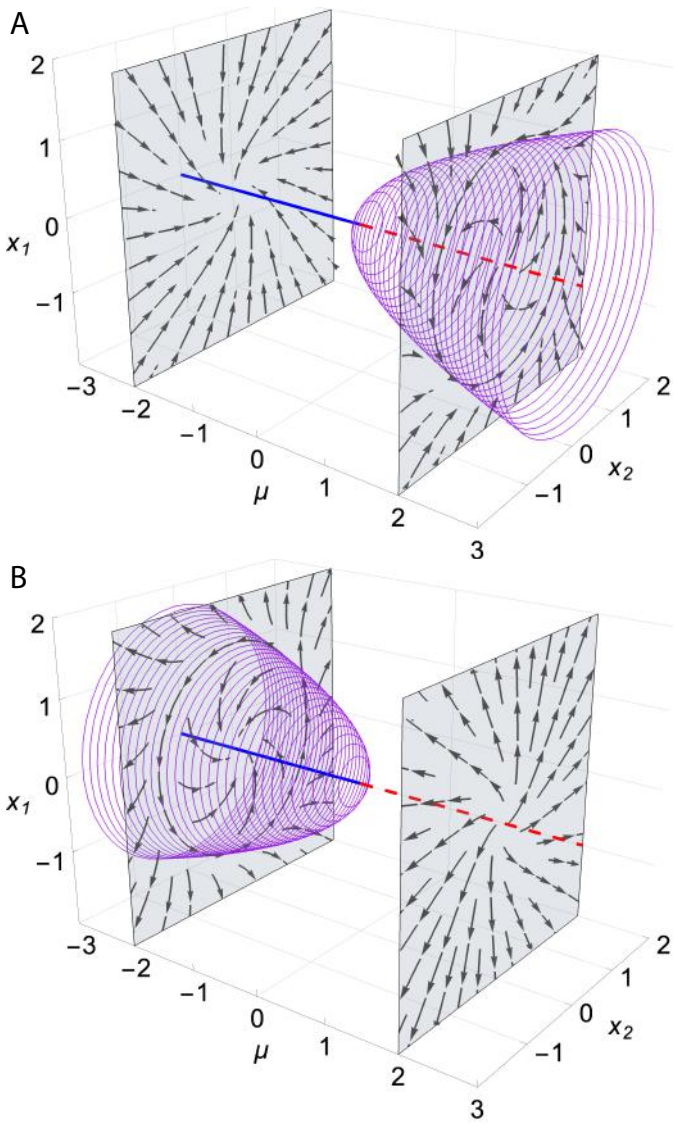

Figure 2.2: Hopf bifurcation diagrams. A. Supercritical. B. Subcritical.

In Chapter 3 we make use of the FitzHugh-Nagumo model, a nonlinear model that exhibits subcritical and supercritical Hopf bifurcations. In Chapters 4 and 5 we study the SIRI epidemiological model in well-mixed and network settings and show that when the infection and reinfection rate are equal, the dynamics exhibit a transcritical bifurcation.

2.3 Bifurcations Without Parameters

Consider an autonomous system of ODEs in $\mathbb{R}^N$ of the form

$$\begin{aligned} \dot{\mathbf{y}} &= \mathbf{g}(\mathbf{y}, \mathbf{z}); & \mathbf{y} &\in \mathbb{R}^{N-n} \\ \dot{\mathbf{z}} &= \mathbf{h}(\mathbf{y}, \mathbf{z}); & \mathbf{z} &\in \mathbb{R}^{n} \end{aligned} \tag{2.11}$$

where $\mathbf{g} : \mathbb{R}^{N-n} \times \mathbb{R}^n \mapsto \mathbb{R}^{N-n}$ and $\mathbf{h} : \mathbb{R}^{N-n} \times \mathbb{R}^n \mapsto \mathbb{R}^n$ are smooth. Suppose further that there is an n-dimensional manifold of equilibria $\mathcal{M} = \{(\mathbf{0}, \mathbf{z}) \in (\mathbb{R}^{N-n}, \mathbb{R}^n) | \mathbf{g}(\mathbf{0}, \mathbf{z}) = \mathbf{0}, \mathbf{h}(\mathbf{0}, \mathbf{z}) = \mathbf{0}\}$.

Suppose $\mathbf{h}(\mathbf{y}, \mathbf{z}) = \mathbf{0}$ for all $\mathbf{y} \in \mathbb{R}^{N-n}$ and all $\mathbf{z} \in \mathbb{R}^n$. Then, the study of (2.11) reduces to the study of the $(N-n)$-dimensional system $\dot{\mathbf{y}} = \mathbf{g}(\mathbf{y}, \mathbf{z})$, where $\mathbf{z}$ plays the role of an n-dimensional parameter and where the manifold of equilibria $\mathcal{M}$ corresponds to the equilibrium $\mathbf{y}^* = \mathbf{0}$.

Assume the stability of $\mathbf{y}^*$ changes with the value of $\mathbf{z}$ through a bifurcation at $(\mathbf{y}^*, \mathbf{z}^*)$. Then, we can study the resulting bifurcation through the "extended" system

$$\begin{aligned} \dot{\mathbf{y}} &= \mathbf{g}(\mathbf{y}, \mathbf{z}), \\ \dot{\mathbf{z}} &= \mathbf{h}(\mathbf{y}, \mathbf{z}) = \mathbf{0}. \end{aligned} \tag{2.12}$$

In the $(\mathbf{y}, \mathbf{z})$ space, the equilibrium $\mathbf{y}^* = \mathbf{0}$ appears as the manifold of equilibria $\mathcal{M}$, with stable and unstable regions separated by the bifurcation point $(\mathbf{y}^*, \mathbf{z}^*) \in \mathcal{M}$. Solutions of (2.12) with $\mathbf{z}$ constant foliate the $(\mathbf{y}, \mathbf{z})$ space. This is the classical bifurcation setup described in Section 2.2.

The theory of *bifurcations without parameters* studies the dynamics of systems of the form (2.11) when $\mathbf{h}(\mathbf{y}, \mathbf{z}) \neq \mathbf{0}$, near the point $(\mathbf{y}^*, \mathbf{z}^*)$. In this case, solutions with

$\mathbf{z}$ constant no longer foliate the $(\mathbf{y}, \mathbf{z})$ space. Instead, the dynamics in the $(\mathbf{y}, \mathbf{z})$ space become considerably more complex.

To better understand the dynamics of (2.11) near $(\mathbf{y}^*, \mathbf{z}^*)$, consider the Jacobian of (2.11) evaluated at an arbitrary point $(\mathbf{y}^*, \mathbf{z}) \in \mathcal{M}$,

$$J_{\mathcal{M}}(\mathbf{y}^*, \mathbf{z}) = \begin{bmatrix} D_{\mathbf{y}}\mathbf{g}(\mathbf{y}^*, \mathbf{z}) & D_{\mathbf{z}}\mathbf{g}(\mathbf{y}^*, \mathbf{z}) \\ D_{\mathbf{y}}\mathbf{h}(\mathbf{y}^*, \mathbf{z}) & D_{\mathbf{z}}\mathbf{h}(\mathbf{y}^*, \mathbf{z}) \end{bmatrix} = \begin{bmatrix} D_{\mathbf{y}}\mathbf{g}(\mathbf{y}^*, \mathbf{z}) & \bar{\mathbf{0}} \\ D_{\mathbf{y}}\mathbf{h}(\mathbf{y}^*, \mathbf{z}) & \bar{\mathbf{0}} \end{bmatrix}. \tag{2.13}$$

The Jacobian $J_{\mathcal{M}}(\mathbf{y}^*, \mathbf{z})$ has at least n zero eigenvalues corresponding to the n dimensions in phase space tangent to $\mathcal{M}$, with the remaining $N - n$ transverse eigenvalues corresponding to the eigenvalues of $D_{\mathbf{y}}\mathbf{g}(\mathbf{y}^*, \mathbf{z})$.

At the point $(\mathbf{y}^*, \mathbf{z}^*)$, the spectrum of $D_{\mathbf{y}}\mathbf{g}(\mathbf{y}^*, \mathbf{z}^*)$ intersects the imaginary axis. If there is only one eigenvalue on the imaginary axis, the eigenvalue is zero and this then implies that $\mathcal{M}$ loses normal hyperbolicity at $(\mathbf{y}^*, \mathbf{z}^*)$. Loss of normal hyperbolicity is a degeneracy that generalizes that of a classical bifurcation in which one of the eigenvalues becomes zero at the bifurcation point. The dynamics near points at which $\mathcal{M}$ loses normal hyperbolicity lead to what Stefan Liebscher and his collaborators call a *bifurcation without parameters* [86].

Similar to the case of classical bifurcations, bifurcations without parameters can be studied by reducing the system dynamics to normal forms. The codimension of a bifurcation without parameters refers to the dimensionality of the manifold of equilibria (i.e., n). For example, if $n = 1$ in (2.11), the system has a line of equilibria with a codimension-one bifurcation without parameters at any isolated points on the z-axis at which the line of equilibria loses normal hyperbolicity. Similarly, if $n = 2$ in (2.11), then the system has a plane of equilibria. Points at which the plane of equilibria loses normal hyperbolicity generically form curves, giving rise to a codimension-two bifurcation without parameters.

Here, we provide a brief description of the two types of codimension-one bifurcation without parameters that we encounter in Chapter 4, their normal forms, and the transversality conditions that control the bifurcation in a system of the form (2.11) with $y, z \in \mathbb{R}$. We refer the reader to [86] for additional details and for a more exhaustive list of codimension-one and codimension-two bifurcations without parameters.

A *transcritical bifurcation without parameters* has normal form

$$\begin{aligned} \dot{y} &= yz \\ \dot{z} &= y. \end{aligned} \tag{2.14}$$

The system has a line of equilibria $(0, z)$. Points on the line are stable if $z < 0$ and unstable if $z > 0$. Trajectories near the bifurcation point $(0, 0)$ form parabolas tangent to the line of equilibria at the origin. Moreover, the direction of the flow is reversed on opposite sides of the equilibrium line. A system with a transcritical bifurcation without parameters at (y^*, z^*) satisfies the following conditions (see Theorem 4.1 in [86]):

(TBWP1) $g(y^*, z) = 0$ and $h(y^*, z) = 0$.

(TBWP2) $\big(\partial g/\partial y\big)(y^*, z^*) = 0$.

(TBWP3) $\big(\partial^2 g/\partial z \partial y\big)(y^*, z^*) \neq 0$.

(TBWP4) $\big(\partial h/\partial y\big)(y^*, z^*) \neq 0$.

Condition (TBWP1) states that there exists a manifold of equilibria. Condition (TBWP2) describes the loss of normal hyperbolicity at the bifurcation point (y^*, z^*) due to the transverse eigenvalue crossing zero. Condition (TBWP3) states that transverse eigenvalue crosses zero transversally with non-vanishing speed. Finally, condition (TBWP4) controls the nondegeneracy of the Jacobian (2.13).

Figure (2.3) compares the classical transcritical bifurcation diagram of the system $\dot{y} = \mu y$ to the phase space for the normal form of the transcritical bifurcation without parameters. In both cases, we denote the stable region of the line of equilibria as a solid blue line and the unstable region as a red dashed line. The black arrows correspond to trajectories of the system. The bifurcation diagram for $\dot{y} = \mu y$ (Figure 2.3A) shows how the line of equilibria $y = 0$ in the extended space (y, μ) is stable for $\mu < 0$ and unstable for $\mu > 0$. In this case, the extended space is foliated by the leaves $\mu = c \in \mathbb{R}$. Compare this to the transcritical bifurcation without parameters (Figure 2.3B), in which the state z controls the stability of the line of equilibria $y = 0$. Solutions now drift to the right due to the nontrivial dynamics of z. The stable and unstable manifolds of the bifurcation point at the origin are represented with solid green and magenta lines, respectively. In the upper half-plane, the stable manifold of the bifurcation point separates the phase space into a region containing solutions that reach a point in the line of equilibria, and a region with solutions that move away from the line of equilibria. In the lower half-plane solutions form manifolds of heteroclinic orbits connecting unstable with stable points in the line of equilibria.

When the system (2.11) depends on parameters, there can be both a classical bifurcation and a bifurcation without parameters.

A *transcritical bifurcation without parameters with drift singularity* has normal form (see Chapter 8.1 in [86]):

$$\begin{aligned} \dot{y} &= yz \\ \dot{z} &= -ay^2 + byz + \mu y \end{aligned} \tag{2.15}$$

where $a, b \neq 0$ and μ are scalar parameters. The system has a line of equilibria $(0, z)$ and an isolated equilibrium at $(\mu/a, 0)$. Points on the line of equilibria are stable if $z < 0$ and unstable if $z > 0$, while the isolated equilibrium is unstable if $\mu < 0$

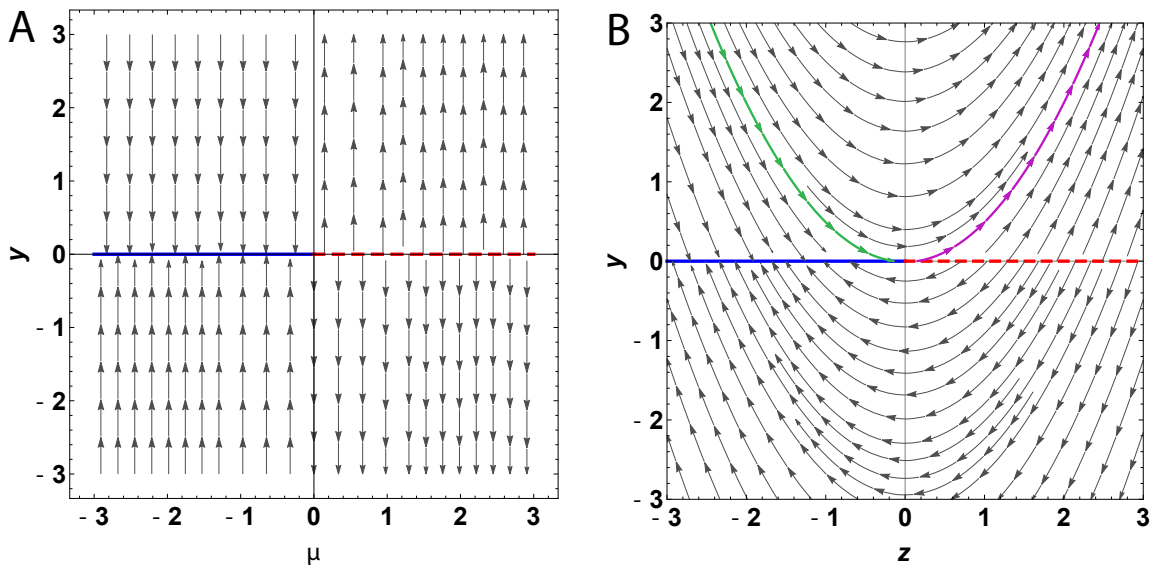

Figure 2.3: A. Bifurcation diagram for $\dot{y} = \mu y$. B. Transcritical bifurcation without parameters phase space.

and stable if $\mu > 0$. For $\mu \neq 0$, trajectories near the bifurcation point $(0, 0)$ form parabolas tangent to the line of equilibria at the origin. A system with a transcritical bifurcation without parameters with singular drift satisfies (TBWP1)–(TBWP3) for all μ and

(TBWP5) At $\mu = 0$, condition (TBWP4) fails, $\big(\partial h/\partial y\big)(y^*, z^*, 0) = 0$.

(TBWP6) $\big(\partial h/\partial \mu \partial y\big)(y^*, z^*) \neq 0$.

Conditions (TBWP5) and (TBWP6) are analogous to conditions (SN1) and (T2) for the classical transcritical bifurcation that describe how the isolated equilibrium loses stability as μ crosses zero. The parameters a and b in (2.15) correspond to the determinant and the trace of the linearization around the equilibrium, and determine the type of trajectories near the equilibrium, i.e., the equilibrium could be a saddle, node, or spiral.

Figure (2.4) shows the transcritical bifurcation without parameters with drift singularity when the isolated equilibrium is a node. We represent the isolated equilibrium as a solid circle when the equilibrium is stable and as an empty circle when the

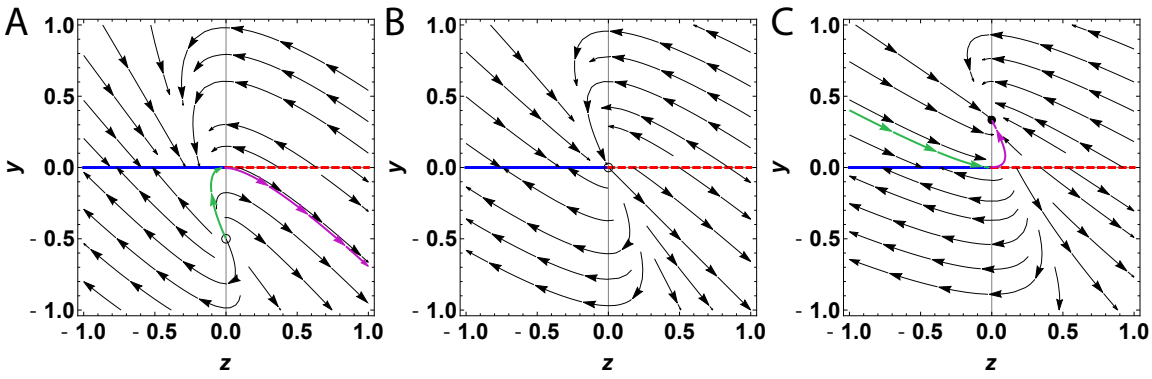

Figure 2.4: Transcritical bifurcation without parameters with bifurcating node drift singularity. A. $\mu < 0$. B. $\mu = 0$. C. $\mu > 0$. For all panels, we set parameters in (2.15) to $a = 1$ and $b = -2.5$.

equilibrium is unstable or degenerate. We illustrate solutions that flow in the stable manifold of the bifurcation point ($y = 0$ and $z = 0$) as green solid curves and solutions that flow in the unstable manifold of the bifurcation point as magenta solid curves. For $\mu < 0$, the node lies in the lower half of the plane and is unstable (Figure 2.4A). All trajectories in the upper half of the plane reach a point in the line of equilibria $y = 0$. For $\mu = 0$, the node lies at $y = 0$ (Figure 2.4B). For $\mu > 0$, the node lies in the upper half of the plane and is stable (Figure 2.4C). Trajectories below the stable manifold of the origin reach a point in the line of equilibria while trajectories above the stable manifold of the origin reach the stable node. Note that when $\mu \neq 0$, the flow close to the origin resembles the transcritical bifurcation without parameters (see Figure 2.3B).

In Chapter 4, we study the SIRI epidemiological model in well-mixed settings and show that the dynamics of the model can be explained through a transcritical bifurcation without parameters with node drift singularity. In Chapter 5, we study the SIRI model in network topologies and discuss how the dynamics of the model extend the idea of a transcritical bifurcation without parameters with node drift singularity to higher-dimensional systems.

2.4 Excitability Dynamics

Consider a nonlinear dynamical system of the form (2.7) in which one of the parameters $\mu_i = I$ for some $i \in \{1, 2, \ldots, K\}$ represents the input to the system. We say the system is *excitable* if there exists a threshold value of I, I_T, such that: 1) subthreshold input perturbations (i.e., short-lived inputs where $I < I_T$) lead to small responses away from a nominal resting state, and 2) suprathreshold perturbations (i.e., short-lived inputs where $I > I_T$) lead to responses that are much larger than the perturbation that elicited it. In an excitable system, the system eventually returns to the same nominal state after responding to a suprathreshold perturbation.

Excitability is a nonlinear phenomena that has been observed in different physical and natural systems across fields, including neural systems, cardiovascular tissue, gene regulatory circuits, lasers, and chemical reactions, to mention a few [92–96]. In each of these cases, the system has a stable *resting* state that attracts solutions when I is at its nominal value I_0, e.g, $I_0 = 0$. In Figure 2.5 we illustrate how, in an excitable system, subthreshold perturbations (light gray) marginally shift the state of the system away from the resting state, eliciting a small transient response as the system returns to rest. In contrast, for suprathreshold perturbations (dark gray), the system exhibits a large excursion in phase space, usually visible as a "spike" in the time-series of one of the state variables called the *output* variable, consisting of an *excited* and *recovery* phase before returning to rest. We make use of an excitable system called the FitzHugh-Nagumo system in Chapter 3 to model the response of a group of forager ants inside the ants' nest entrance chamber in response to interactions with incoming foragers carrying food. We explore the FitzHugh-Nagumo dynamics in more detail in Section 2.4.1.

From a dynamical systems point of view, there are a number of different mechanisms, associated with bifurcations, which can result in a nonlinear system with

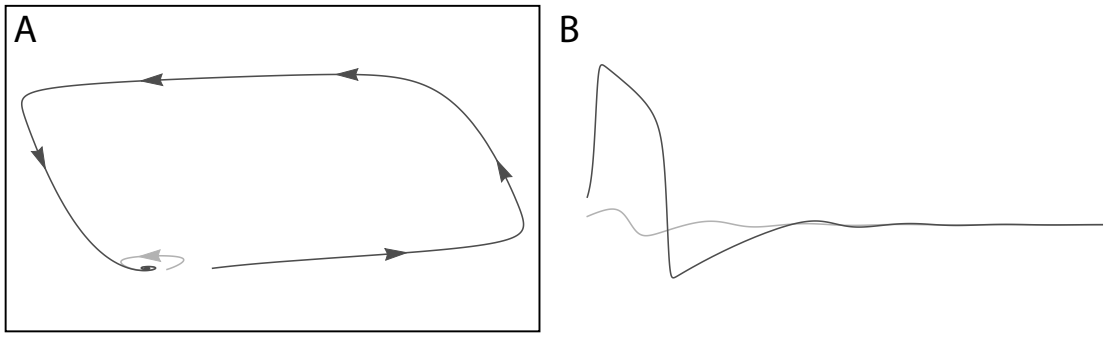

Figure 2.5: Excitability dynamics. A. Small input perturbations (light gray) lead to small amplitude responses in phase space, while large input perturbations (dark gray) lead to trajectories that follow a large amplitude orbit before returning to rest. B. The small amplitude response appears as a small transient in the time-series of the output variable while the large orbits appear as a "spike" consisting of an excited phase and a recovery phase.

excitability. In many cases, the study of excitability is restricted to two dimensional models to facilitate analysis. In these planar models, excitability is commonly achieved by the introduction of a fast positive feedback that is responsible for a rapid initial response in the output variable and a slower time-scale negative feedback responsible for bringing the system back to rest [97, 98].

Planar excitability can be understood and studied geometrically. Short-lasting large amplitude responses can be understood as the system having one or more large amplitude orbits in phase space that starts and ends near either a stable equilibrium or a small amplitude limit cycle representing the resting state. In most cases, the input parameter I is a bifurcation parameter and, for values of I larger than a bifurcation value I_T, the system has a stable limit cycle that leads to periodic behavior and to sequences of "spikes" in the time-series of the output variable. If the spiking frequency of the system increases continuously with I, the system is said to be of *Class I* [98, 99]. A Class I system can generate arbitrarily low spiking frequencies which allows it to integrate the input strength of a step input. If the spiking frequency is more or less

constant across the range of valid suprathreshold values of I, the system is said to be of *Class II* [98, 99]. A Class II system exhibits resonance when the input is a periodic perturbation of a specified frequency.

Class I and Class II excitable systems are characterized by the type of bifurcation that the system undergoes as I increases [98, 100, 101]. In a Class I excitable system, there exists a stable equilibrium that disappears via a *saddle-node on invariant circle* bifurcation (SNIC) (see [98] for details), whereas in a Class II excitable system, there exists a stable equilibrium that loses stability via a subcritical or supercritical Hopf bifurcation.[3] In the remainder of this Section we describe the FitzHugh-Nagumo model, a Class II excitable system which is a key component in the model we develop and analyze in Chapter 3.

2.4.1 FitzHugh-Nagumo Model

The FitzHugh-Nagumo model [103, 104] is a two-dimensional reduction of the Hodgkin-Huxley model [92] for neuronal excitability. The model was developed to capture the essential geometry required for excitability in a small-dimensional system. The dynamics of the FitzHugh-Nagumo model are given by [105]

$$\begin{aligned} \dot{x}_1 &= x_1 - \frac{x_1^3}{3} - a - x_2 + I \\ \dot{x}_2 &= \epsilon(x_1 - bx_2), \end{aligned} \tag{2.16}$$

where x_1 represents the neuronal cell membrane voltage, x_2 is a recovery variable, $0 < \epsilon \ll 1$ is the time separation constant, and a, b are parameters. The x_1-nullcline is cubic, while the x_2-nullcline is linear. This leads to the nullclines having one, two, or three intersections depending on the values of the parameters. For the rest of the

[3] The classification of excitability based on only the SNIC and the Hopf bifurcations is not a comprehensive description of all the mechanisms that can give rise to excitable behavior. We refer the reader to [98, 102] for a more expansive discussion of the topic.

Section we assume $0 < a \leq 1$ and $0 < b < 1$ which restrict the nullclines of (2.16) to intersect at a single point corresponding to the unique equilibrium $\mathbf{x}^* = (x_1^*, x_2^*)$ where

$$x_2^* = \frac{x_1^*}{b}, \qquad \frac{x_1^{*3}}{3} + \frac{(1-b)x_1^*}{b} + a = I. \tag{2.17}$$

As I increases, the cubic x_1-nullcline shifts upwards, changing the location of the equilibrium. This single equilibrium point together with the I dependent cubic nullcline are sufficient to generate excitability.

The linearization of (2.16) at (x_1^*, x_2^*) is given by

$$J = \begin{bmatrix} 1 - x_1^{*2} & -1 \\ \epsilon & -b\epsilon \end{bmatrix} \tag{2.18}$$

and has eigenvalues

$$\lambda_{1,2} = \frac{(1 - x_1^{*2} - b\epsilon) \pm \sqrt{(1 - x_1^{*2} - b\epsilon)^2 - 4\epsilon(1 + bx_1^{*2} - b)}}{2}. \tag{2.19}$$

For $I = 0$, it follows that $\lambda_{1,2} < 0$; the equilibrium is a stable node. The equilibrium loses stability at $x_1^* = \pm\sqrt{1 - b\epsilon}$. Inserting this value into (2.17) gives the expression for the bifurcation values

$$I_{1,2} = a \mp \left(\frac{(1 - b\epsilon)^{3/2}}{3} + \frac{1-b}{b}\sqrt{1 - b\epsilon} \right), \tag{2.20}$$

at which $\lambda_{1,2} = \pm i\omega$ where $\omega^2 = 4\epsilon(1 - b^2\epsilon)$. The first solution I_1 correspond to the bifurcation point at which $\mathbf{x}^*$ loses stability while the second solution I_2 corresponds to a second bifurcation point at which $\mathbf{x}^*$ regains stability.

As the input I increases from zero past point I_1, the dynamics (2.16) go through a complex bifurcation stage within a very small range of I values in which a supercritical saddle-node of limit cycles and a subcritical Hopf bifurcation result in the equilibrium

losing stability, and the appearance of a stable limit cycle (see [106] and references within for details). This stable limit cycle remains for a large range of values of I until the point $I_2 > I_1$ at which dynamics experience the reverse process in which a subcritical saddle-node bifurcation of limit cycles and a supercritical Hopf bifurcation lead to the equilibrium point gaining stability and the disappearance of the limit cycle. Thus, as I varies, the FitzHugh-Nagumo model (2.16) exhibits three states: resting ($I < I_1$), oscillatory ($I_1 < I < I_2$), and saturated ($I > I_2$) (see Figure 2.6). The resting and saturated states are topologically equivalent.

How this complex bifurcation structure gives rise to excitability can be better understood by studying the geometry of the system. It can be shown that the bifurcation points I_1 and I_2 approximately correspond to the values of I at which the linear nullcline intersects the cubic nullcline at the valley or the crest of the cubic nullcline (see solid orange curves in Figure 2.6). These points divide the cubic nullcline into left, center, and right branches. When the linear nullcline intersects the cubic nullcline at a point in the middle branch, $\mathbf{x}^*$ is unstable. If the linear nullcline intersects the cubic nullcline at a point on the right or left branches, then $\mathbf{x}^*$ is stable. This geometry results in small input perturbations leading to trajectories that remain within the region of attraction of the stable equilibrium point, and in larger input perturbations leading to trajectories that cross the middle branch of the cubic nullcline and take a long path in phase space before returning to $\mathbf{x}^*$ (Fig. 2.6A and D). The difference in the response to small and large input perturbations is the result of a soft threshold in which the cubic nullcline acts as a "quasi-threshold" through a *canard trajectory*, where there is a very fast transition within an exponentially small range of a I values [106, 107].

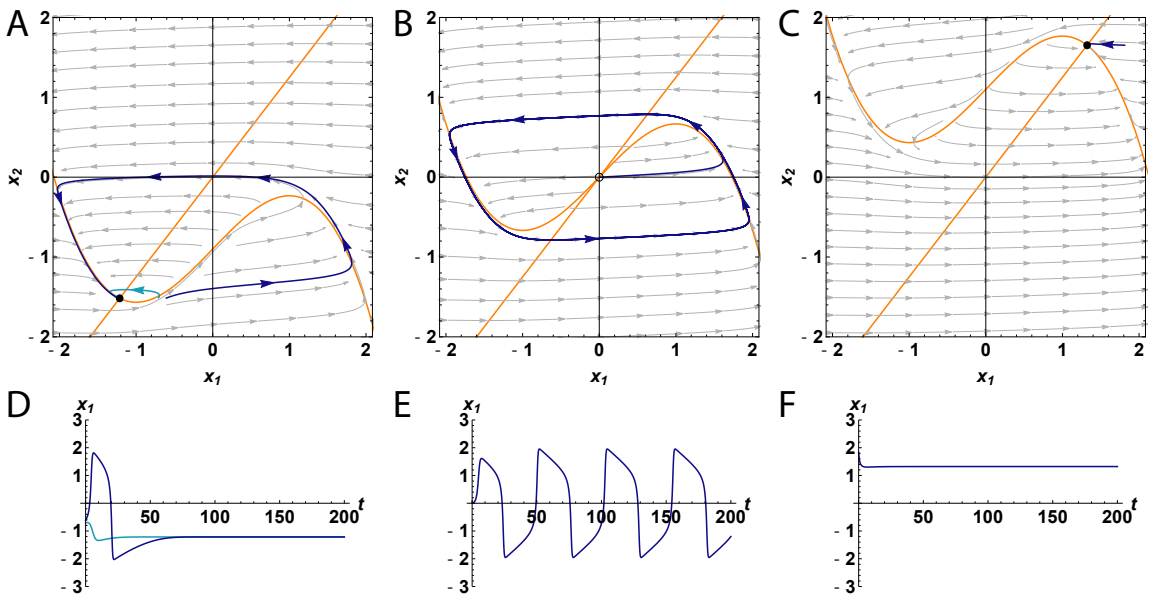

Figure 2.6: Fitzhugh-Nagumo dynamics with $a = 0.9$, $b = 0.8$, and $\epsilon = 0.05$. The bifurcation values $I_1 \approx 0.34$ and $I_2 \approx 1.46$ denote the input values at which the model transitions to a new state. Orange lines represent the system nullclines. A and D. Resting state ($I = 0$) phase space and x_1 time-series. B and E. Oscillatory state ($I = 0.9$) phase space and x_1 time-series. C and F. Saturated state ($I = 2$) phase space and x_1 time-series.

2.5 Queueing Theory

A *queueing system* describes the process through which customers arrive for service, wait for service if no servers are available, and leave the system after being served. The aim of queueing theory is to develop mathematical models that can predict the lengths of queues and waiting times in queueing systems. Although queueing theory was developed with the intention of helping businesses make better decisions about the resources needed to provide a service, queueing theory results apply to any system that can be thought of as a queueing system. In Chapter 3 we use queueing theory to describe how harvester ants forage for seeds outside the nest.

Formally, a queueing system is characterized by the following six descriptors [108]:

1. **Arrival pattern of customers**: probability distribution describing the times between successive customers arrivals (*interarrival* times).

2. **Service pattern of servers**: probability distribution describing customer service times.

3. **Queue discipline**: mechanism through which customers are selected for service when a queue has formed.

4. **System capacity**: maximum possible length of the queue.

5. **Number of service channels**: number of servers that can provide service to customers.

6. **Number of service stages**: number of service stages that a single customer has to go through before service is completed.

A queueing system is usually described in Kendall notation [109], which consists of a series of symbols and slashes of the form $A/B/X/Y/Z$ where A indicates the interarrival-time distribution, B indicates the service time distribution, X the number of servers, Y the system capacity, and Z the queue discipline. It is common to describe queueing systems with only the first three descriptors. In these cases, it is assumed that $Y = \infty$ and $Z = FCFS$, where $FCFS$ denotes the first-come-first-served queueing discipline. For example an $M/G/5$ queue has exponential interarrival-times, arbitrary service times, 5 servers, an infinite system capacity, and a first-come-first-served queue discipline.[4] The most common queueing models assume that interarrival times and service times follow an exponential distribution (i.e., $M/M/n$). This corresponds to cases in which both the arrival process and the service process are described by a homogeneous Poisson process.

[4]The letter M stands for the Markovian or memoryless property of the exponential distribution. The letter G stands for general.

A *homogeneous Poisson process* with rate $r > 0$ describes the probability distribution of a *counting process* $\{N(t) \in \mathbb{N} | t \geq 0\}$,

$$P(N(t) = n) = \frac{(rt)^n}{n!} e^{-rt} \tag{2.21}$$

with mean and variance rt; (2.21) describes the probability of having n events in the counting process by time t. A Poisson process has the *memoryless property* meaning that the time between events is independent of the number of events that have already taken place.

Let $Q(t)$ be the number of customers in the system (either waiting in queue or being served) at time t. For an $M/G/\infty$ queue, the probability of having n customers in the system at steady-state $P(Q(\infty) = n)$ is given by a Poisson process with rate $q = r/\mu$ where r is the mean interarrival-time and $1/\mu$ is the mean service time [108]. Therefore, at steady-state, the expected system size is $\mathbb{E}[Q(\infty)] = q$ and, since there are an infinite number of servers, the expected number of customers waiting in the queue is zero and the average waiting time in the system (from arrival to departure) is equal to the average service time $1/\mu$.

The assumption that the arrival process can be described by a Poisson process with a fixed rate is, in many cases, not realistic as many arrival processes vary with time. A *non-homogeneous Poisson process* with rate function $r(t) > 0$ extends the idea of a homogeneous Poisson process to time-varying rates, and describes the probability distribution of a counting process $N(t)$ as,

$$P(N(t + t') - N(t) = n) = \frac{\Lambda^n}{n!} e^{-\Lambda} \tag{2.22}$$

where

$$\Lambda = \int_t^{t+t'} r(\tau)\, d\tau. \tag{2.23}$$

For an $M(t)/G/\infty$ queue, where $M(t)$ denotes a non-homogeneous Poisson process, with arrival rate function $r(t)$, and where the service time random variable S has the cumulative distribution function $F(s)$, we get the following results (see [110] for details):

- For each time t, $Q(t)$ has a Poisson distribution with mean

$$\mathbb{E}[Q(t)] = \int_0^\infty r(t-u)(1-F(s))\ ds.$$

- The departure process is a non-homogeneous Poisson process with rate function

$$r'(t) = \mathbb{E}[r(t-S)].$$

2.6 Graph Theory

In dynamical systems with interconnected agents, graph theory provides a mathematical framework to study the role of the interconnection topology on individual and group level dynamics.

A *weighted digraph* $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ consists of a set of *nodes* $\mathcal{V}$ and a set of *edges* $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$, where each edge $(j,k) \in \mathcal{E}$ from node $j \in \mathcal{V}$ to node $k \in \mathcal{V}$ has an associated *weight* $w_{jk} > 0$. We denote the set of *neighbors* of node j as $\mathcal{N}_j = \{k \in \mathcal{V} | (j,k) \in \mathcal{E}\}$.

In multi-agent systems, it is common to represent agents as nodes and an interaction between agents as an edge, where the weight of the edge represents the strength of the interaction. The directionality of an edge represents cases in which the information transfer that results from the interaction between agents is not bidirectional. For example, an edge $(j,k) \in \mathcal{E}$ might represent agent j paying attention to agent k but not vice-versa; agent j receives information from agent k but agent k does not receive information from agent j.

The structure of a weighted digraph $\mathcal{G}$ with N nodes is captured by the *weighted adjacency matrix* $A = \{a_{jk}\} \in \mathbb{R}^{N \times N}$ with

$$a_{jk} = \begin{cases} w_{jk} & (j,k) \in \mathcal{E} \\ 0 & \text{otherwise.} \end{cases}$$

A *directed path* between two agents $j, k \in \mathcal{V}$ is a sequence of non-repeating edges in $\mathcal{E}$ that start in node j and end in node k. A digraph is *strongly connected* if there exists a directed path from any node $j \in \mathcal{V}$ to any other node $k \in \mathcal{V}$ (see Figure 2.7A). The adjacency matrix A of digraph $\mathcal{G}$ is *irreducible* if $\mathcal{G}$ is strongly connected.

A *weighted undirected* digraph is a weighted digraph $\mathcal{G}$ with *undirected edges* where if $(j,k) \in \mathcal{E}$, then $(k,j) \in \mathcal{E}$ and $a_{kj} = a_{jk}$. An undirected graph is connected if there exists a sequence of undirected edges from any node $j \in \mathcal{V}$ to any other node $k \in \mathcal{V}$. A digraph is *weakly connected* if replacing every edge in $\mathcal{E}$ with an undirected edge results in an undirected graph that is connected (see Figure 2.7B). Every strongly connected digraph is weakly connected, but the opposite is not true. In a weakly connected digraph there may exist a pair of nodes for which there is no directed path connecting them. A subgraph $\mathcal{G}' = (\mathcal{V}', \mathcal{E}')$ of digraph $\mathcal{G}$ is a digraph with $\mathcal{V}' \subset \mathcal{V}$ and $\mathcal{E}' \subset \mathcal{E}$. A digraph is *disconnected* if it contains at least two isolated subgraphs (see Figure 2.7C).

2.7 Metzler Matrices

A real square matrix $M = \{m_{jk}\}$ is Metzler if its off diagonal entries are nonnegative, $m_{jk} \geq 0$, $j \neq k$.[5] Metzler matrices appear prominently in the study of *positive systems* (systems in which the state variables are always nonnegative). A linear system $\dot{\mathbf{x}} = A\mathbf{x}$

[5]Metzler matrices are also called *quasipositive* or *essentially positive* matrices. If M is a Metlzer matrix, then $-M$ is an *M-matrix*, a matrix with nonpositive off-diagonal entries [111].

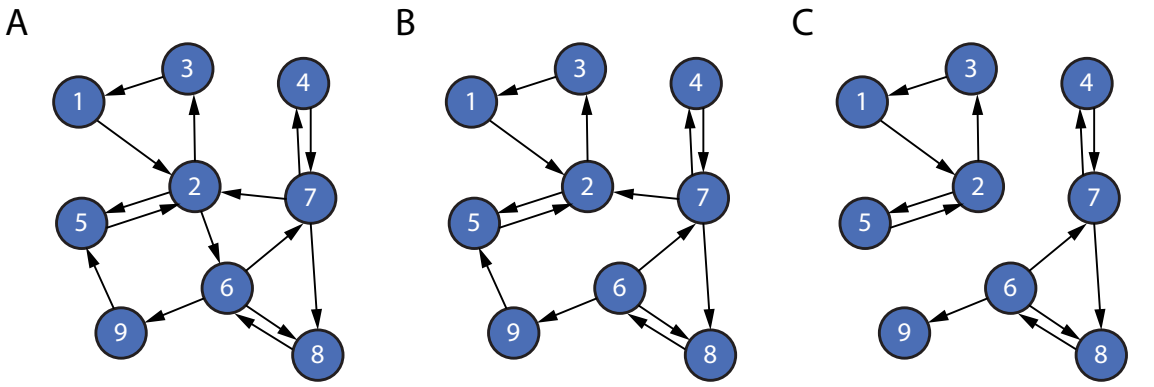

Figure 2.7: A. Strongly connected digraph. There exists a directed path between any two nodes. B. Weakly connected digraph. There exists an undirected path between any two nodes but a directed path may not exist (e.g., nodes 2 and 7). C. Disconnected digraph. There exists no directed or undirected path between some nodes (e.g., nodes 2 and 7).

is a positive linear system if A is Metzler [112]. Positive systems are ubiquitous across many fields, including economics, chemistry, biology, and engineering, as many systems have states that represent physical quantities that cannot be negative. In Chapter 5 we study the SIRI epidemiological model on network topologies, where we make repeated use of the results for Metzler matrices presented in this Section.

A *positive matrix* is a matrix with positive entries $M \gg \bar{\mathbf{0}}$. Similarly, a *nonnegative matrix* is a matrix with nonnegative entries $M \succeq \bar{\mathbf{0}}$. Any positive matrix is nonnegative, but the opposite is not always true. The study of positive and nonnegative matrices is called Perron-Frobenius theory after the contributions of Oskar Perron and Ferdinand Georg Frobenius. In 1907, Perron published his treatment of positive matrices [113, 114] and in 1912, Frobenius extended the theory to nonnegative matrices [115]. Two of the fundamental results of Perron-Frobenius theory are the Perron-Theorem for positive matrices, and the Perron-Frobenius Theorem for nonnegative irreducible matrices [116].

Both theorems state that any positive or nonnegative matrix M has a real eigenvalue $\lambda_{max}(M)$, called the *leading eigenvalue*, with magnitude equal to the spectral radius of the matrix $\lambda_{max}(M) = \rho(M)$. If M is positive, then $\lambda_{max}(M)$ has multiplic-

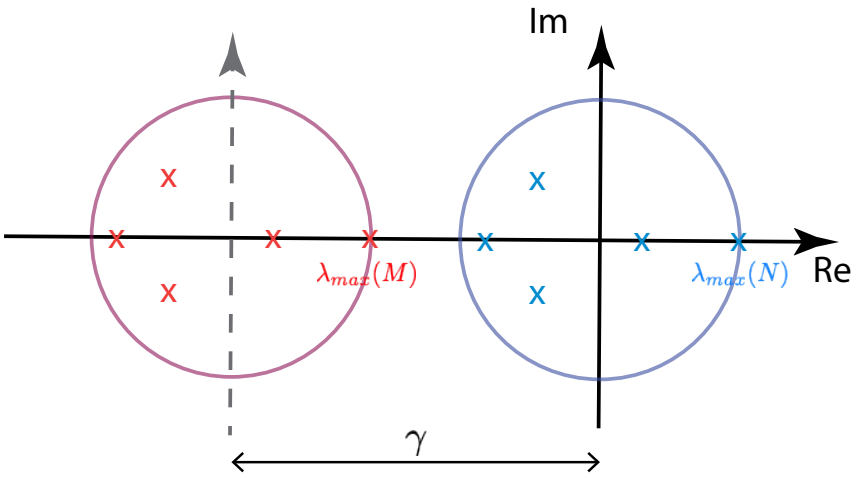

Figure 2.8: Comparison of spectrum of nonnegative matrix N (cyan crosses) and its spectral radius (blue circle) against the spectrum of the Metzler matrix $M = N - \gamma\mathbb{I}$ (red crosses). The red circle represents the γ-shifted spectral radius of N.

ity one and the corresponding eigenvector, called the *leading eigenvector*, is positive $\mathbf{v}_{max} \gg \mathbf{0}$. If M is nonnegative, then the leading eigenvector is nonnegative $\mathbf{v}_{max} \succ \mathbf{0}$.

The matrix M is irreducible if, when interpreted as the adjacency matrix of a digraph $\mathcal{G}$, $\mathcal{G}$ is strongly connected. If M is nonnegative and irreducible, then $\lambda_{max}(M)$ has multiplicity one. In addition, in both the positive and nonnegative cases, M has no other nonnegative eigenvectors besides multiples of $\mathbf{v}_{max}$. A nonnegative irreducible matrix M also has the property that $\lambda_{max}(M)$ increases/decreases as any entry m_{jk} increases/decreases [117].

Any nonnegative matrix N can be transformed into a Metzler matrix $M = N - \gamma\mathbb{I}$ by picking $\gamma \geq |\min_j m_{jj}|$ (see Figure 2.8). As a result, the spectrum of the Metzler matrix M looks like the shifted spectrum of the nonnegative matrix N, and M has a leading eigenvalue $\lambda_{max}(M) = \lambda_{max}(N) - \gamma$. Most results on the spectral properties of nonnegative matrices extend to Metzler matrices (see Theorem 6.2.3 in [111], Chapters 5 and 6 in [112], and Chapter 8 in [116]):

Proposition 1. *Let M be a Metzler matrix. Then,*

1. *$\lambda_{max}(M) \in \mathbb{R}$. If M is irreducible, $\lambda_{max}(M)$ has multiplicity one.*

2. *Let* $\mathbf{w}^T$ *and* $\mathbf{v}$ *be left and right eigenvectors corresponding to* $\lambda_{max}(M)$*. Then,* $\mathbf{w}, \mathbf{v} \succeq \mathbf{0}$*. If* M *is irreducible, then* $\mathbf{w}, \mathbf{v} \gg \mathbf{0}$*, and every other eigenvector of* M *has at least one negative entry.*

3. *Let* M_{min}, M_{max} *be irreducible Metzler matrices where* $M_{min} \prec M \prec M_{max}$*, then*

$$\lambda_{max}(M_{min}) < \lambda_{max}(M) < \lambda_{max}(M_{max}).$$

The existence of the leading eigenvalue and eigenvector of a Metzler matrix M leads to the following necessary and sufficient conditions for M to be Hurwitz (i.e., for all eigenvalues of M to have negative real part).

Proposition 2. *Let* M *be a Metzler matrix. Then, the following statements are equivalent:*

1. M *is Hurwitz.*

2. *There exists a vector* $\mathbf{v} \gg \mathbf{0}$ *such that* $M\mathbf{v} \ll \mathbf{0}$*.*

3. *There exists a vector* $\mathbf{w} \gg \mathbf{0}$ *such that* $\mathbf{w}^T M \ll \mathbf{0}$*.*

The *regular splitting* property of Metzler matrices associates the sign of the leading eigenvalue of a Metzler matrix to the spectral radius of a related matrix.

Definition 2.7.1 (Regular Splitting)**.** *Let* M *be a Metzler matrix.* $M = T + U$ *is a regular splitting of* M *if* $T \succeq \bar{\mathbf{0}}$ *and* U *is a Hurwitz Metzler matrix.*

Proposition 3. *Let* M *be a Metzler matrix and let* $M = T + U$ *be a regular splitting. Then,*

1. $\lambda_{max}(M) < 0$ *if and only if* $\rho(TU^{-1}) < 1$*.*

2. $\lambda_{max}(M) = 0$ *if and only if* $\rho(TU^{-1}) = 1$*.*

3. $\lambda_{max}(M) > 0$ *if and only if* $\rho(TU^{-1}) > 1$.

We use the leading eigenvalue of the regular splitting of Metzler matrices in Chapter 5 to define scalar metrics that determine global transient and steady-state behaviors for the network SIRI dynamics.

Chapter 3

Regulation of Harvester Ant Foraging: Excitability and Feedback

In this Chapter we summarize results from our collaboration with biologist Deborah Gordon from Stanford University, presented in Part II: Chapter 8 which appears as Pagliara, Gordon, and Leonard [43]. Gordon has been studying red harvester ants in the deserts of the Southwestern United States for over 30 years. Her pioneering research on how individuals ants within the colony perform different tasks throughout their lifetime has shown that ant task allocation and performance is a dynamic process that self-regulates according to the colony's needs.

Our collaboration focuses on the robust and flexible collective behaviors observed in the regulation of foraging in red harvester ants. We provide a brief introduction to the problem in Section 1.1 and a detailed introduction in Section 8.1.

In Section 3.1 we discuss important features of the regulation of the foraging problem in red harvester ants. In Section 3.2 we present results from three years of

field observations. In Section 3.3 we introduce a low-dimensional, simply parameterized model, developed in Chapter 8, for the regulation of foraging rates. We use the excitability dynamics of the FitzHugh-Nagumo equations (FN) to model the fast activation mechanism of available foragers inside the nest. We introduce a parameter into the FN model which controls the frequency of oscillations and use it to represent the volatility of available foragers inside the nest to interactions with food-bearing incoming foragers. To model the slow process of foraging for seeds outside the nest, we use a random delay distribution. In Section 3.4 we discuss our use of excitability dynamics. In Section 3.5, we summarize results for the open-loop model in which the incoming foraging rate is the input to the system and the outgoing foraging rate is the output. We show that the volatility parameter controls the nonlinear map between the incoming and the outgoing foraging rates. In Section 3.6, we summarize results for the closed-loop model dynamics in which outgoing foragers return to the nest after foraging. We show the slow feedback of returning ants leads to an equilibration of the foraging rates at a stable rate determined by the volatilty parameter. In Section 3.7 we propose that foragers adapt their volatility after a first exposure to the environmental conditions outside the nest, and discuss how this mechanism, which depends on two volatility values, adjusts the colony foraging rates in response to different conditions. In Section 3.8 we compare simulations of the model with field observations of foraging activity. In Section 3.9 we present a novel discussion, not published in [57] of Chapter 9, on the existence of four distinct behavioral regimes in the closed-loop model with the adaptation mechanism, where each regime is characterized by the two volatility values.

3.1 Ecology, Interactions, and Feedback in Red Harvester Ants

Living in the desert, harvester ant colonies must constantly obtain water to survive. Harvester ants obtain water by metabolizing the fat in the seeds they eat, but foragers lose water through desiccation while outside the nest [32, 33]. Therefore, colonies must balance the trade-off between foraging for seeds to obtain water and conserving water by not foraging. To balance this trade-off, colonies must regulate their foraging activity across different timescales according to food availability and environmental conditions. How they achieve balance and thrive in these extreme environmental conditions is not well understood [14, 118].

The collective behaviors that ant colonies exhibit are highly dependent on the context in which they live [119]. For example, non-harvester ant species that live in competitive environments with clumped food resources, make use of pheromone trails that trace the shortest path towards the food and rapidly increase the number of foragers at the food source [31, 120, 121]. This increases the probability that the colony reaches and consumes the resources before anyone else does. In contrast, harvester ants do not use pheromone trails, probably because they consume resources that are scattered across space and deplete over the course of weeks, not minutes. As such, harvester ants face different evolutionary pressures, where minimizing water loss through desiccation is of the utmost importance.

Colonies regulate the rate of outgoing foragers via interactions with incoming food-bearing foragers [34–38, 122]. The rate of interactions provides available foragers with limited, noisy information on the quality of the foraging conditions outside the nest. Available foragers inside the nest integrate this noisy information in order to decide whether or not to forage [38].

The interactions mechanism regulates the outgoing foraging rate across seconds-to-minutes long timescales in response to food availability, but it does not explain how colonies adjust foraging activity across hour-long timescales in response to different environmental conditions. A fundamental component of the foraging process that has so far not been studied is the feedback that results from outgoing foragers returning to the nest after foraging and becoming available foragers once again. This feedback loop takes place across minutes-to-hour timescales and is, therefore, much slower than the fast activation of foragers inside the nest. Feedback in dynamical systems provides a way to stabilize and control the state of a system, suggesting that the slow feedback of returning foragers may play a fundamental role in the regulation of foraging activity across minutes-to-hours timescales.

Colonies vary in how they regulate foraging activity in response to different environmental conditions [41,123–126]. This variation leads to differences in reproductive success between colonies. Colonies that exhibit large differences in foraging activity between dry and humid days are reproductively more successful than colonies that exhibit smaller differences [41]. We hypothesize that these differences in the collective behavior of colonies result from differences in how individual foragers assess the quality of the environmental conditions. Understanding the mechanisms that colonies use to respond to different environmental conditions would provide valuable insight into how behavioral differences between individuals belonging to different colonies can result in differences in collective behavior between colonies.

3.2 Field Observations: Transients and Quasi Steady-States

In collaboration with Stanford University biologist Deborah Gordon, we performed three years of field observations of red harvester ant colonies from 2015 to 2017.

Observations were made at the site of a long term study near Rodeo, New Mexico, USA, where Prof. Gordon has been studying harvester ant colonies for more than 30 years.

We recorded foragers crossing a line perpendicular to the foraging trail at a distance of about 1 m. from the nest entrance. The recordings were made from the beginning of the foraging period in early morning until around noon. From these recordings, we obtained timestamps for every incoming and outgoing forager (see Section 8.2.1 for details).

Definition 3.2.1. *The sequence of times at which incoming foragers cross the line is* t_i^{in}, $i \in \mathbb{N}$, *and the sequence of times at which outgoing foragers cross the line is* t_j^{out}, $j \in \mathbb{N}$.

Definition 3.2.2. *The incoming foragers indicator function is* $\lambda_{in}(t) = \sum_{i=1}^{n} \delta(t - t_i^{in})$ *and the outgoing foragers indicator function is* $\lambda_{out}(t) = \sum_{j=1}^{m} \delta(t - t_j^{out})$, *where* $\delta(\cdot)$ *is the Dirac delta function and,* n *and* m *are the indices of the last incoming and outgoing forager, respectively, before time* t.

Definition 3.2.3. *The estimated instantaneous incoming and outgoing foraging rates, in units of ants/sec, are*

$$r_{in}(t) = \int_{-\infty}^{\infty} w(\zeta)\lambda_{in}(t-\zeta)\ d\zeta, \qquad r_{out}(t) = \int_{-\infty}^{\infty} w(\zeta)\lambda_{out}(t-\zeta)\ d\zeta, \tag{3.1}$$

where

$$w(t) = \begin{cases} 1/\Delta t & \text{if } -\Delta t/2 \le t \le \Delta t/2 \\ 0 & \text{otherwise.} \end{cases} \tag{3.2}$$

Remark 1. The instantaneous incoming and outgoing foraging rates are computed via a sliding window filter with window $\Delta t = 300$ sec. We selected the size of the sliding window to remove noise but preserve the interesting dynamic features of the foraging rates across tens of minutes to hours.

We collected data for 16 colonies across 25 days for a total of 47 data sets.[1] This was no small feat. Each data set required identifying potential colonies that seemed likely to forage on that given day. These colonies would then be checked every few minutes in order to determine the most likely colony to forage. The recording equipment would then be set up around the most promising colonies and, in many cases, we would have to change the setup once a colony started foraging because they had decided to use a different foraging trail to the one we thought they would use. Sometimes, one of the colonies selected displayed extremely low foraging rates that were not useful for our analysis. At least two people were required to keep track of the camera batteries or manually count ants with a tablet in order to prevent data loss. Each data set took about 6 hours to obtain. In addition, all this work had to be done in extreme temperatures while wearing snake-guards around one's legs for protection against rattlesnakes.[2]

The data sets show that there is large variation in foraging activity across colonies (on the same day), and across days (for the same colony). Despite this variation, we found that across colonies and days, the foraging rates $r_{in}(t)$ and $r_{out}(t)$, where t is time of day, exhibit a transient behavior at the start of the foraging period followed by a *quasi steady-state* (QSS) in which the rates equilibrate, i.e., $r_{in}(t) = r_{out}(t)$ (see Figure 3.1 for two examples of this behavior for different colonies and Section 8.3.1 for more details).

We also found that the variation in transient and QSS responses appears to be consistent with the idea that colonies actively regulate their foraging activity in response to environmental conditions [41,124]. On Figures 3.2A, 3.2C, and 3.2E we show three additional examples of foraging rates for two different colonies across three different days. Figures 3.2B, 3.2D, and 3.2F show the foraging rates in Figures 3.2A, 3.2C,

[1] All raw data and the MATLAB model simulation scripts are available in Princeton University's DataSpace repository.

[2] Being bitten by a rattlesnake would result in a helicopter ride to the nearest hospital. Thankfully, rattlesnakes are nocturnal animals and are not usually observed during daytime.

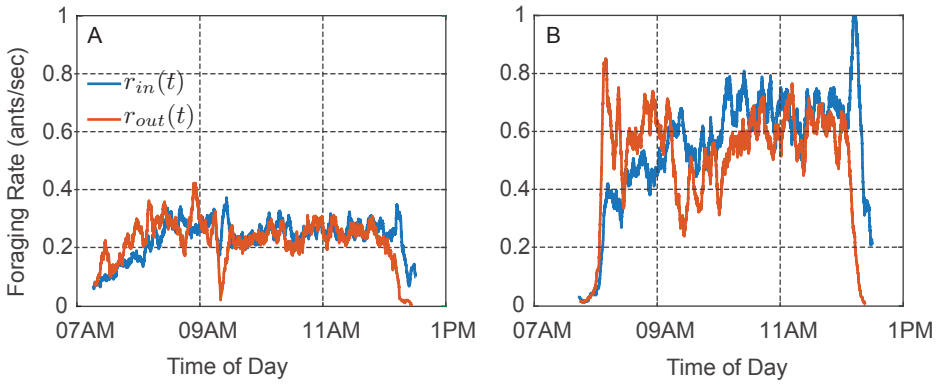

Figure 3.1: Plots of incoming foraging rate r_{in} (blue) and outgoing foraging rate r_{out} (red) versus time of day on August 20, 2016 for A) Colony 1357 and B) Colony 1317. The quasi steady-state (QSS) where incoming and outgoing rates equilibrated to a near-equal value can be observed for both colonies. The QSS rate for Colony 1317 was more than twice as great as it was for Colony 1357. Adapted from Figure 8.4.

and 3.2E, respectively, as *input-output* plots of $r_{in}(t)$ versus $r_{out}(t)$, where the color of the curve represents time of day.

Figures 3.2A to 3.2D show the foraging rates for the same colony on two different days: August 27 and August 31, 2015. The QSS rate on August 27 is more than twice the QSS rate on August 31. The QSS appears in the input-output plot as the region of the curves that remains close to the diagonal. The differences in the value of the QSS provide evidence that colonies increase their foraging activity on cool and humid days. August 27 was slightly cooler and more humid than August 31 (see Section 8.3.1 for details).

Figure 3.2E and 3.2F show the foraging rates for a different colony on a different day. The foraging rates rapidly increase in the morning before turning around and decreasing until the foraging process ends. This appears in the input-output plot as a closed orbit in the form of a loop; half of the curve lies above the diagonal, where $r_{in} = r_{out}$, and the other half lies below the diagonal. We associate this type of foraging behavior with hot and dry days (see Section 8.3.1 for details).

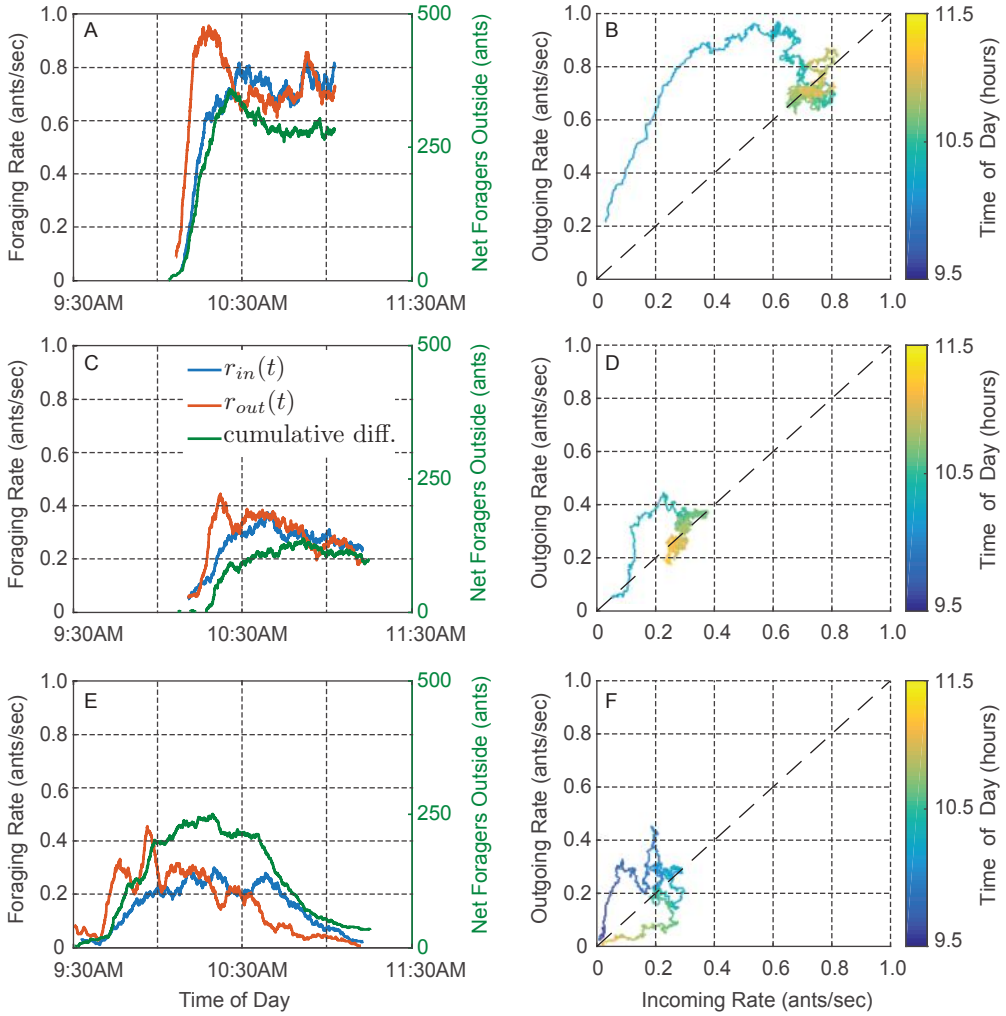

Figure 3.2: Plots of foraging rate data. Time series plots show incoming foraging rate r_{in} (blue), outgoing foraging rate r_{out} (red), and cumulative difference between the number of incoming and outgoing foragers (green) versus time of day. Input-output plots show $r_{out}(t)$ versus $r_{in}(t)$ with the color scale representing time of day t. A) and B) Colony 664 on August 27, 2015. C) and D) Colony 664 on August 31, 2015. E) and F) Colony 863 on September 1, 2015. Adapted from Figure 8.5.

The foraging rates in Figures 3.1 and 3.2 are representative of the data collected and suggest that both early morning transients and the QSS are important features in the regulation of foraging in red harvester ants.

3.3 Closed-Loop Excitable Model

Our model has three components, as shown in Figure 3.3 (see Section 8.2.2 for details):

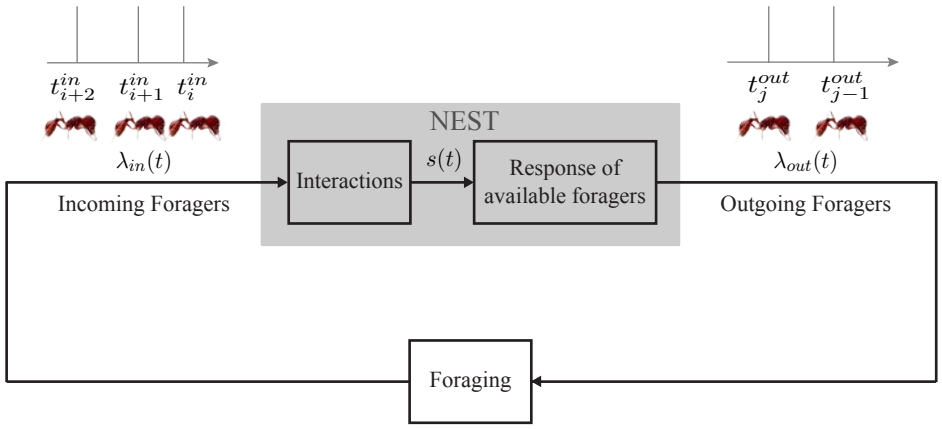

Figure 3.3: Closed-loop model block diagram. Adapted from Figure 8.1.

1. Interactions: Models the stimulus $s(t)$ received by the group of available foragers inside the nest entrance chamber from their interactions with food-bearing incoming foragers. We use leaky-integrator dynamics to model s:

$$\dot{s} = -\frac{s}{\tau} + k\lambda_{in}, \tag{3.3}$$

where $k > 0$ determines the amount by which s increases with every incoming forager and τ is the time constant for the exponential decay of s towards zero at times when there are no incoming foragers.

2. Response of available foragers: Models the activation of available foragers to leave the nest to forage in response to the accumulated evidence. We model this through the variable $v(t)$ representing the activation state of the group of available foragers inside the nest entrance. We interpret every point in time at which v increases above 0.75 from below as a forager leaving the nest. We model

v as the fast timescale variable in the FitzHugh-Nagumo (FN) equations [103, 104]:

$$\epsilon_1\epsilon_2\dot{v} = v - v^3/3 - cu - a + s \tag{3.4}$$

$$\epsilon_1\dot{u} = v - cu, \tag{3.5}$$

where $\epsilon_2 \ll 1$ defines the time separation between the dynamics of the fast and slow states, $\epsilon_1 \ll 1$ defines the time separation between the FN dynamics and the stimulus dynamics (3.3), a provides an offset to s, and c controls the frequency of oscillations in the FN dynamics (see Section 3.4 for details and a discussion of our modeling approach).[3]

3. Foraging: Models how active foragers collect seeds outside the nest. We define the foraging time as the the total time elapsed from when a forager leaves the nest to when it returns with food, and model it through a random variable X with chi-square distribution with parameter $D \in \mathbb{N}$ representing the mean foraging time in minutes. The chi-square distribution describes the distribution of the sum of D random samples from the standard normal distribution [127] and was selected based on field observations of foraging times [39].

Our model describes the process by which the sequence of incoming foragers λ_{in} (Figure 3.4A) gets mapped by the interactions component to the stimulus s that the group of available foragers inside the nest experience (Figure 3.4B). The response of available foragers component then maps s into the activation state v which exhibits oscillations whenever s is in the oscillating region of the FN (Figure 3.4C). We interpret each oscillation as the activation of a forager and threshold v in order to obtain the sequence of outgoing foragers λ_{out} (Figure 3.4D). Each outgoing forager in

[3]The FN equations eqs. (3.4) and (3.5) can be transformed into (2.16) by scaling time by $1/\epsilon$ and setting $v = x_1$, $u = x_2$, $\epsilon_1 = 1$, $\epsilon_2 = \epsilon$, $c = 1$, and $s = I$.

λ_{out} leaves the nest at a time t_j^{out} and becomes an incoming forager in λ_{in} at a time $t_i^{in} = t_j^{out} + X$.

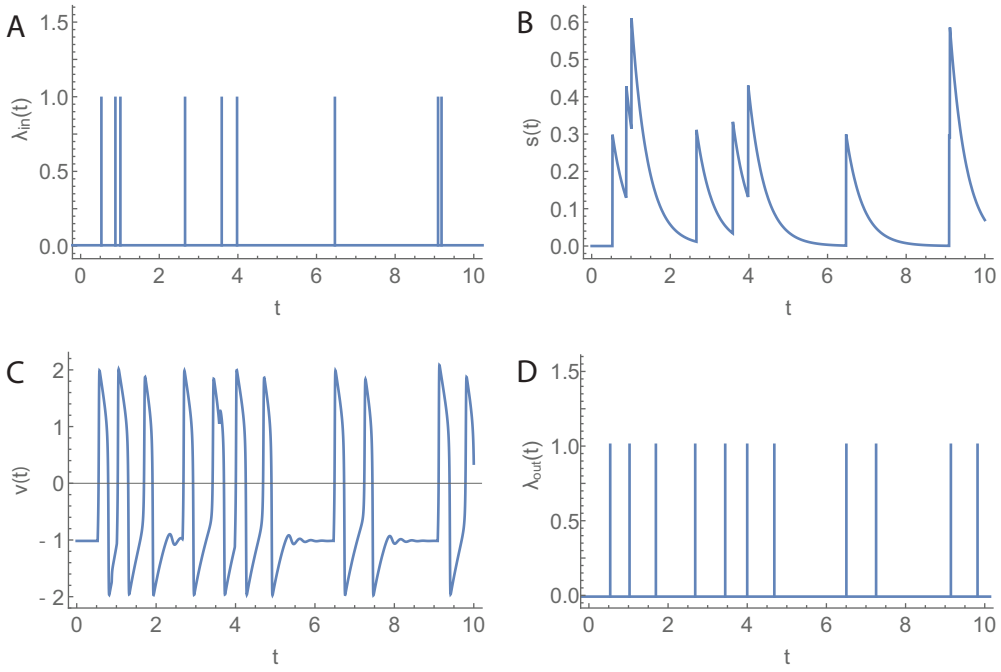

Figure 3.4: Open-loop model. A) Sequence of incoming foragers λ_{in}. B) Stimulus signal s associated with λ_{in}. C) FN output v for input s. D) Sequence of outgoing foragers λ_{out} obtained by thresholding FN output from below at 0.75. Repeated from Figure 8.2.

Throughout the rest of the chapter, we set $k = 0.3$, $\tau = 0.41$, $a = 0.35$, $\epsilon_1 = 0.2$, and $\epsilon_2 = 0.05$. We selected the parameter values to allow for rich dynamical behavior. We refer the reader to Section 8.3.2 for a discussion on our selection of model parameter values.

3.4 Response of Available Foragers as an Excitable System

We use the excitability dynamics of the FN model to represent the process by which available foragers inside the nest entrance chamber become active and leave the nest to forage. The FN model is a low-dimensional model for neuronal excitability, and

our modeling choice is motivated by the many parallels between the ant-to-ant interactions that drive foraging and the neuron-to-neuron interactions that underlie the cognitive abilities of organisms [3,4,38,128,129]. In both cases, system units (i.e., ants or neurons) accumulate evidence via interactions with neighbors. When the evidence reaches a threshold, the unit activates.

The FN system is a well-studied excitable model for weakly-interacting groups of neurons that captures mesoscopic-scale changes in neuronal activity in response to stimuli [130, 131]. In our model, we use the FN model to capture group level changes in the activity of available foragers in response to stimuli in the form of interactions with returning foragers.[4]

As described in Section 2.4.1, the FN equations describe an excitable system that exhibits nonlinear oscillations when the input is larger than the first bifurcation value b_1 and smaller than the second bifurcation value b_2, where $b_{1,2} = a \mp \frac{1}{3}(1 - c\epsilon_2)^{3/2}$. In our model, the stimulus s acts as the input to the FN equations. When $s \in [b_1, b_2]$, the activation state v exhibits short-lived spikes, and we interpret each spike as the activation of an available forager inside the nest entrance chamber. When $s > b_2$, the activation state saturates and the oscillations stop. We use the saturation state of the FN system to model overcrowding effects, the limited size of the nest entrance tunnel, and the difference in timescales between the high outgoing rates, in seconds, and the time required, in minutes, for foragers to move from the deeper chambers of the nest up to the entrance chamber [37, 132].

Remark 2. The non-static input in the FN model yields solutions that dynamically pass through the Hopf bifurcation. When s crosses b_2 from above, solutions are characterized by a *bifurcation delay* in which v remains close to the saturation state for an $\mathcal{O}(1)$ range of s after the Hopf bifurcation at $s = b_2$. Similarly, when s

[4]The activation mechanism via the accumulation of evidence can also be interpreted in terms of a complex contagion process in which the infective state is represented by the active state (i.e., foraging in ants or firing in neurons). We discuss the connection between excitability and complex contagion in more detail in Chapter 6.

crosses b_1 from above, v continues to oscillate for an $\mathcal{O}(1)$ range of s after the Hopf bifurcation at $s = b_1$. We illustrate this delay in Figure 3.5. Analogous bifurcation delays are observed when s crosses b_1 and b_2 from below. The delay is caused by the v dynamics (3.4) being significantly slower near the bifurcation points, with the implication that singular perturbation arguments based on the timescale separation do not hold at these points (see [91] for an example of a bifurcation delay for the pitchfork case).

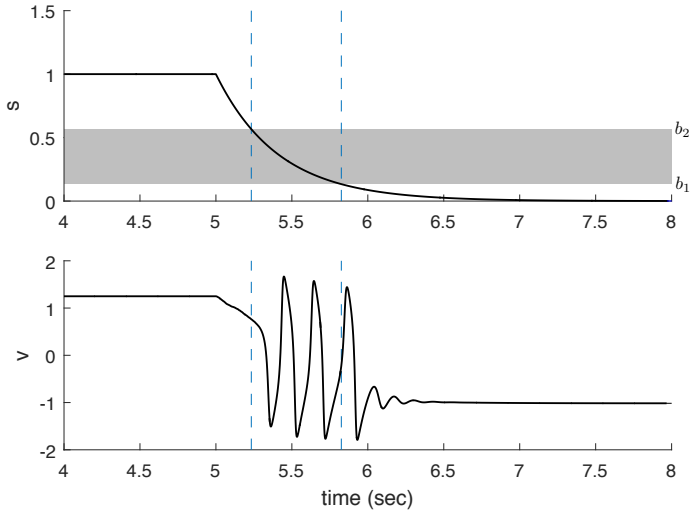

Figure 3.5: Bifurcation delay. As $s(t)$ decays with time constant τ, $v(t)$ crosses the two Hopf bifurcations in the FN model. The activation state v enters the oscillating region when $s(t) = b_2$ and leaves the oscillating region when $s(t) = b_1$. The dynamics of v exhibit a bifurcation delay in which oscillations in v do not start for a range of values $s(t) < b_2$ and continue for a range of values $s(t) < b_1$.

The parameter $c > 0$ in the FN equations (3.4) and (3.5) scales the negative feedback, acting as a control parameter for the frequency of oscillations and the range of values of stimulus s that lead to oscillations (See Section 8.5.1 for details). The parameter c regulates how susceptible the group of available foragers inside the nest is to the stimulus s.[5]

[5]The parameter ϵ_1, ϵ_2, and a also regulates susceptibility, but c does so in a more controlled way.

Definition 3.4.1. *The volatility of the available foragers inside the nest entrance chamber is* $c > 0$.

Remark 3. The volatility parameter c scales the cubic v-nullcline, changing the size of the limit cycle in the oscillating region $s \in [b_1, b_2]$. Small values of c increase the size of the limit cycle, leading to $v(t)$ solutions that take longer to traverse the cycle and, therefore, exhibit smaller oscillation frequencies. Large values of c decrease the size of the limit cycle and lead to higher frequencies. We illustrate this in Figure 3.6 and refer the reader to Section 8.5.1 for details.

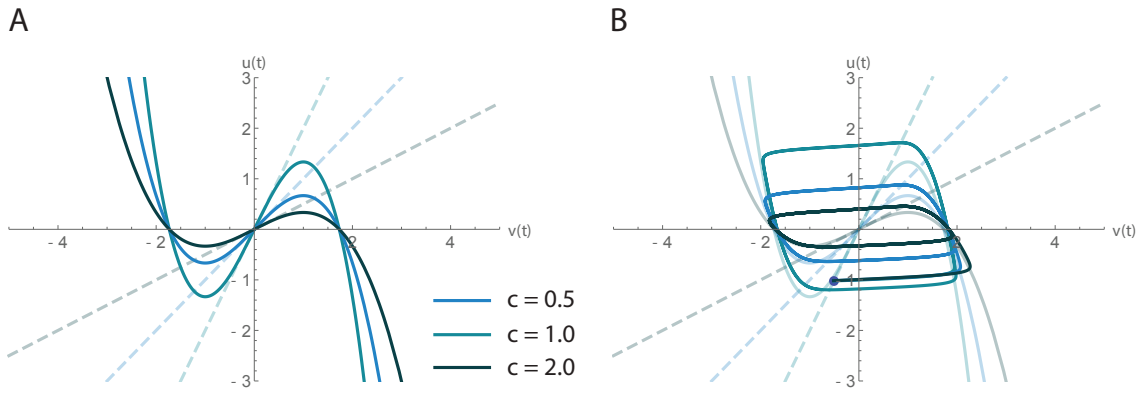

Figure 3.6: Effect of volatility on FN dynamics. A. FN nullclines for $c = 0.5$, $c = 1.0$, and $c = 2.0$. Volatility scales the cubic v-nullcline and modifies the slope of the linear u-nullcline. B. FN limit cycle for for $c = 0.5$, $c = 1.0$, and $c = 2.0$. Volatility scales the limit cycle and modulates the frequency of oscillations in v.

The dynamics of forager activation inside the nest take place in the three-dimensional (s, v, u) phase space. When a forager returns to the nest, s jumps in value by magnitude k and then decays at an exponential rate towards zero with time constant τ. The time that s spends in the oscillating region $[b_1, b_2]$ together with the volatility c determine the number of oscillations in v, and, therefore, the number of outgoing foragers. In Figures 3.7A, 3.7C, and 3.7E, we illustrate an example of a solution in (s, v, u) phase space for an input $s(t)$ that decays from $s = 0.6$ towards zero with time constant τ for three different volatility values: $c = 0.1$, $c = 1.0$, and $c = 2.0$. The planes $s = 0$ (magenta) and $s = 0.6$ (green) illustrate the upward shift

in the v-nullcline as a function of s. Figures 3.7B, 3.7D, and 3.7F show the resulting oscillations in v associated with Figures 3.7A, 3.7C, and 3.7E, respectively.

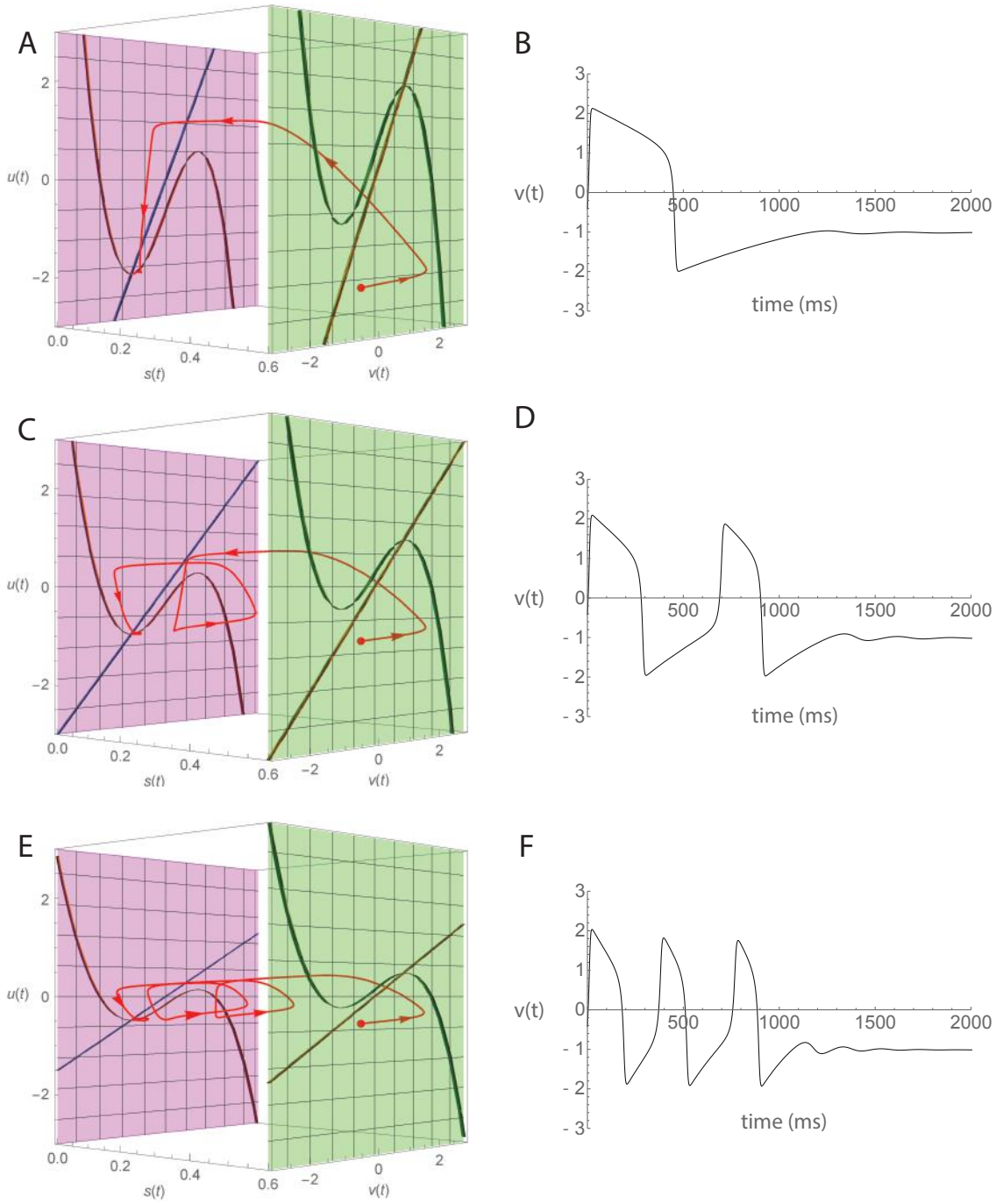

Figure 3.7: Dynamics in the (s, v, u) phase space for a decaying stimulus $\dot{s} = -s/\tau$, $s(0) = 0.6$. A. Phase space dynamics for $c = 0.5$. B. Time-series of v for $c = 0.5$. C. Phase space dynamics for $c = 1.0$. D. Time-series of v for $c = 1.0$. E. Phase space dynamics for $c = 2.0$. F. Time-series of v for $c = 2.0$.

3.5 Volatility and Nest I/O Curves

The Interactions and Response of available foragers components of our model describe an open-loop model (Figure 3.8) which predicts the sequence of outgoing ants λ_{out} with rate $r_{out}(t)$ generated by a given sequence of incoming ants λ_{in} with rate $r_{in}(t)$. This prediction relies on the volatility parameter c, which determines how the activation state v is affected by the stimulus s.

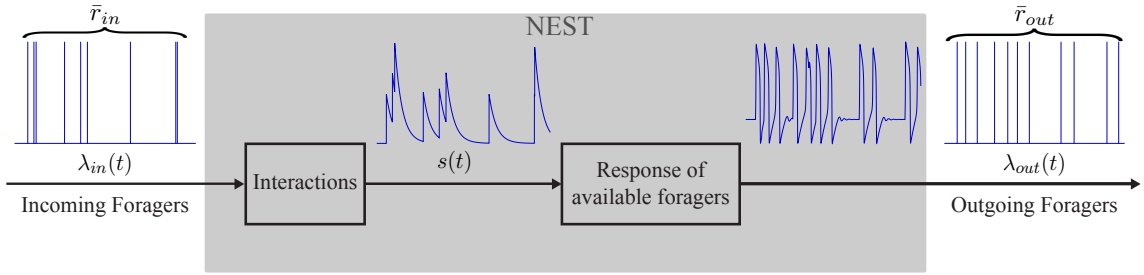

Figure 3.8: Open-loop model block diagram. A sequence of incoming foragers λ_{in} with mean incoming rate $\bar{r}_{in}$ is mapped into the stimulus signal s via leaky-integrator dynamics. The stimulus s then becomes the input to the FitzHugh-Nagumo model, resulting in output oscillations in the activation state v. The oscillations are thresholded from below to obtain the sequence of outgoing foragers λ_{out} with mean outgoing rate $\bar{r}_{out}$.

Definition 3.5.1. *The mean incoming foraging rate $\bar{r}_{in}$ for a time period $t \in [t_1, t_2]$ where $t_2 > t_1 \geq 0$ is equal to the number of incoming foragers $t_i^{in} \in [t_1, t_2]$ divided by the length of the time period $t_2 - t_1$. The mean outgoing foraging rate $\bar{r}_{out}$ is similarly defined.*

We obtain an analytic approximation for $\bar{r}_{out}$ given $\bar{r}_{in}$, parameterized by c:

$$\bar{r}_{out} = \int_{b_1(c)}^{b_2(c)} \frac{p(s, \bar{r}_{in})}{T_{LC}(s, c)} \, \mathrm{d}s, \tag{3.6}$$

where $T_{LC}(s, c)$ is an approximation for the FN limit cycle period and $p(s, \bar{r}_{in})$ is the probability density function of s at steady-state when λ_{in} is a homogeneous Poisson

process with rate $\bar{r}_{in}$. We derive (3.6) in Section 8.5.3, T_{LC} in Section 8.5.1, and p in Section 8.5.4. We refer the reader to Section 8.3.2 for an overview of the derivations.

Using (3.6), we obtain open-loop input-output curves, called *nest I/O curves*, that show how the analytic mapping from $\bar{r}_{in}$ to $\bar{r}_{out}$ depends nonlinearly on c (Figure 3.9A). The resulting curves agree with model simulations of the open-loop system (Figure 3.9B) (See Section 8.3.2 for a discussion of the results).

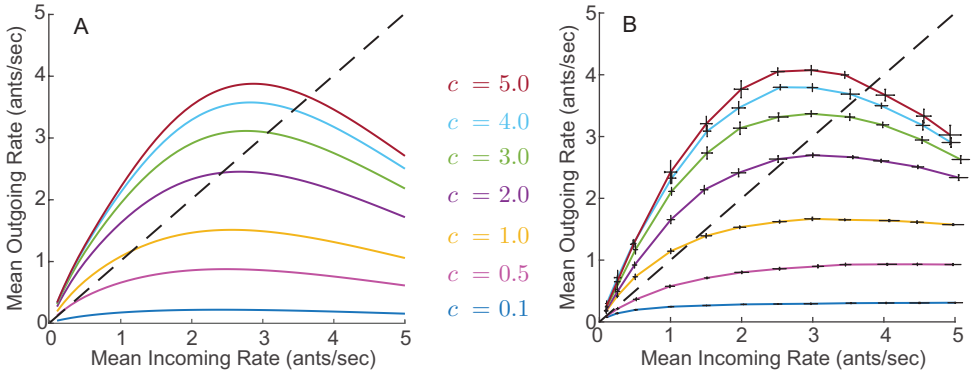

Figure 3.9: A) Analytical approximations for the nest I/O curves. B) Simulated nest I/O curves for different values of c. Each pair of error bars correspond to 10 simulation trials, each 5 minutes long, with a constant expected incoming rate and constant volatility c. The dashed black line represents points at which the mean incoming rate $\bar{r}_{in}$ is equal to the mean outgoing rate $\bar{r}_{out}$. Repeated from Figure 8.6.

We distinguish two different cases for the nest I/O curves. The first case corresponds to small volatility values for which the nest I/O curve lies on or below the black diagonal dashed line where $\bar{r}_{in} = \bar{r}_{out}$ (e.g., $c = 0.1$ and $c = 0.5$). The second case corresponds to high volatility values for which the nest I/O curve lies above the diagonal line for low $\bar{r}_{in}$ values, and below the diagonal line for high $\bar{r}_{in}$ values (e.g., $c = 1.0$ to $c = 5.0$). The nest I/O curves in the second case intersect the diagonal line at either the origin, or at a positive rate bounded away from zero.

3.6 Stable Foraging Rates through Feedback

The closed-loop model (Figure 3.3) describes the feedback process by which λ_{in} maps to λ_{out} (via the open-loop dynamics shown in Figure 3.8), and λ_{out} maps to λ_{in} at a later time via the foraging dynamics outside the nest.

Under the assumption that $\lambda_{out}(t)$ is a non-homogeneous Poisson process, the dynamics outside the nest that map λ_{out} maps to λ_{in} at a later time are analogous to the dynamics of a $M(t)/F/\infty$ queue, where F is the chi-square foraging time distribution (see 8.7 for an expression for F). Foragers play the role of customers in search of a service (finding a seed). The number of servers is infinite since foragers do not wait before they start searching for a seed. The service time is given by the random foraging time X.

We leverage results from Queueing Theory for $M(t)/F/\infty$ queues in order to reduce the study of the closed-loop model dynamics (Figure 3.3), across minutes-to-hours timescales, to the study of the discrete iterated mapping $r_n = G_c(r_{n-1})$ where r_n represents the mean foraging rate at time step n and r_{n-1} represents the mean foraging rate at time step $n-1$ (see Section 8.3.2 for details). The mapping $G_c : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is defined by the c-dependent nest I/O curves shown in Figure 3.9, and describe the process by which r_{in} becomes r_{out} through the dynamics of forager activation inside the nest, and r_{out} becomes r_{in} at a later time.

The points at which G_c (i.e., the nest I/O curves in Figure 8.6) intersect the diagonal line correspond to a stable steady-state of the iterated mapping. Thus, the intersection points correspond to an equilibrated steady-state foraging rate in the closed-loop model dynamics where $r_{in}(t) = r_{out}(t)$. This steady-state is analogous to the QSS observed in the data and we refer to it as the QSS. When G_c lies below the diagonal line where $r_{in} = r_{out}$, iterations of G_c decrease r. When G_c lies above the diagonal line, iterations of G_c increase r (Figure 3.10A). For volatilities where the nest

I/O curve lies on or below the $\bar{r}_{out} = \bar{r}_{in}$ diagonal (e.g., $c = 0.1$ and $c = 0.5$), G_c has a small interval of fixed points close to the origin. This results in the closed-loop model dynamics converging to a trivial steady-state with rate close to zero. In contrast, for volatilities where the nest I/O curve lies above the diagonal for low $\bar{r}_{in}$ values (e.g., $c = 2$ to $c = 5.0$), the closed-loop model dynamics converge to a QSS with rate bounded away from zero. Simulations of the closed-loop system match the analytical predictions (Figure 3.10B). We initialize the foraging dynamics by setting λ_{in} from $t = 0$ to $t = 60 \times (D + 1)$ seconds to be equal to the initial sequence of incoming foragers for Colony 859 on August 20, 2017, which has the very low mean incoming rate of 0.01 ants/sec during the first 15 minutes (see Section 8.3.2 for details).

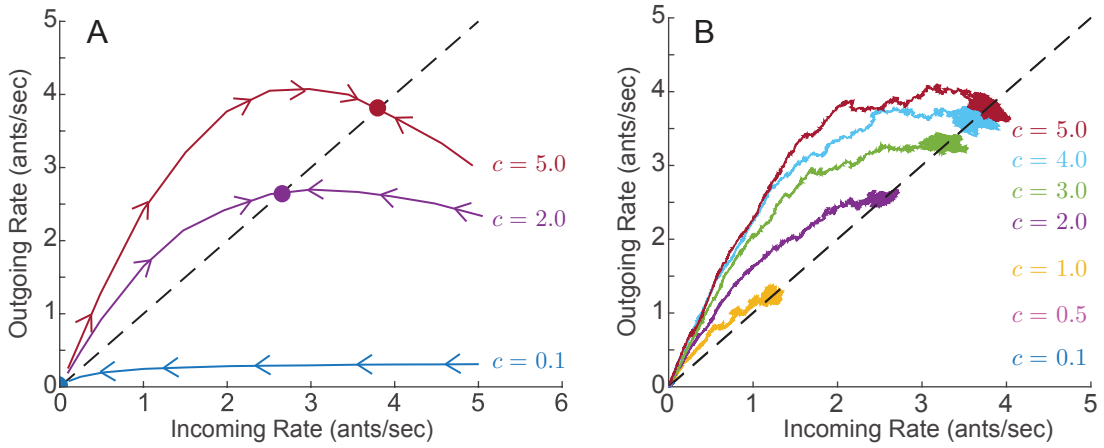

Figure 3.10: Closed-loop model dynamics. A. Red, purple, and blue curves show closed-loop trajectories of $r_{out}(t)$ versus $r_{in}(t)$ for fixed volatility c equal to 5.0, 2.0, and 0.1, respectively. B. Closed-loop model simulations for 7 different values of volatility c. The initial sequence of incoming foragers for all simulations was set equal to the sequence of incoming foragers recorded during the first 11 minutes for Colony 859 on August 20, 2017 which has a mean incoming rate of 0.01 ants/sec. The total time for all simulations was 3 hours. The mean foraging time was set to 10 minutes ($D = 10$). The corresponding curves for $c = 0.1$ and $c = 0.5$ are not visible as they lie near the origin for all time. Adapted from Figures 8.7 and 8.8.

Furthermore, this implies that there exists a critical value c^* such that if $c > c^*$, there is an isolated steady-state away from the origin (i.e., the QSS) and the colony forages, and if $c < c^*$, the steady-state remains close to the origin, which

we interpret as the colony deciding not to forage (Figure 3.11). Through numerical analysis of (3.6), we find the upper bound $\hat{c} > c^*$, where $\hat{c} = 0.5287$ (see Section 8.3.2 for details).

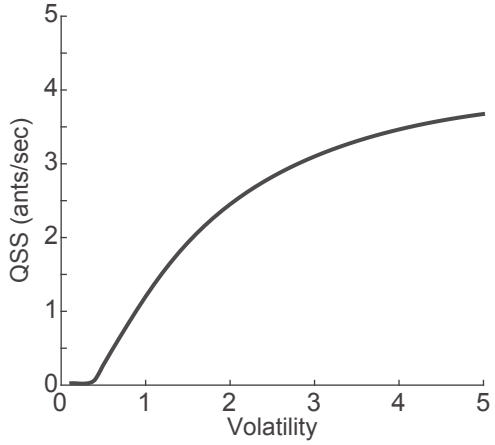

Figure 3.11: Analytical magnitude of the quasi steady-state (QSS) foraging rate as a function of the volatility c obtained from numerically solving eq. (3.6). Repeated from Figure 8.8A.

3.7 Volatility Adjustments as an Adaptation Mechanism

The closed-loop model dynamics show that the fast dynamics of forager activation inside the nest, in combination with the slow dynamics of foraging outside the nest, lead to a stable steady-state foraging rate where $r_{in}(t) = r_{out}(t)$ (i.e., the QSS). The steady-state rate is set by the volatility parameter c. These dynamics capture the initial transient and the QSS observed in field observations, but cannot explain the range of foraging behaviors observed in the data across different environmental conditions. As a step towards the development of refined hypotheses on the mechanisms that colonies use to exhibit flexible collective behavior, we propose a mechanism for colony response to conditions in which the volatility of foragers changes after a first exposure to the environment during their first foraging trip (Figure 3.12). We set

the total number of foragers in the system to N and group available foragers in the nest entrance chamber into two sets: *uninformed* foragers who have not left the nest, and *informed* foragers who have. We represent the uninformed group with an FN model with volatility c_u and the group of informed foragers with an FN model with volatility c_i. We let c_u be the nominal excitability of the colony and fix c_i according to the environmental conditions. High c_i values represent cool and humid conditions while low c_i values represent hot and dry conditions (see Sections 8.2.2 for details).

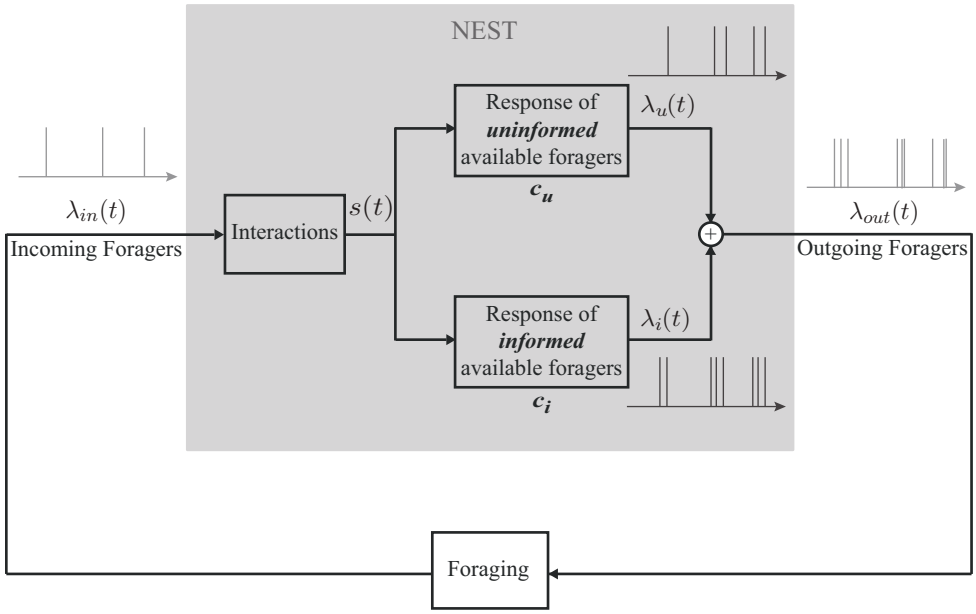

Figure 3.12: Block diagram of proposed mechanism for response of colony to environmental conditions. The available foragers inside the nest comprise two sets: uninformed foragers that have not yet left the nest and so are uninformed about the conditions outside the nest, and informed foragers who have left the nest and therefore are aware of the conditions. The response of each set to s is represented by a different FN model, distinguished by the volatility parameter c_u for the uninformed and c_i for the informed. The outputs of these two oscillator dynamics are weighted probabilistically using thinning to get an outgoing stream of foragers $\lambda_{out}(t)$. Repeated from Figure 8.3.

We extend our open-loop results for $\bar{r}_{out}$ in (3.6) to the open-loop model with the adaptation mechanism:

$$\bar{r}_{out} = x_u \int_{b_1(c_u)}^{b_2(c_u)} \frac{p(s, \bar{r}_{in})}{T_{LC}(s, c_u)} \, \mathrm{d}s + x_i \int_{b_1(c_i)}^{b_2(c_i)} \frac{p(s, \bar{r}_{in})}{T_{LC}(s, c_i)} \, \mathrm{d}s,$$

where x_u is the fraction of uninformed foragers inside the nest and x_i is the fraction of informed foragers, i.e., $x_u + x_i = 1$.

Similarly, we extend our closed-loop results by considering the iterated mapping as it evolves in time $G_c(t)$, from an initial mapping $G_c(t_0) = G_{c_u}$ with volatility c_u to a final mapping $G_c(\infty) = G_{c_i}$ with volatility c_i.

The intuition behind the adaptation mechanism is illustrated in Figure 3.13. At the start of the day, the colony behaves as a colony with volatility c_u. As the fraction of informed foragers in the nest increases, the volatility of the colony slowly shifts towards c_i. The colony behaves as a colony with volatility c_i once all foragers have been outside, i.e., $x_i = 1$. The light and dark green curves show examples of this transition for $c_u = 5.0$. For the light green curve $c_i = 2 \gg c^*$ leading to a solution that converges to a QSS. For the dark green curve, $c_i = 0.1 \ll c^*$ leading to a solution that exhibits a "loop" behavior (similar to that observed in Figure 3.2F) in which the foraging rates reach a maximum before decaying to a trivial rate close to zero, leading to the end of the foraging process.

3.8 Simulations

Figure 3.14 show three example simulations of the closed-loop model with the adaptation mechanism. In each case, we initialized the foraging dynamics as in Figure 3.10B and selected parameter values D, N, c_i and c_u to match the resulting foraging rates to the observed rates in Figure 3.2 (see Section 8.3.2 for details). The fractions of informed foragers x_i and uninformed foragers x_u were updated over time as described in the legend of Figure 3.12.

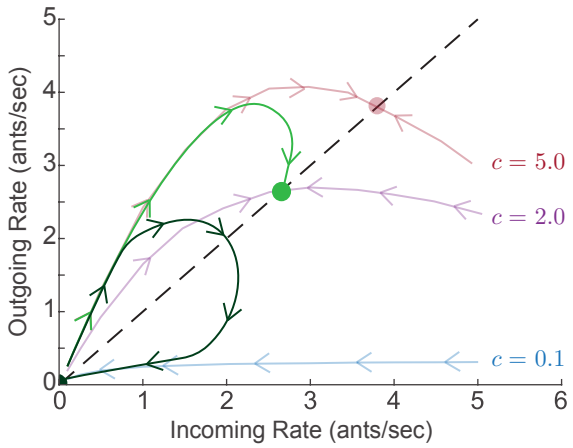

Figure 3.13: Illustration of the response of foraging rates to environmental conditions. Red, purple, and blue semitransparent curves show closed-loop trajectories of $r_{out}(t)$ versus $r_{in}(t)$ for fixed volatility c equal to 5.0, 2.0, and 0.1, respectively. Initially, all available foragers are uninformed about the environment and have volatility $c_u = 5.0$. The light green curve shows the dynamics in the case when foragers exposed to the environment reduce their volatility to $c_i = 2.0$, as might happen on a moderately hot and dry day. The dark green curve shows the dynamics in the case when foragers exposed to the environment reduce their volatility to $c_i = 0.1$, as might happen on a very hot and dry day. Adapted from Figure 8.7.

3.9 Dynamics of the Closed-Loop Model with Adaptation Mechanism

The dynamics of the closed-loop model with adaptation mechanism shown in Figure 3.12 are determined by the system parameters: k, τ, D, N, c_i, and c_u (see Section 8.3.2 and Tables 3.1 and 3.2 for details). We choose k and τ to exhibit rich dynamics in the model, while the parameters D and N can be selected to affect the timing of the dynamics. The volatilities c_u and c_i control the transient and steady-state dynamics.

We identify four foraging behavioral regimes in the closed-loop model dynamics with the adaptation mechanism. Each regime has qualitatively distinct transient and steady-state properties. These behaviors are characterized by the values of the uninformed volatility c_u, the informed volatility c_i, and the initial foraging rates:

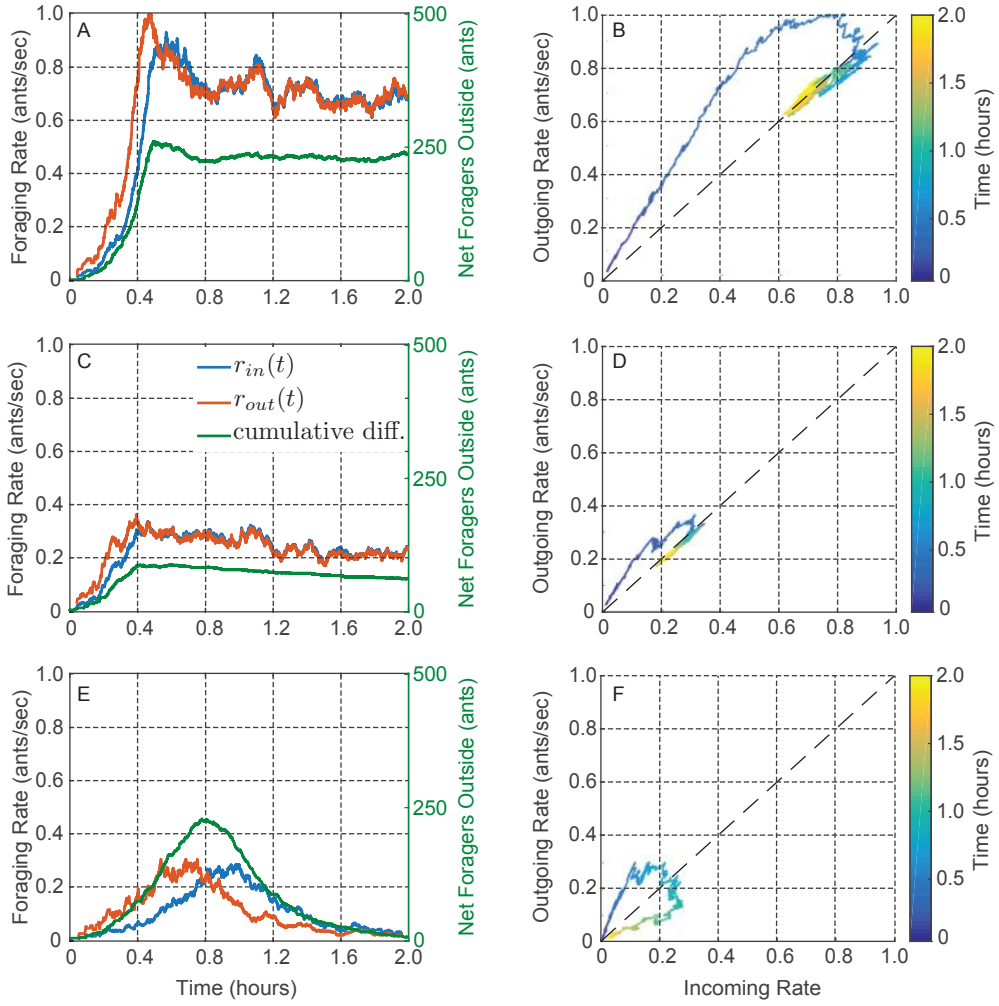

Figure 3.14: Simulations of the closed-loop model with the adaptation mechanism. Plots are of the same form as in Figure 3.2, and qualitative comparisons can be made between A and B here and Figure 3.2A and B, between C and D here and Figure 3.2C and D, and between E and F here and Figure 3.2E and F. A) and B) $c_u = 3$, $c_i = 0.9$, $N = 500$, $D = 5$. C) and D) $c_u = 3$, $c_i = 0.75$, $N = 200$, $D = 5$. E) and F) $c_u = 5$, $c_i = 0.02$, $N = 600$, $D = 15$. Adapted from Figure 8.9.

1. Trivial foraging ($c_u \ll c^*$ and $c_i \ll c^*$): For any initial foraging rates, the rates decrease and converge to a rate close to zero. This follows from the mapping G_c lying below the diagonal for any $\bar{r}_{in}$ bounded away from zero if $c \ll c^*$. Therefore, the mappings G_{c_u} and G_{c_i} lie below the diagonal line and iterations of the mapping $G_c(t)$ for all $t > 0$ decrease r. The Trivial Foraging behavior is analogous to the closed-loop model dynamics without adaptation mechanism

Table 3.1: States of the closed-loop model with adaptation mechanism.

State	Description
λ_{in}	sequence of incoming foragers
λ_{out}	sequence of outgoing foragers
r_{in}	incoming foraging rate
r_{out}	outgoing foraging rate
s	available foragers' stimulus
v	fast FN state (avtivation state)
u	slow FN state
x_i	fraction of informed available foragers
x_u	fraction of uninformed available foragers

Table 3.2: Parameters of the closed-loop model with adaptation mechanism.

Parameter	Description
k	stimulus gain constant
τ	stimulus time constant
ϵ_1	time separation between FN and s dynamics
ϵ_2	time separation between v and u
a	stimulus offset
D	mean foraging time
N	total number of foragers
c_i	informed volatility
c_u	uninformed volatility

when $c < c*$. For example, in Figure 3.10B, the curves for $c = 0.1$ and $c = 0.5$ are not visible because the remain near the origin for all time.

2. Steady-state foraging ($c_u \gg c^*$ and $c_i \gg c^*$): For any initial foraging rates, the rates converge to a steady-state where $r_{in}(t) = r_{out}(t)$ bounded away from zero. This follows from the mapping G_c lying above the diagonal line for low $\bar{r}_{in}$ values, below the line for high $\bar{r}_{in}$ values, and intersecting the line at a rate r^* bounded away from zero, if $c \gg c^*$. Therefore, the mappings G_{c_u} and G_{c_i} are as describe above and iterations of the mapping $G_c(t)$ for all $t > 0$ move r towards the diagonal line. Once all foragers have become informed, the solution

converges to the QSS given by the mapping G_{c_i}. See Figures 3.14A and 3.14C for an example of the Steady-State Foraging regime.

3. Loop foraging ($c_u \gg c^*$, and $c_i \ll c^*$): For low initial foraging rates, the rates initially increase, reach a maximum value, and then decrease to a rate close to zero. This follows from the system initially behaving according to the mapping $G(t_0) = G_{c_u}$, which lies above the diagonal for low $\bar{r}_{in}$ values. This leads to an initial transient that increases the foraging rates. As the number of informed foragers increases, the mapping $G(t)$ goes to G_{c_i}, which lies below the diagonal for any $\bar{r}_{in}$ bounded away from zero, leading to foraging rates that decrease with time and converge to a rate close to zero. See Figure 3.14E for an example of the Loop Foraging regime.

4. Delayed foraging ($c_u \ll c^*$ or $c_u \approx c^*$, and $c_i \gg c^*$): The rates remain close to zero for a period of time before eventually increasing towards a steady-state where $r_{in}(t) = r_{out}(t)$ bounded away from zero. The initial decay in the foraging rates follows from the condition for c_u, which implies that near the initial time, the number of informed foragers in the nest increases at a slow rate. That is, the dynamics of $G_c(t)$ from G_{c_u} to G_{c_i} are very slow. Over a sufficiently long period of time, the number of informed foragers reaches a critical mass and the foraging rates begin to increase. As they increase, the rate at which foragers become informed increases, leading to a positive feedback. Once all foragers have become informed, the solution converges to the QSS given by the mapping G_{c_i}. See Figure 3.15A for an example of the Delayed Foraging regime.

When c_i and/or c_u are close to c^*, the model dynamics are significantly affected by the stochasticity in the sequence of incoming and outgoing foragers. In these cases, the dynamics of the closed-loop model with adaptation may exhibit responses that do not fall into one of these four regimes.

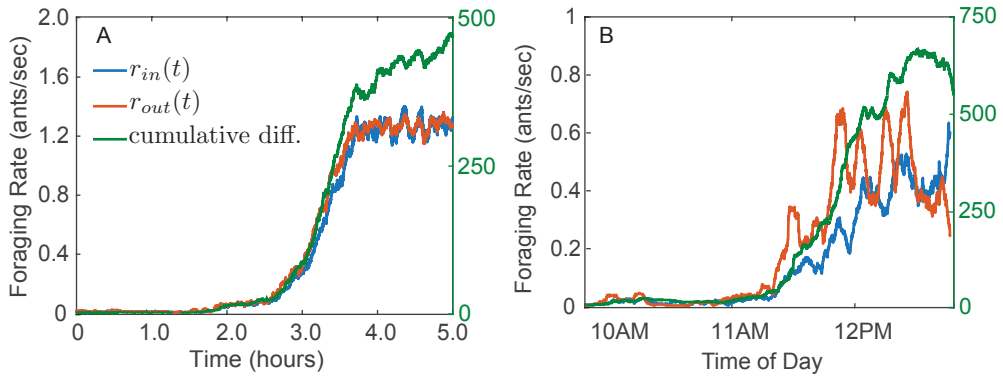

Figure 3.15: Delayed foraging. A. Simulation of the closed-loop model with the adaptation mechanism with $c_u = 0.9$, $c_i = 1.0$, $N = 1000$, and $D = 5$. B. Foraging rates for Colony 859 on August 20, 2017. Adapted from Figures 8.11 and 8.14.

Each one of the four behavioral regimes have analogues in the data. The trivial foraging regime is comparable to cases in which colonies do not forage, the steady-state foraging regime is comparable to cases in which the foraging rates reach a QSS (such as those shown in Figures 3.2A and 3.2C), the loop foraging regime is comparable to cases such as the one shown in Figure 3.2E, and the delayed foraging regime is comparable to cases such as the one shown in Figure 3.15B.

3.10 Concluding Remarks

In this Chapter we derived and analyzed a low-dimensional model for the regulation of foraging in red harvester ants. We proved that the inclusion of feedback from foragers returning to the nest leads to the emergence of a stable equilibrated foraging rate where $r_{in} = r_{out}$. The model captures the range of foraging behaviors observed in the field through a small number of parameters. Particularly, we prove that the volatility parameter c controls the initial transient in r_{in} and r_{out}, parameterizes the magnitude of the equilibrated foraging rate, and predicts the early cessation of foraging or no foraging at all for c below the critical threshold c^*. We showed how a change in volatility as foragers first become exposed to the environmental conditions outside the

nest can account for the observed foraging behavior under different temperature and humidity values. Finally, we categorized the range of observed foraging behaviors into four distinct behavioral regimes characterized by the value of the informed volatility c_i and the uninformed volatility c_u.

Our results suggest that the foraging behavior of the colony can be regulated in the presence of different environmental conditions through the accumulated response of the individual foragers to the conditions. The adjustments in volatility at the forager level slowly change the volatility of the colony, modulating the response of the colony to stimulus and regulating foraging activity. Through the adjustments in volatility, the colony can achieve a robust foraging rate that is flexible in response to different conditions.

Chapter 4

Adaptive Susceptibility in Well-Mixed Reinfection Models

In this chapter we study the SIRI epidemiological model in well-mixed settings. The model considers a large group of individuals that interact with each other, where the susceptibility of individuals change after a first exposure to the infection. Related models have been studied in [133, 134] to study the role of partial immunity in the spread of contagious diseases across populations. We perform an exhaustive analysis of the SIRI model dynamics and consider cases beyond that of partial immunity. We prove the existence of four different dynamical regimes in the model: infection-free, endemic, epidemic, and bistable. As far as we know, this is the first such analysis of its kind.

Our study of the SIRI model is motivated by the results presented in Chapter 3. We have shown that adaptations in susceptibility in response to a first exposure to the environment are sufficient to provide harvester ant colonies with the ability to regulate foraging behavior in response to different environmental conditions. Here, we

generalize this idea to epidemiological settings, where we consider a spreading process and individuals that adjust their susceptibility to the process after a first exposure.

Sections 4.1-4.3 summarize results from Part II: Chapter 9, which appears as Pagliara, Dey, and Leonard [57]. In Section 4.1 we introduce the SIRI model in well-mixed settings. In Section 4.2 we perform a stability analysis of the equilibria and introduce the notion of the basic infection and basic reinfection reproduction numbers. In Section 4.3 we show that the basic reproduction numbers distinguish four distinct behavioral regimes: Infection-free, endemic, epidemic, and bistable. We show that in the bistable regime, there exists a critical initial condition below which the infection dies out and above which it spreads. Solutions that reach the endemic equilibrium exhibit a resurgent epidemic in which the fraction of infected individuals initially decays, reaches a minimum and then grows towards an endemic equilibrium where there are always infected individuals. In addition, we show that the initial conditions can be selected such that the resurgent epidemic spends an arbitrarily long period of time close to zero before resurgence.

In Section 4.4 we prove new results, not published in [57] of Chapter 9, showing that the behavioral regimes in the SIRI model are explained by the existence of a transcritical bifurcation without parameters with drift singularity.

4.1 Well-mixed SIRI model

We consider a large population of individuals, where each individual can be in any of the following three states: *susceptible* (S), *infected* (I), or *recovered* (R). Susceptible individuals become infected with rate β through contact with already infected individuals. Infected individuals recover with rate δ. Recovered individuals become reinfected with rate $\hat{\beta}$ through further contact with already infected individuals.

$$\begin{array}{ccccc} \text{S+I} & \xrightarrow{\beta} & \text{I+I} & \xleftarrow{\hat{\beta}} & \text{R+I} \\ & & \text{I} & \xrightarrow{\delta} & \text{R} \end{array} \tag{4.1}$$

We assume that interactions between any two individuals occur with the same probability, and study the resulting well-mixed model (see Section 1.2.1):

$$\begin{aligned} \dot{x}_S &= -\beta x_S x_I \\ \dot{x}_I &= \beta x_S x_I + \hat{\beta} x_R x_I - \delta x_I \\ \dot{x}_R &= -\hat{\beta} x_R x_I + \delta x_I, \end{aligned} \tag{4.2}$$

where x_S, x_I, and x_R represent the fractions of population that belong to the susceptible, infected, and recovered states, respectively.

The SIRI model dynamics describe a family of models which can be categorized into five different cases, where each case considers individuals that adapt in particular ways after a first exposure to the infection:

1. Setting $\delta = 0$ and redefining x_S as $x_S + x_R$ reduces the SIRI model to the SI model. Infected individuals do not recover.

2. Setting $\hat{\beta} = 0$ reduces the SIRI model to the SIR model. Recovered individuals gain full immunity after a first exposure.

3. Setting $\beta = \hat{\beta}$ and redefining x_S as $x_S + x_R$ reduces the SIRI model to the SIS model. Recovered individuals gain no immunity and become infected at the same rate as susceptible individuals.

4. Setting $0 < \hat{\beta} < \beta$ reduces the SIRI model to a partial immunity model in which individuals become less susceptible to the infection after first exposure.

Table 4.1: Special cases of the SIRI model. Repeated from Table 9.1

Parameter Value	**Equivalent Model**	**Immunity Condition**
$\hat{\beta} = 0$	SIR	Full Immunity
$0 < \hat{\beta} < \beta$	—	Partial Immunity
$\hat{\beta} = \beta$	SIS	No Immunity
$\beta < \hat{\beta} < \infty$	—	Compromised Immunity
$\delta = 0$	SI	No Recovery

5. Setting $\beta < \hat{\beta} < \infty$ reduces the SIRI model to a compromised immunity model in which individuals become more susceptible to the infection after first exposure.

Table 4.1 highlights how the special cases of the SIRI model fall along a spectrum of immunity.

The dynamics of well-mixed epidemiological models usually depend on the basic reproduction number, defined as the average number of infections caused by a single infected individual (see Section 1.2.1). In the SIRI model, the basic reproduction number depends on whether the infected individual interacts with susceptible or with recovered individuals.

When there are no individuals who have been previously exposed to the infection, the SIRI model initially behaves as the SIS model with infection rate β. As infected individuals recover, the system dynamics slowly transition towards a model that behaves as the SIS model with infection rate $\hat{\beta}$.

Definition 4.1.1. *The basic infection reproduction number is* $R_0 = \beta/\delta$. *The basic reinfection reproduction number is* $R_1 = \hat{\beta}/\delta$.

The basic infection reproduction number R_0 represents the average number of infections caused by a single infected individual in a population of susceptible individuals, while the basic reinfection reproduction number R_1 represents the average number of reinfections caused by a single infected individual in a population of recovered individuals (see Section 9.3.1 for details).

4.2 Equilibria and Stability Analysis

The dynamics of the SIRI model (4.2) evolve on the 2-simplex $\Delta_2 = \{(x_S, x_I, x_R) \in [0,1]^3 | x_S + x_I + x_R = 1\}$. Using the substitution $x_R = 1 - x_S - x_I$, we obtain the reduced model

$$\begin{aligned} \dot{x}_S &= -\beta x_S x_I \\ \dot{x}_I &= (\hat{\beta} - \delta) x_I + (\beta - \hat{\beta}) x_S x_I - \hat{\beta} x_I^2. \end{aligned} \tag{4.3}$$

The equilibria of (4.3) consists of a continuum of infection-free equilibria, and an isolated endemic equilibrium point:

1. Infection-Free Equilibria (IFE): $x_S = x_S^*$, $x_I = 0$,

2. Endemic Equilibrium (EE) (only if $\hat{\beta} > \delta$): $x_S = 0$, $x_I = 1 - \delta/\hat{\beta}$.

The IFE forms a line of equilibria in which, at every point on the line, there are no infected individuals. The EE corresponds to an equilibrium in which every individual is either in the infected or recovered state.

The Jacobian at the EE and at an arbitrary point in the IFE are given by

$$J_{EE} = \begin{bmatrix} -\beta(\hat{\beta} - \delta)/\hat{\beta} & 0 \\ (\beta - \hat{\beta})(\hat{\beta} - \delta)/\hat{\beta} & -(\hat{\beta} - \delta) \end{bmatrix}, \quad J_{IFE} = \begin{bmatrix} 0 & -\beta x_S^* \\ 0 & (\beta - \hat{\beta}) x_S^* + \hat{\beta} - \delta \end{bmatrix}. \tag{4.4}$$

The Jacobian J_{EE}, when the EE exists, has two negative eigenvalues, from which we get the following result:

Lemma 1 ([57]). *The EE is an equilibrium point of* (4.3) *if and only if* $R_1 \geq 1$. *Moreover, the EE is locally stable.*

At every point $x_S = x_S^*$, $x_I = 0$ in the IFE, J_{IFE} has a zero eigenvalue with eigenvector $[1, 0]^T$, corresponding to the direction of the line of equilibria $x_I = 0$, and a second eigenvalue $(\beta - \hat{\beta}) x_S^* + \hat{\beta} - \delta$ with eigenvector $[0, 1]^T$ corresponding

to a direction transverse to $x_I = 0$. As long as this transverse eigenvalue is not equal to zero, the point in the IFE is normally hyperbolic. Normal hyperbolicity implies that the stability of the point $x_S = x_S^*, x_I = 0$ is given by the sign of the transverse eigenvalue (see Section 2.1 for details). At $x_S^* = (1 - R_1)/(R_0 - R_1)$, the IFE subspace $x_I = 0$ loses normal hyperbolicity through a bifurcation without parameters. We explore this in more detail in Section 4.4.

Definition 4.2.1. *The IFE bifurcation point is* $M = (1 - R_1)/(R_0 - R_1)$.

Lemma 2 ([57])**.** *The following holds true for the IFE:*

1. *If* $R_0 < 1$ *and* $R_1 < 1$*, then all points in the IFE are locally stable.*
2. *If* $R_0 > 1$ *and* $R_1 > 1$*, then all points in the IFE are unstable.*
3. *If* $R_0 > 1$ *and* $R_1 \leq 1$*, points in the IFE with* $x_S^* < M$ *are locally stable and points with* $x_S^* > M$ *are unstable.*
4. *If* $R_0 \leq 1$ *and* $R_1 > 1$*, points in the IFE with* $x_S^* > M$ *are locally stable and points with* $x_S^* < M$ *are unstable.*

4.3 Global Dynamics and Basic Reproduction Numbers

The dynamics of x_S in (4.3) are monotonic decreasing, preventing the SIRI model from exhibiting non-trivial periodic orbits.

Lemma 3 ([57])**.** *The SIRI model does not exhibit non-trivial periodic orbits on* Δ_2.

Lemma 3 implies that any solution starting on Δ_2 must end at either a point in the IFE or the EE. As shown in Section 4.2, the local stability of points in the IFE and the EE are determined by the values of R_0 and R_1. It follows that the dynamics of the SIRI model (4.3) are determined by the values of R_0 and R_1.

Theorem 1 (Behavioral Regimes of SIRI [57]). *Given an initial condition $x_I = x_{I0}$, $x_S = 1 - x_{I0}$, where $0 < x_{I0} < 1$, the SIRI model (4.3) exhibits four different dynamical behaviors:*

1. *Infection-Free: If $R_0 < 1$ and $R_1 < 1$, then all solutions reach a point in the IFE as $t \to \infty$, and x_I decays monotonically to zero.*

2. *Endemic: If $R_0 > 1$ and $R_1 > 1$, then all solutions reach the EE as $t \to \infty$.*

3. *Epidemic: If $R_0 > 1$ and $R_1 \leq 1$, then all solutions reach a point in the IFE as $t \to \infty$ and, at equilibrium, $x_S^* < M$. For initial conditions where $x_{I0} \geq (\beta - \delta)/\beta$, x_I decays monotonically to zero. While for initial conditions where $x_{I0} < (\beta - \delta)/\beta$, x_I grows initially and reaches a maximum value:*

$$x_I^{max} = \frac{R_0 - R_1}{R_1 (R_0^{R_0} (1 - x_{I0})^{R_1})^{1/(R_0 - R_1)}} + \frac{R_1 - 1}{R_1},$$

   before decaying to zero as $t \to \infty$.

4. *Bistable: If $R_0 \leq 1$, $R_1 > 1$, then x_I decays initially. Moreover, there is a critical initial fraction of infected individuals*

$$x_{IC} = 1 - M(R_0 M)^{-\frac{R_0}{R_1}}. \tag{4.5}$$

   Solutions with initial condition $x_{I0} < x_{IC}$ reach a point in the IFE as $t \to \infty$ and x_I decays monotonically to zero. Solutions with initial conditions $x_{I0} > x_{IC}$ reach the EE as $t \to \infty$.

Figure 4.1 summarizes the results of Theorem 1. In each quadrant of the R_0, R_1 parameter space, we show a simulation of the corresponding dynamics on Δ_2. The bottom boundary of Δ_2 represents the IFE. The solid blue and dashed red lines correspond to locally stable and unstable points in the IFE, respectively, as described

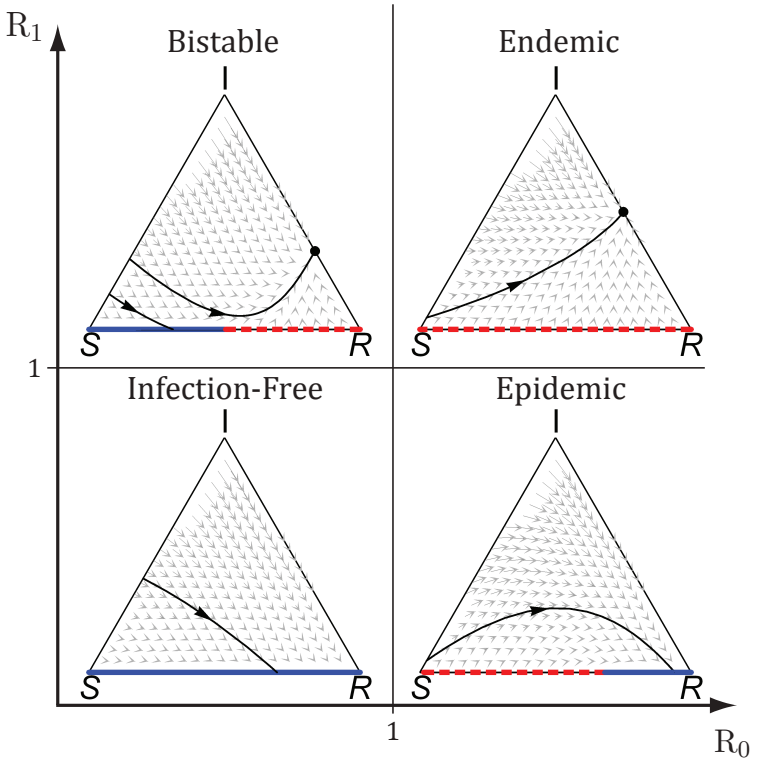

Figure 4.1: The four different behavioral regimes of the SIRI model plotted on Δ_2. The four plots are arranged in the R_0, R_1 parameter space to illustrate the four corresponding regimes. Bistable regime: $R_0 = 0.5$, $R_1 = 1.5$, $x_I(0) = 0.15, 0.30$. Endemic regime: $R_0 = 1.5$, $R_1 = 2.0$, $x_I(0) = 0.05$. Infection-free regime: $R_0 = 0.75$, $R_1 = 0.25$, $x_I(0) = 0.45$. Epidemic regime: $R_0 = 2.5$, $R_1 = 0.25$, $x_I(0) = 0.05$. Repeated from Figure 9.1.

in Lemma 2. The thinner black lines are example trajectories of solutions with initial conditions where $x_S(0) = 1 - x_I(0)$. We highlight the bistability phenomenon by showing two trajectories, corresponding to a trajectory with $x_{I0} = 0.15$ that reaches a point in the IFE and a trajectory with $x_{I0} = 0.3$ that reaches the EE.

Remark 4. The transient dynamics in the infection-free and endemic regimes depend on the ratio R_0/R_1. When $R_0/R_1 > 1$ recovered individuals become partially immune to the infection and as x_R increases, the rate at which the infection spreads through the population decreases, leading to concave trajectories in Δ_2. In contrast, when $R_0/R_1 < 1$, recovered individuals become more susceptible to the infection and as x_R increases, the rate of change of x_I accelerates, leading to convex trajectories in

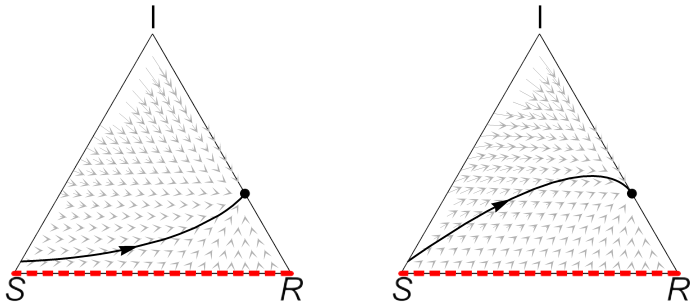

Figure 4.2: Two distinct cases of transient dynamics in the endemic regime (see Remark 4)

.

Δ_2. Figure 4.2 provides an example for these two types of trajectories in the endemic regime.

Remark 5. The behavioral regimes defined in Theorem 1 provide analytical predictions on the behavior of solutions in any of the cases described in Table 4.1. Figure 4.3 illustrates this point by coloring the different regions of the R_0, R_1 parameter space according to the type of immunity that a model in that region would exhibit. Note that the SIS case is represented by a dashed diagonal line where $R_0 = R_1$, while the SIR case is represented with a dashed line where $R_1 = 0$.

4.3.1 Resurgent Epidemics

When $R_0 < 1$ the infection cannot spread in a population of susceptible individuals. As a result, the fraction of infected individuals x_I decays near the initial time $t = t_0$. However, if $R_1 > 1$, the infection can subsequently spread in the population of recovered individuals. Here we show that solutions in the bistable regime ($R_0 \leq 1$, and $R_1 > 1$) that reach the EE as $t \to \infty$ exhibit a *resurgent epidemic* in which x_I reaches a minimum value at some time $t = t_{min}$ before increasing towards the EE. Moreover, we show that as the difference between the initial condition $x_I(0) = x_{I0}$

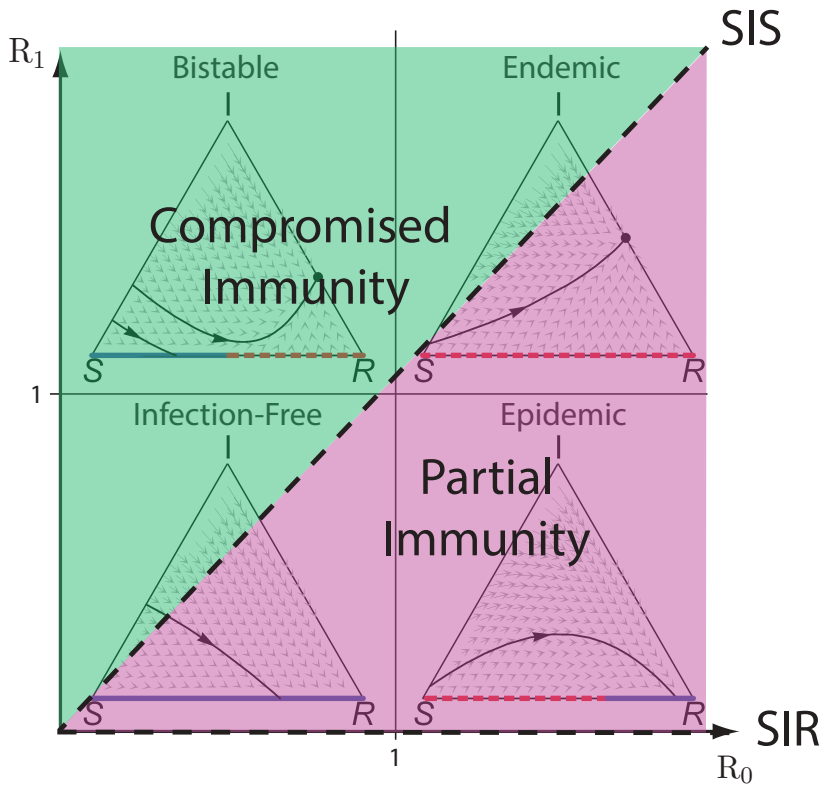

Figure 4.3: The SIRI behavioral regimes describe the possible behaviors of solutions in models with different types of immunity. Adapted from Figure 9.1.

and the critical initial condition x_{IC} goes to zero, x_I spends an arbitrarily long period of time near zero before rapidly increasing towards the EE.

Theorem 2 (Resurgent Epidemic [57])**.** *Consider a solution in the bistable regime with initial condition $x_{IC} < x_{I0} < 1$ such that the solution reaches the EE as $t \to \infty$. For that solution, the fraction of infected individuals decreases initially, reaches a minimum value*

$$x_I^{min} = \frac{R_0 - R_1}{R_1(R_0^{R_0}(1-x_{I0})^{R_1})^{1/(R_0-R_1)}} + \frac{R_1 - 1}{R_1},$$

and then increases until it reaches the EE.

Theorem 3 (Time to Resurgence [57])**.** *Consider a solution in the bistable regime that exhibits a resurgent epidemic. The time to resurgence $t_{RS} = t_{min} - t_0$ satisfies*

the lower bound

$$t_{RS} \geq \frac{\log x_{I0} - \log x_I^{min}}{\delta - \beta},$$

where t_0 *is the initial time and* t_{min} *is the time at which* $x_I = x_I^{min}$. *Moreover,* $t_{RS} \to \infty$ *as* $x_{I0} - x_{IC} \to 0_+$.

Figure 4.4 shows a simulation that exhibits resurgent epidemics with $\beta = 0.5$, $\delta = 1$, and $\hat{\beta} = 1.5$. The initial fraction of infected individuals x_{I0} was set to 0.207. The infection decays at first, reaching a value close to zero after 20 time units. The infection stays close to zero for over 350 time units before increasing towards an endemic state where $x_I = 0.33$. The critical initial conditions for this system is $x_{IC} = 0.206$. In Figure 4.5 we show the time to resurgence and the bound presented in Theorem 3 against x_{I0} for the same system.

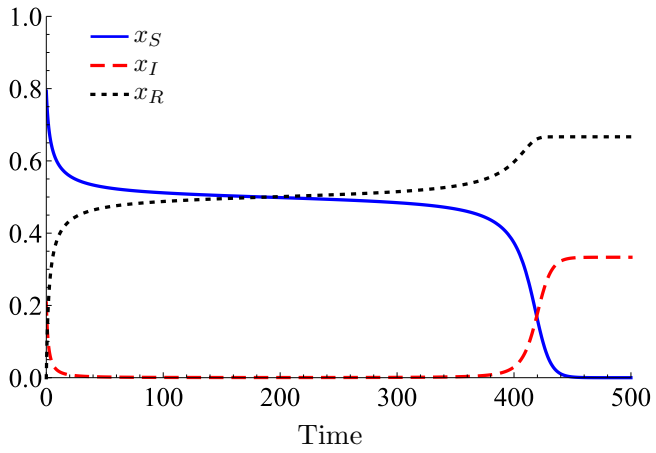

Figure 4.4: Resurgent epidemic for $\beta = 0.5$, $\delta = 1$, $\hat{\beta} = 1.5$, and $x_{I0} = 0.207$. Repeated from Figure 9.2.

4.4 Well-Mixed SIRI Bifurcations

In this section we show an alternative analysis of the SIRI model dynamics (4.3) using both classical bifurcations and bifurcations without parameters (see Sections 2.2 and 2.3). We first analyze the SIS case $R_0 = R_1$ and show that the SIS model exhibits

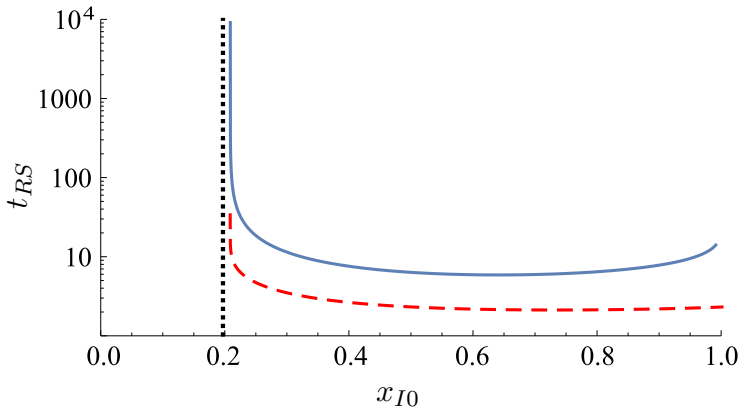

Figure 4.5: Numerical simulations (solid) and lower bound in Theorem 3 (dashed) for time to resurgence t_{RS} versus initial condition x_{I0} for $\beta = 0.5$, $\delta = 1$, $\hat{\beta} = 1.5$. The dotted line shows the critical initial condition $x_{IC} = 0.206$. Repeated from Figure 9.3

a transcritical bifurcation at $R_0 = 1$. We then analyze the case $R_0 \neq R_1$ and show that, at the bifurcation point $x_S^* = M$, $x_I = 0$, the SIRI model dynamics (4.2) exhibit a transcritical bifurcation without parameters with drift singularity.

4.4.1 Transcritical Bifurcation

Assume $R_0 = R_1$. Renaming $x_S + x_R$ as x_S, and using $x_S = 1 - x_I$, the SIRI model (4.2) reduces to the SIS model:

$$\dot{x}_I = (\beta - \delta)x_I - \beta x_I^2.$$

Letting $\mu = \beta - \delta = \delta(R_0 - 1)$, puts (4.4.1) in the transcritical bifurcation normal form $x_I = \mu x_I - x_I^2$ (see Section 2.2). Therefore, the system exhibits a transcritical bifurcation at $x_I = 0$, $R_0 = 1$.

4.4.2 Transcritical Bifurcation Without Parameters

Assume $R_0 \neq R_1$. Then the SIRI model dynamics (4.2) satisfy conditions (TBWP1)–(TBWP6) for the transcritical bifurcation without parameters with drift singularity at the point $x_S = M$, $x_I = 0$, with bifurcation parameter $\mu = R_0(R_1 - 1)$:

- (TBWP1) The system has a line of equilibria at $x_I = 0$.
- (TBWP2) The transverse eigenvalue of J_{IFE} (see (4.4)) is zero at $x_S^* = M$.
- (TBWP3) For all parameter values, $\beta, \hat{\beta}, \delta > 0$ where $\beta \neq \hat{\beta}$, the transverse eigenvalue crosses zero with nonvanishing speed as x_S^* increases,

$$\frac{\partial^2 \dot{x}_I}{\partial x_S \partial x_I}(M, 0) = \beta - \hat{\beta} = \delta(R_0 - R_1). \tag{4.6}$$

- (TBWP4) and (TBWP5) The non-degeneracy condition fails when $\mu = 0$:

$$\frac{\partial \dot{x}_S}{\partial x_I}(M, 0) = -\beta \frac{\delta - \hat{\beta}}{\beta - \hat{\beta}} = \delta \frac{R_0(R_1 - 1)}{R_0 - R_1} = \delta \frac{\mu}{R_0 - R_1}. \tag{4.7}$$

- (TBWP6) The non-degeneracy condition for the bifurcating equilibrium is satisfied:

$$\frac{\partial^2 \dot{x}_S}{\partial \mu \partial x_I}(M, 0) = \frac{1}{\beta - \hat{\beta}} = \frac{1}{R_0 - R_1}.$$

Moreover, we can put the SIRI dynamics (4.3) into the normal form for the transcritical bifurcation without parameters with drift singularity in two steps. First, we shift the coordinates of (4.3) to place the origin at the bifurcation point $(M, 0)$,

$$\begin{aligned} \dot{\bar{x}}_S &= -x_I \delta R_0(\bar{x}_S + M) \\ \dot{x}_I &= x_I \delta\big((R_0 - R_1)\bar{x}_S - R_1 x_I\big). \end{aligned}$$

where $\bar{x}_S = x_S - M$. Second, we scale time by δ by introducing the variable $\tau = t\delta$, and apply the coordinate transformation $y = x_I$, $z = \big((R_0 - R_1)\bar{x}_S - R_1 x_I\big)$ to obtain the normal form

$$\begin{aligned}\frac{\mathrm{d}y}{\mathrm{d}\tau} &= yz \\ \frac{\mathrm{d}z}{\mathrm{d}\tau} &= -R_0R_1y^2 - (R_0 + R_1)yz + R_0(R_1 - 1)y.\end{aligned} \tag{4.8}$$

In the normal form (4.8), the IFE line of equilibria is now given by $(0, z)$, while the EE is given by the equilibrium at $(0, R_0(R_1 - 1)/R_0R_1)$. As R_1 increases, the location of the EE shifts along the z-axis. Figure 4.6 summarizes the results of this section. In each quadrant of the R_0, R_1 parameter space, we show a simulation of the corresponding phase portrait of (4.8). The gray triangles represent the region of the phase portrait corresponding to Δ_2. Similar to Figure 4.1, solid blue lines and dashed red lines correspond to locally stable and unstable points in the IFE, respectively. The isolated equilibrium, corresponding to the EE, is represented as a hollow circle if the equilibrium is unstable and as a black circle if the equilibrium is stable. For $\mu < 0$, or equivalently for $R_1 < 1$, the equilibrium is an unstable node and lies below the y-axis. For $\mu > 0$, or equivalently for $R_1 > 1$, the equilibrium is a stable node and lies above the y-axis. For $\mu = 0$, or equivalently for $R_1 = 1$, the equilibrium is degenerate and lies directly on the y-axis. In the bistable regime, we denote the stable manifold of the bifurcation point $(M, 0)$ in green and the unstable manifold in magenta.

4.5 Concluding Remarks

In this Chapter we analyzed the SIRI model in well-mixed settings. We prove that the SIRI model has a line of infection-free equilibria and an isolated endemic equilibrium. Depending on the system parameters β, $\hat{\beta}$, and δ, the infection-free equilibria is di-

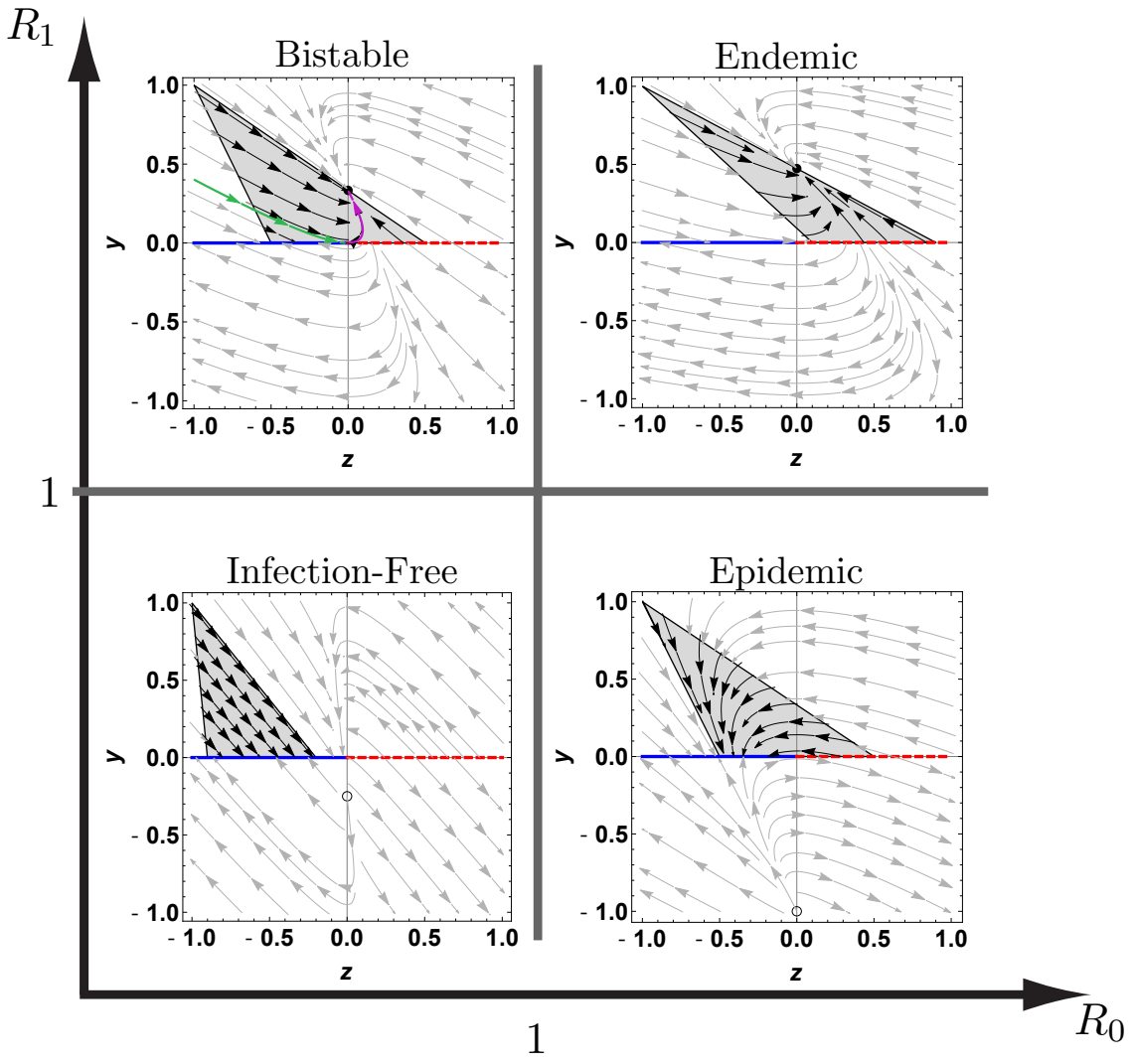

Figure 4.6: The four different behavioral regimes of the SIRI model plotted in the phase space diagram of the transcritical bifurcation without parameters with drift singularity normal form (4.8). The four plots are arranged in the R_0, R_1 parameter space to illustrate the four corresponding regimes. The position of Δ_N (dark gray triangles) within the phase-space depends on the values of R_0 and R_1, which changes the size of the stable (blue) and unstable (dashed red) regions of the IFE set. The critical initial condition x_{IC} in the bistable regime corresponds to the stable manifold (green) of the bifurcation point $(M, 0)$. The unstable manifold (magenta) of the bifurcation point converges to the EE as $t \to \infty$. Bistable regime: $R_0 = 0.5$ and $R_1 = 1.5$. Endemic regime: $R_0 = 1.1$ and $R_1 = 1.9$. Infection-free regime: $R_0 = 0.1$ and $R_1 = 0.8$. Epidemic regime: $R_0 = 1.5$ and $R_1 = 0.5$.

vided into a stable and an unstable region. The transient and steady-state dynamics of the SIRI model are characterized by the values of the infection reproduction num-

ber R_0 and the reinfection reproduction number R_1 and can be categorized into four behavioral regimes: infection-free, endemic, epidemic, and bistable. We proved that in the bistable regime there exists a critical initial condition that separates solutions where the infection dies out from solutions where the infection spreads and remains endemic. Moreover, we showed that solutions in the bistable regime that go endemic exhibit a resurgent epidemic where the fraction of infected individuals remains close to zero for arbitrarily long periods of time. Finally, we showed that the four behavioral regimes in the SIRI model can be explained by the existence of a transcritical bifurcation without parameters with drift singularity.

Our results show that the transient and steady-state dynamics that describe how the contagion process spreads through a population are captured by the two scalar parameters R_0 and R_1. These parameters describe how individuals within the population adapt after a first exposure to the contagion and can be used to provide straightforward predictions on the short and long term behavior of the group. The existence of the bistable regime in the SIRI model suggests that it is possible to design a contagion process such that it initially decays, avoiding detection, before rapidly spreading through the population at a later time. Control strategies that seeks to eradicate the contagion process from the population can prevent the resurgent epidemic by stopping the control effort only once it is absolutely certain that the contagion has been completely eradicated.

Chapter 5

Contagion Processes on Networks of Heterogeneous Agents with Adaptive Susceptibility

In this chapter we study the SIRI epidemiological model on network topologies. The model considers a contagion process that spreads across a population of two or more interconnected individuals where the susceptibility of each individual changes after a first exposure to the infection. We generalize results from Chapter 4 and show how the interaction topology, represented by a strongly connected digraph, affects the transient and steady-state dynamics of the contagion. Unlike the well-mixed SIRI model studied in Chapter 4, individuals in the network SIRI model are not assumed to be homogeneous. Different individuals may react differently in response to a first exposure to the infection. Through the network SIRI model, we examine the role of network structure as well as heterogeneity among individuals in infection, re-infection, and recovery rates, in the spread of contagion processes. Our analytic results show the effect that individual adaptations in susceptibility have on group behavior and

provide design opportunities for multi-agent systems capable of robust and flexible collective behavior.

Sections 5.1-5.6 summarize results from Part II: Chapter 10, which appears as Pagliara and Leonard [58]. In Section 5.1 we introduce the network SIRI model. In Section 5.2 we classify the network SIRI model into six cases and introduce the notion of the basic and extreme basic reproduction numbers. In Section 5.3 we present results for the stability of equilibria, with an emphasis on the stability of infection-free points, which form a local manifold of equilibria. In Section 5.4 we show that the basic and extreme basic reproduction numbers characterize four distinct behavioral regimes: infection-free, endemic, epidemic, and bistable. These regimes are analogous to the behavioral regimes of the well-mixed SIRI model in Chapter 4. In Section 5.5 we describe properties of the network SIRI model near points at which the infection-free set of equilibria loses normal hyperbolicity. In Section 5.6 we present analytical and numerical results for the bistable regime in the case of d-regular digraphs with N agents and complete digraphs with two agents ($N = 2$). In Section 5.7 we discuss how our results suggest the existence of a codimension-N transcritical bifurcation without parameters with drift singularity in the network SIRI model.

5.1 From Markov Process to ODEs: Individual-Based Mean-Field Approach

In this section we derive the network SIRI model, which considers a contagious process with reinfection in a population of N agents (i.e, individuals). We consider a digraph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with N nodes and adjacency matrix $A = \{a_{jk}\}$. Each node $j \in \mathcal{V}$ represents an agent and each directed edge $(j, k) \in \mathcal{E}$ represents an interaction channel between agents j and k, where j is susceptible to k but k is not susceptible to j. I.e., k can infect j but j cannot infect k. If both (j, k) and (k, j) are in $\mathcal{E}$, then agents

j and k are susceptible to each other. The state of each agent is given by the random variable $X_j(t) \in \{S, I, R\}$, where S, I and R correspond to "susceptible", "infected", and "recovered", respectively. Agents change state according to independent Poisson processes with rates defined as follows: susceptible agent j becomes infected through contact with infected neighbor k at the rate $\beta_{jk} = f_{jk}(a_{jk}) \geq 0$. Infected agent j recovers from the infection at the rate $\delta_j \geq 0$. Recovered agent j becomes reinfected through contact with infected neighbor k at the rate $\hat{\beta}_{jk} = \hat{f}_{jk}(a_{jk}) \geq 0$.[1] These transitions are summarized as:

$$S_j + I_k \xrightarrow{\beta_{jk}} I_j + I_k \xleftarrow{\hat{\beta}_{jk}} R_j + I_k$$

$$I_j \xrightarrow{\delta_j} R_j.$$

Agents can become infected or reinfected by interacting with infected neighbors only. That is, $\beta_{jk} = 0$, and $\hat{\beta}_{jk} = 0$ if $(j, k) \notin \mathcal{E}$ (i.e., $a_{jk} = 0$). We assume that if $(j, k) \in \mathcal{E}$ (i.e., $a_{jk} \neq 0$), then $\beta_{jk} > 0$ but $\hat{\beta}_{jk} \geq 0$; we allow the susceptibility of j to k to change after a first infection (by any infected neighbor of j) such that j is no longer susceptible to k.

The contagion process as described above corresponds to a Markov process with 3^N states. Here, we use the individual-based mean-field approach (IBMF) introduced in Section 1.2.2 to derive a reduced model consisting of $3N$ continuous-time ODEs for the probability p_j^{Ω} of node j belonging to compartment $\Omega \in \{S, I, R\}$ at time t. The IBMF approach is a mean-field approximation that assumes that the state of every node in the network is statistically independent of the state of its nearest neighbors. We refer the reader to [135, 136] for a discussion and numerical exploration of the accuracy of mean-field approximations in network dynamics. We follow the IBMF approach as described in the derivation of the network SIS model in [75,77]. We begin

[1]The functions f_{jk} and $\hat{f}_{jk}$ describe the relationship between the adjacency matrix entry a_{jk} and the infection rate β_{jk} and reinfection rate $\hat{\beta}_{jk}$, respectively. For example, the relationship could be linear: $f_{jk} = c_{jk}a_{jk}$, $\hat{f}_{jk} = \hat{c}_{jk}a_{jk}$ for some $c_{jk}, \hat{c}_{jk} \in \mathbb{R}$.

by examining the Markov process in more detail. We highlight that if solutions in an IBMF model converge to an infection-free equilibrium, then the stochastic Markov model reaches the infection-free absorbing state in sublinear time with respect to the size of the network in expectation [74].

Definition 5.1.1 (Markov process transition rate [137]). *For an agent j, the transition rate between any two states $\Omega_1, \Omega_2 \in \{S, I, R\}$ where $\Omega_1 \neq \Omega_2$ is*

$$q^j_{\Omega_1,\Omega_2} = \lim_{\epsilon \to 0} \frac{\mathcal{P}[X_j(t+\epsilon) = \Omega_2 \,|\, X_j(t) = \Omega_1]}{\epsilon},$$

where $\mathcal{P}[\cdot]$ denotes probability.

For the SIRI transitions summarized in 5.1, the possible transition rates are:

- $q^j_{SI} = \sum_{k=1}^{N} \beta_{jk} \mathbb{1}_{X_j(t)=I}$, where $\mathbb{1}_x : x \mapsto \{0, 1\}$ is the indicator function associated with the event x:
$$\mathbb{1}_x = \begin{cases} 1 & x \text{ is true} \\ 0 & \text{otherwise.} \end{cases}$$
- $q^j_{IR} = \delta_j$.
- $q^j_{RI} = \sum_{k=1}^{N} \hat{\beta}_{jk} \mathbb{1}_{X_j(t)=I}$.

Definition 5.1.2 (Infinitesimal Generator). *For an agent j, the infinitesimal generator is*

$$Q_j = \begin{bmatrix} -q^j_{SI} & q^j_{SI} & 0 \\ 0 & -q^j_{IR} & q^j_{IR} \\ 0 & q^j_{RI} & -q^j_{IR} \end{bmatrix}.$$

The transition rates in the infinitesimal generator Q_j for node j are themselves random variables that depend on the random state of the neighbors of j, which themselves are coupled to the random state of their neighbors, etc. To simplify

the transition rates, we replace the random infection transition rates $q^j_{S,I}$ and the reinfection transition rates $q^j_{R,I}$ with "effective" infection and reinfection transition rates

$$\mathbb{E}[q^j_{S,I}] = \sum_{k=1}^{N} \beta_{jk} \mathcal{P}[X_k = I] \tag{5.1}$$

$$\mathbb{E}[q^j_{R,I}] = \sum_{k=1}^{N} \hat{\beta}_{jk} \mathcal{P}[X_k = I] \tag{5.2}$$

which results in the *effective infinitesimal generator*

$$\bar{Q}_i = \begin{bmatrix} -\mathbb{E}[q^j_{SI}] & \mathbb{E}[q^j_{SI}] & 0 \\ 0 & -q^j_{IR} & q^j_{IR} \\ 0 & \mathbb{E}[q^j_{RI}] & -\mathbb{E}[q^j_{RI}]. \end{bmatrix} \tag{5.3}$$

The effective $\bar{Q}_i$ allows us to proceed with Markov theory and define dynamic equations for the time evolution of the probabilities p^S_j, p^I_j, and p^R_j:

$$\begin{aligned} \dot{p}^S_j &= - p^S_j \sum_{k=1}^{N} \beta_{jk} p^I_k \\ \dot{p}^I_j &= - \delta_j p^I_j + p^S_j \sum_{k=1}^{N} \beta_{jk} p^I_k + p^R_j \sum_{k=1}^{N} \hat{\beta}_{jk} p^I_k \\ \dot{p}^R_j &= - p^R_j \sum_{k=1}^{N} \hat{\beta}_{jk} p^I_k + \delta_j p^I_k. \end{aligned} \tag{5.4}$$

We can reduce the number of equations from $3N$ to $2N$ by using the substitution $p^R_j = 1 - p^S_j - p^I_j$ in (5.4):

$$\begin{aligned} \dot{p}^S_j &= - p^S_j \sum_{k=1}^{N} \beta_{jk} p^I_k \\ \dot{p}^I_j &= - \delta_j p^I_j + \sum_{k=1}^{N} \big((1 - p^S_j)\hat{\beta}_{jk} + p^S_j \beta_{jk}\big) p^I_k - p^I_j \sum_{j=1}^{N} \hat{\beta}_{jk} p^I_k. \end{aligned} \tag{5.5}$$

The network SIRI dynamics (5.5) can be written in matrix form by defining the vectors $\mathbf{p}^{\Omega} = [p_1^{\Omega}, \cdots, p_N^{\Omega}]^T$ for $\Omega \in \{S, I\}$:

$$\begin{aligned} \dot{\mathbf{p}}^S &= -P^S B \mathbf{p}^I \\ \dot{\mathbf{p}}^I &= \big((\mathbb{I} - P^S)\hat{B} + P^S B\big)\mathbf{p}^I - \Gamma \mathbf{p}^I - P^I \hat{B} \mathbf{p}^I, \end{aligned} \tag{5.6}$$

where $P^{\Omega} = \text{diag}(\mathbf{p}^{\Omega})$, $B = \{\beta_{jk}\} \succ \bar{\mathbf{0}}$ is the infection matrix, $\hat{B} = \{\hat{\beta}_{jk}\} \succeq \bar{\mathbf{0}}$ is the reinfection matrix, and $\Gamma = \text{diag}(\delta_1, \ldots, \delta_N) \succeq \bar{\mathbf{0}}$ is the recovery matrix.

In the remainder of this Chapter, we assume that Γ is non-singular and that $\hat{B}$ is irreducible. An irreducible $\hat{B}$ implies that the graph $\mathcal{G}_{\hat{B}}$, generated by interpreting $\hat{B}$ as its adjacency matrix, is strongly connected. As a result, at any stage in the spread of the contagion, there is a directed path connecting any two agents in the network through which the infection can spread.

The reducible $\hat{B}$ case depends on the results of the irreducible case. If $\hat{B}$ is reducible, then $\mathcal{G}_{\hat{B}}$ is weakly connected or disconnected. If $\mathcal{G}_{\hat{B}}$ is weakly connected, the adjacency matrix of $\mathcal{G}_{\hat{B}}$ can always be written as an upper block triangular matrix with K diagonal irreducible blocks which describe the K strongly connected subgraphs of $\mathcal{G}$ (see [138] for a study of the reducible case in the network SIS model). If $\mathcal{G}_{\hat{B}}$ is disconnected, then each connected subgraph of $\mathcal{G}_{\hat{B}}$ can be independently analyzed.

5.2 Classification and Reproduction Numbers

The network SIRI model can be classified as the following six different families of models (summarized in Table 5.1) [quoted from [58]]:

- Case 1 (SI): When $\Gamma = \bar{\mathbf{0}}$ the network SIRI model specializes to the network SI model.

- Case 2 (SIR): When $\hat{B} = \bar{\mathbf{0}}$, the network SIRI model specializes to the network SIR model.
- Case 3 (SIS): When $B = \hat{B}$ the network SIRI model specializes to the network SIS model with $\mathbf{p}^S \mapsto \mathbf{p}^S + \mathbf{p}^R$.
- Case 4 (Partial Immunity): When $B \succ \hat{B} \succ \bar{\mathbf{0}}$, every recovered agent acquires partial (or no) immunity to each of its infected neighbors.
- Case 5 (Compromised Immunity): When $\hat{B} \succ B \succ \bar{\mathbf{0}}$, every recovered agent acquires compromised (or no) immunity to each of its infected neighbors.
- Case 6 (Mixed Immunity): Models not in Cases 1-5. Notably, there is at least one pair of edges (j,k) and (l,m) such that $\beta_{jk} \geq \hat{\beta}_{jk}$ and $\beta_{lm} < \hat{\beta}_{lm}$. We classify mixed immunity into two sub-cases:
  - Case 6a (Weak Mixed Immunity): For every agent j, $\beta_{jk} - \hat{\beta}_{jk} \geq 0$ for all $k \in \mathcal{N}_j$ or $\beta_{jk} - \hat{\beta}_{jk} \leq 0$ for all $k \in \mathcal{N}_j$.
  - Case 6b (Strong Mixed Immunity): Mixed immunity that is not weak.

Note that the first five cases (Cases 1-5) are analogous to the cases of the well-mixed SIRI model described in (4.1). Case 6 showcases how, unlike the well-mixed SIRI model, the network SIRI model can be used to study contagion processes in networks of heterogeneous agents, where agents adapt differently to the infection after a first exposure.

Previous work on the network SIS model [77, 79, 81]

$$\dot{\mathbf{p}}^I = (B - \Gamma)\mathbf{p}^I - P^I B \mathbf{p}^I, \tag{5.7}$$

has shown that the steady-state behavior of solutions depends on the value of the basic reproduction number $R_0 = \rho(B\Gamma^{-1})$ where $\rho(\cdot)$ is the spectral radius. If $R_0 \leq$

Table 5.1: Network SIRI model cases. Repeated from Table 10.1

Case	Parameter Value	Equivalent Model
1	$\Gamma = \bar{\mathbf{0}}$	SI
2	$\hat{B} = \bar{\mathbf{0}}$	SIR
3	$B = \hat{B}$	SIS
4	$B \succ \hat{B} \succ \bar{\mathbf{0}}$	Partial Immunity
5	$\hat{B} \succ B \succ \bar{\mathbf{0}}$	Compromised Immunity
6	Otherwise	Mixed Immunity

1, solutions converge to an infection-free steady-state, while if $R_0 > 1$, solutions reach an endemic steady-state where the probability of infection of every agent is nonzero [81, 139].

Based on the network SIS results and the well-mixed SIRI model results presented in Chapter (4), we define four scalar quantities that extend the notion of the basic reproduction number to contagion processes in which agents adapt their susceptibility to the infection after a first exposure.

Definition 5.2.1 (Basic Reproduction Numbers)**.** *The basic infection reproduction number is* $R_0 = \rho(B\Gamma^{-1})$ *and the basic reinfection reproduction number is* $R_1 = \rho(\hat{B}\Gamma^{-1})$.

Definition 5.2.2 (Extreme Basic Reproduction Numbers)**.** *Let* $B^*(\mathbf{p}^S) = (\mathbb{I} - P^S)\hat{B} + P^S B$. *The maximum basic reproduction number is* $R_{max} = \max_{\mathbf{p}^S} \rho(B^*(\mathbf{p}^S)\Gamma^{-1})$ *and the minimum basic reproduction number is* $R_{min} = \min_{\mathbf{p}^S} \rho(B^*(\mathbf{p}^S)\Gamma^{-1})$.

Remark 6 ([58]). Each reproduction number corresponds to $K = B^*(\mathbf{p}^S) - \Gamma$ for a particular value of $\mathbf{p}^S$, where K is the linear term of the dynamics of $\mathbf{p}^I$ in (5.6). K is a Metzler matrix and each reproduction number R is the spectral radius associated with this regular splitting, i.e., $R = \rho(B^*(\mathbf{p}^S)\Gamma^{-1})$ as defined in Proposition 3. R_0 is R for $P^S = \mathbb{I}$, R_1 is R for $P^S = \bar{\mathbf{0}}$, R_{max} is R for $\mathbf{p}^S = \text{argmax}_{\mathbf{p}^S} \rho(B^*(\mathbf{p}^S)\Gamma^{-1})$ and R_{min} is R for $\mathbf{p}^S = \text{argmin}_{\mathbf{p}^S} \rho(B^*(\mathbf{p}^S)\Gamma^{-1})$.

5.3 Manifold of Infection-Free Equilibria

The network SIRI model (5.6) dynamics evolve on a $2N$-dimensional space corresponding to N copies of the the 2-simplex $\Delta := \{(p_j^S, p_j^I, p_j^R) \in [0,1]^3 |\, p_j^S + p_j^I + p_j^R = 1\}$. On this space, the only equilibria of the network SIRI model (5.6) with $\hat{B}$ irreducible are an invariant set of infection-free equilibria (IFE) $\mathcal{M} = \{(\mathbf{p}^{S*}, \mathbf{0}) \in \Delta_N |\, \mathbf{0} \preceq \mathbf{p}^{S*} \preceq \mathbf{1}\}$ and a unique isolated endemic equilibrium (EE) where $\mathbf{p}^{S*} = \mathbf{0}$ and $\mathbf{p}^{I*} \gg \mathbf{0}$ satisfy

$$p_j^{I*} = \frac{\sum_{k=1}^{N} \hat{\beta}_{jk} p_k^{I*}}{\delta_j + \sum_{k=1}^{N} \hat{\beta}_{jk} p_k^{I*}}. \tag{5.8}$$

The EE is locally stable and exists if and only if $R_1 > 1$ (see Proposition 9 for details). To study the stability of points in the IFE set $\mathcal{M}$, we consider the network SIRI dynamics on $\mathbb{R}^{2N}$. In this extended space, $\mathcal{M}$ is a subset of the manifold of equilibria $\mathcal{M}' = \{(\mathbf{p}, \mathbf{0}) | \mathbf{p} \in \mathbb{R}^N, \mathbf{0} \in \mathbb{R}^N\}$.

Definition 5.3.1. *The boundary of* $\mathcal{M}$ *is* $\partial\mathcal{M} = \{\mathbf{x} = (\mathbf{p}^{S*}, \mathbf{0}) \in \mathcal{M} \,|\, \exists j,\, p_j^{S*} \in \{0,1\}\}$. *The corner set of* $\mathcal{M}$ *is* $\hat{\mathcal{M}} = \{\mathbf{x} = (\mathbf{p}^{S*}, \mathbf{0}) \in \partial\mathcal{M} \,|\, p_j^{S*} \in \{0,1\},\, \forall j\}$. *The interior of* $\mathcal{M}$ *is* $\text{int}\, \mathcal{M} = \mathcal{M} \setminus \partial\mathcal{M}$.

5.3.1 Stability of points in the IFE set $\mathcal{M}$

For an arbitrary point $\mathbf{x} = (\mathbf{p}^{S*}, \mathbf{0}) \in \mathcal{M}$, the Jacobian is given by

$$J_{\mathcal{M}}(\mathbf{x}) = \begin{bmatrix} \bar{\mathbf{0}} & -P^{S*}B \\ \bar{\mathbf{0}} & J_T(\mathbf{p}^{S*}) \end{bmatrix} \tag{5.9}$$

where $J_T(\mathbf{p}^{S*}) = B^*(\mathbf{p}^{S*}) - \Gamma$. The Jacobian $J_{\mathcal{M}}(\mathbf{x})$ has N zero eigenvalues corresponding to the N dimensions tangent to $\mathcal{M}$. The N transverse eigenvalues of $J_{\mathcal{M}}(\mathbf{x})$ are the eigenvalues of $J_T(p^{S*})$.

Remark 7. $J_T(\mathbf{p}^{S*})$ is Metzler irreducible since B and $\hat{B}$ are Metzler irreducible.

Definition 5.3.2. *The leading transverse eigenvalue of* $J_{\mathcal{M}}(\mathbf{x})$ *is* $\lambda_{Tmax}(J_{\mathcal{M}}(\mathbf{x})) = \lambda_{max}(J_T(\mathbf{p}^{S*})) \in \mathbb{R}$.

The stability of points in the IFE set $\mathcal{M}$ depends on the sign of $\lambda_{Tmax}(J_{\mathcal{M}}(\mathbf{x}))$.

Lemma 4 (Local Stability of Points in the IFE set $\mathcal{M}$. Adapted from [58])**.** *Let* $\mathbf{x} = (\mathbf{p}^{S*}, \mathbf{0}) \in \mathcal{M}$. *Then the following hold.*

- *Suppose* $\lambda_{Tmax}(J_{\mathcal{M}}(\mathbf{x})) < 0$. *Then,* $\mathbf{x}$ *is locally stable. I.e., given a neighborhood* U *of* $\mathbf{x}$ *on* $\mathcal{M}'$ *such that* $\lambda_{Tmax}(J_{\mathcal{M}}(\mathbf{u})) < 0$ *for all* $\mathbf{u} \in U$, *there exists* $V \subset \Delta_N$ *and* $\mathbf{x} \in V$ *such that any solution starting in* V *converges exponentially to a point in* $U \cap \Delta_N$.

- *Suppose* $\lambda_{Tmax}(J_{\mathcal{M}}(\mathbf{x})) > 0$. *Then,* $\mathbf{x}$ *is unstable. I.e., there exists* $W \subset \Delta_N$ *and* $\mathbf{x} \in W$, *such that any solution starting in* W *leaves* W.

5.3.2 Stable, Unstable, and Center Subsets of $\mathcal{M}$

Based on the results of Lemma 4, we partition the IFE set $\mathcal{M}$ into three subsets.

Definition 5.3.3 (Stable, unstable, and center IFE subsets)**.** *The stable IFE subset is* $\mathcal{M}_- = \{\mathbf{x} \in \mathcal{M} \,|\, \lambda_{Tmax}(J_{\mathcal{M}}(\mathbf{x})) < 0\}$. *The unstable IFE subset is* $\mathcal{M}_+ = \{\mathbf{x} \in \mathcal{M} \,|\, \lambda_{Tmax}(J_{\mathcal{M}}(\mathbf{x})) > 0\}$. *The center IFE subset is* $\mathcal{M}_0 = \{\mathbf{x} \in \mathcal{M} \,|\, \lambda_{Tmax}(J_{\mathcal{M}}(\mathbf{x})) = 0\}$.

Proposition 4 ([58])**.** $\mathcal{M}_- \cup \mathcal{M}_+ \cup \mathcal{M}_0 = \mathcal{M}$. *Every point in* $\mathcal{M}_-$ *is locally stable and every point in* $\mathcal{M}_+$ *is unstable.*

We prove that the size and location of the subsets $\mathcal{M}_-$, $\mathcal{M}_+$, and $\mathcal{M}_0$ depend on the value of the extreme basic reproduction numbers R_{max} and R_{min}.

Theorem 4 (Stability of the IFE set $\mathcal{M}$ [58])**.**

(A) If $R_{max} < 1$, *then* $\mathcal{M}_- = \mathcal{M}$.

(B) If $R_{min} > 1$, then $\mathcal{M}_+ = \mathcal{M}$.

(C) If $R_{max} = R_{min} = 1$, then $\mathcal{M}_0 = \mathcal{M}$.

(D) If $R_{min} < R_{max} = 1$, then $\mathcal{M}_- = \mathcal{M} \setminus \mathcal{M}_0$ and $\mathcal{M}_0 \subset \partial\mathcal{M}$.

(E) If $R_{max} > R_{min} = 1$, then $\mathcal{M}_+ = \mathcal{M} \setminus \mathcal{M}_0$ and $\mathcal{M}_0 \subset \partial\mathcal{M}$.

(F) If $R_{max} > 1$ and $R_{min} < 1$, then $\mathcal{M}_-, \mathcal{M}_+, \mathcal{M}_0 \neq \emptyset$ and each subset consists of n_-, n_+, n_0 connected sets, respectively. Each of the center connected sets $\mathcal{M}_0^j$, $j = 1, \ldots, n_0$, is an $N-1$-dimensional smooth hypersurface with boundary $\partial\mathcal{M}_0^j \subset \partial\mathcal{M}$. Each $\mathcal{M}_0^j$ separates an N-dimensional stable connected hypervolume from an N-dimensional unstable connected hypervolume.

Remark 8 ([58]). Theorem 4 applies to the six different cases of the network SIRI model (see Table 4.1) as follows. (A) applies to Cases 2, 3, 4, 5, and 6. (B) applies to Cases 3, 4, 5, and 6. (C) applies to Case 3. (D) applies to Cases 2, 4, 5, and 6. (E) applies to Cases 4, 5, and 6. (F) applies to Cases 2, 4, 5, and 6. We specialize (F) in Theorem 5 to provide the key to characterizing global behavior in Cases 2, 4, 5, and 6a.

Theorem 5 (Uniqueness of stable, unstable, and center subsets [58])**.** *If $R_{max} > 1$ and $R_{min} < 1$, then for Cases 2, 4, 5, 6a, $\mathcal{M}_0$ consists of a unique $N-1$-dimensional surface with boundary $\partial\mathcal{M}_0 \subset \partial\mathcal{M}$ dividing $\mathcal{M}$ into $\mathcal{M}_-$ and $\mathcal{M}_+$.*

Remark 9 ([58]). We conjecture that Theorem 5 can be extended to Case 6b. Extensive simulations of the IFE subset $\mathcal{M}$ for a three agent network (i.e., $N = 3$), with different network configurations and parameter values, consistently show a unique connected surface $\mathcal{M}_0$ dividing $\mathcal{M}$ into $\mathcal{M}_-$ and $\mathcal{M}_+$.

Remark 10. The proofs of Theorems 4 and 5 rely on the description of $\lambda_{Tmax}(J_{\mathcal{M}}(\mathbf{x}))$, for $\mathbf{x} = (\mathbf{p}, \mathbf{0})$ in a subset E of $\mathcal{M}'$ as a potential function with respect to $\mathbf{p}$, which

we call the λ_{Tmax} function $\Lambda(\mathbf{p}, \mathbf{0})$. We introduce and make use of these results in Section 5.5 to analyze the dynamics of the network SIRI model near points in $\mathcal{M}_0$.

5.4 Transient and Steady-State Dynamics

The analysis of the global dynamics of the network SIRI model (5.6) is facilitated by the decreasing monotonicity of the solutions along the $\mathbf{p}^S$ directions, which implies that solutions either converge to a point in the IFE set $\mathcal{M}$ or to the EE as $t \to \infty$.

Definition 5.4.1 ([87]). *A point* $\mathbf{x}$ *is an* ω*-limit point of a solution* $\mathbf{y}(t, \mathbf{y}_0) = \mathbf{y}(t)$ *if there is a sequence* $\{t_n\}$*, with* $t_n \to \infty$ *as* $n \to \infty$*, such that* $\mathbf{y}(t_n) \to \mathbf{x}$ *as* $n \to \infty$*. The set of all* ω*-limit points is called the* ω*-limit set* $\Omega(\mathbf{y}_0)$ *of* $\mathbf{y}(t)$*.*

Lemma 5 ([58]). *Let* $\mathbf{y}(t, \mathbf{y}_0) = (\mathbf{p}^S(t), \mathbf{p}^I(t))$ *be the solution of* (10.3) *with initial condition* $\mathbf{y}_0 \in \Delta_N$*. Then the following hold:*

- *The* ω*-limit set* $\Omega(\mathbf{y}_0)$ *of* $\mathbf{y}(t, \mathbf{y}_0)$ *is either a point in the IFE set* $\mathcal{M}$ *or the EE.*
- $\mathbf{y}(t, \mathbf{y}_0)$ *cannot exhibit non-trivial periodic orbits.*
- $\mathbf{y}(t, \mathbf{y}_0)$ *converges to a point in* $\mathcal{M}$ *as* $t \to \infty$ *if* $R_1 \leq 1$*.*

The transient and steady-state dynamics of the network SIRI model (5.6) can be categorized into four distinct behavioral regimes, analogous to those of the well-mixed SIRI model described in Theorem 1:

Theorem 6 (Behavioral Regimes [58]). *Let* $\mathbf{p}^I(0) \succ \mathbf{0}$ *and* $\mathbf{w}_m^T$ *be the leading left-eigenvector of* $B^*(\mathbf{p}^S)\Gamma^{-1}$*, where* $\mathbf{p}^S = \text{argmax}_{\mathbf{p}^S} \rho(B^*(\mathbf{p}^S)\Gamma^{-1})$*. Then the network SIRI model* (5.6) *exhibits four qualitatively distinct behavioral regimes:*

1. *Infection-Free Regime: If* $R_{max} \leq 1$ *then the following hold.*
   (a) *All solutions converge to a point in* $\mathcal{M}$ *as* $t \to \infty$*.*

Table 5.2: Behavioral Regimes of the Network SIRI Special Cases. Repeated from 10.2

Case	Model	Infection-Free	Endemic	Epidemic	Bistable
1	SI		✓		
2	SIR	✓		✓	
3	SIS	✓	✓		
4	Partial	✓	✓	✓	
5	Compromised	✓	✓		✓
6	Mixed	✓	✓	✓	✓

(b) If $B \succeq \hat{B}$ or $\hat{B} \succeq B$, the weighted infected average $p^I_{\text{avg}}(t) = \mathbf{w}_m^T \Gamma^{-1} \mathbf{p}^I(t)$ decays monotonically to zero.

2. *Endemic Regime: If $R_{min} > 1$ then all solutions converge to the EE as $t \to \infty$.*

3. *Epidemic Regime: If $R_{min} < 1, R_{max} > 1$ and $R_1 \leq 1$, then the following hold.*

   (a) All solutions converge to a point in $\mathcal{M}$ as $t \to \infty$.

   (b) There exists $H \in \Delta_N$ and $H \supset \mathcal{M}_+$ that is foliated by families of heteroclinic orbits, each orbit connecting two points in $\mathcal{M}$.

4. *Bistable Regime: If $R_{min} < 1, R_{max} > 1$ and $R_1 > 1$, then, depending on the initial conditions, solutions converge to a point in $\mathcal{M}$ or converge to the EE as $t \to \infty$.*

Remark 11 ([58]). Theorem 6 applies to the six different cases of the network SIRI model as follows (summarized in Table 5.2).[2] The infection-free regime applies to Cases 2, 3, 4, 5, and 6. The endemic regime applies to Cases 3, 4, 5, and 6. The epidemic regime applies to Cases 2, 4, and 6. The bistable regime applies to Cases 5 and 6.

[2]The basic and extreme basic reproduction numbers are not well-defined for the network SI model (Case 1). Therefore the results of Theorem 6 do not apply. However, we include the SI case in Table 5.2 under the endemic regime because all solutions in the model reach the EE $\mathbf{p}^{I*} = \mathbf{1}$ as $t \to \infty$ [69].

5.4.1 Examples

Figures 5.1 to 5.4 illustrate the results of Theorem 6 on a strongly connected digraph $\mathcal{G}$ with four agents ($N = 4$). We show how knowledge of the basic and extreme basic reproduction numbers is sufficient to predict the effects of changes in graph structure, system parameters, and initial conditions on the transient and steady-state dynamics of the network SIRI model. Panels in the figures show time-series for $p_j^I, j \in \{1, 2, 3, 4\}$ (solid colored lines).

Definition 5.4.2 (Stubborn agents)**.** *An agent $j \in \mathcal{V}$ is stubborn if $(\beta_{jk} - \hat{\beta}_{jk}) = 0$ for all $k \in \mathcal{N}_j$.*

Effects of the recovery rates

Figure 5.1 shows the effects of heterogeneity in the recovery rates on the dynamics of a partial immunity case (Case 4) network with adjacency matrix A, where agent 1 acquires partial immunity to all its infected neighbors, and agents 2, 3, and 4 are stubborn. We set $B = 0.9A$ and $\hat{B} = \text{diag}([0.5, 0.9, 0.9, 0.9])A$. We use the initial condition $\mathbf{p}^S(0) = \mathbf{1} - \mathbf{p}^I(0)$, $\mathbf{p}^I(0) = [0.1, 0.1, 0.1, 0.1]^T$. In Figure 5.1A, all agents have the same recovery rate ($\delta_j = 1$) and the reproduction numbers are $R_{max} = R_0 = 1.19$ and $R_{min} = R_1 = 0.94$. This places the network in the epidemic regime: the probability of infection of every agent in the network initially increases, reaches a maximum and decays to zero. In Figure 5.1B, we halve the recovery rate of agent 1 ($\delta_1 = 0.5$) and double the recovery rate of agent 4 ($\delta_4 = 2$). The new reproduction numbers are $R_{max} = R_0 = 1.37$ and $R_{min} = R_1 = 1.09$. This places the network in the endemic regime: the probability of infection of all agents is non-zero at steady-state. Note that the steady-state probability of infection $\mathbf{p}^{I*}$ is given by (5.8), which describes how p_j^{I*} is inversely proportional to the recovery rate δ_j.

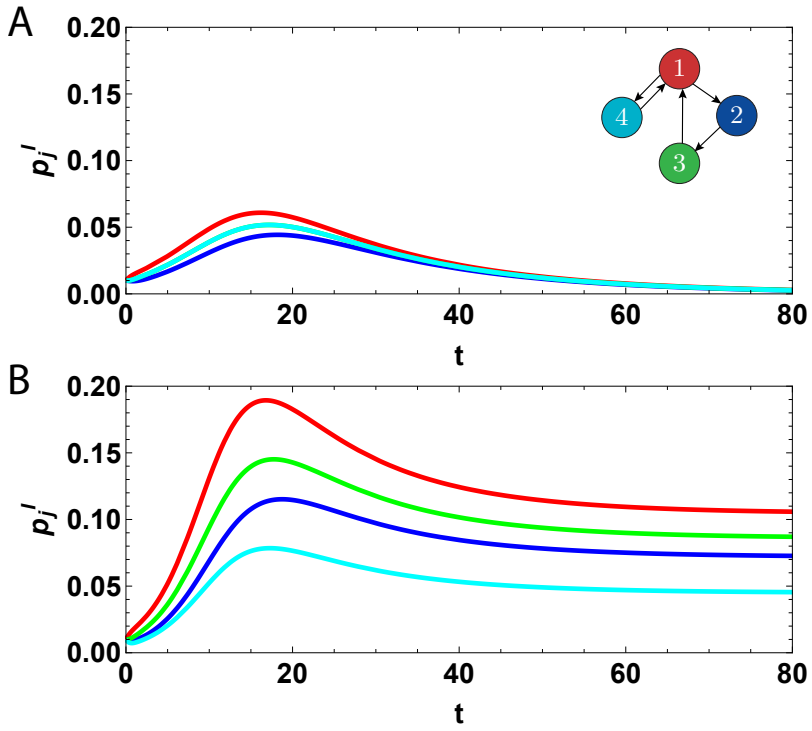

Figure 5.1: Effects of heterogeneity in the recovery rates on the steady-state behavior. Partial immunity case (Case 4) with four agents. Agent 1 acquires partial immunity to all its infected neighbors, and agents 2, 3, and 4 are stubborn. A. $\Gamma = \text{diag}([1, 1, 1, 1])$, $R_{max} = R_0 = 1.19$, and $R_{min} = R_1 = 0.94$. B. $\Gamma = \text{diag}([0.5, 1, 1, 2])$, $R_{max} = R_0 = 1.37$, and $R_{min} = R_1 = 1.09$. In both panels we set $B = 0.9A$ and $\hat{B} = \text{diag}([0.5, 0.9, 0.9, 0.9])A$, and use the initial condition $\mathbf{p}^S(0) = \mathbf{1} - \mathbf{p}^I(0)$, $\mathbf{p}^I(0) = [0.1, 0.1, 0.1, 0.1]^T$.

Effects of the reinfection rates

Figure 5.2 shows the effects of heterogeneity in the reinfection rates on the steady-state behavior. In Figure 5.2A, we consider again the partial immunity case (Case 4) in which all agents acquire partial immunity to all their infected neighbors: $B = 1.5A$, $\hat{B} = 0.5A$, and $\Gamma = \mathbb{I}$. The corresponding reproduction numbers are $R_{max} = R_0 = 1.99$ and $R_{min} = R_1 = 0.66$. This places the network in the epidemic regime: the probability of infection of every agent in the network initially increases, reaches a maximum and decays to zero. In Figure 5.2B, we increase the susceptibility to reinfection of agent 4 to $\hat{\beta}_{41} = 2.5$. The network is now in the weak mixed immunity case (Case 6a) with agent 4 acquiring compromised immunity to all its infected neigh-

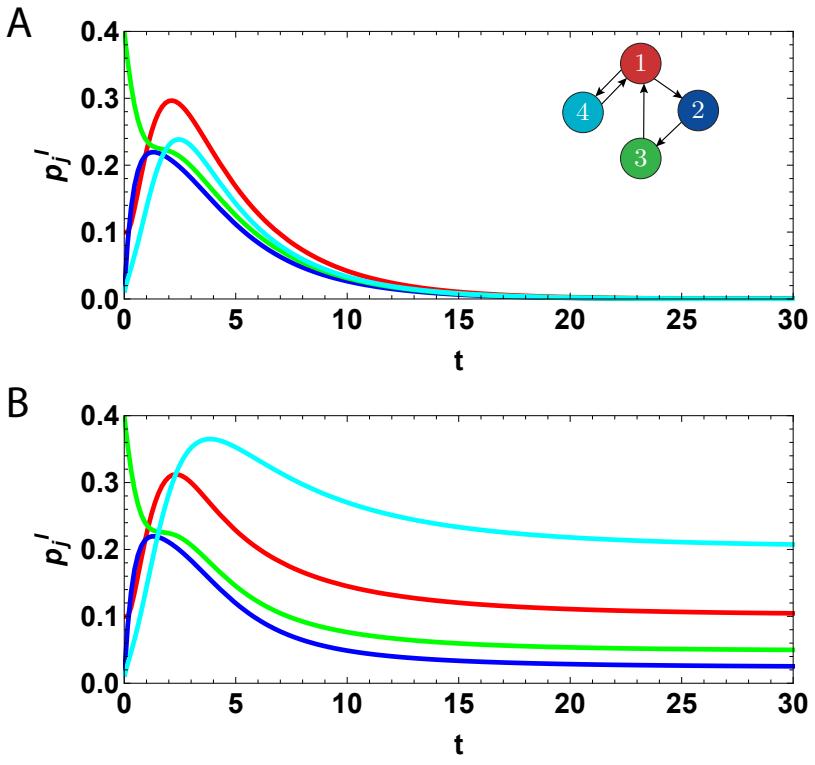

Figure 5.2: Effects of heterogeneity in the reinfection rates on the steady-state behavior. A. Partial Immunity (Case 4). All agents acquire partial immunity to all their infected neighbors: $B = 1.5A$ and $\hat{B} = 0.5A$. $R_{max} = R_0 = 1.99$, and $R_{min} = R_1 = 0.66$. B. Weak mixed immunity (Case 6a). Agents 1,2, and 3 acquire partial immunity to all their infected neighbors and agent 4 acquires compromised immunity to all its infected neighbors: $B = 1.5A$ and $\hat{B} = \text{diag}([0.5, 0.5, 0.5, 2.5])A$. $R_{max} = R_0 = 1.99$, and $R_{min} = R_1 = 0.66$. In both panels we set $\Gamma = \mathbb{I}$, and use the initial condition $\mathbf{p}^S(0) = \mathbf{1} - \mathbf{p}^I(0)$, $\mathbf{p}^I(0) = [0.1, 0.01, 0.4, 0.01]^T$.

bors and all other agents acquiring partial immunity to all their infected neighbors. The new basic reproduction numbers are $R_{max} = 2.29$, $R_{min} = 0.94$, $R_0 = 1.99$ and $R_1 = 1.17$. This places the network in the bistable regime. With the given initial conditions, the probability of infection of all agents is non-zero at steady-state. The large reinfection rate $\hat{\beta}_{41}$ results in a high probability of infection at steady-state for agent 4.

Effects of graph structure

Figure 5.3 shows the effect of network structure on the steady-state behavior. We consider the partial immunity case (Case 4) where agent 1 acquires partial immunity to all its infected neighbors, and agents 2, 3, and 4 are stubborn. In Figure 5.3A, agents 1 and 4 are susceptible to each other and agent 4 is not susceptible to any other agent. While, in Figure 5.3B, agents 2 and 4 are susceptible to each other and agent 4 is not susceptible to any other agent. In both cases, we set $\Gamma = \mathbb{I}$, $B = 0.9A$ and, $\hat{B} = \text{diag}([0.5, 0.9, 0.9, 0.9])A$. We set the initial condition to $\mathbf{p}^S(0) = \mathbf{1} - \mathbf{p}^I(0)$, $\mathbf{p}^I(0) = [0.1, 0.1, 0.1, 0.1]^T$.

For Figure 5.3A, $R_{max} = R_0 = 1.19$ and $R_{min} = R_1 = 0.94$, placing the system in the epidemic regime. The probability of infection of every agent initially increases, reaches a maximum and decays to zero. For Figure 5.3B, $R_{max} = R_0 = 1.19$ and $R_{min} = R_1 = 1.09$, placing the system in the endemic regime: the probability of infection of all agents is non-zero at steady-state. Differences in p_j^{I*} values are the result of both the location of agents within the graph and heterogeneities in the reinfection rates (see (5.8)).

Bistability and Resurgent Epidemic

Figure 5.4 illustrates the bistability and resurgent epidemic phenomena in the compromised immunity case (Case 5). For this case, we set $\Gamma = \mathbb{I}$, $B = 0.7A$, and $\hat{B} = \text{diag}([1.5, 0.7, 0.7, 0.7])A$. That is, agents $2, 3$, and 4 are stubborn and agent 1 acquires compromised immunity to all its infected neighbors. For this system, $R_{max} = R_1 = 1.28$ and $R_{min} = R_0 = 0.85$, placing the system in the bistable regime. To illustrate the bistability phenomenon, we set the initial conditions to $p^I = [0, 0.05, 0.1, 0]$ in Figure 5.4A and to $p^I = [0, 0.08, 0.1, 0]$ in Figure 5.4B. In Figure 5.4A, the solution converges to zero within the first 50 time units; the infection readily dies out. In contrast, in Figure 5.4B there is a resurgent epidemic in which

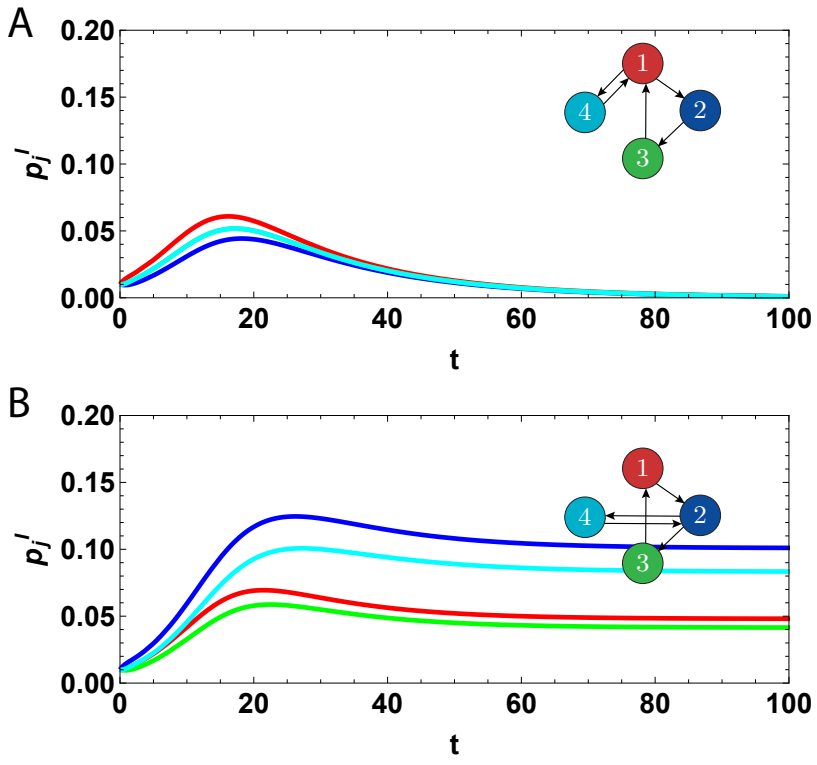

Figure 5.3: Effects of graph structure on steady-state behavior. Partial immunity case (Case 4) with four agents. Agent 1 becomes partially immune to all its infected neighbors and agents 2,3, and 4 are stubborn. A. Agents 1 and 4 are susceptible to each other. B. Agents 2 and 4 are susceptible to each other. In both panels we set $B = 0.9A$ and $\hat{B} = \text{diag}([0.5/0.9, 1, 1, 1])B$, and use the initial condition $\mathbf{p}^S(0) = \mathbf{1} - \mathbf{p}^I(0)$, $\mathbf{p}^I(0) = [0.1, 0.1, 0.1, 0.1]^T$.

the probabilities of infection reach a minimum and stay close to zero for close to 100 time units before increasing and converging to the EE.

Effect of agents that acquire full immunity

Figure 5.5 shows the effects of agents that acquire full immunity on the steady-state behavior. Recall that $\mathcal{G}_B$ and $\mathcal{G}_{\hat{B}}$ are the networks described by interpreting B and $\hat{B}$ as an adjacency matrix, respectively. Therefore, $\mathcal{G}_B$ describes the *effective infection network*, which represents the possible paths in $\mathcal{G}$ through which the contagion process can spread when all agents are susceptible. Similarly, $\mathcal{G}_{\hat{B}}$ describes the *effec-*

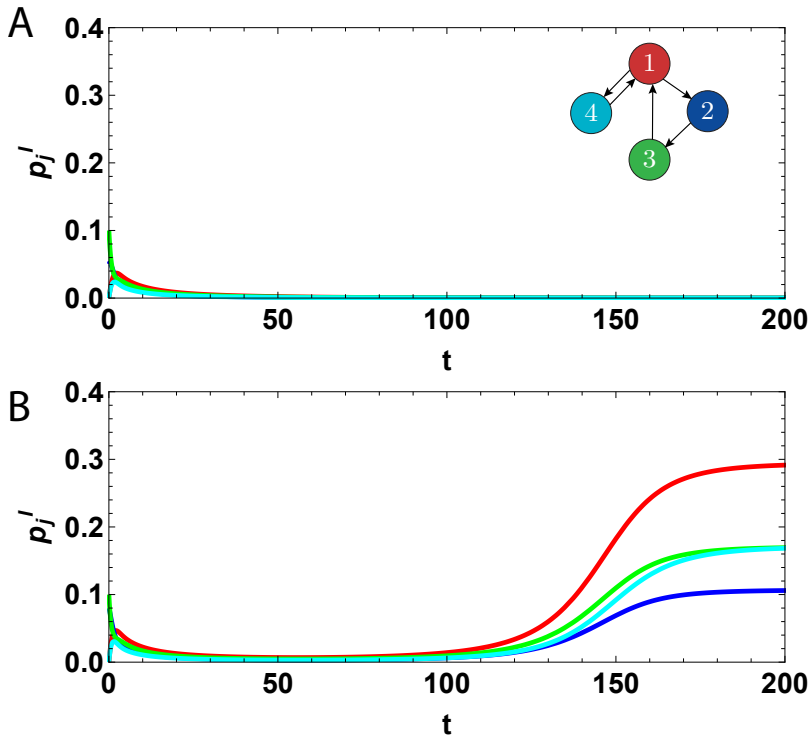

Figure 5.4: Bistability and resurgent epidemic. Compromised immunity digraph with four agents. All agents become compromised. A. initial condition $\mathbf{p}^S(0) = \mathbf{1} - \mathbf{p}^I(0)$, $\mathbf{p}^I(0) = [0, 0.05, 0.1, 0]^T$. B. initial condition $\mathbf{p}^S(0) = \mathbf{1} - \mathbf{p}^I(0)$, $\mathbf{p}^I(0) = [0, 0.08, 0.1, 0]^T$. In both panels we set $B = 0.7A$, $\hat{B} = 1.5A$, and $\Gamma = \mathbb{I}$. Adapted from Figure 10.2.

tive reinfection network, which represents the possible paths in $\mathcal{G}$ through which the contagion process can spread when all agents have been infected at least once.

In Figure 5.5, we show the SIS case (Case 3) for the complete four agents graph. We set $B = \hat{B} = A$, where A is the adjacency matrix, and $\Gamma = \text{diag}([0.25, 0.5, 0.75, 1])$. Therefore, $\mathcal{G}_B = \mathcal{G}_{\hat{B}}$, and the network can be in the infection-free regime or in the endemic regime depending on the value of $R_{max} = R_{min} = R_0 = R_1 = \rho(A\Gamma^{-1})$ (see Table 5.2). For this case, $\rho(A\Gamma^{-1}) = 5.79$, placing the network in the endemic regime, leading to solutions that converge to the EE.

In Figure 5.5B we modify the network in Figure 5.5A, by letting the agents acquire full immunity in a selective way, such that $\mathcal{G}_{\hat{B}}$ describes a directed cycle graph. The matrix $\hat{B}$ is Metzler irreducible and all our results hold. For this case, $R_{max} = R_0 =$

$\rho(A\Gamma^{-1}) = 5.79$ and $R_{min} = R_1 = \rho(\hat{B}\Gamma^{-1}) = 1.80$. This places the network in the endemic state, but the resulting value of $\mathbf{p}^{I*}$ is significantly smaller at each entry. Thus, through a selective removal of links in the effective reinfection network $\mathcal{G}_{\hat{B}}$, we can achieve a significant reduction in the probability of infection at steady-state.

Moreover, if we maintain the cycle structure in $\mathcal{G}_{\hat{B}}$ but decrease the susceptibility to reinfection between all pairs of connected agents, we can eradicate the infection. In Figure 5.5C we modify the network in Figure 5.5B, by setting $\hat{B} = 0.2A$. For this case, $R_{max} = R_0 = \rho(A\Gamma^{-1}) = 5.79$ and $R_{min} = R_1 = \rho(\hat{B}\Gamma^{-1}) = 0.36$. This places the network in the epidemic regime: solutions converge to an infection-free steady-state.

5.5 Dynamics near $\mathcal{M}_0$

In this section we extend the results of Chapter 10 by investigating the dynamics near the center IFE subset $\mathcal{M}_0$ in the epidemic and bistable regimes.

5.5.1 Center Eigenspace of points in $\mathcal{M}_0$

At an arbitrary point $\mathbf{m_0} = (\mathbf{m}, \mathbf{0})$ in $\mathcal{M}_0$, the leading eigenvalue $\lambda_{Tmax}(J_{\mathcal{M}}(\mathbf{m_0}))$ is zero, and by Proposition 1, all other transverse eigenvalues have negative real part. As a consequence, the Jacobian $J_{\mathcal{M}}(\mathbf{m_0})$ has $N+1$ zero eigenvalues and $N-1$ eigenvalues with negative real part. The first N zero eigenvalues correspond to the tangent space of the IFE set $\mathcal{M}$ while the last remaining zero eigenvalue corresponds to $\lambda_{Tmax}(J_{\mathcal{M}}(\mathbf{m_0}))$. Therefore, points in $\mathcal{M}_0$ correspond to points at which the IFE set $\mathcal{M}$ loses normal hyperbolicity.

The loss of normal hyperbolicity at these points is characterized by the deficiency of $J_{\mathcal{M}}(\mathbf{m_0})$: there are only N linearly independent eigenvectors corresponding to the $N+1$ zero eigenvalues. Therefore, the right eigenvectors of $J_{\mathcal{M}}(\mathbf{m_0})$ do not span its

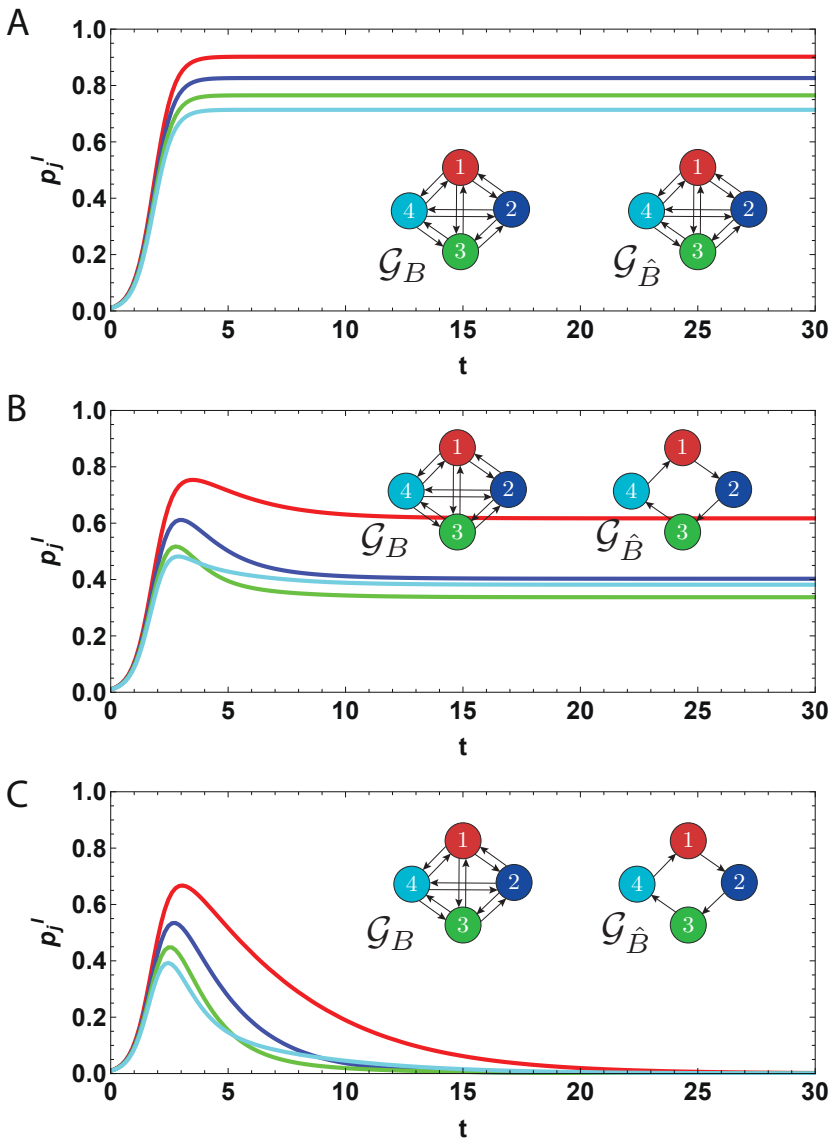

Figure 5.5: Effective Infection and Effective Reinfection Networks. A. SIS case (Case 3) for the complete four agents graph with $B = \hat{B} = A$, where A is the adjacency matrix. B. Partial immunity case (Case 4) where B is adjacency matrix for the complete graph ($B = A$) and $\hat{B}$ is the adjacency matrix for the directed cycle graph. C. Same as in B, but the entries of $\hat{B}$ have magnitude 0.2 or 0. In all cases, $\Gamma = [0.25, 0.5, 0.75, 1]^T$ and $\mathbf{p}^S(0) = \mathbf{1} - \mathbf{p}^I(0)$, $\mathbf{p}^I(0) = [0.1, 0.1, 0.1, 0.1]^T$.

center eigenspace. To see this, note that any eigenvector of $J_{\mathcal{M}}(\mathbf{m_0})$ corresponding to a zero eigenvalue must be of the form $[\hat{\mathbf{v}}_S^T, \mathbf{0}^T]^T$ where $\mathbf{v}_S \in \mathbb{R}^N \neq \mathbf{0}$.

To complete the center eigenspace basis, we need one right generalized eigenvector. Let $\hat{\mathbf{v}}_S = -\text{diag}(\mathbf{m})B\mathbf{v}$ where $\mathbf{v} \gg \mathbf{0}$ is the right eigenvector of $\lambda_{max}(J_T(\mathbf{m}))$. If $\mathbf{m} \neq$

$\mathbf{0}$, then, $[\hat{\mathbf{v}}_S^T, \mathbf{0}^T]^T$ is an eigenvector for zero and $J_{\mathcal{M}}(\mathbf{m_0})[\mathbf{0}^T, \mathbf{v}^T]^T = [\hat{\mathbf{v}}_S^T, \mathbf{0}^T]^T$. Therefore, $\kappa = [\mathbf{0}^T, \mathbf{v}^T]^T$ is a generalized eigenvector for zero, and the center eigenspace is spanned by the set of vectors

$$\left\{ \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{e}_2 \\ \mathbf{0} \end{bmatrix}, \ldots, \begin{bmatrix} \mathbf{e}_{N-1} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \hat{\mathbf{v}}_S \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{0} \\ \mathbf{v} \end{bmatrix} \right\}. \tag{5.10}$$

We note that the generalized eigenvector κ is the only vector in (5.10) that is not tangent to $\mathcal{M}$. By the Center Manifold Theorem [30], there exists an $(N+1)$-dimensional center manifold $\mathcal{W}^c(\mathbf{m_0})$ tangent to the center eigenspace of $J_{\mathcal{M}}(\mathbf{m_0})$ at $\mathbf{m_0}$. Moreover, since the remaining $N-1$ eigenvalues have negative real part, $\mathcal{W}^c(\mathbf{m_0})$ is attractive.

5.5.2 Towards the Center Manifold Dynamics

In this section we discuss the procedure by which the dynamics on the center manifold $\mathcal{W}^c(\mathbf{m_0})$ could be derived. We begin by shifting the origin of (5.6) to a point $\mathbf{m_0} = (\mathbf{m}, \mathbf{0})$ in $\mathcal{M}_0$,

$$\begin{bmatrix} \dot{\bar{\mathbf{p}}}^S \\ \dot{\mathbf{p}}^I \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{0}} & -MB \\ \bar{\mathbf{0}} & B^*(\mathbf{m}) - \Gamma \end{bmatrix} \begin{bmatrix} \bar{\mathbf{p}}^S \\ \mathbf{p}^I \end{bmatrix} + \begin{bmatrix} -\bar{P}^S B \mathbf{p}^I \\ \bar{P}^S (B - \hat{B}) \mathbf{p}^I - P^I \hat{B} \mathbf{p}^I \end{bmatrix}, \tag{5.11}$$

where $\bar{P}^S = \mathrm{diag}(\bar{\mathbf{p}}^S)$ and $M = \mathrm{diag}(\mathbf{m})$. Let $\bar{\mathbf{y}} = [(\bar{\mathbf{p}}^S)^T, (\mathbf{p}^I)^T]^T$ and $\dot{\bar{\mathbf{y}}} = \mathbf{f}(\bar{\mathbf{y}})$. We can rewrite (5.11) as $\dot{\bar{\mathbf{y}}} = J_{\mathcal{M}}(\mathbf{m_0})\bar{\mathbf{y}} + \bar{F}(\bar{\mathbf{y}})$. The nonlinear term $\bar{F}$, corresponding to the last term in (5.11), can also be expressed in terms of the Hessian $D^2\mathbf{f}$:

$$\bar{F}(\bar{\mathbf{y}}) = \frac{1}{2} D^2 \mathbf{f}(\bar{\mathbf{y}}, \bar{\mathbf{y}}) = \frac{1}{2} \begin{bmatrix} \bar{\mathbf{y}}^T \bar{H}_1 \bar{\mathbf{y}} \\ \vdots \\ \bar{\mathbf{y}}^T \bar{H}_{2N} \bar{\mathbf{y}} \end{bmatrix}. \tag{5.12}$$

where $\bar{H}_j^T = \bar{H}_j = \{h_{kl}\}$ and $h_{kl} = \partial \bar{f}_j / \partial \bar{y}_k \partial \bar{y}_l$.

Let $K \in \mathbb{R}^{2N \times 2N}$ drive $J_{\mathcal{M}}(m_0)$ to Jordan form,

$$J = K^{-1} J_{\mathcal{M}}(\mathbf{m_0}) K = \begin{bmatrix} J_0 & \bar{\mathbf{0}} \\ \bar{\mathbf{0}} & J_- \end{bmatrix} \tag{5.13}$$

where J_0 consists of $N-1$, 1×1 blocks and one 2×2 block (corresponding to the $N+1$ zero eigenvalues), and J_- consists of the corresponding Jordan blocks for the N eigenvalues with negative real part. Without loss of generality, we can pick the columns $\mathbf{k}^j$ of K such that they correspond to the (generalized) right eigenvectors of $J_{\mathcal{M}}(\mathbf{m_0})$ and the rows $\bar{\mathbf{k}}^j$ of K^{-1}, such that they correspond to the left (generalized) eigenvectors of $J_{\mathcal{M}}(\mathbf{m_0})$. Using the similarity transformation $\bar{\mathbf{y}} = K\hat{\mathbf{y}}$, we get $\dot{\hat{\mathbf{y}}} = J\hat{\mathbf{y}} + F(\hat{\mathbf{y}})$, where

$$F(\hat{\mathbf{y}}) = \frac{1}{2} K^{-1} D^2 \mathbf{f}(\hat{\mathbf{y}}, \hat{\mathbf{y}}) = \frac{1}{2} \begin{bmatrix} \bar{\mathbf{y}}^T H_1 \bar{\mathbf{y}} \\ \vdots \\ \bar{\mathbf{y}}^T H_{2N} \bar{\mathbf{y}} \end{bmatrix} \tag{5.14}$$

where $H_j^T = H_j = K^T (\sum_{i=1}^{2N} \bar{\mathbf{k}}_i^j \bar{H}_i) K$.

The dynamics $\dot{\hat{\mathbf{y}}} = J\hat{\mathbf{y}} + F(\hat{\mathbf{y}})$ can be written in the block diagonal form

$$\begin{aligned} \dot{\mathbf{x}}_\mathbf{0} &= J_0 \mathbf{x_0} + \mathbf{g}(\mathbf{x_0}, \mathbf{x}_-); & \mathbf{x_0} &\in \mathbb{R}^{N+1} \\ \dot{\mathbf{x}}_- &= J_- \mathbf{x}_- + \mathbf{h}(\mathbf{x_0}, \mathbf{x}_-); & \mathbf{x}_- &\in \mathbb{R}^{N-1}. \end{aligned} \tag{5.15}$$

The system (5.15) can be used to derive an approximation for the center manifold dynamics up to any desired order by seeking series solutions to the center manifold equation ((3.2.16) in [30]). However, the high-dimensionality of the system makes this approach analytically challenging. As an alternative, we study the dynamics

near points in $\mathcal{M}_0$ by looking for transversality conditions as described in the next section.

5.5.3 Transversality of Solutions and $\mathcal{M}_0$

In this section we gain additional insight into the transient behaviors of solutions in the epidemic and bistable regimes by studying the projection of solutions of the network SIRI model onto the IFE manifold $\mathcal{M}$. We find conditions for these projected solutions to cross the center IFE subset $\mathcal{M}_0$ transversally. To do so, we make use of the following definitions and lemmas.[3]

Definition 5.5.1 (Neighborhood $E \subset \mathcal{M}'$ of $\mathcal{M}$)**.** *Neighborhood $E \subset \mathcal{M}'$ of $\mathcal{M}$ is the union of $\mathcal{M}$ and the neighborhoods $\bar{U} \subset \mathcal{M}'$ of every $\bar{\mathbf{x}} \in \partial\mathcal{M}$ for which $J_{\mathcal{M}}(\bar{\mathbf{u}})$ is Metzler irreducible for any $\bar{\mathbf{u}} \in \bar{U}$, (see the proof of Lemma 13 for details).*

Definition 5.5.2 (J_T and Λ)**.** *Let $J_T(p) = B^*(p) - \Gamma$ for $(p, \mathbf{0}) \in E$. The λ_{Tmax} function is $\Lambda : E \to \mathbb{R}$, $(p, \mathbf{0}) \mapsto \lambda_{max}(J_T(p))$. For ease of notation we use $\Lambda(p)$ for $\Lambda(p, \mathbf{0})$.*

Lemma 6 (IFE subsets as level surfaces of Λ [58])**.** *Let $\Lambda_c = \{\Lambda^{-1}(c) \,|\, c \in \mathbb{R}\}$ be the level surface of Λ on $E \subset \mathcal{M}'$ corresponding to $c \in \mathbb{R}$. Then, $\mathcal{M}_0 = \Lambda_0 \cap \mathcal{M}$, $\mathcal{M}_- = \bigcup_{c<0} \Lambda_c \cap \mathcal{M}$ and $\mathcal{M}_+ = \bigcup_{c>0} \Lambda_c \cap \mathcal{M}$.*

Lemma 7 (Gradient of Λ [58])**.** *For $\acute{x} = (p, \mathbf{0}) \in E$, let $w^T, v \in \mathbb{R}^N$ be left and right eigenvectors of $J_T(p)$ for $\Lambda(p)$. Then, Λ is smooth on E, i.e., $\Lambda(\cdot) \in C^\infty(E)$, with partial derivatives*

$$\frac{\partial \Lambda}{\partial p_j}(p) = w_j \sum_{k=1}^{N} (\beta_{jk} - \hat{\beta}_{jk}) v_k \tag{5.16}$$

and gradient

$$\nabla \Lambda(p) = \mathrm{diag}(w)(B - \hat{B})v. \tag{5.17}$$

[3]Many of the definitions and lemmas in this section are used in the proof of Theorem 4 (see Section 5.3 for details).

In addition, the following hold.

- *If $B = \hat{B}$, all points $\acute{\mathbf{x}} \in E$ are critical points of Λ.*
- *If j is a stubborn agent, $(\partial\Lambda/\partial p_j)(\mathbf{p}) = 0$ for all $\acute{\mathbf{x}} \in E$.*
- *If $B \succ \hat{B}$, and there are no stubborn agents in $\mathcal{G}$, $\nabla\Lambda(\mathbf{p}) \gg \mathbf{0}$ for all $\acute{\mathbf{x}} \in E$, and Λ has no critical points.*
- *If $\hat{B} \succ B$, and there are no stubborn agents in $\mathcal{G}$, $\nabla\Lambda(\mathbf{p}) \ll \mathbf{0}$ for all $\acute{\mathbf{x}} \in E$ and Λ has no critical points.*
- *If $B \neq \hat{B}$ then either Λ has no critical points or all points $\acute{\mathbf{x}} \in E$ are critical points of Λ.*

Lemma 8 (Maximum and minimum values of Λ in $\mathcal{M}$ [58])**.** *Let $(\mathbf{p}^S, \mathbf{0}) \in \mathcal{M}$. Then $\Lambda(\mathbf{p}^S)$ achieves its global maximum and minimum c_{max}, c_{min} at one or more points in $\partial\mathcal{M}$. In Cases 2, 4, 5, and 6a, there exist unique corner points $(\mathbf{p}^{max}, \mathbf{0}) \in \hat{\mathcal{M}}$, $(\mathbf{p}^{min}, \mathbf{0}) \in \hat{\mathcal{M}}$ such that $\Lambda(\mathbf{p}^{max}) = c_{max}$, $\Lambda(\mathbf{p}^{min}) = c_{min}$, $R_{max} = \rho(B^*(\mathbf{p}^{max})\Gamma^{-1})$, and $R_{min} = \rho(B^*(\mathbf{p}^{min})\Gamma^{-1})$. Moreover, $\mathbf{p}^{max}$ and $\mathbf{p}^{min}$ are the respective unique global maximum and minimum points of Λ in $\mathcal{M}$ if, and only if, there are no stubborn agents in $\mathcal{G}$.*

Definition 5.5.3 (Transversal crossing of Λ_c)**.** *Let $\mathbf{y}(t) = (\mathbf{p}^S(t), \mathbf{p}^I(t)) \in \Delta_N$, $t \geq 0$, be a solution of* (10.3)*. We say that $\mathbf{y}$ crosses Λ_c transversally if $\mathbf{p}^S$, the projection of $\mathbf{y}$ onto $\mathcal{M}$, crosses Λ_c transversally. This holds if there exists a time $t' > 0$ and $\mathbf{m} \in \Lambda_c$ such that $\mathbf{p}^S(t') = \mathbf{m}$ and $\dot{\mathbf{p}}^S(t')^T\nabla\Lambda(\mathbf{m}) \neq 0$.*

Proposition 5 (Transversal crossing direction [58])**.** *Let $\mathbf{y}(t) = (\mathbf{p}^S(t), \mathbf{p}^I(t)) \in \Delta_N$, $t \geq 0$, be a solution of* (10.3) *that crosses Λ_c transversally at the point $(\mathbf{m}, \mathbf{0}) \in \mathcal{M}$ and time $t = t'$, where $c = \Lambda(\mathbf{m})$. If $\dot{\mathbf{p}}^S(t')^T\nabla\Lambda(\mathbf{m}) < 0$, then Λ decreases as $\mathbf{p}^S$ crosses Λ_c, and if $\dot{\mathbf{p}}^S(t')^T\nabla\Lambda(\mathbf{m}) > 0$, then Λ increases as $\mathbf{p}^S$ crosses Λ_c. Suppose*

$(\mathbf{m}, \mathbf{0}) \in \mathcal{M}_0$, *then if* $\dot{\mathbf{p}}^S(t')^T \nabla \Lambda(\mathbf{m}) < 0$, $\mathbf{p}^S$ *crosses* $\mathcal{M}_0$ *from* $\mathcal{M}_+$ *to* $\mathcal{M}_-$ *and if* $\dot{\mathbf{p}}^S(t')^T \nabla \Lambda(\mathbf{m}) > 0$, *then* $\mathbf{p}^S$ *crosses* $\mathcal{M}_0$ *from* $\mathcal{M}_-$ *to* $\mathcal{M}_+$.

Definition 5.5.4. *The stable manifold of* $\mathcal{M}_0$ *is* $\mathcal{W}^s(\mathcal{M}_0)$ *and the unstable manifold of* $\mathcal{M}_0$ *is* $\mathcal{W}^u(\mathcal{M}_0)$.

Theorem 7 (Transversality of solutions [58])**.** *Consider Cases 2 and 4 in the epidemic regime* ($R_0 > 1$, $R_1 < 1$) *and Case 5 in the bistable regime* ($R_0 < 1$, $R_1 > 1$). *Assume no stubborn agents. Let* $\mathbf{y}(t) = (\mathbf{p}^S(t), \mathbf{p}^I(t)) \in \Delta_N$, $t \geq 0$, *be a solution of* (10.3) *for which there exists a time* $t' > 0$ *and* $(\mathbf{m}, \mathbf{0}) \in$ *int*$(\mathcal{M})$, *such that* $\mathbf{p}^S(t') = \mathbf{m}$ *and* $\mathbf{p}^I(t') \succ \mathbf{0}$. *Let* $c = \Lambda(\mathbf{m})$ *so that* $\mathbf{m} \in \Lambda_c$. *Then* $\mathbf{y}$ *crosses* Λ_c *transversally. Suppose* $(\mathbf{m}, \mathbf{0}) \in$ *int*$(\mathcal{M}_0)$. *In the epidemic regime of Cases 2 and 4,* $\mathbf{p}^S$ *crosses* $\mathcal{M}_0$ *from* $\mathcal{M}_+$ *to* $\mathcal{M}_-$, *and the stable and unstable manifolds of* $\mathcal{M}_0$ *lie outside* Δ_N. *In the bistable regime of Case 5,* $\mathbf{p}^S$ *crosses* $\mathcal{M}_0$ *from* $\mathcal{M}_-$ *to* $\mathcal{M}_+$, *and the stable and unstable manifolds of* $\mathcal{M}_0$ *lie inside* Δ_N.

Corollary 1 ([58])**.** *In the epidemic regime of Cases 2 and 4, every heteroclinic orbit in* Δ_N *connects a point in* $\mathcal{M}_+$ *to a point in* $\mathcal{M}_-$.

Corollary 2 ([58])**.** *Consider Case 5 in the bistable regime. Let* $\mathbf{y}(t) = (\mathbf{p}^S(t), \mathbf{p}^I(t))$ *be a solution of* (10.3). *Then it holds that*

- *If* $\mathbf{y}$ *crosses* $\mathcal{M}_0$ *transversally or* $(\mathbf{p}^S(0), \mathbf{0}) \in \mathcal{M}_+$, *then* $\mathbf{y}$ *converges to the EE as* $t \to \infty$. *Moreover, the EE lies on the unstable manifold of* $\mathcal{M}_0$.
- *The stable manifold of* $\mathcal{M}_0$ *intersects the boundary of* Δ_N *where* $\mathbf{p}^S = \mathbf{1} - \mathbf{p}^I$.

In Figure 5.6, we illustrate the results of Theorem 17 for the compromised immunity case in the bistable regime, and show that, in the mixed immunity case, solutions do not necessarily cross $\mathcal{M}_0$ transversally. The figure shows the IFE set $\mathcal{M}$ for the complete two agent graph (i.e., $N = 2$) in two different cases. In Figure 5.6A, we set $\beta_{12} = \beta_{21} = 0.8$, $\hat{\beta}_{12} = \hat{\beta}_{21} = 1.3$, and $\delta_1 = \delta_2 = 1$. This corresponds to the

compromised immunity case in the bistable regime; the corresponding reproduction numbers are $R_{min} = R_0 = 0.8$, $R_{max} = R_1 = 1.3$. That is, both agents become more susceptible to reinfection. In Figure 5.6B, we set $\beta_{12} = \hat{\beta}_{21} = 0.8$, $\hat{\beta}_{12} = \beta_{21} = 1.3$, and $\delta_1 = \delta_2 = 1$. This corresponds to the weak mixed immunity case in the bistable regime; the corresponding reproduction numbers are $R_0 = R_1 = 1.02$, $R_{min} = 0.8$, $R_{max} = 1.3$. That is, agent 1 becomes more susceptible to reinfection and agent 2 becomes less susceptible to reinfection. The advantage of studying the case $N = 2$ is that the IFE manifold $\mathcal{M}$ is 2-dimensional, which can be visualized as the (p_1^{S*}, p_2^{S*}) plane. The center IFE subset $\mathcal{M}_0$ in this case corresponds to the dashed white curve that partitions the plane into two regions, with the blue region corresponding to $\mathcal{M}_-$ and the red striped region corresponding to $\mathcal{M}_+$. The remaining solid white curves correspond to level-surfaces of Λ, with the white curves in the blue region corresponding to Λ_c for $c < 0$ and the white curves in the red region corresponding to Λ_c for $c > 0$. The black arrows represent the vector field for the $\mathbf{p}^S$ dynamics $\dot{\mathbf{p}}^S = -P^S B \mathbf{p}^I$ when $\mathbf{p}^I = c\mathbf{1}$ for $c = 0.1$. Note that, in both cases, the vector field is identical, but the orientation of the level-surfaces of Λ is not. As a result of this, the vector field in Figure 5.6A is transverse to $\mathcal{M}_0$, as proved in Theorem 17. While, the vector field in Figure 5.6B appears to be parallel to $\mathcal{M}_0$ at multiple points.

Figure 5.7 illustrates the results of Theorems 6 and 17. In the figure, we illustrate the local dynamics near $\mathcal{M}_0$ for Cases 2-5 in the R_0, R_1 parameter space. For simplicity of illustration, we represent the N-dimensional manifold $\mathcal{M}$ as a plane in a 3-dimensional space. The solid blue region of the plane represents $\mathcal{M}_-$ and the striped red region represents $\mathcal{M}_+$. The line dividing the blue and red regions represents $\mathcal{M}_0$. Black arrows are used to illustrate the behavior of solutions near $\mathcal{M}$. When $R_0 \leq 1$ and $R_1 \leq 1$ (infection-free regime), all solutions decay to a point in $\mathcal{M}$. When $R_0 > 1$ and $R_1 > 1$ (endemic regime), all solutions grow away from $\mathcal{M}$ and eventually reach the EE. When $R_0 > 1$ and $R_1 \leq 1$ (epidemic regime), there

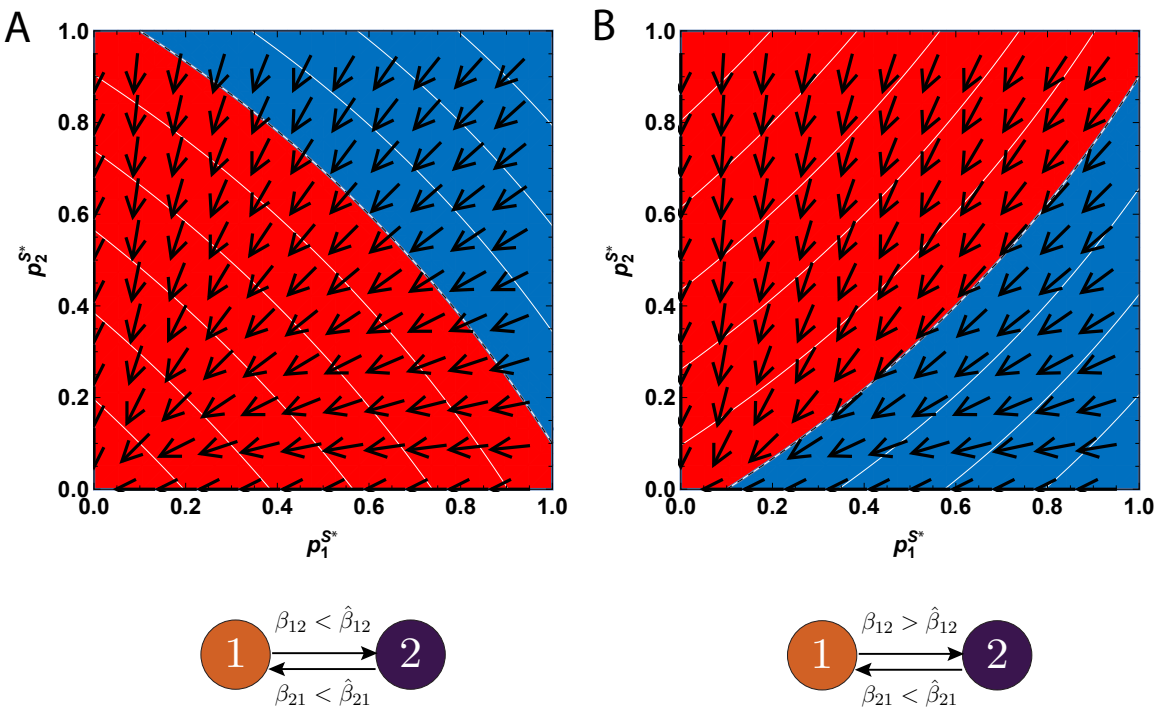

Figure 5.6: IFE set $\mathcal{M}$ for two interconnected agents with $\delta_1 = \delta_2 = 1$. A. Compromised immunity case with $\beta_{12} = \beta_{21} = 0.8$, $\hat{\beta}_{12} = \hat{\beta}_{21} = 1.3$. B. Weak mixed immunity case with $\beta_{12} = \hat{\beta}_{21} = 0.8$, $\hat{\beta}_{12} = \beta_{21} = 1.3$. The dashed white curve represents $\mathcal{M}_0$ and the solid white curves represent level-surfaces Λ_c for $c \neq 0$. The black arrows represent the component of the vector field (5.6) tangent to $\mathcal{M}$, $\dot{\mathbf{p}}^S = -P^S B \mathbf{p}^I$, for $\mathbf{p}^I = [0.1, 0.1]^T$.

exists a family of solutions near $\mathcal{M}$ that form heteroclinic orbits connecting points in $\mathcal{M}_-$ and $\mathcal{M}_+$. Finally, when $R_0 \leq 1$ and $R_1 > 1$ (bistable regime), there exists a stable manifold $\mathcal{W}^s(\mathcal{M}_0)$ (green translucent surface) and unstable manifold $\mathcal{W}^u(\mathcal{M}_0)$ (magenta translucent surface) of $\mathcal{M}_0$.

Remark 12. The behavioral regimes defined in Theorem 6 provide analytical predictions on the behavior of solutions as described in Table 5.2. Figure 5.8 illustrates this point by coloring the different regions of the R_0, R_1 parameter space according to the type of immunity that a model in that region would exhibit. Note that the SIS case is represented by a dashed diagonal line where $R_0 = R_1$, while the SIR case is represented with a dashed line where $R_1 = 0$.

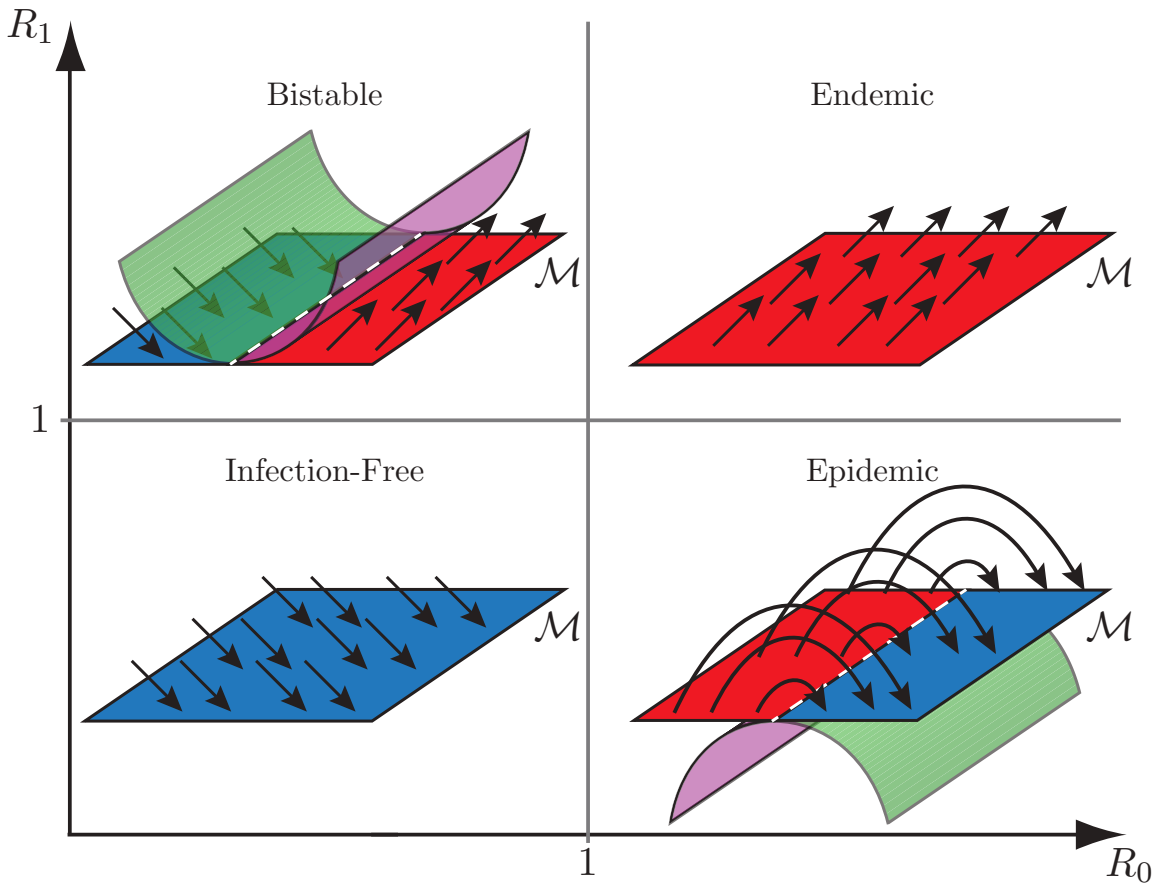

Figure 5.7: Illustration of local dynamics near $\mathcal{M}$ for the four different behavioral regimes of the network SIRI model (5.6) when $B \succeq \hat{B}$ or $\hat{B} \succeq B$. The diagrams are arranged where they exist in the R_0, R_1 parameter space according to Theorem 6. $\mathcal{M}_-$ is blue, $\mathcal{M}_+$ is red, and $\mathcal{M}_0$ is the black dashed line. The stable and unstable manifolds of $\mathcal{M}_0$ are green and magenta, respectively. Repeated from 10.1.

5.6 Bistability, Critical Manifold, and Vaccination

In this section we examine the bistability regime in the compromised immunity case $\hat{B} \succ B$ (Case 5) in two special cases. In the first case, we consider d-regular digraphs and obtain analytical results for initial conditions of the form $\mathbf{p}^S(0) = \mathbf{1} - \mathbf{p}^I(0)$ and $\mathbf{p}^I(0) = p_{ic}\mathbf{1}$ for $p_{ic} \in [0, 1]$. In the second case, we consider complete digraphs with two agents ($N = 2$). We use numerical simulations to study the bistable regime

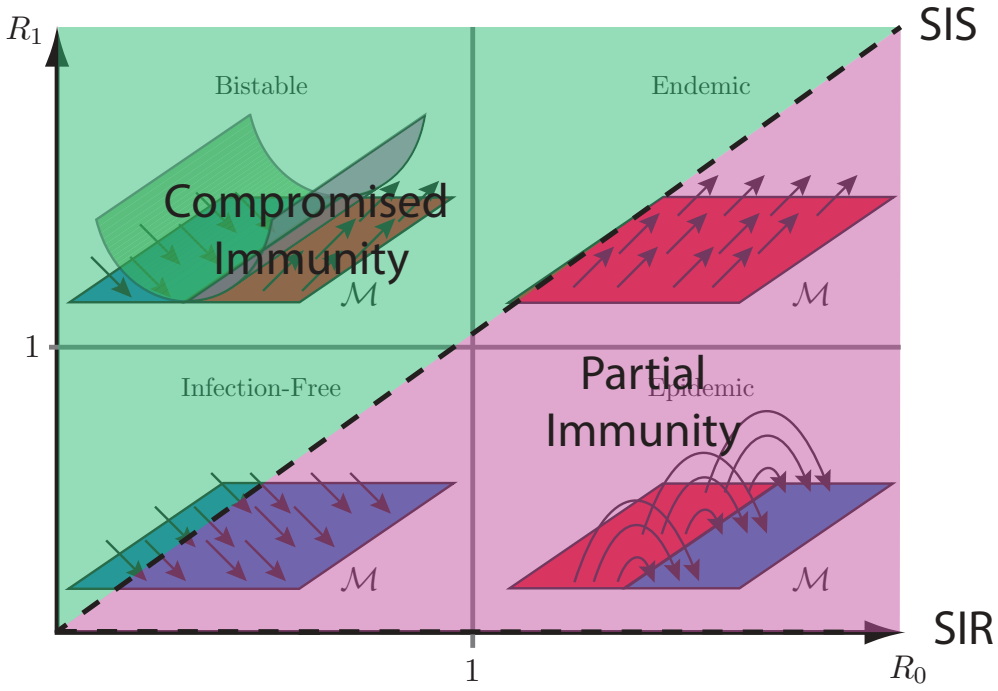

Figure 5.8: Illustration of the network SIRI behavioral regimes and how they apply to Cases 2, 3, 4, and 5. See Table 5.2.

with initial conditions of the form $\mathbf{p}^S(0) = \mathbf{1} - \mathbf{p}^I(0)$, and how it is affected by initial conditions in which agent j has been previously exposed to the infection (i.e., $p_j^S(0) = 0$).

5.6.1 d-Regular Digraphs

The degree of node j is $d_j = \sum_{k=1}^{N} a_{jk}$, where a_{jk} are the entries of the adjacency matrix A. A digraph $\mathcal{G}$ is *d-regular* if $d_j = d$ for all $j = 1, \ldots, N$.

We present the following analysis for d-regular digraphs (quoted from [58]):

> Consider a d-regular digraph with global recovery, infection, and reinfection rates: $\Gamma = \delta\mathbb{I}$, $B = \beta A$, and $\hat{B} = \hat{\beta} A$, where A is the adjacency

matrix. The network SIRI dynamics (5.5) are

$$\begin{aligned}
\dot{p}_j^S &= -\beta d p_j^S p_j^I - \beta p_j^S \sum_{k=1}^{N} a_{jk}(p_k^I - p_j^I) \\
\dot{p}_j^I &= -\delta p_j^I + (\beta - \hat{\beta}) d p_j^S p_j^I + \hat{\beta} d p_j^I - \hat{\beta} d (p_j^I)^2 \\
&\quad + \Big((\beta - \hat{\beta}) p_j^S + \hat{\beta}(1 - p_j^I) \Big) \sum_{k=1}^{N} a_{jk}(p_k^I - p_j^I),
\end{aligned} \tag{5.18}$$

where we have used the identity $\sum_{k=1}^{N} a_{jk} p_k^I = p_j^I d + \sum_{k=1}^{N} a_{jk}(p_k^I - p_j^I)$. Let $p^I(0) = p_{ic}\mathbf{1}$. Then (5.18) reduce to

$$\begin{aligned}
\dot{p}_j^S &= -\beta d p_j^S p_j^I \\
\dot{p}_j^I &= -\delta p_j^I + (\beta - \hat{\beta}) d p_j^S p_j^I + \hat{\beta} d p_j^I - \hat{\beta} d (p_j^I)^2.
\end{aligned} \tag{5.19}$$

(5.19) describes identical and uncoupled dynamics for every agent j. Further, the dynamics (5.19) are equivalent to the dynamics of the well-mixed SIRI model [57] with infection rate βd and reinfection rate $\hat{\beta} d$. Following [57], we find the critical initial condition $p_{crit} = 1 - \xi(R_0 d \xi)^{-\beta/\hat{\beta}}$, where $\xi = (R_1 - 1/d)/(R_1 - R_0)$. If $p_{ic} < p_{crit}$ solutions converge to a point in the IFE as $t \to \infty$. If $p_{ic} > p_{crit}$ solutions converge to the EE, $p^{I*} = (1 - \delta/(\hat{\beta} d))\mathbf{1}$, as $t \to \infty$.

If $p_{ic} = p_{crit}$, the solution flows along the stable manifold of the point $\xi\mathbf{1} \in \mathcal{M}_0$ and converges to $\xi\mathbf{1}$. For the well-mixed SIRI model [57], we proved that the stable manifold of the point at which the IFE set loses normal hyperbolicity separates solutions that converge to an infection-free steady-state from solutions that converge to the EE. The case of d-regular graphs suggests how these results generalize to the network SIRI model.

5.6.2 Complete Digraphs with $N = 2$

We can gain insight into the role of the stable and unstable manifolds of $\mathcal{M}_0$ in the bistable regime by studying the compromised immunity and weak mixed immunity cases for complete digraphs with $N = 2$. Consider the complete digraphs with two agents introduced in Figure 5.6. Figure 5.6A considers the case of two homogeneous agents whose susceptibility increases after a first exposure to the infection. In contrast, Figure 5.6B considers the case of two heterogeneous agents, where agent 1 increases its susceptibility after a first exposure to the infection and agent 2 decreases its susceptibility after a first exposure to the infection.

In Figure 5.9 we show numerical simulations results for the set of initial conditions that yield solutions that converge to a point in the IFE set $\mathcal{M}$ or solutions that converge to the EE. Figures 5.9A and 5.9B illustrate the space of initial conditions of the form $\mathbf{p}^S(0) = \mathbf{1} - \mathbf{p}^I(0)$ for the two agent networks shown in Figures 5.6A and 5.6B, respectively. In each of these two figures, we represent an initial condition $p_1^S(0) = 1 - p_1^I(0)$, $p_2^S(0) = 1 - p_2^I(0)$ as a point in the $p_1^I(0), p_2^I(0)$ plane. Green points correspond to initial conditions that lead to solutions that converge to a point in the IFE set $\mathcal{M}$, while magenta points correspond to initial conditions that lead to solutions that converge to the EE. In the compromised immunity case (Figure 5.9A), there is a small region of initial conditions, near the origin, that lead to an infection-free steady-state. The curve separating the green and magenta regions in this case corresponds to the intersection of $\mathcal{W}^s(\mathcal{M}_0)$ with the boundary of Δ_N where $\mathbf{p}^S = \mathbf{1} - \mathbf{p}^I$. In contrast, in the weak immunity case (Figure 5.9B), all initial conditions lead to solutions that converge to the EE.

At first sight, this result appears to contradict Theorem 6, which guarantees the existence of initial conditions in the bistable regime that lead to solutions that converge to a point in $\mathcal{M}$. However, note that in Figure 5.9B we only consider solutions

with initial conditions of the form $p_1^S(0) = 1 - p_1^I(0)$, $p_2^S(0) = 1 - p_2^I(0)$. Figures 5.9C and 5.9D illustrate the space of initial conditions of the form $p_1^S(0) = 1 - p_1^I(0)$, $p_2^S(0) = 0$. This corresponds to initial conditions for which agent 2 in the network has been previously exposed to the infection. This is analogous to vaccinating agent 2. In the compromised immunity case (Figure 5.9C), vaccinating agent 2 has the effect of making the network more susceptible to the infection; agent 2 is more susceptible to reinfection ($\beta_{21} < \hat{\beta}_{21}$), facilitating the spread of the infection at time $t = 0$. As a result, all initial conditions lead to solutions that converge to the EE. In contrast, for the weak mixed immunity case (Figure 5.9D), vaccinating agent 2 has the effect of making the network less susceptible to the infection; agent 2 is less susceptible to reinfection ($\beta_{21} > \hat{\beta}_{21}$), impeding the spread of the infection at time $t = 0$. As a result, there exists a region of initial conditions (yellow) that leads to solutions that converge to a point in the IFE set $\mathcal{M}$.

Remark 13. Vaccinating agent 1 instead of agent 2 changes the results for the weak mixed immunity case but not for the compromised immunity case. In the former case, vaccinating agent 1 increases the susceptibility of the network to the infection since agent 1 is more susceptible to reinfection ($\beta_{12} < \hat{\beta}_{12}$). In the latter case, the results do not change because of the symmetry in the system.

5.7 Bifurcation Without Parameters

The results in Sections 5.5 and 5.6 suggest that the dynamics of the network SIRI model (5.6) exhibit a codimension-N transcritical bifurcation without parameters with drift singularity (TBWPDS). In Section 2.3 we described the normal form and transversality conditions for the codimension-one TBWPDS, and in Section 4.4 we proved that the dynamics of the well-mixed SIRI model (9.7) exhibit this bifurcation.

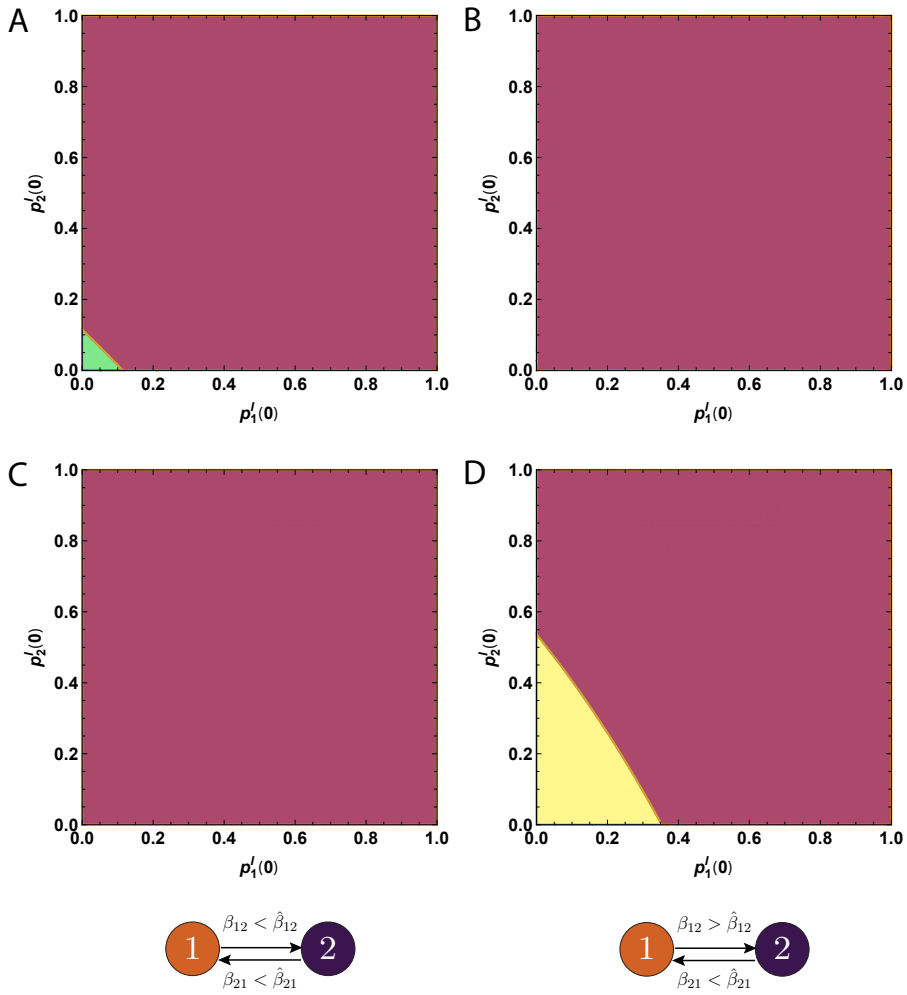

Figure 5.9: Bistability and space of initial conditions for two interconnected agents. Compromised immunity case with $\beta_{12} = \beta_{21} = 0.8$, $\hat{\beta}_{12} = \hat{\beta}_{21} = 1.3$. B. Weak mixed immunity case with $\beta_{12} = \hat{\beta}_{21} = 0.8$, $\hat{\beta}_{12} = \beta_{21} = 1.3$.
A. Compromised immunity case with $\mathbf{p}^S(0) = \mathbf{1} - \mathbf{p}^I(0)$. Green points correspond to initial conditions that converge to a point in the IFE set $\mathcal{M}$ as $t \to \infty$ and magenta points correspond to initial conditions that converge to the EE as $t \to \infty$. B. Weak mixed immunity case with $\mathbf{p}^S(0) = \mathbf{1} - \mathbf{p}^I(0)$. All initial conditions converge to the EE. C. Compromised immunity case where agent 2 has been previously exposed to the infection: $p_1^S(0) = 1 - p_1^I(0)$ and $p_2^S(0) = 0$. All initial conditions converge to the EE. D. Weak mixed immunity case where agent 2 has been previously exposed to the infection: $p_1^S(0) = 1 - p_1^I(0)$ and $p_2^S(0) = 0$. Yellow points correspond to initial conditions that converge to a point in the IFE set $\mathcal{M}$ as $t \to \infty$ and magenta points correspond to initial conditions that converge to the EE as $t \to \infty$. For the compromised immunity case, we set $\beta_{12} = \beta_{21} = 0.8$, $\hat{\beta}_{12} = \hat{\beta}_{21} = 1.3$. For the weak mixed immunity case, we set $\beta_{12} = \hat{\beta}_{21} = 0.8$, $\hat{\beta}_{12} = \beta_{21} = 1.3$. In both cases, $\delta_1 = \delta_2 = 1$.

In both the well-mixed SIRI model (9.7) and the network SIRI model (5.6), we proved that if $R_1 < 1$ all solutions converge to an infection-free equilibrium, and that if $R_1 > 1$ solutions converge a point on the IFE or the EE depending on the system parameters and the initial conditions. Moreover, in the well-mixed case we proved that the stable manifold of the point at which the IFE set loses normal hyperbolicity separates solutions that converge to an infection-free steady-state from solutions that converge to the EE. In Section 5.6, we showed that a comparable situation arises in the case of d-regular graphs. Our results in Section 5.5 on the transversality of solutions suggests that these results may generalize to the network SIRI model under certain parameter conditions.

Transversality conditions for the codimension-N transcritical bifurcation without parameters with drift singularity do not currently exist. However, based on our results, we can conjecture on what they would look like.

Extending conditions (TBWP1) and (TBWP2) to the codimension-N case is straightforward. Condition (TBWP1) corresponds to the existence of a manifold of equilibria and condition (TBWP2) corresponds to the loss of normal hyperbolicity as $\lambda_{Tmax}(J_{\mathcal{M}}(\mathbf{x}))$ crosses zero. Condition (TBWP3) corresponds to the transversal crossing of solutions at points in $\mathcal{M}_0$. This condition is analogous to the condition $\dot{\mathbf{p}}^S \nabla \Lambda(\mathbf{p}^S) \neq 0$ we used in Theorem 17. Conditions (TBWP4) and (TBWP5) are related to the non-degeneracy and degeneracy of the top right block of $J_{\mathcal{M}}(\mathbf{x})$ in (5.9). Further analysis of the network SIRI model could provide insight into how to generalize conditions (TBWP4)-(TBWP6).

5.8 Concluding Remarks

In this Chapter we studied the network SIRI model. We proved that, similar to the well-mixed case, the model exhibits four behavioral regimes: infection-free, en-

demic, epidemic, and bistable. However, unlike the well-mixed case, the network SIRI model contains an explicit description of the communication network structure which facilitates the study of heterogeneities between agents (in how they recover and/or respond to a first infection by their neighbors). To characterize the dynamics, we generalized the basic reproduction numbers R_0 and R_1 to network topologies and introduced the notion of extreme basic reproduction numbers: R_{max} and R_{min}. We studied the geometry of the IFE set $\mathcal{M}$ and found conditions on R_0, R_1, R_{max} and R_{min} for the existence of stable, center, and unstable IFE subsets $\mathcal{M}_-, \mathcal{M}_0$ and $\mathcal{M}_+$, respectively. Understanding how these subsets change with the system parameters facilitates the study of transient and steady-state dynamics in the model and was fundamental in our derivation of conditions for the four behavioral regimes. We illustrated the connection between these subsets and the resulting dynamics in the bistable case through analytical results for d-regular digraphs and though numerical results for $N = 2$. Furthermore, we examined possible control strategies through example simulations in which we modified the recovery rates, reinfection rates, and network structure. Finally, we studied vaccination as a control strategy and suggested that the dynamics of the network SIRI model exhibit a codimension N transcritical bifurcation without parameters with drift transversality.

Our results provide analytical results on the effects of network structure and agent heterogeneity on the transient and steady-state dynamics of the group. Moreover, we characterize these effects with a small number of scalar parameters: R_0, R_1, R_{max} and R_{min}. Knowledge of these four reproductive numbers facilitates the design of control strategies for the eradication or spread of the contagion process. Finally, the existence of the bistable regime, which is absent in the SIS and SIR models, presents opportunities for the analysis and design of multi-agent systems that exhibit more complex collective behaviors.

Chapter 6

Final Remarks

In this dissertation, we derived and analyzed dynamic models for groups of interconnected individuals that exhibit robust and flexible collective behaviors that are controlled by a small number of critical system parameters. In Chapter 3, motivated by field observations of foraging red harvester ants, we showed that adaptations in volatility at the level of the individual, in response to a first exposure to the environment, lead to the emergence of complex collective behaviors that are robust to perturbations and flexible to changes in environmental conditions [43]. In Chapters 4 and 5, we generalized our results to contagion processes in which individuals adjust their susceptibility after a first exposure to the infection. We examined both well-mixed settings (Chapter 4) and network topologies (Chapter 5), demonstrating that adaptations in susceptibility in response to a first exposure to the infection lead to a distinct number of transient and steady-state group behaviors that depend on initial conditions and communication topology. Moreover, we showed that these behaviors are characterized by the values of a small number of critical system parameters [57,58].

Our results suggest that the ability of individuals to adjust their susceptibility in response to interactions, and the feedback across multiple timescales that results from it, constitutes a general principle for robust and flexible collective behavior. Groups

that operate without central control rely on complex patterns of interactions to function. The response of a group to stimulus depends on the effects of the stimulus on the individuals within the group, and on how they interact with one another and their environment. Adaptations in the susceptibility of individuals introduce a feedback mechanism at the level of the group that is significantly slower than the dynamics at the level of the individual. This fast-slow process modulates the sensitivity of the group to stimulus and results in group behaviors that evolve in time to reflect the collective assessment of the conditions by the individuals in the group.

The abstraction provided by our modelling approach allows for generalizations and comparisons with other systems that exhibit collective behaviors that display a balance of stability and plasticity. Our use of feedback across multiple timescales to regulate group sensitivity has a number of parallels with the homeostatic regulation of excitability in neurons, in which neurons adjust their sensitivity to stimulus based on their previous activity [21,22]. These changes in the excitability of neurons have been linked with changes in the number and density of receptors and ion-gate channels that modulate the level of stimulus required for activation. Comparably, in our models, group sensitivity to stimulus is regulated as a result of the accumulated changes in susceptibility of the individuals in the group based on previous activity.

Our results can be applied to suggest improvements to existing engineered distributed systems. The regulation of bandwidth congestion in TCP-IP communication networks has many parallels with the regulation of foraging in harvester ants [40,140]. In TCP-IP networks, source nodes transfer files to destination nodes via packets of information. When the recipient node receives each packet, it sends an acknowledgment to the source node. This feedback loop is used to prevent congestion in the network [140]. If the rate of acknowledgements is slow, then the sender node interprets this as a signal for low available bandwidth. The sender reacts to this signal by decreasing the data transmission rate accordingly. Similarly, higher rates of ac-

knowledgements lead to higher transmission speeds. This is similar to the process of interactions by which harvester ant colonies regulate the rate of outgoing foragers [34–38, 122]. The acknowledgement process provides information on the quality of the available bandwidth in the network, similar to how the interaction process in harvester ants provides information on the quality of the foraging conditions. In both cases, the information is used to regulate the rate of the output "packets". Our model for the regulation of foraging in harvester ants suggests that adjusting how source nodes respond to different acknowledgement rates could lead to more flexible communication networks. Furthermore, introducing heterogeneities between nodes in how they respond to different acknowledgement rates could prioritize the flow of information along selected parts of the network when bandwidth is limited.

The mechanism by which foragers in the nest become active is comparable to a complex contagion process in which the active foragers play the role of infected individuals. Individuals change state from inactive to active via the complex contagion process of interactions with active individuals (incoming foragers), and active individuals become inactive after an arbitrary delay (the foraging time). This analogy allows us to compare the behavioral regimes for the closed-loop foraging ant model with the adaptation mechanism, described in Section 3.9, to the behavioral regimes of the well-mixed SIRI model, described in Section 4.3. The trivial foraging regime is comparable to the infection-free regime, the steady-state foraging regime is comparable to the endemic regime, the loop foraging is comparable to the epidemic regime, and the delayed foraging regime is comparable to the resurgent epidemic phenomenon of the bistable regime. The equivalence of the two sets of behavioral regimes, one in the complex contagion process of foraging harvester ants, and the other in the simple contagion process of the SIRI model, support our hypothesis that the adaptations in susceptibility mechanism is a general principle for robust and flexible collective behavior. In addition, our results show that these behavioral regimes are determined by

the value of two critical system parameters that capture the sensitivity of the group to stimulus before and after the adjustments in susceptibility.

In Chapter 5 we developed multi-agent system dynamics that incorporate our insights on robust and flexible collective behavior that we developed in Chapters 3 and 4. We extended the results of the previous chapters by demonstrating how the behavior of the group is affected by network structure and agent heterogeneities. Our results highlight the use of a small number of critical system parameters to understand the complex mapping between the actions of individual agents and the resulting group outcomes. Knowledge of the analytical form of these critical parameters yields exact predictions on how changes in network structure and/or individual behavior affect group outcomes.

6.1 Future Work

Many exciting avenues of research remain in the study of robust and flexible collective behavior. The most promising of these would develop simple rules at the level of the individual that provide analytic guarantees on the performance of the group in new, unforeseen environments. Generating design principles and control methodologies for resilient collective behavior requires the development of new models and theories to answer questions concerning the relationship between robustness and communication topology, the benefits and/or costs of agent heterogeneity in the flexibility of group behaviors, and the importance of agent learning and adaptation in group level strategies for an ever changing environment. Here we present a selected number of ideas that would extend the results of this dissertation along these avenues.

6.1.1 Refined Models for Harvester Ant Foraging

The results in this dissertation were inspired by the regulation of foraging problem in harvester ants. A better understanding of the regulation of foraging in harvester ants would lead to better principles for robust and flexible collective behavior.

The closed-loop excitable model developed in Chapter 3 is a phenomenological model that captures the main components in the regulation of foraging. The model can be used as a starting point for the development of mechanistic models that resemble the biological system more closely. In our model, we have assumed that foragers adapt their volatility once after a first exposure to the environmental conditions. In an ongoing collaboration, we are analyzing data for individual foraging ants to test whether foragers adapt their sensitivity to interactions after every foraging trip. Results on the continuous effects of temperature and humidity on the sensitivity of foragers could then be used to develop models that describe how colonies adapt to changing environments and decide when to stop foraging. Supplementary improvements to the model include the addition of dynamics that describe the movement of foragers between the entrance chamber and the deeper tunnels of the nest, dynamics that describe the effects of foraging time on volatility, and the inclusion of forager heterogeneity. All of these will require further experimental efforts to support the modeling decisions.

6.1.2 Exact Bistability Conditions in Networks

In the bistable regime of the network SIRI model, described in Section 5.4, solutions converge to an infection-free equilibrium or to an endemic equilibrium depending on the initial conditions. Descriptions of the set of initial conditions that lead to an infection-free equilibrium could be used to design control methodologies that prevent or exploit the resurgent epidemic phenomenon. We conjecture that analytic results

for the exact bistability conditions in the network SIRI model might be feasible in the case of two agents. In this case, the center manifold dynamics are three-dimensional, allowing for the description of transversality conditions for the resulting co-dimension two bifurcation without parameters. These results could then be used to gain insight into the N agent case and the resulting co-dimension N bifurcation without parameters.

6.1.3 Centrality Measures for Contagion Processes

Centrality measures provide a metric that ranks nodes within a network in terms of how "well-positioned" they are within the graph. These measures are used in a variety of contexts to better understand the roles of individual agents in multi-agent system dynamics. Deriving a centrality measure for contagion processes would facilitate the design of control strategies that seek to eradicate or spread the infection by ranking the nodes in the network in terms of their effect on the contagion dynamics.

Usually, centrality measures are static metrics that depend on properties of the graph structure only [29]. Understanding how the placement of agents within the network influences the spread of the contagion process requires a different notion of centrality that takes into account the dynamics of the infection and the way in which agents adapt to it [141–145]. We propose that the equation for the endemic equilibrium (EE) in the network SIRI model (5.8) could be used as a basis for the development of natural centrality measures for contagion processes on networks. The equation provides a natural ranking of nodes according to their probability of infection at steady-state, acting as a centrality measure that identifies agents at the most risk of being reinfected once the infection has become endemic. However, since (5.8) is only valid once the infection has spread through the network, this centrality measure is not helpful in the derivation of control laws that seek to prevent the spread of the infection. Future centrality measures should incorporate information on the state of

the system at any point in time, providing a ranking of nodes that could then be used to develop centralized or distributed control strategies. Moreover, any centrality measure derived from the network SIRI model dynamics could potentially be used for the stochastic SIRI Markov process described in Section 5.1. This would provide a valuable resource for the development of control strategies for contagion processes in stochastic settings which more closely resemble real-world scenarios [146].

6.1.4 Control of Contagion Processes on Networks

Previous results on the network SIS model (5.7) have shown that whether the infection dies or spreads depends on the value of the basic reproduction number $R_0 = \rho(B\Gamma^{-1})$ [77,81]. Many proposed control methodologies for contagion processes focus on mechanisms that affect R_0 via changes to the network structure and the system parameters [74]. Example mechanisms include optimal node removal, optimal link removal, and budget-constrained allocation [73, 74, 147, 148]. Under the assumption of perfect knowledge of the network structure and the system parameters, these same mechanisms could be used to derive control strategies for the network SIRI model that affect the value of the basic and extreme basic reproduction numbers: R_{max}, R_{min}, R_0, and R_1. We note that optimal node and link removal have been shown to be NP-complete and NP-hard problems, respectively [74, 149]. Alternatives include the use of centrality measures to select the nodes that should be removed first. Moreover, if knowledge of the network structure or system parameters is uncertain or not available, these control strategies are not viable. Instead, distributed control strategies that affect the system parameters are required. Our network SIRI model results invite an interpretation of the changes in susceptibility by the individuals in the network as a distributed heuristic feedback policy: individuals assess the quality of the process spreading through the network and adjust their susceptibility accordingly, leading to the eradication or spread of the infection. Future studies could exploit our results for

the network SIRI model and use centrality measures for contagion processes to derive necessary and sufficient conditions on the reinfection and/or recovery rates of a small number of selected agents that would guarantee that the infection dies out or spreads through the network.

Part II

Published Work

Chapter 7

Outline and Contributions

Part II of this dissertation consists of two chapters containing work that has been published in peer-reviewed journals and a third chapter containing work that has been submitted for publication. Each chapter has been slightly adapted from their published or submitted version to conform with the format and notation used in Part I of this dissertation.

7.1 Outline

In Chapter 8 we present a dynamic model for the regulation of foraging in desert harvester ants [43]. The model implements feedback across different timescales and is based on field observations of foraging rates. The model is low-dimensional and retains sufficient analytical tractability to perform analysis on the effects of critical system parameters on transient and steady-state foraging rates. We show how a simple mechanism in which foragers adapt after first exposure to the environment is sufficient to capture most of the foraging behaviors observed in the data.

In Chapter 9 we study the SIRI epidemic model for reinfection in well-mixed settings in which the susceptibility of individuals changes irreversibly after a first

exposure to the infection [57]. We show how the dynamics of the model exhibit four distinct dynamical regimes: infection-free, endemic, epidemic, and bistable. The regimes are determined by two parameters that are related to the basic reproduction number. We prove that in the bistable regime, there exists a critical initial condition below which solutions reach an infection-free equilibrium and above which solutions reach an endemic equilibrium. Moreover, we prove that solutions in the latter case exhibit a resurgent epidemic in which the solution initially decays and gets close to zero for arbitrarily long lengths of time before rapidly increasing.

In Chapter 10 we study the SIRI epidemic model for reinfection in network topologies. We consider the heterogeneous case in which each individual has a recovery rate and each pair of individuals have an infection rate and a reinfection rate. We extend the results from the well-mixed model and show that the dynamics of the network model also exhibit four distinct dynamical regimes: infection-free, endemic, epidemic, and bistable. We prove that the regimes are determined by two parameters that capture the effect of heterogeneities and network structure.

7.2 Contributions

I am the lead author and lead contributor to the conceptualization, formal analysis, investigation, illustrations, simulations, and writing presented in all three papers. My advisor, Professor Naomi Ehrich Leonard, provided advice and guidance on all aspects of this work. Journal reviewers provided valuable suggestions that improved the quality of the manuscripts, and listed co-authors provided advice on different aspects of the research, as described below.

In Chapter 8, Deborah Gordon provided a first data set which guided my initial questions and analysis. Data from 2015 and 2016 was collected by myself and Deborah Gordon's team, but I did not participate in the 2017 data collection process.

Members of Deborah Gordon's team are acknowledged at the end of the chapter. I performed the data cleaning and analysis for all data sets. Naomi Leonard and I, with guidance from Deborah Gordon, framed the main questions addressed in the work. I developed and refined the model with guidance from Naomi Leonard. I proposed the adaptation mechanism and wrote all simulations. Naomi Leonard and Deborah Gordon revised and edited the manuscript. Chapter 8 appeared in PLoS Computational Biology in December of 2018 [43]. I presented early versions and partial results at the Princeton 2017 MAE Research Day, the SIAM 2017 Conference on Applications of Dynamical Systems, the Janelia 2018 Conference on Distributed, Collective Computation in Biological and Artificial Systems, and the International Center for Theoretical Physics 2018 Conference on Collective Behavior.

In Chapter 9, I framed the questions, derived the model, and performed all analysis. Biswadip Dey suggested the time to resurgence definition and checked the mathematical results. Both Biswadip Dey and Naomi Leonard revised and edited the manuscript. I am grateful to Zahra Aminzare for providing helpful suggestions for the structure of the paper that improved early stages of the manuscript. Chapter 9 appeared in IEEE Control Systems Letters in April 2018 and as a contributed paper in the Proceedings of the 2018 IEEE Conference on Decision and Control, where I presented the work.

In Chapter 10, I framed the questions with guidance from Naomi Leonard. I performed all of the analysis with the exception of the reduction of the model for d-regular digraphs, which I performed with the help of Naomi Leonard. Naomi Leonard made a number of suggested changes to the structure of the manuscript and results which vastly improved the presentation of the material. I am very grateful to Naomi Leonard for her invaluable help in revising the results and editing the manuscript. Chapter 10 has been submitted for publication. I presented early versions and partial results at the 2018 Dynamic Days Conference, the International Center for Theoretical

Physics 2018 Conference on Collective Behavior, and the 2018 IEEE Conference on Decision and Control.

Chapter 8

Regulation of Harvester Ant Foraging as a Closed-Loop Excitable System*.

Renato Pagliara, Deborah M. Gordon, and Naomi Ehrich Leonard

Ant colonies regulate activity in response to changing conditions without using centralized control. Desert harvester ant colonies forage for seeds, and regulate foraging to manage a tradeoff between spending and obtaining water. Foragers lose water while outside in the dry air, but ants obtain water by metabolizing the fats in the seeds they eat. Previous work shows that the rate at which an outgoing forager leaves the nest depends on its recent rate of brief antennal contacts with incoming foragers carrying food. We examine how this process can yield foraging rates that are robust to uncertainty and responsive to temperature and humidity across minute-to-hour timescales. To explore possible mechanisms, we develop a low-dimensional

[1]This chapter was published as Pagliara, Gordon, and Leonard in PLOS Computational Biology on December 4, 2018 [43]. All raw data and the MATLAB model simulation scripts are available in the repository in Princeton Universitys DataSpace at the permanent address: http://arks.princeton.edu/ark:/88435/dsp01qv33s040d

analytical model with a small number of parameters that captures observed foraging behavior. The model uses excitability dynamics to represent response to interactions inside the nest and a random delay distribution to represent foraging time outside the nest. We show how feedback from outgoing foragers returning to the nest stabilizes the incoming and outgoing foraging rates to a common value determined by the volatility of available foragers. The model exhibits a critical volatility above which there is sustained foraging at a constant rate and below which foraging stops. To explain how foraging rates adjust to temperature and humidity, we propose that foragers modify their volatility after they leave the nest and become exposed to the environment. Our study highlights the importance of feedback in the regulation of foraging activity and shows how modulation of volatility can explain how foraging activity responds to conditions and varies across colonies. Our model elucidates the role of feedback across many timescales in collective behavior, and may be generalized to other systems driven by excitable dynamics, such as neuronal networks.

Author Summary: We investigate the collective behavior that allows colonies of desert harvester ants to regulate foraging activity in response to environmental conditions. We develop an analytical model connecting three processes: 1) the interactions between foragers returning to the nest and available foragers waiting inside the nest, 2) the effect of these interactions on the likelihood of available foragers to leave the nest to forage, and 3) the return of foragers to the nest after finding seeds. We propose a mechanism in which available foragers modify their response to interactions after their first exposure to the environment. We show how this leads to colony foraging rates that adjust to environmental conditions over time scales from minutes to hours. Our model may prove useful for studying other classes of systems with excitability dynamics that exhibit

both stability in behavior and flexibility with respect to environmental conditions.

8.1 Introduction

Social insect colonies operate without central control. Colonies maintain coherence and plasticity in the face of perturbation and change, even though individuals have limited and uncertain information on the state of the group and the state of the environment. Collective behavior emerges from the response of individuals to social interactions and their assessment of the local environment [7–9, 24]. The study of social insects provides opportunities to investigate open, fundamental questions on how collective behavior adjusts to different conditions and how small differences in these adjustments can lead to large differences in behavior across groups.

The regulation of foraging activity in colonies of the harvester ant (*Pogonomyrmex barbatus*) is a well-studied example of collective behavior [41]. Harvester ants live in the hot and dry Southwestern US desert where they forage for seeds scattered by wind and flooding on the timescale of weeks and months. Foragers do not use pheromone trails; instead, they spread out across the foraging area in search of seeds [39]. Thus the regulation of foraging in harvester ant colonies, unlike in honey bees and in ant species that use pheromone trails, does not allocate workers among spatially fixed resources that differ in quality and availability [150–153]. A harvester ant's foraging trip time may vary with food availability. However, on the scale of the colonys foraging area, food is not depleted in the course of the foraging period on a given day, and hot and dry conditions, rather than lack of food, can cause colonies to stop foraging.

The regulation of foraging activity manages a tradeoff between spending water and obtaining water and food: foragers lose water while outside in the dry air, but colonies obtain water by metabolizing the fats in the seeds that they eat [32, 33].

Foraging is initiated each morning by a distinct group of workers, the patrollers [154, 155], who leave the nest before the foragers emerge and explore the nest mound and foraging area. It is the safe return of the patrollers that initiates foraging, through encounters inside the nest between foragers and returning patrollers [122]. Once foraging has begun, harvester ant colonies regulate the rate at which foragers leave the nest using the incoming rate of successful foragers returning with food [34–38]. When an ant contacts another ant with its antennae, it perceives the other ant's cuticular hydrocarbon (CHC) profile [34]. Because conditions outside the nest change the chemistry of the cuticular hydrocarbons, CHC profiles are task-specific [156], so that in the course of antennal contact, one ant can detect whether another is a forager. An available forager, waiting in the entrance chamber inside the nest, is stimulated to leave the nest by antennal contact with foragers carrying food [36–38]. The rate of interactions experienced by an available forager inside the nest entrance chamber correlates with the local density of ants [38]. Thus a higher rate of incoming foragers leads to a higher rate of interactions [37]. Because each forager searches until it finds a seed, the rate of interaction serves as a noisy measurement of the current foraging conditions [39, 40]. A higher rate of forager return, which reflects a greater food supply, increases the likelihood that available foragers will leave the nest to forage [37, 38, 132].

In the integrator model of [38], each available forager inside the nest collects evidence from incoming foragers by integrating its recent experience of antennal contacts. When the integrated stimulus passes a threshold, the available forager is likely to leave the nest; in the absence of interactions the forager is likely to descend from the entrance chamber to the deeper nest [37, 132], protecting the colony from the inherently noisy signal that results from limited and uncertain interactions [157]. The integrator model has been used to study regulation of the outgoing foraging rate on short timescales of minutes [40].

Colonies regulate their foraging activity on longer timescales, such as from hour to hour, from day to day [123,125], and across years [41,123,124,126] as colonies grow older and larger. Over timescales from tens of minutes to hours, ants that start as available foragers inside the nest leave the nest to forage, find seeds, return to the nest, and become available foragers again. Thus, the activation of available foragers inside the nest through interactions with incoming foragers is connected in a "closed loop" to the foraging activity outside the nest through feedback of the ants themselves: the stream of foraging ants out of the nest is the input to the foraging activity, and the output of the foraging activity is the stream of foraging ants into the nest (see Figure 8.1). However, little is known about the role of feedback in the regulation of foraging activity at the timescale of hours and as foraging activity is adjusted to changing environmental conditions. By mid-day in the summer, temperature is high and humidity is low (Figure 8.10). Foraging activity increases from its start in early morning and then levels off, often remaining at a steady rate for tens of minutes to hours. It declines to no activity during the heat of the afternoon.

How a colony regulates foraging in response to environmental conditions, especially temperature and humidity, is ecologically important. Colonies live for 20-30 years, the lifetime of the single founding queen who produces short-lived workers year after year. At about five years of age the queen begins to produce reproductives that mate with those of other colonies, and the daughter queens found offspring colonies [42]. Colonies differ in the regulation of foraging and these differences persist from year to year, including variation in how often colonies are active [123] and in how they respond to changing temperature and humidity conditions [41, 124, 125]. The persistence of foraging behavior across years, in subsequent cohorts of workers, and comparison of parent and offspring colonies [42] suggests that foraging behavior is heritable. It appears that colony differences in the regulation of foraging arise from

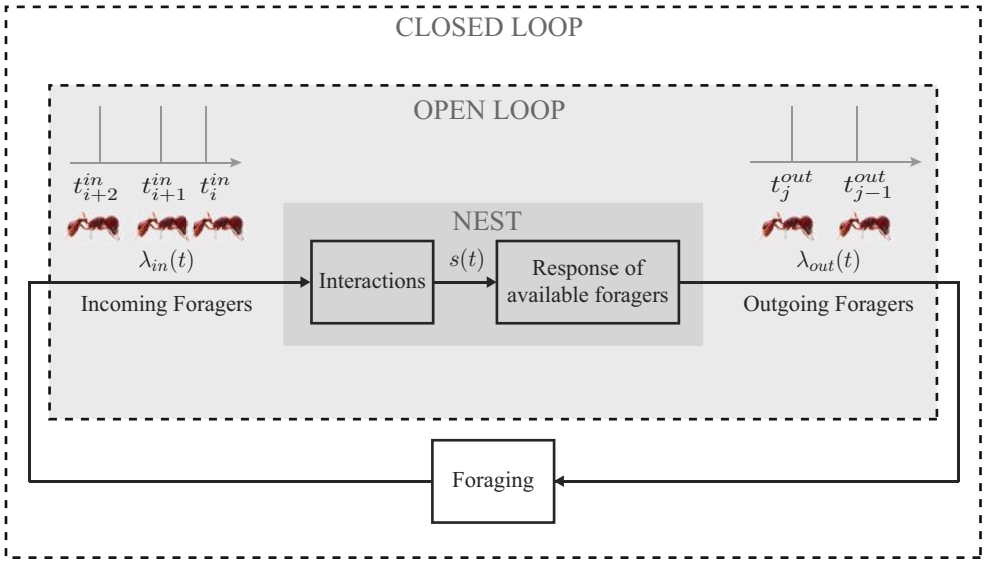

Figure 8.1: Diagram of the closed-loop model with two components inside the nest and one component outside the nest. The "Interactions" component maps the sequence of incoming foragers λ_{in} to a stimulus s to represent the result of interactions of available foragers inside the nest entrance chamber with incoming food-bearing foragers. The mapping uses a leaky integrator that increases by a fixed magnitude with every incoming forager and has a natural decay rate. The "Response of available foragers" component maps s to the sequence of outgoing foragers λ_{out} using the nonlinear FitzHugh-Nagumo oscillator dynamics. Each oscillation represents an ant leaving the nest to forage. The "Foraging" component maps λ_{out} to λ_{in} using a random time delay with an associated probability distribution to represent the time an ant spends outside the nest foraging.

differences in how individuals respond to interactions with incoming foragers, that is, in the rates of interaction required to stimulate a forager to leave the nest [38].

How a colony adjusts foraging activity to low humidity and high temperature is crucial for reproductive success: colonies that conserve water are more likely to have offspring colonies [41]. We hypothesize that these differences among colonies in their response to different environmental conditions are the result of differences in how their foragers assess humidity and in how this influences their response to interactions with incoming foragers [125]. Recent work suggests this depends on variation in the neurophysiology of biogenic amines such as dopamine [158]. Here we model how colonies adjust to environmental conditions by adjusting their "volatility" defined as their sensitivity to interactions with returning foragers. Our goal is to suggest testable

hypotheses about the sources of variation among colonies upon which natural selection can act to shape collective behavior.

Previous modeling work has elucidated how the outgoing foraging rate across timescales of minutes depends on the incoming foraging rate [40], and how individuals assess interaction rate [38]. But we do not know how these are combined to adjust foraging activity across minute-to-hour timescales, how steady foraging rates are maintained, how the adjustments may depend on environmental conditions, and how they may differ from colony to colony.

Here we propose a closed-loop model (Figure 8.1) to address these questions by examining how an incoming forager's assessment of external conditions provides additional feedback to the colony and in turn adjusts the colony foraging rate. Our model is motivated in part by the frequent use of excitability dynamics to model neurons, and the parallels between ant-to-ant interactions that drive foraging and neuron-to-neuron interactions that underlie the cognitive abilities of organisms [3, 4, 38, 128, 129]. Using well-studied excitability dynamics of a weakly interacting collective, we introduce feedback at multiple time-scales and explore general questions concerning stability and responsiveness to a changing environment.

Drawing on theory and tools from dynamical and control systems, we study the relationship, in the model, between the fast activation of foragers inside the nest and the slow feedback from incoming foragers to describe, with a small number of parameters, how the incoming and outgoing foraging rates adjust to changing conditions on the timescale of tens of minutes to hours. We show how the foraging rates are stabilized, and we suggest how small differences in parameter values can lead to variation in the regulation of foraging for different environmental conditions and for different colonies.

8.2 Methods

8.2.1 Field Observations of Foraging Activity

We performed field observations of red harvester ant colonies at the site of a long term study near Rodeo, New Mexico, USA. Observations were made in August and September of 2015, 2016, and 2017. Foragers leave the nest in streams or trails that can extend up to 20 m from the nest [159]. Each forager leaves the trail to search for seeds, and once it finds food, it returns to the nest [39, 159]. Data on foraging rates were recorded from the beginning of the foraging period in early morning until around noon. We recorded the times at which foragers crossed a line perpendicular to the trail at a distance of about 1 m from the nest entrance, as in previous work (e.g. [40, 124, 160]). The timestamps for each forager crossing the line were recorded either manually in real-time with the assistance of an electronic tablet and custom software, or from video recordings, processed with computer vision software (AnTracks Computer Vision Systems, Mountain View, CA). In some cases we used both tablet and video to ensure that both data collection methods provided similar results.

We denote by t_i^{in}, $i \in \mathbb{N}$, the sequence of times incoming foragers cross the line and by t_j^{out}, $j \in \mathbb{N}$, the sequence of times outgoing foragers cross the line. Sequences of incoming and outgoing foragers are represented as sums of infinitesimally narrow, idealized spikes in the form of Dirac-delta functions:

$$\lambda_{in}(t) = \sum_{i=1}^{n} \delta(t - t_i^{in}), \qquad \lambda_{out}(t) = \sum_{j=1}^{m} \delta(t - t_j^{out}), \tag{8.1}$$

where n and m are the indices of the last incoming and outgoing forager, respectively, before time t. We estimated the instantaneous incoming and outgoing foraging rates,

in units of ants/sec, using a sliding window filter with window $\Delta t = 300$ sec:

$$r_{in}(t) = \int_{-\infty}^{\infty} w(\zeta)\lambda_{in}(t-\zeta)\ \mathrm{d}\zeta, \qquad r_{out}(t) = \int_{-\infty}^{\infty} w(\zeta)\lambda_{out}(t-\zeta)\ \mathrm{d}\zeta \tag{8.2}$$

where

$$w(t) = \begin{cases} 1/\Delta t & \text{if } -\Delta t/2 \leq t \leq \Delta t/2 \\ 0 & \text{otherwise.} \end{cases} \tag{8.3}$$

We selected the size of the sliding window to be sufficiently long to remove noise but sufficiently short to preserve the interesting dynamic features of the foraging rates across tens of minutes to hours.

8.2.2 Model

We propose a low-dimensional dynamic model with a small number of parameters that has sufficiently rich dynamics to capture the range of observed foraging behavior across minute-to-hour timescales and yet retains tractability for analysis. We use the model to systematically investigate the effects of model parameters and environmental conditions, notably temperature and humidity, on foraging rates.

Our model has three components as shown in Figure 8.1: 1) the *Interactions* component models the accumulation of evidence by available foragers inside the nest entrance chamber from their interactions with incoming foragers carrying food, 2) the *Response of available foragers* component models the activation of available foragers to leave the nest to forage in response to accumulated evidence, and 3) the *Foraging* component models the collecting of seeds outside the nest by active foragers. We assume the total number of foragers N (active foragers outside the nest plus available foragers inside the nest) remains constant throughout the foraging day, although this assumption could be relaxed in a generalization of the model.

Interactions

We use leaky-integrator dynamics to model the stimuli s that the group of available foragers inside the nest entrance chamber experience from their interactions with incoming food-bearing foragers:

$$\dot{s} = -\frac{s}{\tau} + k\lambda_{in}. \tag{8.4}$$

The leaky-integrator (8.4) integrates information from the sequence of incoming foragers (Figure 8.2A), but the information "leaks," i.e., decays slowly over time. Here, we use the leaky-integrator dynamics to estimate the instantaneous rate of incoming foragers which is proportional to the overall rate of interactions experienced by available foragers inside the nest [37]. The continuous-time signal s increases by a fixed amount k with every incoming forager in λ_{in} and decays exponentially back to zero with a time constant of τ (Figure 8.2B).

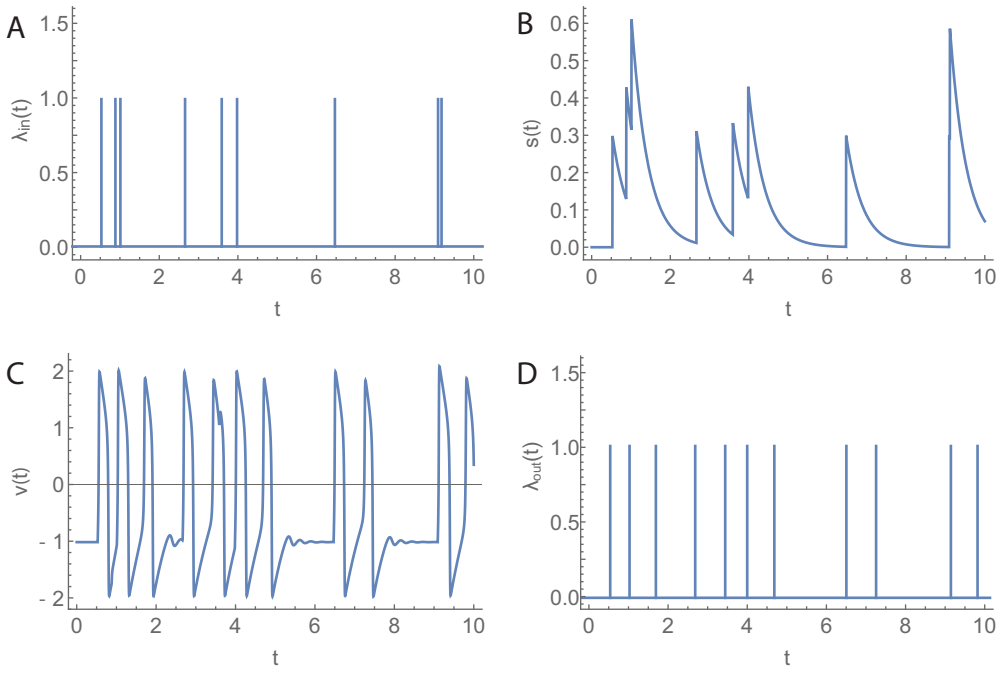

Figure 8.2: Open-loop model. A) Sequence of incoming foragers λ_{in}. B) Stimulus signal s associated with λ_{in}. C) FN output v for input s. D) Sequence of outgoing foragers λ_{out} obtained by thresholding FN output from below at 0.75.

The leaky-integrator dynamics work as an evidence accumulator that gradually forgets past evidence. These dynamics have been used to model chemical synapses [161] and have been used as the integrate-and-fire neuronal model when there is no reset boundary [162–164].

Response of Available Foragers

We use the FitzHugh-Nagumo (FN) excitability dynamics [103, 104], often used to model neuronal excitability, as a phenomenological model for the activation of available foragers inside the nest entrance chamber. Our choice of dynamics is motivated by the similarities between the ant-to-ant interactions that activate available foragers to leave the nest to forage and the neuron-to-neuron interactions that drive cognition in organisms [3, 4, 38, 128, 129]. In the neuronal setting, the FN dynamics model the membrane voltage response of a neuron to an electrical stimulus. The state v is the voltage and a second state u is a recovery variable that models the flow of ions across the membrane and provides a relatively slow negative feedback on the rate of change of v. For low values of stimulus, the voltage remains at rest; for intermediate values, the voltage oscillates; and for large values, the voltage saturates. In our phenomenological model, an oscillation represents the activation of an outgoing forager.

We first consider a homogeneous colony and model the dimensionless, scalar activation state v of available foragers in the nest entrance chamber as the fast timescale variable in the FN equations [103, 104]:

$$\epsilon_1\epsilon_2\dot{v} = v - v^3/3 - cu - a + s \tag{8.5}$$

$$\epsilon_1\dot{u} = v - cu. \tag{8.6}$$

These equations describe nonlinear oscillator dynamics with the stimulus s of (8.4) as the input and v as the output. Oscillations result from a balance between positive feedback in v (first term on the right of (8.5)) and negative feedback in the dimensionless, slow timescale, recovery variable u. The parameter c, which scales the negative feedback, modulates the frequency of oscillations and the range of values of stimulus s that lead to oscillations. We introduce c in both (8.5) and (8.6) so that a change in c determines the frequency of oscillations and the range of values of stimulus s that lead to oscillations, but it leaves other features of the dynamics unchanged. Since c regulates responsiveness, and does so better than parameters ϵ_1, ϵ_2, and a, we let c represent the *volatility* of the available foragers.

The parameter ϵ_2 defines the time separation between the dynamics of the fast and slow states, and the parameter ϵ_1 defines the time separation between the FN dynamics and the stimulus dynamics (8.4). The parameter a provides an offset to s and its value is selected based on the value of k, which is the increase in stimulus s resulting from an interaction with an incoming forager. Parameter a is chosen so that k is greater than the threshold above which input s elicits an oscillation. Eliciting at least one oscillation per isolated incoming forager in the model allows for a rapid increase in the foraging rates during the first few minutes of foraging, when the initial incoming foraging rate is low.

The activation dynamics ((8.5) and (8.6)) of the available foragers yield three qualitatively distinct dynamical regimes, determined by the magnitude of input s, and bifurcation values b_1 and b_2 (Figure 8.2C). In the first regime, the system remains in a *resting* state for $0 < s < b_1$. This reflects the situation in which the stimulus for available foragers to leave the nest is low because there are few incoming foragers. In the second regime, which takes place when $b_1 < s < b_2$, the system is in an excited state with oscillations in v. This reflects the situation in which incoming foragers are sufficiently frequent to stimulate the available foragers. The transition from resting

to oscillatory behavior as s increases corresponds to a Hopf bifurcation and b_1 is the corresponding bifurcation point. The oscillations appear as short-lived spikes, and we define each spike for which v increases above 0.75 as a forager leaving the nest (Figure 8.2D). The shortest possible time between foragers leaving the nest is determined by the volatility c (see Section 8.5.1).

In the third regime, corresponding to very large values of $s > b_2$, there are no oscillations and the system is fixed in a *saturation* state. The transition from oscillatory to saturated regime is a second Hopf bifurcation with bifurcation point b_2. This means that a high instantaneous incoming rate that produces a high value of s will lead to saturation in the FN dynamics resulting in a decrease in instantaneous outgoing rate. Conditions represented by the effect of saturation include 1) overcrowding effects, which reduce the percentage of interactions experienced by each available forager relative to the incoming foraging rate, 2) the limited size of the nest entrance tunnel, which constrains how many foragers can enter and leave the nest in a short amount of time, and 3) the difference in timescales between the high outgoing rates, in seconds, and the time required, in minutes, for foragers to move from the deeper chambers of the nest up to the entrance chamber [37, 132].

Foraging

We treat the process of foraging for seeds outside the nest as a random time delay. We model the interval between the time that a forager leaves the nest and the time when it returns with food as a chi-square random variable X, with parameter D representing the mean foraging time in minutes. The distribution of foraging times $F(X, D)$ is

$$F(X, D) = 1 - \frac{\Gamma(D/2, X/2)}{\Gamma(X/2)}, \tag{8.7}$$

where $\Gamma(X)$ and $\Gamma(a, X)$ are the Gamma function and the upper incomplete Gamma function, respectively. This right-skewed distribution is based on field observations of

the duration of foraging trips, measured as the total time elapsed from when a forager leaves the nest to when it returns with food [39]. For $D = 2$, $F(X, 2) = 1 - e^{-X/2}$.

Our model for the foraging process is equivalent to a queueing system [165] in which arriving customers, represented by outgoing foragers λ_{out}, find a seed after a given random "service time". The number of servers in this analogy of the foraging process as a queue is assumed to be infinite because foragers do not need to wait before they start looking for a seed (i.e., before receiving the service). In queueing theory, queues with random service time and infinite number of servers can elucidate the effects of service time on the expected number of customers being serviced at any time.

Proposed Mechanism for Response to Environmental Conditions

We propose a mechanism for colony response to environmental conditions, illustrated in Figure 8.3, in which the volatility of a forager changes after it has been on a foraging trip and exposed to the conditions outside the nest. The proposed mechanism is based on measurements showing that the temperature and humidity inside the nest remain constant throughout the foraging activity period (see Figure 8.10). This means that foragers have no information about conditions outside until they leave the nest.

As a first approximation, the model changes the volatility of each forager after it leaves the nest to forage for the first time. Available foragers that have not yet been outside the nest, and are therefore uninformed about the current temperature and humidity outside the nest, have volatility c_u. Available foragers that have been outside at least once to forage, and are therefore informed about the current temperature and humidity, have volatility c_i.

The values of c_u and c_i, representing an average uninformed and an average informed available forager, can be any positive real numbers. These values can vary across colonies and across days. The uninformed volatility c_u can vary across days for

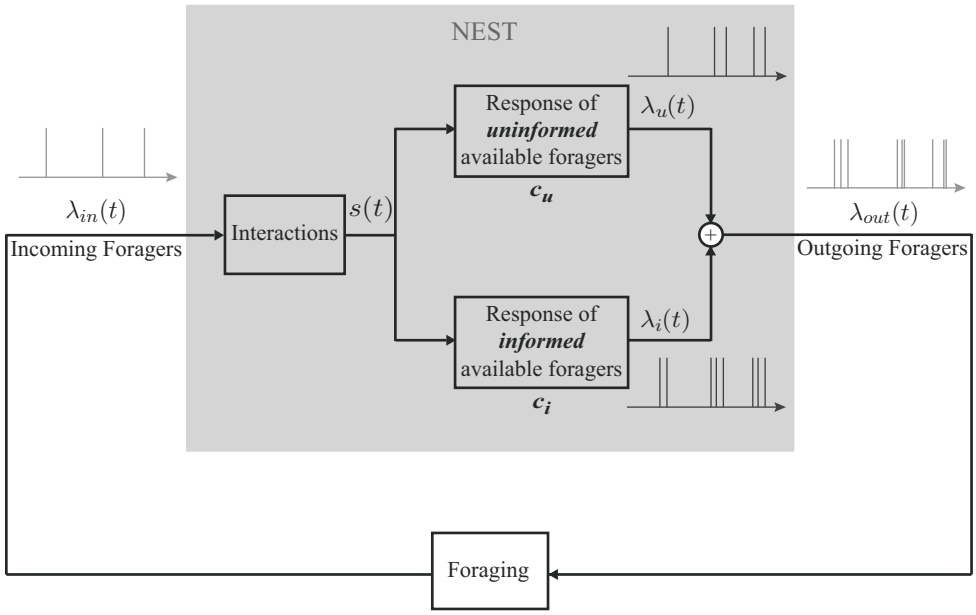

Figure 8.3: Block diagram of proposed mechanism for response of colony to environmental conditions. The available foragers inside the nest comprise two sets: f_u corresponds to those that have not yet left the nest and so are uninformed about the conditions outside the nest, and f_i corresponds to those informed during a previous foraging trip. The response of each set to s is represented by a different FN model, distinguished by the volatility parameter c_u for the uninformed and c_i for the informed. The outputs of these two oscillator dynamics are weighted probabilistically using thinning to get an outgoing stream of foragers $\lambda_{out}(t)$.

a given colony, and across colonies, in response to colony needs, such as the amount of brood to be fed and the amount of stored food, colony size, or neurophysiological factors such as biogenic amines (e.g. [158]). The informed volatility c_i reflects response to conditions that change both on the hourly and daily timescales, such as humidity and temperature outside the nest. For example, the hotter and drier it is outside, the smaller the c_i, so the foragers become less volatile and thus less likely to make subsequent foraging trips; the cooler and more humid it is outside, the larger the c_i, so the foragers become more volatile and thus more likely to make subsequent foraging trips. c_i can vary across colonies due to physiological differences among colonies in response to conditions. There is currently no evidence that c_i depends on how long it takes for a forager to find a seed, but further work is needed to investigate this.

Let f_u be the set of n_u *uninformed* available foragers that have not yet left the nest during the day and thus have no information about the environmental conditions and f_i the set of n_i *informed* available foragers that have been exposed to the environmental conditions during one or more previous foraging trips that day. We assume that once a forager becomes informed, it remains informed for the rest of the foraging day. The ants in f_u have volatility c_u and the ants in f_i have volatility c_i. Let $x_u = n_u/(n_u + n_i)$ and $x_i = n_i/(n_u + n_i)$ be the fraction of available foragers that are uninformed and informed, respectively, where we assume that $n_u + n_i > 0$. Then $x_u + x_i = 1$.

Initially, $x_i = 0$ and the colony is completely uninformed ($x_u = 1$). As foragers return to the nest after their first trip, x_i begins to increase and can continue to increase until $x_i = 1$ ($x_u = 0$), when all N foragers have been outside the nest at least once. How many minutes (or hours) it takes for x_i to transition from 0 to 1 depends on N, D, and the changing foraging rates. To model the changing foraging rates, we use two sets of FN oscillator dynamics: one to represent the response to s of the uninformed ants in f_u with volatility c_u and a second to represent the response to s of informed ants in f_i with volatility c_i. Let the corresponding sequences of output from the two oscillator dynamics be λ_u and λ_i, respectively. We define the sequence of outgoing foragers λ_{out} as a probabilistic sum of λ_u and λ_i, using a method called *thinning* [166]: Every event in λ_i is kept in λ_{out} with probability x_i, and every event in λ_u is kept in λ_{out} with probability $1 - x_i$. When $x_i = 0$ the foraging rate is determined by c_u, and when $x_i = 1$ the foraging rate is determined by c_i. When $0 < x_i < 1$, the effective c will be a nonlinear combination of c_u and c_i. The higher the effective c, the higher the outgoing foraging rate.

Here foragers adjust their volatility only once after their first foraging trip outside. We find that even with this adjustment at first exposure, the model provides the range of foraging behavior observed. However, the model can be generalized and predictions

refined by allowing for adjustments on subsequent foraging trips, and by allowing for other kinds of adjustments. For example, more than two sets of available foragers with different values of volatility can be used to model effects of repeated exposures to the environment, changing conditions on successive trips, or decay of information about the external environment over time. A decrease in N (total number of foragers outside and available inside the nest) can be used to model active foragers that return to the deeper nest after exposure to hot and dry outside conditions [37].

8.3 Results

8.3.1 Observations of Regulation of Foraging in Red Harvester Ants

Observations of instantaneous foraging rates computed from the 2015, 2016, and 2017 data show that across colonies and days, the incoming and outgoing foraging rates $r_{in}(t)$ and $r_{out}(t)$, where t is time of day, undergo a transient (i.e., a temporary pattern of change) early in the foraging period followed by an equilibration to a near-equal value, i.e., $r_{in}(t) \approx r_{out}(t)$, during the middle part of the foraging period.

The equilibration of the incoming and outgoing foraging rates to a near-equal value lasts for intervals from tens of minutes to several hours, and so we refer to it as a quasi steady-state (QSS). We show the data for two colonies in Figure 8.4. We plot the incoming rate r_{in} (blue) and the outgoing rate r_{out} (red) computed from the data for Colony 1357 (Figure 8.4A) and Colony 1317 (Figure 8.4B) versus time of day on August 20, 2016. For Colony 1357, the rates equilibrated to a near-equal value early in the day, i.e., between 8:00 and 8:30 am. This was followed by a couple of dynamic adjustments, but then by 9:30 am until just before noon, when all the ants returned to the nest, the incoming and outgoing rates were very closely equilibrated

at a QSS rate of around 0.25 ants/sec. Colony 1317 also was observed to reach a QSS. Its incoming and outgoing rates equilibrated to a near equal value shortly after 10:00 am, which lasted until just before noon, when all the ants returned to the nest. Colonies vary greatly in foraging rate [124], and that was true of these two as well. For Colony 1317, the QSS rate was approximately 0.65 ants/sec, more than twice the QSS rate for Colony 1357 on the same day.

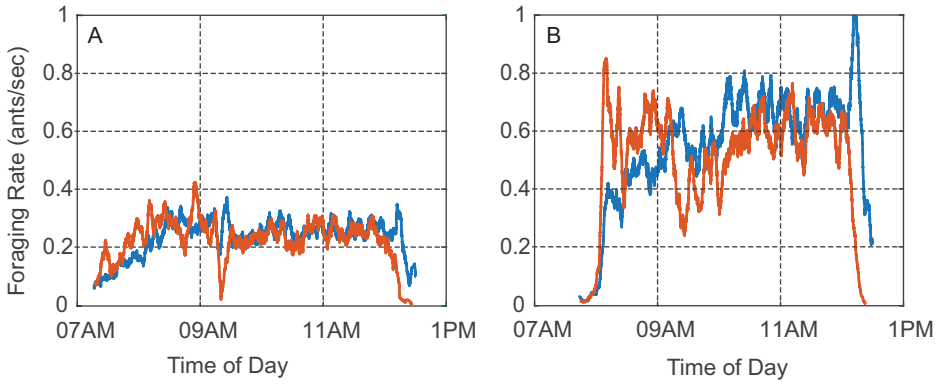

Figure 8.4: Plots of incoming foraging rate r_{in} (blue) and outgoing foraging rate r_{out} (red) versus time of day on August 20, 2016 for A) Colony 1357 and B) Colony 1317. The quasi steady-state (QSS) where incoming and outgoing rates equilibrated to a near-equal value can be observed for both colonies. The QSS rate for Colony 1317 was more than twice as great as it was for Colony 1357.

We show data for two other colonies in Figure 8.5. Figure 8.5A and C show r_{in} (blue) and r_{out} (red) versus time of day for a single colony, Colony 664, on two different days: August 27, 2015 and August 31, 2015. In each plot, the rates can be seen to come to a near-equal value sometime after 10:30 am. We plot in green the cumulative difference between number of incoming and number of outgoing foragers versus the time of day. The rates are at a QSS when the green curve is approximately horizontal. These data show, as has been observed previously [167], that a given colony varies in foraging rate from day to day, demonstrating that foraging is regulated by processes other than the number of foragers in a colony, which remains relatively constant on the timescale of months. From Figure 8.5A and C it can be seen that Colony 664 reached a QSS rate on August 27, 2015 that is more than twice the QSS rate it reached on

August 31, 2015. We note that August 27, 2015 was slightly cooler and more humid than August 31, 2015. On August 27 the average temperature and humidity were 25.3 C° and 58% while on August 31 they were 26.8 C° and 53%. Moreover, at 11:00 am on August 27, they were 27.5 C° and 52% while at 11:00 am on August 31, they were 28.8 C° and 45%.

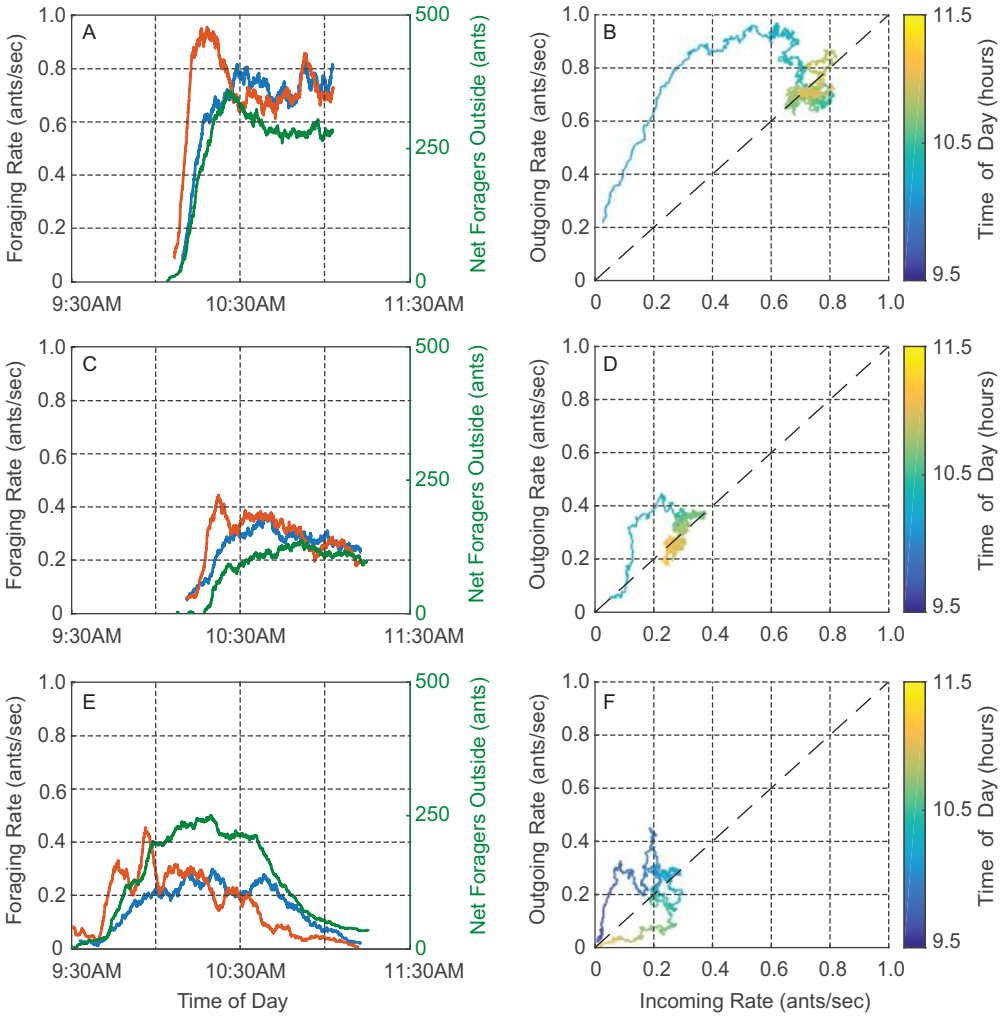

Figure 8.5: Plots of foraging rate data. Time series plots show incoming foraging rate r_{in} (blue), outgoing foraging rate r_{out} (red), and cumulative difference between the number of incoming and outgoing foragers (green) versus time of day. Input-output plots show $r_{out}(t)$ versus $r_{in}(t)$ with the color scale representing time of day t. A) and B) Colony 664 on August 27, 2015. C) and D) Colony 664 on August 31, 2015. E) and F) Colony 863 on September 1, 2015.

Figure 8.5E shows the data for Colony 863 on September 1, 2015, which were recorded manually. No QSS is observed, i.e., the ants went out but then returned to the nest by 11:00 am without maintaining a steady-state of foragers outside of the nest. Colony 863 did reach a QSS at a reasonably high foraging rate at 11:00 am on September 5, 2015 (see Figure 8.11A). These observations are consistent with measurements showing that September 1, 2015 was much hotter and drier than September 5, 2015. On September 1 the average temperature and humidity were 25.2 C° and 53% while on September 5 they were 22.6 C° and 77%. See Table 8.1 of the SI for more details.

Figure 8.4 and Figure 8.5 are representative of observations that suggest the equilibration of incoming and outgoing foraging rates to a near-equal rate to be an important feature in the regulation of foraging in red harvester ant colonies. Further, the equilibrated rate, and the possibility of early cessation of foraging, depend on factors that differ among colonies (Figure 8.4) and from day to day (Figure 8.5). We examine the transient in foraging rates further in Figure 8.5. Early in the foraging day, both r_{in} and r_{out} increased rapidly with r_{out} increasing more rapidly than r_{in}. This led to a rapid increase in the number of active foragers outside the nest. The rapid increase in both rates was followed by a decrease in r_{out} to the equilibrated near-equal value of the QSS (Figure 8.5A and C) or to an early return of the ants to the nest (Figure 8.5E).

Input-output plots show the relation between incoming and outgoing foraging rates Figure 8.5B, D, and F. These figures show the same data as Figure 8.5A, C, and E, respectively, but plot $r_{out}(t)$ versus $r_{in}(t)$ with time of day t in hours indicated by the color scale. The transient in rates during the early part of the foraging day appear as curved trajectories above the diagonal, because $r_{out}(t)$ is typically higher than $r_{in}(t)$. In Figure 8.5B and D, the curve rises and then falls to the QSS value where the trajectory then equilibrates around a point on the diagonal corresponding

to equal incoming and outgoing rates. This rise and fall of the curve in the input-output plot is typical, even when the trajectory returns to the origin as in the case of Figure 8.5F.

The data shown in Figs 8.4 and 8.5 as well as in Figure 8.11 are representative of the data collected in 2015, 2016, and 2017. Temperature and humidity for these data sets are given in Table 8.1. Figure 8.11B shows another example of a very early cessation of foraging. Figure 8.11C and D show two different examples of long transients. Figure 8.11E and F show two examples of a burst in the outgoing foraging rate at the start of the foraging day. See Section 8.5.2 for details.

8.3.2 Model Dynamics

Foraging Dynamics Inside the Nest

Given a sequence of incoming ants λ_{in}, our open-loop model of foraging dynamics inside the nest (Figure 8.1) predicts a corresponding sequence of outgoing ants λ_{out}. We find an analytic approximation for the mapping from mean incoming foraging rate $\bar{r}_{in}$ to mean outgoing foraging rate $\bar{r}_{out}$, parametrized by volatility c. To do so, we assume λ_{in} is a Poisson process with (constant) mean incoming rate $\bar{r}_{in}$; this is justified for observations of incoming and outgoing sequences of foragers for short periods of time [40].

We assign model parameter values to be $k = 0.3$, $\tau = 0.41$, $a = 0.35$, $\epsilon_1 = 0.2$, and $\epsilon_2 = 0.05$, which allow for rich dynamical behavior. While the qualitative behavior is unchanged for different values of $\epsilon_2 \ll 1$, very high or low values of a, k, and/or τ yield dynamics in which the stimulus s is either too low or too high to produce oscillations. So the values for a, k, and τ are selected to balance their opposing effects on s and the FN oscillating region.

The oscillating region of the FN dynamics corresponds to the range of values of s between the FN bifurcation points b_1 and b_2, computed as $b_{1,2} = a \mp \frac{1}{3}(1 - c\epsilon_2)^{3/2}$. The offset a does not affect the size of the oscillating region whereas the volatility c can control it: as c increases, the size of the oscillating region decreases (Figure 8.12). The offset a modifies the lower threshold value of s, i.e., the lower bifurcation point b_1, above which the FN oscillates. To ensure that every isolated incoming forager elicits at least one outgoing forager, given k, which is the increase in s for an interaction with an incoming forager, we choose a such that $k > b_1$ for all $c \in [0, 5]$. We choose k and τ to produce sensible values of s for the range of incoming foraging rates observed in the data. One of the strengths of the model is that, while there is some flexibility in selecting a, k, and τ such that they satisfy these desired conditions, the qualitative behavior of the system is not affected by the specific values selected.

For very low $\bar{r}_{in}$, $\bar{r}_{out}$ is low because s is low and the FN system remains in the resting state with occasional short-lasting periods of oscillatory behavior (Figure 8.2). For very high $\bar{r}_{in}$, $\bar{r}_{out}$ is also low because s is high and the FN system remains most of the time in the saturated state. In contrast, $\bar{r}_{out}$ is high for $\bar{r}_{in}$ that yields an s that keeps the FN system inside the oscillating region. In the oscillating region, $\bar{r}_{out}$ is equal to the frequency of the oscillations, which is inversely proportional to the volatility c as we show in Section 8.5.1.

To get an expression for the natural frequency of the oscillations in the FN, we compute an approximation for its period $T_{LC}(s, c)$ that uses the time-scale separation between the dynamics of v and u, see Section 8.5.1 and Figure 8.13. Under the assumption of a Poisson incoming rate, the process s is ergodic (see Section 8.5.3). Thus, over sufficiently long periods of time, suitable time statistics converge to ensemble statistics, allowing us to approximate the fraction of time that s spends in the oscillating region using $p(s, \bar{r}_{in})$, the probability density function of s at steady-state. We compute $p(s, \bar{r}_{in})$ in Section 8.5.4 as a piecewise function where the piecewise ele-

ments satisfy recurrence equations and depend on k and τ. From this we can construct an analytical expression for $\bar{r}_{out}$ as a function of both $\bar{r}_{in}$ and c (see Section 8.5.3):

$$\bar{r}_{out} = \int_{b_1(c)}^{b_2(c)} \frac{p(s, \bar{r}_{in})}{T_{LC}(s, c)} \, \mathrm{d}s. \tag{8.8}$$

In Figure 8.6A we plot $\bar{r}_{out}$ versus $\bar{r}_{in}$ using (8.8) for different values of c. The resulting open-loop input-output curves, which we call *nest I/O curves* show that the analytic mapping from $\bar{r}_{in}$ to $\bar{r}_{out}$ depends nonlinearly on c. The increasing steepness of the curve at low $\bar{r}_{in}$ becomes more pronounced for higher c because the frequency of oscillations is proportional to c. Similarly, the decreasing steepness of the curves for high $\bar{r}_{in}$ also becomes more pronounced for higher c. This is because as c increases b_2 decreases, causing the FN to saturate at lower $\bar{r}_{in}$ values. The maximum value of $\bar{r}_{out}$ takes place at the $\bar{r}_{in}$ that yields an s that keeps the FN system inside the oscillating region. Because of this, the maximum $\bar{r}_{out}$ must be less than or equal to the natural frequency of the oscillations at the given value of c.

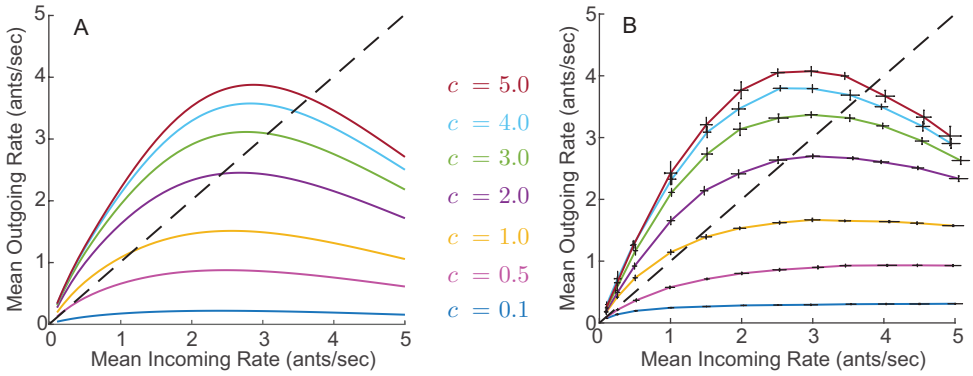

Figure 8.6: A) Analytical approximations for the nest I/O curves. B) Simulated nest I/O curves for different values of c. Each pair of error bars correspond to 10 simulation trials, each 5 minutes long, with a constant expected incoming rate and constant volatility c. The dashed black line represents points at which the mean incoming rate $\bar{r}_{in}$ is equal to the mean outgoing rate $\bar{r}_{out}$.

In Figure 8.6B we show the nest I/O curves obtained by simulating the open-loop system for different constant Poisson incoming rates at a fixed volatility. We

measured the resulting mean outgoing rate in each case. We set λ_{in} to a five-minutes-long Poisson process, and, in each of 10 simulation trials, we recorded the output λ_{out}. We computed the mean outgoing foraging rate $\bar{r}_{out}$ by dividing the total number of outgoing foragers in the trial by the 300 seconds that the trial lasted. We used the average of the 10 trials as a point estimate for $\bar{r}_{out}$ as a function of $\bar{r}_{in}$ given the volatility parameter c. We constructed nest I/O curves by repeating this point estimation process for twelve different values of $\bar{r}_{in} \in [0.1, 5]$ while keeping c constant.

The simulated I/O curves in Figure 8.6B are in good agreement with the analytical I/O curves in Figure 8.6A. The simulation curves are slightly higher than the analytical curves because T_{LC} underestimates the period of the FN oscillations (see Figure 8.13) and because (8.8) relies on s being ergodic (see Section 8.5.3). The fact that there is good agreement between the simulation curves computed from 5-minute-long input sequences and the analytical curves derived under the assumption of an infinite time period suggests that time statistics of $s(t)$ converge rapidly to ensemble statistics. This rapid convergence indicates that our analytical approximation is also valid across short timescales. We make use of this in our analysis of the closed-loop model dynamics.

The points at which the nest I/O curves in Figure 8.6 intersect the black dashed diagonal line correspond to $\bar{r}_{in} = \bar{r}_{out}$, which are predictive of the (quasi) steady-state solutions at an equal incoming and outgoing foraging rate observed in the data. Figure 8.6 suggests that for sufficiently high values of c, the equal foraging rate is positive and bounded away from zero, capturing a nontrivial steady-state foraging rate as in Figure 8.4 and Figure 8.5B and D. However, Figure 8.6 suggests that for low values of c, the equal foraging rate is nearly zero, capturing a steady-state with negligible foraging as in Figure 8.5F.

To understand how c affects the equal foraging rate, consider that, because $k > b_1$, every isolated incoming forager elicits at least one oscillation in the FN output and

so at low $\bar{r}_{in}$, $\bar{r}_{out}$ is equal to or larger than $\bar{r}_{in}$. At high c values, the frequency of oscillations in the FN is high and an isolated incoming forager will elicit more than one oscillation, resulting in nest I/O curves with an initial slope higher than one and an intersection with the diagonal line at a single point away from the origin. In contrast, for low c values, an isolated incoming forager will elicit exactly one oscillation, resulting in nest I/O curves with an initial slope of one, i.e., the curve lies on the diagonal line close to the origin and intersects nowhere else.

This analysis implies the existence of a critical value c^* such that if $c > c^*$, there is an isolated steady-state solution away from the origin, and if $c < c^*$, the steady-state solution remains close to the origin. We find an upper bound $\hat{c} > c^*$ so that $c > \hat{c}$ is sufficient for the existence of an isolated steady-state solution away from the origin. For $b_1 < k < b_2$, it can be shown that the number of oscillations caused by a single incoming forager is at most $(-\tau \log b_1/k)/T_{LC}$. We numerically solved this expression for c using the asymptotic expansion of T_{LC} described in Section 8.5.1 and found that for $c > \hat{c} = 0.5287$ the FN oscillates at least two times per every incoming forager. Therefore, $c > \hat{c}$ is a sufficient condition for the nest I/O curve to lie above the diagonal line at low $\bar{r}_{in}$ and to intersect the diagonal line at an isolated point, corresponding to a nontrivial steady-state foraging rate.

Foraging Dynamics Outside the Nest

Given a sequence of outgoing foragers λ_{out} with rate r_{out}, the foraging dynamics outside the nest predict a corresponding delayed sequence of incoming foragers λ_{in} with rate r_{in}. We use results from queueing theory to find analytic expressions relating r_{out} to r_{in} and the expected number of active foragers outside the nest.

To facilitate the analysis we assume that λ_{out} is a non-homogeneous Poisson process (i.e., a Poisson process with time-varying rate) [40]. Applying known results for

queues with a non-homogeneous Poisson distribution of arrival times [110] we obtain the following three results:

1. Let $Q(t)$ represent the number of active foragers outside the nest, then, for each time $t' = t/60$ where t is seconds, $Q(t')$ has a Poisson distribution with mean

$$\mathbb{E}[Q(t')] = \int_0^\infty r_{out}(t' - x)(1 - F(x, D)) \, \mathrm{d}x. \tag{8.9}$$

2. The output process describing how foragers leave the queueing system, that is, the process λ_{in} describing how foragers return to the nest, is a non-homogeneous Poisson process with mean

$$\mathbb{E}[\lambda_{in}(t)] = \int_0^t r_{in}(x) \, \mathrm{d}x. \tag{8.10}$$

3. r_{in} is related to r_{out} by

$$r_{in}(t') = \int_0^\infty r_{out}(t' - x) \, \mathrm{d}F(x, D) = \mathbb{E}[r_{out}(t' - X)]. \tag{8.11}$$

(8.9) and (8.7) show how the number of active foragers outside the nest depends on the history of outgoing foragers. (8.10) shows that if the outgoing foraging process is a non-homogeneous Poisson process, then the incoming foraging process is also a non-homogeneous Poisson process. And (8.11) shows how the incoming foraging rate r_{in} depends on the history of the outgoing foraging rate r_{out}.

At steady-state, the outgoing foraging rate is constant, i.e., $r_{out}(t) = r^*$, and (8.11) reduces to $r_{in} = r_{out} = r^*$, i.e., the incoming foraging rate is also constant and equal to the outgoing foraging rate. Moreover, (8.9) reduces to

$$\mathbb{E}[Q] = r_{out}\mathbb{E}[X] = r^* D, \tag{8.12}$$

i.e., the mean number of active foragers outside the nest is given by the steady-state foraging rate r^* multiplied by the average foraging trip time D.

The relaxation time for the queue output process to reach steady-state can be analyzed by considering the step-function arrival rate $r_{out}(t') = 0$ for $t' < 0$ and $r_{out}(t') = r^*$ for $t' \geq 0$. Then, from (8.11), $r_{in}(t') = r^* \int_0^{t'} dF(x, D) = r^* F(t', D)$ for $t' \geq 0$. The difference between the queue input and output rates as a function of time is

$$||r_{out}(t') - r_{in}(t')|| = r^* \frac{\Gamma(D/2, t'/2)}{\Gamma(t'/2)} \tag{8.13}$$

for $t' \geq 0$. To illustrate, we compute for $D = 2$ that the right-hand side of (8.13) simplifies to $r^* e^{-t'/2}$ and the foraging queue converges exponentially in time towards a steady-state where the input and output rates are equal.

Closed-loop Model Dynamics

In our model, outgoing foragers from the nest go out to forage, return to the nest as incoming foragers after finding a seed, and then go back out to forage again if sufficiently excited (Figure 8.1). Here we show that adding the feedback connection from outgoing to incoming foragers to the open-loop dynamics in the nest yields long-term dynamics with a stable and attracting equilibrium where the incoming and outgoing rates are equal. Stability of an isolated equilibrium implies robustness: the steady-state equilibrated foraging rate is maintained even in the presence of disturbances, e.g., small changes in the rates of incoming foragers. When the volatility $c > c^*$, the critical value defined earlier, the steady-state foraging rate is nontrivial, whereas if $c < c^*$, the steady-state foraging rate is negligible.

For the dynamics inside the nest, we have shown that c parametrizes a family of nest I/O curves, described by (8.8), which map $\bar{r}_{in}$ to $\bar{r}_{out}$ across short timescales of a few minutes. For $c < c^*$, the nest I/O curve always has slope less than or equal to 1, such that it lies on or below the diagonal line where $r_{in} = r_{out}$. For $c > c^*$, the

nest I/O curve has initial slope greater than 1 and then intersects the diagonal line $r_{in} = r_{out}$ at a nontrivial point. For the dynamics outside the nest, we have shown that the mapping from $r_{out}(t)$ to $r_{in}(t)$ is described across longer timescales of tens of minutes by a time delay given by (8.11).

We study the closed-loop model dynamics for timescales ranging from tens of minutes to hours by investigating the behavior of a discrete iterated mapping $r_n = G_c(r_{n-1})$ where r_n represents the mean foraging rate at time step n and r_{n-1} represents the mean foraging rate at time step $n-1$. We can interpret r_n and r_{n-1} as mean incoming rate or mean outgoing foraging rate since the mean incoming rate becomes the mean outgoing rate after a time delay. The mapping $G_c : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is defined by the c-dependent nest I/O curves shown in Figure 8.6. G_c describes the process by which the incoming foraging rate becomes the outgoing foraging rate through the dynamics of forager activation inside the nest, which then becomes the incoming foraging rate at a later time.

When G_c lies below (above) the diagonal line where $r_{in} = r_{out}$, the average number of outgoing foragers per every incoming forager is less (greater) than one, and iterations of G_c decrease (increase) r (Figure 8.7). For $c > c^*$, G_c has one unstable fixed point at the origin and one attractive stable fixed point where $r_{in} = r_{out}$. For $c < c^*$, G_c has a small interval of fixed points close to the origin. Thus, the closed-loop model dynamics evolve in time towards either a finite steady-state foraging rate $r_{in} = r_{out} = r^*$ (Figure 8.7, $c = 2$ and $c = 5$) or to negligible foraging (Figure 8.7, $c = 0.1$).

The stability of the steady-state equilibrated foraging rate and the implications for robustness result from the balance between positive feedback from incoming ants activating a larger number of outgoing ants, and negative feedback from saturation effects. The magnitude and variance of the steady-state foraging rate increase with c. The magnitude also depend on k and τ, as these affect s in (8.8), which can be

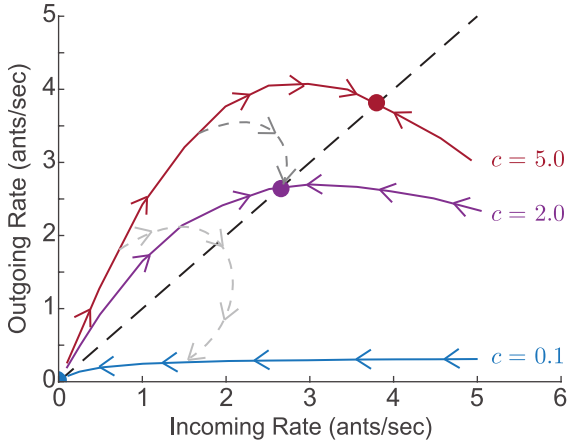

Figure 8.7: Model dynamics illustrating response of foraging rates to environmental conditions. Red, purple, and blue curves show closed-loop trajectories of $r_{out}(t)$ versus $r_{in}(t)$ for fixed volatility c equal to 5.0, 2.0, and 0.1, respectively. Initially, all available foragers are uninformed about the environment and have volatility $c_u = 5.0$. The darker gray dashed curve shows the dynamics in the case when foragers exposed to the environment reduce their volatility to $c_i = 2.0$, as might happen on a moderately hot and dry day. The lighter gray dashed curve shows the dynamics in the case when foragers exposed to the environment reduce their volatility to $c_i = 0.1$, as might happen on a very hot and dry day.

numerically solved to find how the magnitude changes with c (Figure 8.8A). We refer to the steady-steady foraging rate as the QSS.

As shown in Figure 8.8B, simulations of the closed-loop model validate the predictions of the iterated mapping model (Figure 8.7). We initialize the foraging dynamics by setting λ_{in} from $t = 0$ to $t = 60 \times (D+1)$ seconds to be equal to the initial sequence of incoming foragers for Colony 859 on August 20, 2017, which has the very low mean incoming rate of 0.01 ants/sec during the first 15 minutes (Figure 8.11 C). Using the additional minute, i.e., $D + 1$ instead of D, for the initial sequence of incoming foragers helps ensure that the sequence of incoming foragers does not abruptly end before the first few outgoing foragers return to the nest.

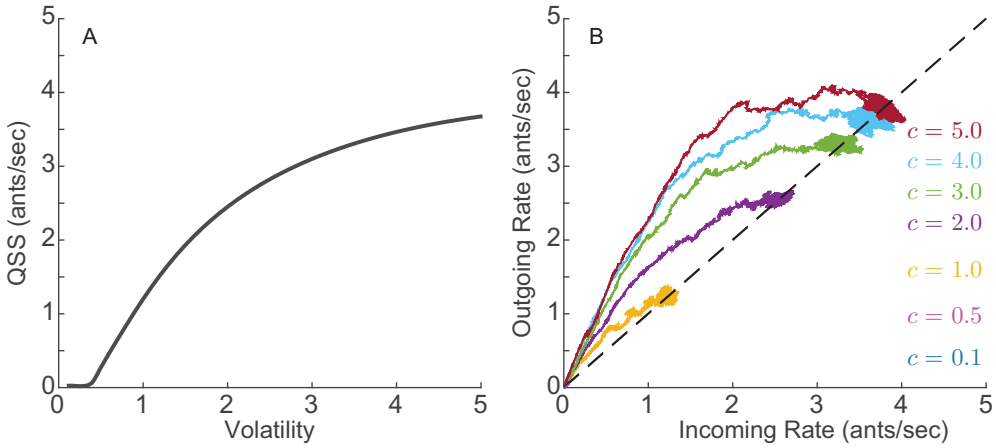

Figure 8.8: A. Analytical magnitude of the quasi steady-state (QSS) foraging rate obtained from numerically solving (8.8). B. Closed-loop model simulations for 7 different values of volatility c. The initial sequence of incoming foragers for all simulations was set equal to the sequence of incoming foragers recorded during the first 11 minutes for Colony 859 on August 20, 2017 which has a mean incoming rate of 0.01 ants/sec. The total time for all simulations was 3 hours. The mean foraging time was set to 10 minutes ($D = 10$).

Closed-loop Dynamics with Response to Environmental Conditions

For Poisson sequences of incoming foragers, the mean outgoing foraging rate of the colony is given as the weighted sum of the outputs of the uninformed and informed:

$$\bar{r}_{out} = x_u \int_{b_1(c_u)}^{b_2(c_u)} \frac{p(s, \bar{r}_{in})}{T_{LC}(s, c_u)}\, \mathrm{d}s + x_i \int_{b_1(c_i)}^{b_2(c_i)} \frac{p(s, \bar{r}_{in})}{T_{LC}(s, c_i)}\, \mathrm{d}s. \tag{8.14}$$

The closed-loop dynamics can still be studied as an iterated mapping, but we allow the mapping to evolve in time, $G_c = G_c(t)$, from an initial mapping $G_c(t_0) = G_{c_u}$ with volatility c_u to a final mapping $G_c(\infty) = G_{c_i}$ with volatility c_i. The dark and light gray curves in Figure 8.7 provide an illustration for how the map $G_c(t)$ changes with time when $c_u = 5.0$, and $c_i = 2.0$ or $c_i = 0.1$. The dynamics first evolve along G_{c_u} (red), but as x_i increases, the dynamics shift increasingly to G_{c_i}, and the trajectory on the plot of $r_{out}(t)$ versus $r_{in}(t)$ moves towards the c_i curve. In the case $c_i = 2.0$, the trajectory converges to the fixed point associated with $c = 2.0$ (darker gray dashed

curve). In the case $c_i = 0.1$, the trajectory converges to the only fixed point of G_{c_i}, which is the origin, leading to a cessation of foraging (lighter gray dashed curve).

Figure 8.9 shows the resulting time-series and input-output plots for three different simulations of the model with the mechanism for response to environmental conditions. The simulations are distinguished by the set of four parameters: c_u, c_i, N, and D. The simulated trajectories qualitatively resemble the trajectories from the field observations shown in Figure 8.5. We set the initial sequence of incoming foragers as in Figure 8.8B.

Figure 8.9A and B show the results for $c_u = 3$, $c_i = 0.9$, $N = 500$, and $D = 5$. In this case, c_u is much higher than c_i, leading to a system with an overshoot behavior in which the outgoing foraging rate increases more rapidly than the incoming rate and then decreases before settling around a steady-state where the rates are approximately equal to 0.7 ants/sec. This is qualitatively similar to the observations of Colony 664 on August 27, 2015 of Figure 8.5A and B. The result of a relatively small number of total foragers N and short mean foraging time D is that the fraction of informed foragers increases rapidly, leading to a quick convergence towards the steady-state. The net number of foragers outside the nest at steady-state fluctuates with low variability at around 230, close to the prediction given by (8.12).

Figure 8.9C and D show the results for $c_u = 3$, $c_i = 0.75$, $N = 200$, and $D = 5$. This case simulates the same colony as in Figure 8.9A and B but on a hotter and drier day, when the total number of ants N that engage in foraging may be reduced and the volatility of the informed ants c_i may be reduced. The overshoot behavior is followed by the foraging rates settling around a steady-state of about 0.25 ants/sec. The result of low values of N and D is that the fraction of informed foragers increases very rapidly, leading to a rapid convergence towards the steady-state. This is qualitatively similar to the observations of Colony 664 on August 31, 2015 of Figure 8.5C and D.

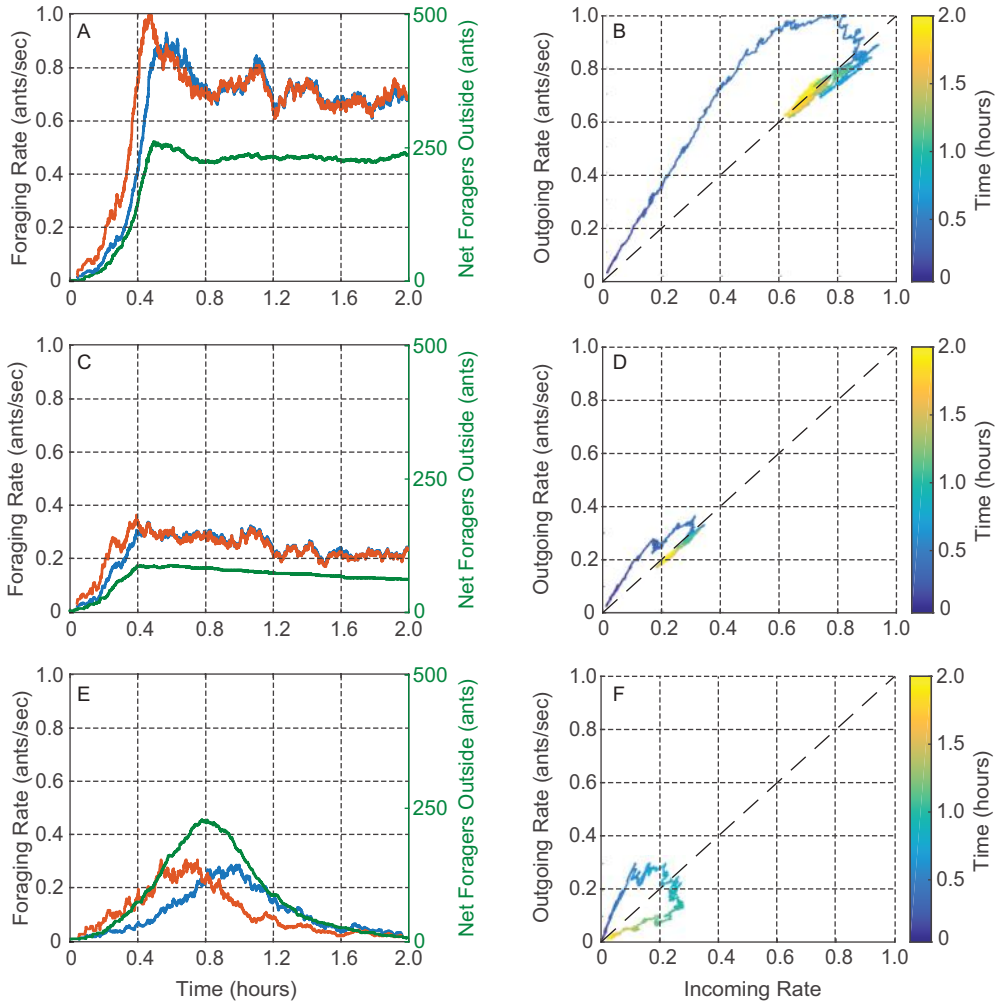

Figure 8.9: Simulations of the closed-loop model with the adaptation mechanism. Plots are of the same form as in Figure 8.5, and qualitative comparisons can be made between A and B here and Figure 8.5A and B, between C and D here and Figure 8.5C and D, and between E and F here and Figure 8.5E and F. A) and B) $c_u = 3$, $c_i = 0.9$, $N = 500$, $D = 5$. C) and D) $c_u = 3$, $c_i = 0.75$, $N = 200$, $D = 5$. E) and F) $c_u = 5$, $c_i = 0.02$, $N = 600$, $D = 15$.

Figure 8.9E and F show the simulation results for $c_u = 5$, $c_i = 0.02$, $N = 600$, $D = 15$. In this case, c_i is close to zero, leading to a colony that goes out to forage but then returns to the nest without sustained foraging. The result of the long mean foraging time D is that the fraction of informed foragers increases at a slow rate, leading to longer lasting transient dynamics towards the steady-state. This is qualitatively similar to the observations of Colony 664 on August 27, 2015 of Figure 8.5E and F.

The time it takes for the colony to transition from fully uninformed to fully informed about outside conditions is dictated by c_u, c_i, D, N, and the initial conditions for r_{in} and r_{out}. Low values of c_u result in initially low outgoing foraging rates, so that the corresponding rate at which foragers become informed is low too, even if c_i is high (Figure 8.14 A). Low values of c_i can cause long transients, because once a critical number of foragers has become informed, low volatility makes it difficult for the remaining foragers to become informed. High D and N can also result in long transients because the time it takes for the transition to a fully informed state depends on the number of available foragers and on how long it takes for informed foragers to return to the nest (Figure 8.14 B). Finally, initially high values of r_{out} produce a rapidly increasing number of active foragers, reducing the time it takes to reach the informed state with foraging rates that reach a QSS (Figure 8.14 C). Qualitative comparisons can be made between the simulations in Figure 8.14 A, B, and C, and the data in Figure 8.11 C, D, and, E and F, respectively.

8.4 Discussion

We have derived and analyzed a low-dimensional analytical model of foraging dynamics that requires only a small number of parameters to qualitatively capture a wide range of transient and steady-state features observed in the foraging rates of red harvester ant colonies. Our model extends previous work by using feedback at multiple timescales to account for how foraging rates to and from the nest change over long timescales, from tens of minutes to hours.

Importantly, the long timescales allow for a model-based investigation into how a colony, with no centralized control and little individual information about the state of the colony or environment, can stably regulate its foraging rates and be responsive to temperature and humidity outside the nest across minute-to-hour timescales. Sta-

bility implies robustness of the steady-state foraging rate to small disturbances, e.g., small changes in the rate of incoming foragers. Further, because the model is analytically tractable, it can be used to systematically derive empirically testable predictions of foraging behavior as a function of critical model parameters, including number of foragers N, mean foraging trip time D, and volatility c. In our model, these parameters determine the steady-state foraging rate, independent of initial foraging rates. The transient and convergence time to the steady-state, however, do depend on initial rates; higher initial rates lead to faster convergence. The model suggests that a change in volatility as the foragers become exposed to the temperature and humidity outside the nest can account for the observed foraging behavior under different environmental conditions. Further, the model suggests that differences among colonies in volatility, in response to temperature and humidity, can produce the observed variation among colonies in the regulation of foraging.

Our model and analysis highlight the importance of feedback across multiple timescales in the regulation of foraging activity. Previous work isolates the open-loop dynamics inside the nest, which maps incoming ants to outgoing ants on very short timescales. We address the minute-to-hour timescales by examining analytically the closed-loop dynamics that connect the slow foraging activity outside the nest to the fast activation of foragers inside the nest through feedback generated by the ants themselves and their interactions with others. The stream of foraging ants out of the nest is the input to the foraging activity, and the output of the foraging activity is the stream of foraging ants into the nest. The effective volatility of the colony also changes in the model at the timescale of minutes to hours, as foragers leave the nest for the first time and become exposed to the outside humidity and temperature, yielding flexibility in the regulation of foraging activity at minute-to-hour timescales.

In the model, volatility c approximates the average sensitivity of available foragers in the nest to interactions with incoming foragers: the higher the c the fewer inter-

actions needed to activate available foragers to go out and forage. The relationship between c and the activation of foragers is nonlinear, and the subtleties of our model reflect some of the complexities of the system. We use analytical predictions to show how c determines three important features of the foraging model dynamics (see Figure 8.8): 1) the initial transient in incoming and outgoing foraging rates, parametrized by c, 2) the equilibration of incoming and outgoing foraging rates to a stable, and thus robust, quasi steady-state rate, parametrized by c, and 3) the prediction of an early cessation of foraging or no foraging at all if $c < c^*$, a critical volatility value c^*.

The behavior of different colonies on the same day or the same colony on different days correspond in the model to different values of c. Lower values of c result in model dynamics that are consistent with data for hotter and drier days, because lower c means that available foragers are less volatile and thus less likely to go out and forage. Higher values of c result in model dynamics that are consistent with data for cooler and more humid days, because higher c means that available foragers are more volatile and thus more likely to go out and forage. The model distinguishes the volatility c_u of available foragers in the nest who have yet to to go on a foraging trip from the volatility c_i of available foragers in the nest who have already been outside the nest and been exposed to the environment (Figure 8.3). The result is a transition from the foraging activity of ants with volatility c_u to the foraging activity of ants with volatility c_i, which can last from minutes to hours as each of the total N ants goes out at a different time on its first foraging trip and returns to the nest after foraging for an average of D minutes (Figure 8.7).

Differences among colonies in the values of parameters c_u, c_i, N, and D could lead to the differences among colonies in foraging behavior that we observe. Indeed, over a range of values for the four parameters c_u, c_i, N, and D, the model describes the range of transient and quasi steady-state foraging rate behavior observed in the data collected for red harvester colonies in August and September of 2015, 2016,

and 2017. The model thus suggests hypotheses about the physiological processes that would lead to different parameter values, such as differences among colonies in how outgoing foragers respond to interactions with returning foragers, and differences among colonies in how foragers respond to conditions such as humidity.

The model represents the case in which foragers make an adjustment to their volatility only after their first foraging trip. To include more variability within a colony the model could be generalized to $M > 2$ groups of available foragers in the nest, distinguished by M values of volatility $c_1, \ldots, c_M$. For example, the generalization could be used to account for foragers that make adjustments to how they respond to interactions in the nest after subsequent foraging trips due to repeated exposure or changing temperature and humidity. The generalization could also be used to account for decay of information for those foragers who stay in the nest for a long period after a foraging trip, or to represent foragers that return to the deeper nest after exposure to hot and dry outside conditions.

Foraging models that consider the regulation of foraging activity tend to fall into two categories: multi-agent models that keep track of every individual [150–152], and compartmental models that keep track of the time evolution of fractions of individuals engaged in a specific task [67, 153]. Multi-agent models allow for a detailed modeling of foraging dynamics, often relying on simulations due to their complexity and poor analytical tractability. In contrast, compartmental models provide high tractability in many cases but assume very large group sizes which affect predictions when the group size is small. Our model considers two idealized processes, the activation of foragers inside the nest and the collection of seeds outside the nest, to generate a dynamical system with a small number of equations and parameters. The model accommodates any group size and retains sufficient tractability to generate predictions on the impact of critical model parameters.

Our biologically informed, low-dimensional, and simply parameterized model allows for systematic exploration of mechanisms and sensitivities that can explain collective behavior and guide further theoretical and experimental investigations. Our use of well-studied excitability dynamics opens the way for comparison with other complex systems, such as neuronal networks, that are driven by excitable dynamics. The model together with our analysis based on dynamics and control theory contribute to a better understanding of the role of feedback across multiple timescales in collective behavior.

Acknowledgments

We thank many dedicated undergraduate field assistants for their help: Sam Crow, Eleanor Glockner, Christopher Jackson, Sarah Jiang, Ga-Il Lee, Arthur Mestas, Becca Nelson, and especially Rebia Khan.

8.5 Supporting Information

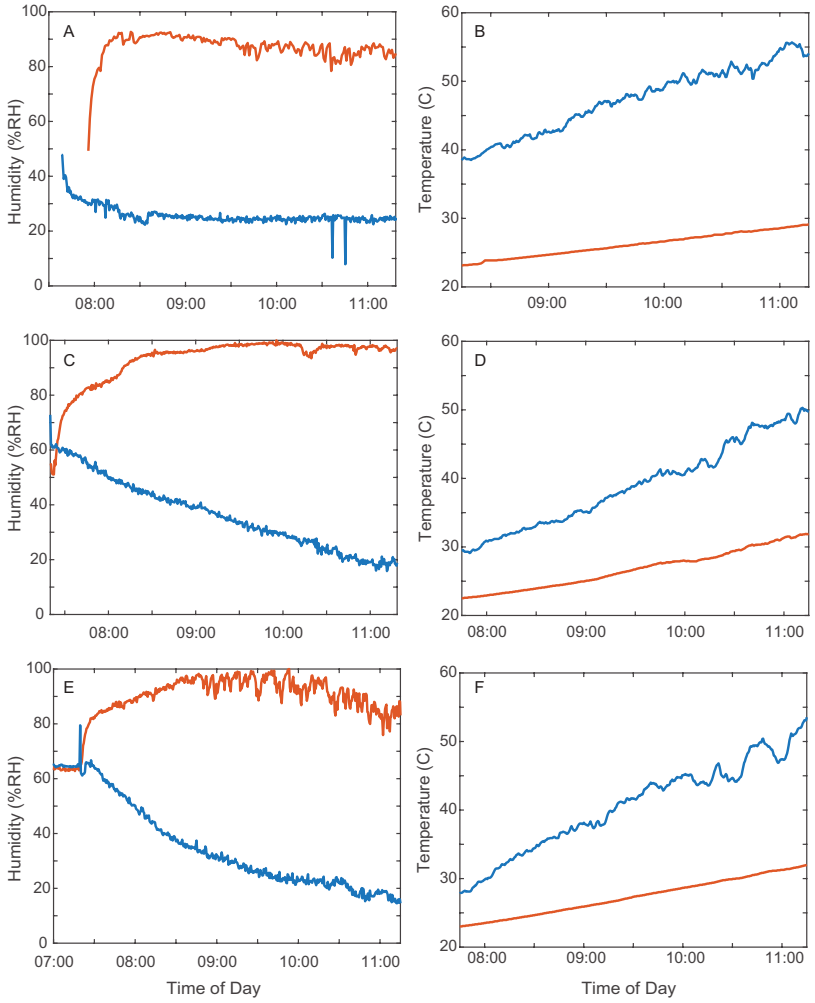

Figure 8.10: Humidity and temperature readings recorded on the surface of the desert soil (blue) and inside the nest entrance chamber (red). Temperature and humidity ibutton sensors were placed outside but close to the nest entrance on the desert soil, unshaded, and inside in the nest in an excavated hole, which had been uncovered by excavation and then covered with glass on top and shaded. The humidity and temperature outside the nest changed significantly throughout the morning hours while the humidity and temperature inside the nest entrance chamber remained relatively constant. The measured moderate rise in temperature inside the nest is likely due to the light coming into the nest entrance chamber through the glass. A) Humidity on August 29, 2014 (Colony E). B) Temperature on August 29, 2014 (Colony E). C) Humidity on August 31, 2015 (Colony 10). D) Temperature on August 31, 2015 (Colony 10). E) Humidity on September 1, 2015 (Colony 10). F) Temperature on September 1, 2015 (Colony 10).

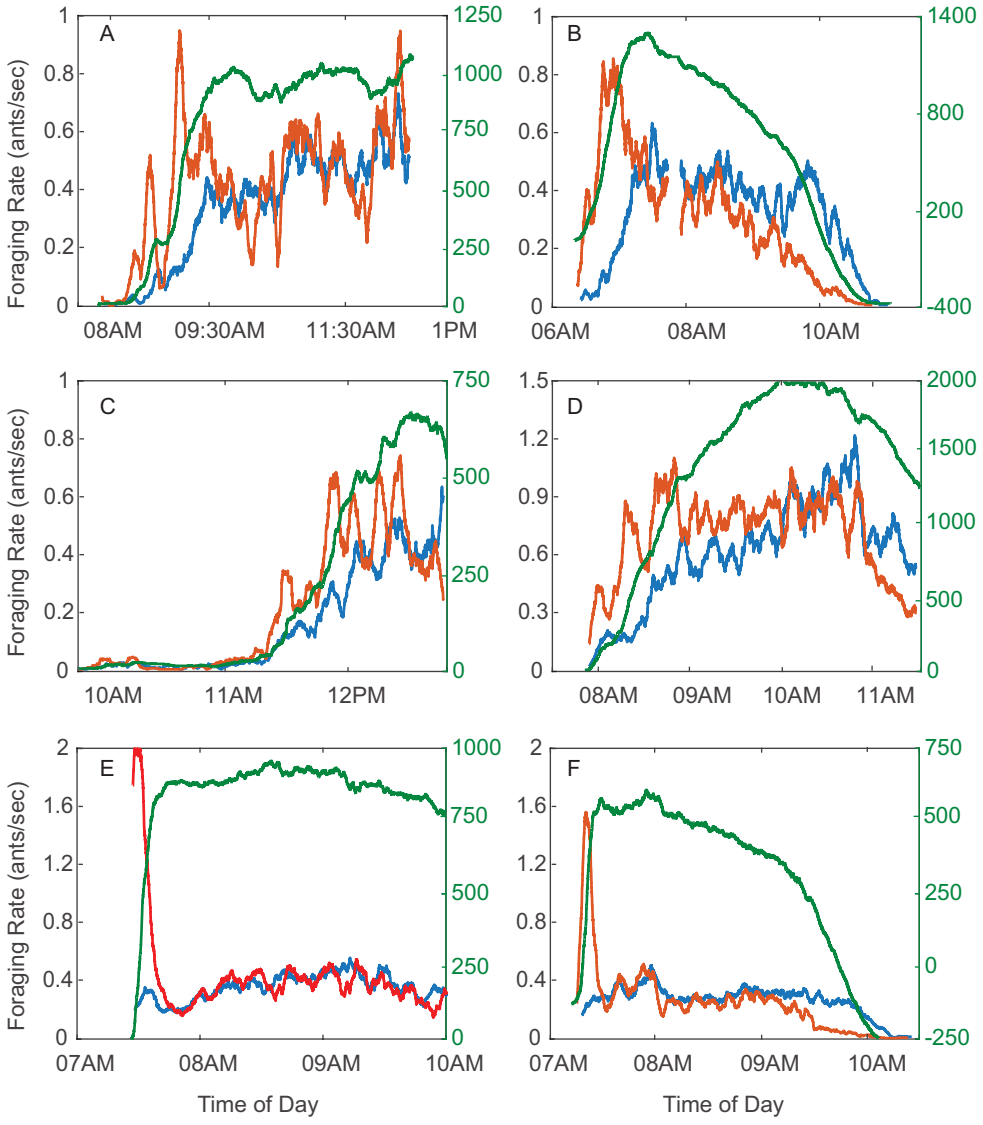

Figure 8.11: Additional field observations of foraging rates. Incoming foraging rate r_{in} (blue), outgoing foraging rate r_{out} (red), and difference between number of incoming and outgoing foragers (green) versus time of day. A) Colony 863 September 5, 2015 reached a QSS at a high rate; compare to Figure 8.5E when on the much hotter and drier day, September 1, 2015, Colony 863 returned to the nest early. B) Colony D19 August 08, 2016 returned to the nest early; the day was very hot and dry. C) Colony 859 August 20, 2017; the transient started late in the morning. The day was cool and humid. D) Colony 1107 August 16, 2017; the transient was slow. The day was dry. E) Colony 1017 August 23, 2016; the initial transient was more like a burst of outgoing foragers. The day was dry. F) Colony 1015 August 18, 2016; another initial burst of outgoing foragers. The day was very dry.

		Daily Average		At 11 am	
Figure	**Date**	**Temperature**	**Humidity**	**Temperature**	**Humidity**
3A, 3B	Aug. 20, 2016	25.9C°	57%	24.8C°	57%
4A, 4B	Aug. 27, 2015	25.3C°	58%	27.5C°	52%
4C, 4D	Aug. 31, 2015	26.8C°	53%	28.8C°	45%
4E, 4F	Sept. 1, 2015	25.2C°	53%	27.5C°	52%
S1A	Sept. 5, 2015	22.6C°	77%	23.3C°	77%
S1B	Aug. 8, 2016	29.7C°	48%	39.9C°	43%
S1C	Aug. 20, 2017	23.0C°	71%	22.7C°	73%
S1D	Aug. 16, 2017	26.0C°	48%	27.4C°	41%
S1E	Aug. 23, 2016	24.1C°	43%	28.8C°	36%
S1F	Aug. 18, 2016	25.5C°	27%	31.2C°	23%

Table 8.1: Average temperature, average relative humidity, temperature at 11 am, and relative humidity at 11 am in Rodeo, New Mexico, USA for days with data plotted in Figures 8.4, 8.5, and Figure 8.11. Data collected by the Citizen Weather Observer Program station E8703 and accessed through Weather Underground [168]. The station is located 1.7 miles from the study site.

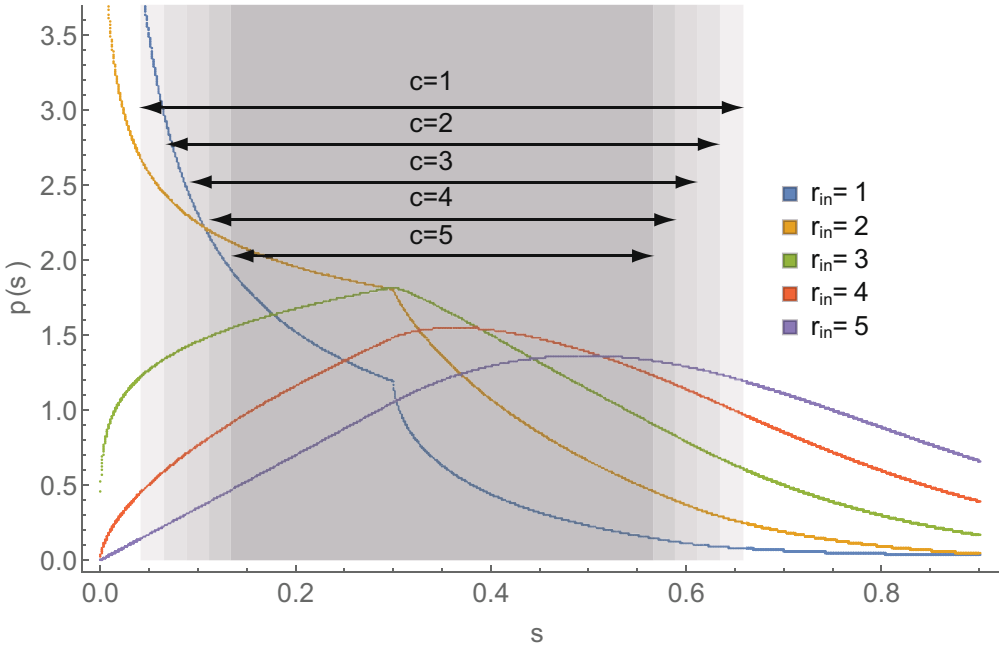

Figure 8.12: Probability density function for the stimulus function. Each curve represents the PDF p of the stimulus function s for different values of incoming rate r_{in}. The gray rectangles represent the size of the oscillatory region in the FN system (b_1, b_2) for $a = 0.35$ and different values of volatility c. For all curves, $k = 0.3$, $\tau = 0.41$.

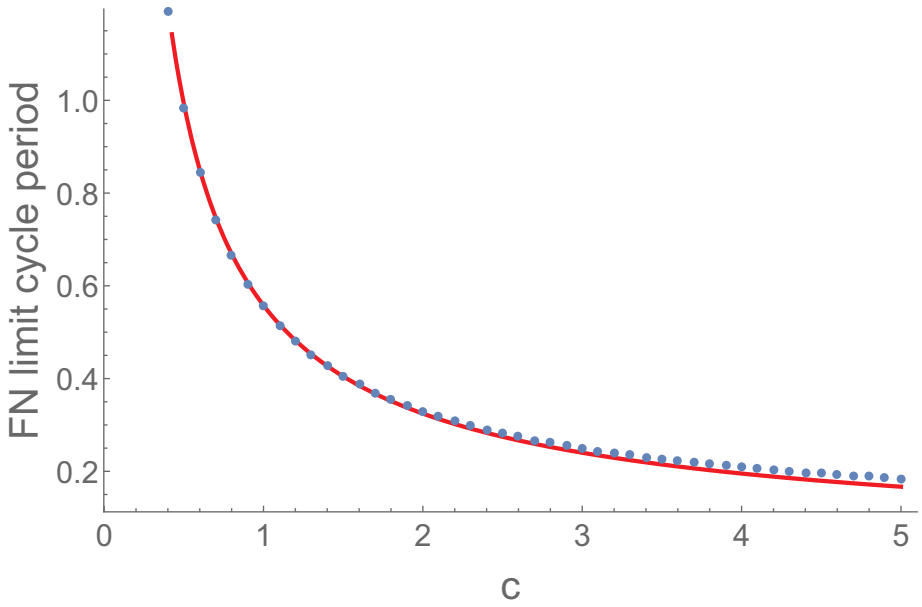

Figure 8.13: Period of FN Limit Cycle when s=0.35. Blue dots represent numerical simulations for the period of the FN limit cycle. The red curve represents the analytical approximation in S1 Text. In both cases we set $s = 0.35$.

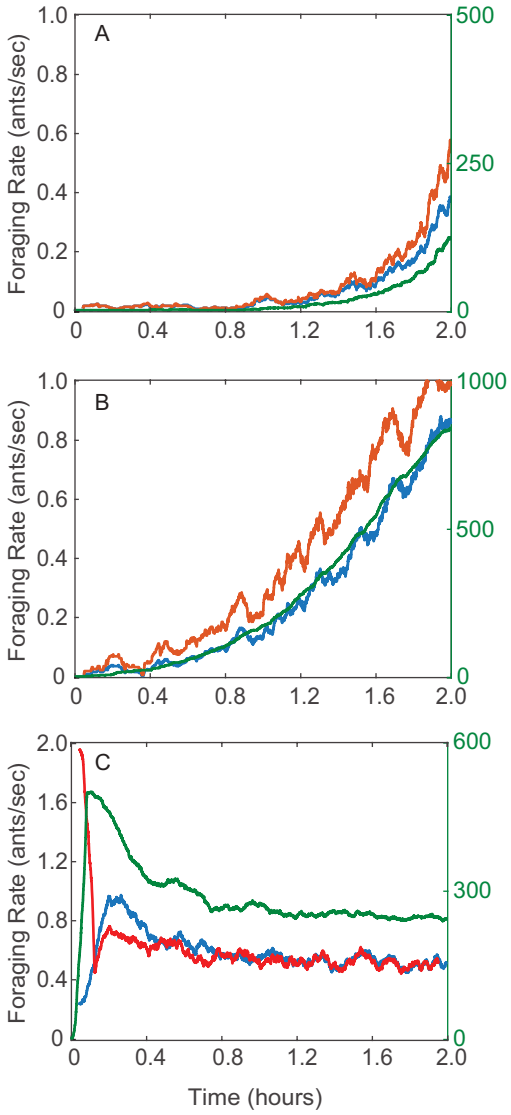

Figure 8.14: Additional simulations of the closed-loop system with the adaptation mechanism. Plots resemble observed foraging behaviors in Figure 8.11. Qualitative comparisons can be made between A here and Figure 8.11C, between B here and Figure 8.11D, and between C here and in Figure 8.11E and Figure 8.11F. A) $c_u = 0.9$, $c_i = 2.2$, $N = 500$, $D = 5$. Setting $c_u < c_i$ where c_u is close to c^* results in a long period before the rates ramp up. B) $c_u = 1$, $c_i = 1$, $N = 1000$, $D = 15$. Setting the mean foraging trip time D to be large results in long lasting transients. C) $c_u = 0.7$, $c_i = 0.9$, $N = 1000$, $D = 7$. Setting the initial λ_{in} equal to the sequence from the first 5 minutes of λ_{in} for Colony 1017 on Aug. 23, 2016 yields the behavior shown in Figure 8.11E and Figure 8.11F that follows an initial burst of foragers.

8.5.1 Effect of Volatility on the Frequency of Oscillations in the FN

Here we use results from [1] Chapter III, Theorem 3 to obtain an asymptotic expansion for the period of the limit cycle T_{LC} as ϵ_2 goes to zero. We show that the period of the FN limit cycle is inversely proportional to c by computing the leading term in the expansion.

The limit cycle of the FN is comprised of four components: two fast components that stretch along the v direction between the crest and valley of the cubic nullcline, and two slow components that stretch along the sides of the cubic nullcline. Because it takes much longer to traverse the slow components of the limit cycle than the fast components of the cycle, the period can be approximated by the time it takes trajectories to traverse the two slow components. These slow components run from $v = -2$ to $v = -1$ and from $v = 2$ to $v = 1$, respectively, and are proportional to the length of the sides of the cubic nullcline, which we show are proportional to c.

Theorem 8. *The limit cycle of the FN system*

$$\epsilon_1\epsilon_2\dot{v} = v - v^3/3 - cu - 0.35 + s$$
$$\epsilon_1\dot{u} = v - cu$$

has the asymptotic representation:

$$T_{LC} = T_0 + C_1\epsilon_2^{2/3} + C_2\epsilon_2 \ln\frac{1}{\epsilon_2} + \mathcal{O}(\epsilon_2),$$

as $\epsilon_2 \to 0$, where C_1 and C_2 are constants and where

$$T_0 = \frac{3\epsilon_1}{c}\left\{\int_{-2}^{-1}\frac{(v^2-1)}{(3z-v^3)}dv + \int_{2}^{1}\frac{(v^2-1)}{(3z-v^3)}dv\right\}.$$

Proof. By Chapter III, Theorem 3 of [169], the limit cycle of the FN model has the asymptotic representation

$$T_{LC} = T_0 + C_1\epsilon_2^{2/3} + C_2\epsilon_2 \ln\frac{1}{\epsilon_2} + C_3\epsilon_2 + \mathcal{O}(\epsilon_2^{4/3}).$$

Or, equivalently,

$$T_{LC} = T_0 + C_1\epsilon_2^{2/3} + C_2\epsilon_2 \ln\frac{1}{\epsilon_2} + \mathcal{O}(\epsilon_2).$$

Let time be scaled by $1/\epsilon_1$ and let $z = s - 0.35$, leading to the new system

$$\epsilon_2\dot{v} = v - v^3/3 - cu + z$$
$$\dot{u} = v - cu$$

The critical manifold of this fast-slow system is $M_0 := \{(v,u) \in \mathbb{R}^2 | u = (v - v^3/3 + z)/c\}$. In the limit $\epsilon_2 \to 0$, the slow manifold is equal to the critical manifold. Let Ψ_0 denote the limit cycle in this limit.

Using the description of M_0 as a graph $u = h(v)$, the dynamics of the system on the slow flow can be written as

$$\dot{u} = v^3/3 - z.$$

We get a second expression for $\mathrm{d}u/\mathrm{d}t$ by differentiating M_0 with respect to t

$$\dot{u} = \frac{1 - v^2}{c}\dot{v}.$$

Equating the two expressions we obtain

$$\dot{v} = \frac{c(v^3 - 3z)}{3(1 - v^2)}.$$

Multiplying both sides by dt and integrating over Ψ_0,

$$T_0 = \frac{3}{c}\int_{\Psi_0} \frac{(1-v^2)}{(v^3-3z)}\mathrm{d}v.$$

In Ψ_0, the fast components of the orbit take place instantaneously and the time taken to complete the orbit is equal to the time taken to traverse the slow components. The slow components of the trajectory take place on the slow manifold between $v \in [-2,-1]$ and $v \in [1,2]$, yielding the expression

$$T_0 = \frac{3\epsilon_1}{c}\left\{\int_{-2}^{-1} \frac{(1-v^2)}{(v^3-3z)}dv + \int_{2}^{1} \frac{(1-v^2)}{(v^3-3z)}dv\right\}$$

where time has been scaled back to its original form. This expression is inversely proportional to c. Furthermore, this integral has a short closed form solution when the slow components are symmetric (i.e. $z = 0$, or $s = 0.35$),

$$T_0 = 2\,\frac{3\epsilon_1}{c}\int_{-2}^{-1} \frac{(1-v^2)}{v^3}dv = \frac{\epsilon_1}{4c}(-9+8\log 8).$$

We compute the constants C_1 and C_2 by applying formulas 7.9 and 7.10 of [1] Chapter III, Theorem 3. When $s = 0.35$, the flow along the system is symmetric and the constants C_1 and C_2 are given by

$$C_1 = \frac{3.79366\epsilon_1}{c^{1/3}}$$

and

$$C_2 = \frac{-\epsilon_1}{2}.$$

When $s \neq 0.35$, $z \neq 0$ and the flow along the system is not symmetric. In this case the constants C_1 and C_2 can each be represented as the sum of two constants,

and the asymptotic representation for T_0 becomes

$$T_{LC} = T_0 + (C_{11} + C_{12})\epsilon_2^{2/3} + (C_{21} + C_{22})\epsilon_2 \ln \frac{1}{\epsilon_2} + \mathcal{O}(\epsilon_2)$$

where

$$C_{11} = \frac{3.37214|1 - 3z| \text{ Sgn}(1/3 - z)}{(-8/9 + z(-7/3 + z))(c/\epsilon_1^2)^{1/3}},$$

$$C_{12} = \frac{\epsilon_1(-2 + 6z + |1 - 3z|)}{2(1 - 3z)^2},$$

$$C_{21} = \frac{3.37214|1 + 3z| \text{ Sgn}(1/3 + z)}{(-8/9 + z(-7/3 + z))(c/\epsilon_1^2)^{1/3}},$$

$$C_{22} = \frac{\epsilon_1(-2 - 6z + |1 + 3z|)}{2(1 + 3z)^2}$$

and Sgn represents the sign function. □

8.5.2 Additional Field Observations of Foraging Rates

Here we present additional details for the field observations of foraging rates shown in Figure 8.11.

Panel A of Figure 8.11 shows the foraging rates for Colony 863 on September 5, 2015. The rates increased before reaching a QSS at around 10:30 am. The same colony on September 1, 2015 did not exhibit a QSS and had stopped foraging by 11 am (Figure 8.5E). These observations are consistent with measurements showing that September 5, 2015 was a particularly cool and humid day while September 1, 2015 was much hotter and drier (see Table 8.1).

Panel B of Figure 8.11 shows the data for Colony D19 on August 8, 2016 (from video recording) and provides an example of a very early cessation of foraging where both outgoing and incoming rates reached zero well before 11:00 am, similar to Colony 863 on September 1, 2015 (Figure 8.5E). Both August 8, 2016 and September 1, 2015 were very hot and dry days (see Table 8.1).

Panel C of Figure 8.11 show the data for Colony 859 on August 20, 2017 (from manual recording) and provides an example where the initial transient took a long time before ramping up. The initial transient for Colony 859 on August 20, 2017 remained at around 0.01 ants/sec from 10 am to 11:15 am before increasing to about 0.4 ants/sec by 12:30 pm. August 20, 2017 was a cool and humid day (see Table 8.1). Colonies might prefer different ranges of temperature and humidity; on cool and humid days, colonies that prefer warmer temperatures might forage at slightly later times of the day than colonies that prefer more cool temperatures.

Panel D of Figure 8.11 show the data for Colony 1107 on August 16, 2017 (from manual recording) and provides a different example of a slow transient; it took from 8 am to 10:30 am for the foraging rates to increase from 0.3 ants/sec to around 0.9 ants/sec. During this period the number of foragers outside the nest reached almost 2000. In this case August 16, 2017 was not a particularly cool or humid day (see Table 8.1). The long transient and large numbers of active foragers suggests that the average time it took a forager to find a seed was long. Long foraging trip times can result in slow transients and high numbers of active foragers because when foragers take a long time to find a seed, it takes longer for foragers to return to the nest and interact with available foragers who then become active foragers. As well, when the average foraging trip time is long, more foragers might be required to cover larger and less dense foraging areas.

Panel E and F of Figure 8.11 show the data for Colony 1017 on August 23, 2016 (from manual recording) and for Colony 1015 on August 18, 2016 and provide two examples of a burst in the outgoing foraging rate at the start of the foraging day that rapidly increases the number of active foragers outside the nest; it took from 7:30 am to 7:45 am for Colony 1017 on August 23, 2016 to increase the number of active foragers from 0 to 800 and it took from 7:15 am to 7:30 am for Colony 1015 on August 18, 2016 to increase the number of active foragers by 600. In both cases,

the foraging rates reached a QSS that lasted tens of minutes. Both August 23 and August 18, 2016 were very dry days (see Table 8.1).

The burst kick starts the foraging process very rapidly and appears to be different from the mechanism that activates available foragers to leave the nest through interactions between incoming successful foragers and the available foragers. The rapid increase in the number of active foragers outside the nest might be advantageous on hot and dry days on which there will be only a short period of time in the early morning with acceptable foraging conditions.

8.5.3 Analytical Approximation for $\bar{\mathbf{r}}_{\mathbf{out}}$ in terms of $\bar{\mathbf{r}}_{\mathbf{in}}$ and c

Under the assumption that λ_{in} is a Poisson process with constant rate $\bar{r}_{in}$, Eq. (4) is equivalent to a Poisson shot-noise process with exponential decay:

$$s(t) = s(0)h(t) + \sum_{i=1}^{N(t)} k\ h(t - t_i)$$

where t_i are the jump times of the Poisson process, and

$$h(t) = \begin{cases} e^{-t/\tau}, & t \geq 0 \\ 0, & t < 0. \end{cases}$$

The mean and variance of this random process for an initial condition $s(0) = 0$ are given by $\bar{r}_{in}\tau k(1 - e^{-t/\tau})$ and $\frac{1}{2}\bar{r}_{in}\tau k^2(1 - e^{-2t/\tau})$ respectively [64]. Shot-noise processes are Markovian and it can be shown that for finite jump sizes, $k < \infty$, s is ergodic [170], meaning that as $t \to \infty$, $s(t)$ converges in total variation to a unique stationary probability distribution $\pi(s)$ for any initial condition $s(0)$. In other words, s has the property that time averages converge in time to statistical averages. The ergodicity of s allows us to find an asymptotic expression as $t \to \infty$ for the expected

fraction of time that any single outcome of the random process spends in a region (b_1, b_2) by looking at its stationary probability density function.

Let $S_f = \{t_f \in [t_0, t_0 + T] \,|\, b_1 < s < b_2\}$ be the set of all times over the time interval $[t_0, t_0 + T]$ for which the stimulus is in the (b_1, b_2) region. Then $S_f \subseteq S$ where $S = \{t \in [t_0, t_0 + T]\}$. We define $\mathbb{1}_{S_f} : S \to \{0, 1\}$ to be the indicator function associated with the subset S_f:

$$\mathbb{1}_{S_f}(t) = \begin{cases} 1 & t \in S_f \\ 0 & \text{otherwise.} \end{cases}$$

Let T_f be the amount of time that s is between b_1 and b_2:

$$T_f = \int_{t_0}^{t_0+T} \mathbb{1}_{S_f}(t)\ \mathrm{d}t.$$

From the ergodic properties of s, and by the strong law of large numbers,

$$\lim_{T\to\infty} \frac{1}{T} \int_T \mathbb{1}_{S_f}(s)\ \mathrm{d}s = \int_{b_1}^{b_2} p(s)\ \mathrm{d}s,$$

where $p(s)$ is the density associated with $\pi(s)$, i.e. the stationary probability density function (PDF) of s:

$$\pi(s) = \int_0^s p(y)\ \mathrm{d}y.$$

The PDF (see Section 8.5.4) is given as a piecewise function $p_n(s)$ for $(n-1)k \leq s < nk$ where the piecewise elements satisfy recurrence equations that depend on $\bar{r}_{in}$, τ, and k.

Let b_1 and b_2 be the FN bifurcation values of the input to the FN that takes the system from quiescence into the oscillatory regime and from this regime into saturation respectively:

$$b_{1,2} = 0.35 \mp \frac{1}{3}(1 - c\epsilon_2)^{3/2}.$$

The size of the oscillatory region is given by the difference between b_2 and b_1 and it decreases with increasing volatility c (see Figure 8.12). For constant s where $b_1 < s < b_2$, the output rate is a constant given by the oscillation frequency of the FN when driven by a constant input s.

For s not constant, the FN transitions between quiescence, oscillatory behavior, and saturation as s varies. For $\epsilon_1 \ll 1$, the FN dynamics are much faster than the dynamics of s, and the number of foragers leaving the nest in a given time period $[t_0, t_0 + T]$ is proportional to T_f, the amount of time spent by s in the oscillatory region.

For $T \to \infty$, nonlinear effects in the oscillations become negligible and the mean outgoing rate becomes

$$\bar{r}_{out} = \lim_{T\to\infty} \frac{1}{T} \int_T f_{\epsilon 2}(s) \cdot \mathbb{1}_{S_f}(s)\, \mathrm{d}s$$

where $f_{\epsilon 2}$ is the mean oscillation frequency of the FN when the driving input is constant and equal to s. We approximate $f_{\epsilon 2} = 1/T_{LC}$ through the asymptotic representation [169]:

$$T_{LC} = T_0 + C_1 \epsilon_2^{2/3} + C_2 \epsilon_2 \ln \frac{1}{\epsilon_2} + \mathcal{O}(\epsilon_2),$$

where T_0, C_1, and C_2 are given in Section 8.5.1 to obtain an approximate expression for how $\bar{r}_{out}$ changes as a function of both $\bar{r}_{in}$ and c:

$$\bar{r}_{out} = \int_{b_1}^{b_2} \frac{p(s, \bar{r}_{in})}{T_{LC}(s, c)}\, \mathrm{d}s.$$

8.5.4 Probability Density Function of s(t)

Here we find an analytical description of the probability density function of the stimulus function $s(t)$ under the assumption that the incoming rate is a Poisson process. Under this assumption $s(t)$ takes the form of a Poisson shot-noise process. Before we state our results, we state a result by Gilbert and Pollak (1959) [171]:

Lemma 9. *The amplitude distribution function $F_s(\xi) = Pr[s(t) \leq \xi]$ for the Poisson shot-noise process*

$$s(t) = \sum_{i=-\infty}^{N(t)} h(t - t_i),$$

where $h(t)$ is called the impulse shape function, satisfies the integral equation

$$sF_s(s) = \int_{-\infty}^{s} F_s(x)\ dx + \bar{r}_{in} \int_{-\infty}^{\infty} F_s(s - h(t))\ h(t)\ dt$$

where $\bar{r}_{in}$ is the rate parameter of the underlying Poisson process.

Proof. We refer the reader to [171] for the proof. □

Theorem 9. *Consider the shot-noise process with exponential decay and impulses arriving with rate $\bar{r}_{in}$ given by*

$$s(t) = \sum_{i=1}^{N(t)} k\ h(t - t_i),$$

where $k > 0$ is a constant and

$$h(t) = \begin{cases} e^{-t/\tau}, & t \geq 0 \\ 0, & t < 0. \end{cases}$$

Then, the steady-state probability density function of $s(t)$ can be written as a piecewise function $p(s)$ where the piecewise elements $p_n(s)$ for $(n-1)k \leq s < nk$ satisfy the

recurrence equations:

$$p_n(s) = p_{n-1}(s) + \alpha s^{\bar{r}_{in}\tau - 1}(-\bar{r}_{in}\tau)^{n-1} g_n(s, \bar{r}_{in}, k, \tau)$$

$$g_n(s, \bar{r}_{in}, k, \tau) = \int_k^s (x-k)^{\bar{r}_{in}\tau - 1} x^{-\bar{r}_{in}\tau} g_{n-1}(x-k, \bar{r}_{in}, k, \tau) dx$$

with

$$\alpha = \frac{(ke^{\gamma})^{-\bar{r}_{in}\tau}}{\Gamma(\bar{r}_{in}\tau)} s^{\bar{r}_{in}\tau - 1}, \quad p_0(s) = 0, \quad g_1(s, \bar{r}_{in}, k, \tau) = 1,$$

where $\gamma = 0.5772...$ *is Euler's constant and* Γ *is the gamma function.*

Proof. For a Poisson shot-noise process with impulse shape function:

$$h(t) = \begin{cases} e^{-t/\tau}, & t \geq 0 \\ 0, & t < 0, \end{cases}$$

the integral equation in Lemma 9 can be rewritten as

$$sp(s) = \bar{r}_{in}\tau \int_0^k p(s-\xi)\ \mathrm{d}\xi = \bar{r}_{in}\tau \int_{s-k}^s p(x)\ \mathrm{d}x,$$

where $p(s) = \mathrm{d}F_s/\mathrm{d}s$ is the density function of s.

Differentiating with respect to s, we obtain

$$s\frac{dp}{ds} + p(s)[1 - \bar{r}_{in}\tau] = -\bar{r}_{in}\tau p(s-k). \tag{8.15}$$

When $0 \leq s \leq k$, then $p(s-k) = 0$ and

$$s\frac{dp}{ds} + p(s)[1 - \bar{r}_{in}\tau] = 0.$$

Picking $p(s) = \alpha\ s^{\bar{r}_{in}\tau - 1}$ satisfies the equation above. Thus, we have obtained a solution for $p(s) = p_1(s)$ when $0 \leq s \leq k$. For $s > k$, the differential equation eq. (8.15)

may be converted to an integral form:

$$p(s) = s^{\bar{r}_{in}\tau - 1}\left[\alpha - \bar{r}_{in}\tau \int_k^s p(x-k)x^{-\bar{r}_{in}\tau}\ \mathrm{d}x\right]. \tag{8.16}$$

Since the integrand is known for $k < x < 2k$, we can determine $p(s) = p_2(s)$ for $k < s < 2k$. This in turn enables us to integrate further to get $p(s) = p_3(s)$ for $2k < s < 3k$, etc. Let $m = \bar{r}_{in}\tau$, then the results for the first three jump regions $p_n(s)$, $n = 1, 2, 3$ are given by

$$\begin{aligned}
p_1(s) &= \alpha\ s^{m-1} \\
p_2(s) &= p_1(s) - \alpha\ s^{m-1} m \int_k^s (x-k)^{m-1} x^{-m}\ \mathrm{d}x \\
p_3(s) &= p_2(s) + \alpha\ s^{m-1} m^2 \int_k^s (x-k)^{m-1} x^{-m} \int_k^{x-k} (\xi - k)^{m-1} \xi^{-m} \mathrm{d}\xi\ \mathrm{d}x.
\end{aligned}$$

We now show by induction that $p(s) = p_n(s)$, for $(n-1)k \le s < nk$, satisfies the following recurrence equations:

$$\begin{aligned}
p_n(s) &= p_{n-1}(s) + \alpha s^{m-1}(-m)^{n-1} g_n(s) \\
g_n(s) &= \int_k^s (x-k)^{m-1} x^{-m} g_{n-1}(x-k)\ \mathrm{d}x
\end{aligned}$$

with $p_0 = 0$, $g_1(s) = 1$. For $n = 1$,

$$p_1(s) = p_0(s) + \alpha s^{m-1}(-m)^0 g_1(s) = \alpha s^{m-1}$$

as expected. Now, assume that for $n = j$,

$$\begin{aligned}
p_j(s) &= p_{j-1}(s) + \alpha s^{m-1}(-m)^{j-1} g_j(s) \\
g_j(s) &= \int_k^s (x-k)^{m-1} x^{-m} g_{j-1}(x-k)\ \mathrm{d}x.
\end{aligned}$$

Then, for $n = j + 1$,

$$\begin{aligned} p_{j+1}(s) &= s^{m-1}\left[\alpha - m\int_k^s p_j(x-k)x^{-m}\,\mathrm{d}x\right] \\ &= s^{m-1}\left[\alpha - m\int_k^s x^{-m}p_{j-1}(x-k)\,\mathrm{d}x\right. \\ &\quad \left. + \alpha(-m)^j\int_k^s (x-k)^{m-1}x^{-m}g_j(x-k)\,\mathrm{d}x\right] \\ &= s^{m-1}\left[\alpha - m\int_k^s x^{-m}p_{j-1}(x-k)\,\mathrm{d}x\right] + \alpha s^{m-1}(-m)^j g_{j+1}(s) \\ &= p_j(s) + \alpha s^{m-1}(-m)^j g_{j+1}(s). \end{aligned}$$

Finally, the constant α must be determined by the condition

$$\int_0^\infty p(s)\,\mathrm{d}s = 1.$$

To compute the constant, we first note that the characteristic equation of s is given by

$$C(\zeta) = \exp\left[-\bar{r}_{in}\int_{-\infty}^{\infty}(1-\exp[-\zeta F_s(t)])\,\mathrm{d}t\right]$$

(see [64] for derivation). The characteristic function is the Laplace transform $\hat{p}$ of p,

$$\hat{p}(\zeta) = \exp\left[-\bar{r}_{in}\tau\int_0^{\zeta k}\frac{1-e^{-y}}{y}\,\mathrm{d}y\right].$$

Using partial integration, this can be rewritten as

$$\begin{aligned} \hat{p}(\zeta) &= \exp\left[-\bar{r}_{in}\tau(1-e^{-\zeta k})\log\zeta k + \bar{r}_{in}\tau\left(\int_0^\infty e^{-y}\log y\,\mathrm{d}y - \int_s^\infty e^{-y}\log y\,\mathrm{d}y\right)\right] \\ &= (\zeta k)^{-\bar{r}_{in}\tau}e^{-\bar{r}_{in}\tau\gamma}(1+\mathcal{O}[e^{-\zeta k(1-\varepsilon)}]) \quad \text{for any } \varepsilon > 0. \end{aligned}$$

Thus, for $0 \le s \le k$,

$$\alpha = \frac{(ke^\gamma)^{-\bar{r}_{in}\tau}}{\Gamma(\bar{r}_{in}\tau)}s^{\bar{r}_{in}\tau-1}$$

where $\gamma = 0.5772...$ is Euler's constant and Γ is the gamma function. $\square$

Chapter 9

Bistability and Resurgent Epidemics in Reinfection Models*

Renato Pagliara, Biswadip Dey, and Naomi Ehrich Leonard

Spreading processes that propagate through local interactions have been studied in multiple fields (e.g., epidemiology, complex networks, social sciences) using the SIR (Susceptible-Infected-Recovered) and SIS (Susceptible-Infected-Susceptible) frameworks. SIR assumes individuals acquire full immunity to the infection after recovery, while SIS assumes individuals acquire no immunity after recovery. However, in many spreading processes individuals may acquire only partial immunity to the infection or may become more susceptible to reinfection after recovery. We study a model for reinfection called SIRI (Susceptible-Infected-Recovered-Infected). The SIRI model generalizes the SIS and SIR models and allows for study of systems in which the susceptibility of agents changes irreversibly after first exposure to the infection. We show that when the rate of reinfection is higher than the rate of primary infection, the SIRI model exhibits bistability with a small difference in the initial fraction of infected indi-

[1]This chapter was published as Pagliara, Dey, and Leonard in IEEE Control Systems Letters on May 1, 2018 [57].

viduals determining whether the infection dies out or spreads through the population. We find this critical value and show that when the infection does not die out there is a resurgent epidemic in which the number of infected individuals decays initially and remains at a low level for an arbitrarily long period of time before rapidly increasing towards an endemic equilibrium in which the fraction of infected individuals is non-zero.

9.1 Introduction

Epidemiological models [47] have been widely studied and successfully applied in many settings, including mobile networks [172], rumor spreading [51], and even viral video dynamics [53]. These compartmental models typically describe how the group sizes of different types of individuals evolve over time. The main appeal for these models is their high analytic tractability, which makes them a powerful framework for studying transient and steady-state system behaviors.

These models are also central to the understanding of contagious processes [74, 173,174]. and to the development of control and optimal resource allocation strategies that seek to inhibit or promote the spread of the process [175–177].

Two of the most successful and well-studied epidemiological models are the SIS and SIR models. In the SIS model individuals can be either susceptible or infected. Susceptible individuals become infected through contact with already infected individuals, and return to the susceptible state after recovering from the infection. The SIR model is similar to the SIS model except for the fact that recovered individuals acquire full immunity to the infection, meaning they cannot become infected again.

While the SIS (no immunity) and the SIR (full immunity) models have been extensively used and studied, they do not address many of the applicable real-world situations in which the susceptibility of individuals to primary infections is different

from the susceptibility to secondary infections (i.e., reinfections). For instance, in the case of infectious diseases, a lower probability of reinfection corresponds to the development of partial immunity in which primary infections are more likely than secondary infections, such as in the case of influenza [178]. Alternatively, a higher probability of reinfection might correspond to a compromised immune system in which secondary infections are more likely, such as in the case of tuberculosis in particular populations [179].

In the spread of social behaviors, past experiences may lead to differences between primary and secondary infections. A lower probability of reinfection could be the result of a negative experience that reduces the propensity of an individual to further engage in the behavior, while a higher probability of reinfection could result from a positive experience that increases the propensity of an individual to engage in the behavior.

In this paper we study the role of susceptibility to reinfections by considering the spread of a contagious process using the SIRI (Susceptible-Infected-Recovered-Infected) model in which the rate of primary infections is different from the rate of secondary infections. The SIRI model contains the SIS and SIR models as special cases and allows for the study of systems in which individuals become more or less susceptible to the infection after first exposure.

In the theoretical biology literature reinfection models have been used to study the role of partial immunity and waning immunity across populations [134], while in the physics community spatial reinfection models have garnered attention due to their critical behavior connecting directed percolation and dynamic percolation [180]. In [181] the authors study the Markovian SIRI model on arbitrary networks and show through numerical simulations on random networks that the model exhibits bistability in which a low number of initially infected individuals leads to an infection-free steady-state while a much larger number of initially infected individuals leads to an endemic

steady-state. We have found no other work in the literature that examines this bistability. In the present paper we formalize the observation of [181] in the case of a well-mixed population by proving new results on the critical initial condition below which the infection dies out and above which solutions reach an endemic equilibrium.

Our contribution in the present paper is a rigorous analysis of the SIRI model dynamics over its entire phase space. We identify and prove conditions for each of four different dynamical regimes exhibited by the SIRI model: infection-free, endemic, epidemic, and bistable. As far as we know this is the first such analysis of its kind. We prove that the bistability phenomenon occurs when secondary infections are more likely than primary infections. We prove that when the bistability condition leads to an endemic steady-state, the system exhibits a resurgent epidemic in which the number of infected individuals initially decreases before ramping up after an arbitrarily long delay.

The paper is organized as follows. Section 9.2 introduces the SIRI model. In Section 9.3 we analyze the dynamics of the SIRI model over the entire phase space and prove conditions for the four dynamical regimes. In Section 9.4 we study the bistable regime in more detail and introduce the concept of a resurgent epidemic with an arbitrarily long delay. We provide closing remarks and discuss future work in Section 9.5.

9.2 Model Description

Consider a large population in which an individual can be in any of the following three states: *susceptible* (S), *infected* (I), or *recovered* (R). Susceptible and recovered individuals become infected through contact with already infected individuals at respective rates $\beta \geq 0$ and $\hat{\beta} \geq 0$, while infected individuals recover at a fixed rate $\delta \geq 0$:

$$\begin{array}{ccccc} \text{S+I} & \xrightarrow{\beta} & \text{I+I} & \xleftarrow{\hat{\beta}} & \text{R+I} \\ & & \text{I} & \xrightarrow{\delta} & \text{R} \end{array} \tag{9.1}$$

By assuming that interactions between any two individuals occur with the same probability (i.e., under homogeneous mixing conditions), we can model the system dynamics as

$$\begin{aligned} \dot{x}_S &= -\beta x_S x_I \\ \dot{x}_I &= \beta x_S x_I + \hat{\beta} x_R x_I - \delta x_I \\ \dot{x}_R &= -\hat{\beta} x_R x_I + \delta x_I, \end{aligned} \tag{9.2}$$

where x_S, x_I, and x_R represent the fractions of population that belong to the susceptible, infected, and recovered states, respectively. Note that $x_S + x_I + x_R = 1$, and this constraint is preserved under (9.2).

Table 9.1 shows the special cases of the SIRI model. Setting $\hat{\beta} = 0$ reduces the SIRI model to the SIR model, while setting $\hat{\beta} = \beta$ and redefining x_S as $x_S + x_R$ reduces the SIRI model to the SIS model. In between the SIR (full immunity) and the SIS (no immunity) models, the rate of secondary infections is larger than zero but lower than the rate of primary infections ($0 < \hat{\beta} < \beta$), and we say that recovered individuals have developed *partial immunity* to the infection, i.e., they are less likely to become reinfected. For example, in the spread of rumors, partial immunity might represent the scenario wherein individuals become less likely to spread new rumors, possibly due to negative consequences of an initially spread rumor. When the rate of secondary infections is larger than that of primary infections ($\beta < \hat{\beta} < \infty$), recovered individuals become reinfected more easily, and we say that recovered individuals have developed *compromised immunity* to the infection. In the example, this could represent the scenario wherein individuals become more willing to spread new rumors, possibly due to benefits from spreading a previous rumor. In the limit $\hat{\beta} \to \infty$, the rate of recovery is negligible compared to the rate of reinfection, and the SIRI model approximates

Table 9.1: Special cases of the SIRI model.

Parameter Value	Equivalent Model	Immunity Condition
$\hat{\beta} = 0$	SIR	Full Immunity
$0 < \hat{\beta} < \beta$	—	Partial Immunity
$\hat{\beta} = \beta$	SIS	No Immunity
$\beta < \hat{\beta} < \infty$	—	Compromised Immunity
$\delta = 0$	SI	No Recovery

the SI model where infected individuals can never recover. An exact equivalence with the SI model can be achieved by setting the recovery rate δ to zero and redefining x_S as $x_S + x_R$. This could represent the scenario wherein individuals cannot stop spreading rumors once they hear a rumor.

9.3 Model Analysis

9.3.1 Epidemic Analysis

Here we derive conditions on model parameters that guarantee an epidemic, i.e., growth in number of infected individuals for a small initial number of infected individuals.

The dynamics for x_I in (9.2) can be written as

$$\dot{x}_I = \delta((R_0 x_S + R_1 x_R) - 1)x_I = \delta(R(x_S, x_R) - 1)x_I \tag{9.3}$$

where $R_0 \triangleq \beta/\delta$, $R_1 \triangleq \hat{\beta}/\delta$, and $R(x_S, x_R) \triangleq R_0 x_S + R_1 x_R$. The infection decays when $R < 1$, grows when $R > 1$, and neither grows nor decays when $R = 1$.

If $R_0 > 1$, a small fraction of initially infected individuals can spread the infection in a population with no recovered individuals ($x_R = 0$). To see this, consider the dynamics for x_I around a point where $x_R = 0$ and $x_I \approx 0$:

$$\dot{x}_I = (\beta - \delta)x_I = \delta(R_0 - 1)x_I. \tag{9.4}$$

Similarly, if $R_1 > 1$ a small fraction of initially infected individuals can spread the infection in a population with no susceptible individuals ($x_S = 0$). To see this, consider the dynamics for x_I around a point where $x_R \approx 1$ and $x_I \approx 0$:

$$\dot{x}_I = (\hat{\beta} - \delta)x_I = \delta(R_1 - 1)x_I. \tag{9.5}$$

We can investigate the effect of introducing a small fraction of infected individuals in a population with both susceptible and recovered individuals by looking at the linearized dynamics of x_I around an infection-free point $x_S = 1 - \bar{x}_R$, $x_I = 0$, and $x_R = \bar{x}_R$,

$$\dot{x}_I = \delta(R_0 + (R_1 - R_0)\bar{x}_R - 1)x_I. \tag{9.6}$$

For initial conditions where $\bar{x}_R = 0$, by (9.4) the initial infection spreads if $R_0 > 1$. If all individuals become infected, i.e., $\bar{x}_R = 1$, by (9.5) the infection spreads through the recovered population if $R_1 > 1$.

When $0 < \bar{x}_R < 1$, the infection spreads if $R_0 + (R_1 - R_0)\bar{x}_R > 1$. So, if $R_0 > R_1$, i.e., $\beta > \hat{\beta}$, as in the case of partial immunity, the effective spreading power of the infection decreases with the fraction of recovered individuals $\bar{x}_R$. That is, recovered individuals are less prone to the infection than susceptible individuals, which makes it harder for the infection to spread. And if $R_0 < R_1$, i.e., $\beta < \hat{\beta}$, as in the case of compromised immunity, recovered individuals are likely to become reinfected and facilitate the spread of the infection through the population.

9.3.2 Equilibrium Points and Stability Analysis

The dynamics of the SIRI model (9.2) evolve on the 2-simplex $\Delta_2 \triangleq \{(x_S, x_I, x_R) \in [0,1]^3 | x_S + x_I + x_R = 1\}$, and the corresponding reduced dynamics can be expressed

as

$$\begin{aligned} \dot{x}_S &= -\beta x_S x_I \\ \dot{x}_I &= (\hat{\beta} - \delta) x_I + (\beta - \hat{\beta}) x_S x_I - \hat{\beta} x_I^2. \end{aligned} \tag{9.7}$$

This reduced model (9.7) has one continuum of equilibria and an isolated equilibrium point:

1. Infection-Free Equilibria (IFE): $x_S = x_S^*$, $x_I = 0$,
2. Endemic Equilibrium (EE): $x_S = 0$, $x_I = 1 - \delta/\hat{\beta}$.

The IFE is a continuum of equilibria corresponding to the boundary of Δ_2 where $x_I = 0$ and $x_S^* \in [0, 1]$, while the EE corresponds to the case in which every individual is either in the infected or recovered state. The SIRI model does not have an equilibrium point where all three states S, I, and R coexist.

We now show how the steady-state solution x_S^* at a point in the IFE depends on the initial conditions.

Theorem 10. *The fraction of susceptible individuals x_s^* at a point in the IFE is given by the implicit equation*

$$\left(\frac{x_S^*}{x_{S0}}\right)^{R_1/R_0} \left(x_{I0} + x_{S0} - \frac{R_1 - 1}{R_1}\right) - x_S^* + \frac{R_1 - 1}{R_1} = 0,$$

where x_{I0} and x_{S0} are the initial fractions of infected and susceptible individuals, respectively.

Proof. Dividing the two equations in (9.7) we get an expression for dx_I/dx_S:

$$\frac{dx_I}{dx_S} = \frac{(\hat{\beta} - \beta)}{\beta} - \frac{\hat{\beta} - \delta}{\beta x_S} + \frac{\hat{\beta} x_I}{\beta x_S}$$

with solution

$$x_I = -x_S + \frac{\hat{\beta} - \delta}{\hat{\beta}} + k x_S^{\hat{\beta}/\beta} \tag{9.8}$$

where the value of k can be found by setting $t = 0$, yielding

$$\frac{x_I + x_S - (\hat{\beta} - \delta)/\hat{\beta}}{x_S^{\hat{\beta}/\beta}} = \frac{x_{I0} + x_{S0} - (\hat{\beta} - \delta)/\hat{\beta}}{x_{S0}^{\hat{\beta}/\beta}}. \tag{9.9}$$

In the limit $t \to \infty$, $x_I(\infty) = 0$. Simplifying and making the substitution $x_S^* = x_S(\infty)$ we get the implicit equation

$$\left(\frac{x_S^*}{x_{S0}}\right)^{\hat{\beta}/\beta}\left(x_{I0} + x_{S0} - \frac{\hat{\beta} - \delta}{\hat{\beta}}\right) - x_S^* + \frac{\hat{\beta} - \delta}{\hat{\beta}} = 0.$$

Substituting $R_1 = \hat{\beta}/\delta$ completes the proof. □

Corollary 3. *Given an initial condition $x_I = x_{I0}$, $x_S = 1 - x_{I0}$, where $0 < x_{I0} < 1$, the fraction of susceptible individuals x_s^* at the IFE is given by the implicit equation*

$$\frac{1}{R_1}\left(\frac{x_s^*}{1 - x_{I0}}\right)^{R_1/R_0} - x_s^* + \frac{R_1 - 1}{R_1} = 0. \tag{9.10}$$

Proof. The proof follows by setting $x_{S0} = 1 - x_{I0}$ in Theorem 10. □

Before we state the main theorem of the paper, we define the quantity $M \triangleq (1 - R_1)/(R_0 - R_1)$ which we use throughout the rest of this section.

Theorem 11 (Behavioral Regimes of SIRI)**.** *Given an initial condition $x_I = x_{I0}$, $x_S = 1 - x_{I0}$, where $0 < x_{I0} < 1$, the SIRI model* (9.7) *exhibits four different dynamical behaviors:*

1. *Infection-Free: If $R_0 < 1$ and $R_1 < 1$, then all solutions reach a point in the IFE as $t \to \infty$, and x_I decays monotonically to zero.*

2. *Endemic: If $R_0 > 1$ and $R_1 > 1$, then all solutions reach the EE as $t \to \infty$.*

3. *Epidemic: If $R_0 > 1$ and $R_1 \leq 1$, then all solutions reach a point in the IFE as $t \to \infty$ and, at equilibrium, $x_S^* < M$. For initial conditions where*

$x_{I0} \geq (\beta - \delta)/\beta$, x_I decays monotonically to zero. While for initial conditions where $x_{I0} < (\beta - \delta)/\beta$, x_I grows initially and reaches a maximum value:

$$x_I^{max} = \frac{R_0 - R_1}{R_1(R_0^{R_0}(1 - x_{I0})^{R_1})^{1/(R_0 - R_1)}} + \frac{R_1 - 1}{R_1},$$

before decaying to zero as $t \to \infty$.

4. *Bistable: If $R_0 \leq 1$, $R_1 > 1$, then x_I decays initially. Moreover, there is a critical initial fraction of infected individuals*

$$x_{IC} = 1 - M(R_0 M)^{-\frac{R_0}{R_1}}. \tag{9.11}$$

Solutions with initial condition $x_{I0} < x_{IC}$ reach a point in the IFE as $t \to \infty$ and x_I decays monotonically to zero. Solutions with initial conditions $x_{I0} > x_{IC}$ reach the EE as $t \to \infty$.

Before proving Theorem 11, we prove three lemmas.

Lemma 10. *The EE is an equilibrium point of* (9.7) *if and only if $R_1 \geq 1$. Moreover, the EE is locally stable.*

Proof. To show necessity, note that at the EE we have $x_I = 1 - \delta/\hat{\beta} = 1 - 1/R_1$ which is nonnegative only if $R_1 \geq 1$.

Sufficiency follows from the fact that $x_I = 1 - 1/R_1$, $x_s = 0$ and $x_R = 1 - x_I$ is an equilibrium point of (9.7).

The Jacobian of (9.7) around the EE is given by

$$J = \begin{bmatrix} -\beta(\hat{\beta} - \delta)/\hat{\beta} & 0 \\ (\beta - \hat{\beta})(\hat{\beta} - \delta)/\hat{\beta} & -(\hat{\beta} - \delta) \end{bmatrix}$$

which is Hurwitz if $R_1 > 1$. □

The following lemma shows that in the epidemic and bistable regimes, the IFE contains both locally stable and unstable equilibrium points.

Lemma 11. *The following holds true for the IFE:*

1. *If $R_0 < 1$ and $R_1 < 1$, then all points in the IFE are locally stable.*

2. *If $R_0 > 1$ and $R_1 > 1$, then all points in the IFE are unstable.*

3. *If $R_0 > 1$ and $R_1 \leq 1$, points in the IFE with $x_S^* < M$ are locally stable and points with $x_S^* > M$ are unstable.*

4. *If $R_0 \leq 1$ and $R_1 > 1$, points in the IFE with $x_S^* > M$ are locally stable and points with $x_S^* < M$ are unstable.*

Proof. The Jacobian for the linearized system about $x_I = 0$, $x_S = x_S^*$ is

$$J = \begin{bmatrix} 0 & -\beta x_S^* \\ 0 & (\beta - \hat{\beta})x_S^* + \hat{\beta} - \delta \end{bmatrix} \tag{9.12}$$

The zero eigenvalue has eigenvector $[1, 0]^T$ corresponding to the invariant subspace $x_I = 0$. The second eigenvalue $J_a = (\beta - \hat{\beta})x_s^* + (\hat{\beta} - \delta)$ determines the local stability of points in the IFE.

To prove 1, assume $R_0 < 1$ and $R_1 < 1$. If $R_0 > R_1$, then $\beta > \hat{\beta}$ and $J_a < 0$ for any $0 \leq x_S^* \leq 1$ and all points in the IFE are locally stable. If $R_0 < R_1$ then $\beta < \hat{\beta}$ and $\max J_a = \beta - \delta < 0$ and all points in the IFE are locally stable.

To prove 2, assume $R_0 > 1$ and $R_1 > 1$. If $R_0 > R_1$, $J_a > 0$ for any $0 \leq x_S^* \leq 1$ and all points in the IFE are unstable. If $R_0 < R_1$ then $\max J_a = \beta - \delta > 0$ and all points in the IFE are unstable.

To prove 3, assume $R_0 > 1$ and $R_1 \leq 1$. It follows that $J_a < 0$ if $0 < x_S^* < M$ and $J_a > 0$ if $M < x_S^* < 1$, which is equivalent to 3.

To prove 4, assume $R_0 \leq 1$ and $R_1 > 1$. It follows that $J_a < 0$ if $M < x_S^* < 1$ and $J_a > 0$ if $0 < x_S^* < M$, which is equivalent to 4. □

We rule out the existence of periodic orbits in the SIRI model as this has the implication that any solution starting on Δ_2 must end at either a point in the IFE or the EE.

Lemma 12. *The SIRI model does not exhibit non-trivial periodic orbits on* Δ_2.

Proof. We rule out the existence of periodic orbits by contradiction. Suppose there is a periodic solution of (9.7) on Δ_2. Then $x_S(t) = x_S(t')$ for some $t' > t$. Since x_S is nonincreasing in Δ_2, this implies that $\dot{x}_S \equiv 0$ on $[t, t']$ which holds if and only if $x_I x_S \equiv 0$ on $[t, t']$.

If $x_I = 0$ at any time $\bar{t} \in [t, t']$, then the system is at a point in the IFE at time $\bar{t}$. Then $x_I \equiv 0$ on $[\bar{t}, t']$ and the solution is not a non-trivial periodic orbit. If $x_S \equiv 0$ on $[t, t']$, then $x_R = 1 - x_I$ and the dynamics of (9.7) can be reduced to a single first order ODE. Because periodic orbits cannot take place in a first order system, we have a contradiction. □

Proof of Theorem 11. Assume $R_0 < 1$ and $R_1 < 1$. By Lemma 11 the IFE are the only equilibria of the system and by Lemma 12 there are no periodic solutions. These two statements imply that all solutions reach a point in the IFE as $t \to \infty$. Moreover, solutions decrease monotonically since $R < 1$ for any $x_S, x_I \in [0, 1]$. From (9.3) it follows that $\dot{x}_I < 0$ for $x_I \neq 0$. This completes the proof for 1.

Assume $R_0 > 1$ and $R_1 > 1$. By Lemmas 11 and 10 all points in the IFE are unstable and the EE is locally stable. By Lemma 12 there are no periodic solutions. These statements imply that all solutions reach the EE as $t \to \infty$. This completes the proof for 2.

Assume $R_0 > 1$ and $R_1 \leq 1$. Following the same argument as in the proof of 1, all solutions reach a point in the IFE as $t \to \infty$. From Lemma 11, equilibrium points

in the IFE for which $x_S^* > M$ are unstable. Therefore solutions reach points in the IFE where $x_S^* < M$.

At the initial condition $x_I = x_{I0}$, $x_S = 1 - x_{I0}$, with $0 < x_{I0} < 1$, the initial rate of change of x_I by (9.7) is $((\beta - \delta) - \beta x_{I0})x_{I0}$. It follows that x_I grows initially if $x_{I0} < (\beta - \delta)/\beta$ and decays initially if $x_{I0} > (\beta - \delta)/\beta$.

Points along the solution where $\dot{x}_I = 0$ belong to the x_I-nullcline and satisfy $R = 1$ if $x_I \neq 0$. Since all solutions reach a point in the IFE, solutions that grow initially must reach a maximum value x_I^{max} where $R = 1$. Since solutions cannot intersect, this implies that solutions that decay initially do not change sign and continue to decay monotonically until they reach a point in the IFE.

By setting $\dot{x}_I = 0$ in (9.7), we can express x_S^{max} (i.e. the maximum value x_S) in terms of x_I^{max} as

$$x_S^{max} = \frac{\delta + \hat{\beta}(x_I^{max} - 1)}{\beta - \hat{\beta}}. \tag{9.13}$$

Then substituting (9.13) into (9.9) and simplifying, we get

$$x_I^{max} = \frac{R_0 - R_1}{R_1(R_0^{R_0}(1 - x_{I0})^{R_1})^{1/(R_0 - R_1)}} + \frac{R_1 - 1}{R_1}. \tag{9.14}$$

This completes the proof for 3.

Assume $R_0 \leq 1$, $R_1 > 1$. It follows that at the initial condition $\dot{x}_I < 0$ for any $0 < x_{I0} \leq 1$ and x_I decays initially. A necessary condition for the solution to reach a point in the IFE is for the fraction of susceptible individuals at steady-state x_S^* to satisfy the implicit equation (9.10) and for the IFE point $(x_S^*, 0)$ to be locally stable.

By Lemma 11 any point in the IFE with $x_S^* < M$ is unstable. Therefore we require $x_S^* > M$. Solving (9.10) for x_{I0} shows that the necessary condition is satisfied if $x_{I0} < x_{IC} \triangleq 1 - M(R_0 M)^{-R_0/R_1}$. Thus, $x_{I0} < x_{IC}$ is a necessary condition for the solution to reach a point in the IFE.

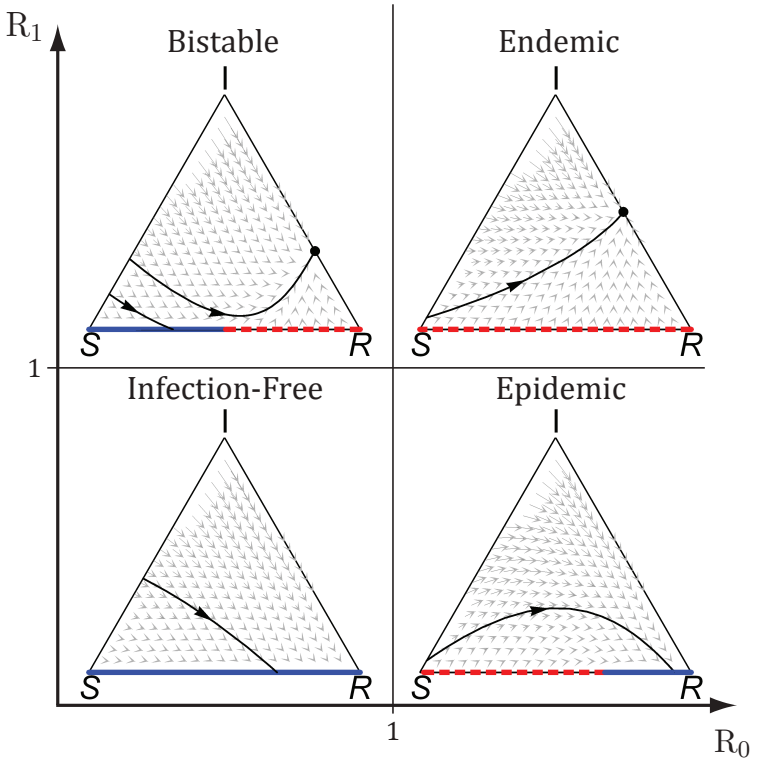

Figure 9.1: The four different behavioral regimes of the SIRI model plotted on Δ_2. The four plots are arranged in the R_0, R_1 parameter space to illustrate the four corresponding regimes.

To prove sufficiency, we show that $\dot{x}_I$ does not change sign when $x_{I0} < x_{IC}$ and therefore solutions with $x_{I0} < x_{IC}$ decrease monotonically.

If $\dot{x}_I$ changes sign, then x_I has a minimum value where $R = 1$. To be a valid minimum of $x_I \in [0, 1]$, we require a minimum value $x_I^{min} \in [0, 1]$. If $x_I^{min} = 0$, then we must have $R = 1$ at $(x_S^*, 0)$. This condition is satisfied when $x_S^* = M$, or equivalently when $x_{I0} = x_{IC}$. Thus, any solution with $x_{I0} < x_{IC}$ cannot have $x_I^{min} > 0$ and $x_I \to 0$ monotonically as $t \to \infty$.

From the discussion above, it follows that a necessary and sufficient condition for the solution not to reach a point in the IFE is $x_{I0} > x_{IC}$. Due to the invariance of Δ_2 and the impossibility of periodic orbits, this implies that $x_{I0} > x_{IC}$ is a necessary and sufficient condition for solutions to reach the EE. This completes the proof. □

Figure 9.1 summarizes the results of Theorem 11. In each quadrant of the R_0, R_1 parameter space, we show a simulation of the corresponding dynamics on Δ_2. The bottom boundary of Δ_2 represents the IFE. The solid blue and dashed red lines correspond to locally stable and unstable points in the IFE, respectively. The thinner black lines are example trajectories. We show two trajectories in the bistable case corresponding to a trajectory with $x_{I0} = 0.15$ that reaches a point in the IFE and a trajectory with $x_{I0} = 0.3$ that reaches the EE.

Remark 14. The transient dynamics in the infection-free and endemic regimes depend on the ratio R_0/R_1. When $R_0/R_1 > 1$ recovered individuals inhibit the spread of the infection, leading to concave trajectories in Δ_2. In contrast, when $R_0/R_1 < 1$, recovered individuals facilitate the spread, leading to convex trajectories in Δ_2.

9.4 Resurgent Epidemics

In this section we study the bistable regime in more detail and show that when the initial condition is above the critical value solutions exhibit a resurgent epidemic in which the infection initially decreases before increasing after an arbitrarily long period of time.

Theorem 12 (Resurgent Epidemic)**.** *Consider a solution in the bistable regime with initial condition $x_{IC} < x_{I0} < 1$ such that the solution reaches the EE as $t \to \infty$. For that solution, the fraction of infected individuals decreases initially, reaches a minimum value*

$$x_I^{min} = \frac{R_0 - R_1}{R_1(R_0^{R_0}(1 - x_{I0})^{R_1})^{1/(R_0 - R_1)}} + \frac{R_1 - 1}{R_1},$$

and then increases until it reaches the EE.

Proof. Assume $R_0 < 1$, $R_1 > 1$, and $x_{I0} > x_{IC}$. From (9.4) we get that the initial fraction of infected individuals decays exponentially while from result 4 of Theorem 11 we get that the solution reaches the EE as $t \to \infty$.

Similar to the analysis for x_I^{max} in the epidemic case, any minimum of x_I must satisfy $R = 1$. This is only satisfied along the portion of the x_I-nullcline between the points $(M, 0)$, which separates the IFE into locally stable and unstable sets, and $(\frac{R_1-1}{R_1}, 0)$, which corresponds to the EE. We refer to this portion of the x_I-nullcline as Λ.

To show that all trajectories reach a minimum, note that if a trajectory passes through a point in Λ, that point will correspond to x_I^{min}, the minimum value of x_I along the trajectory. Solving (9.9) for x_{I0} we get

$$x_{I0} = 1 - Q(R_0 Q)^{-R_0/R_1} \tag{9.15}$$

where $Q = (1 + R_1(x_I - 1))/(R_0 - R_1)$.

Setting $x_I = 0$ in (9.15) yields a lower bound on the initial condition x_{I0} that results in a trajectory with a minimum value $x_I^{min} \in [0, 1]$, while setting $x_I = (R_1 - 1)/R_1$ in (9.15) yields an upper bound on the initial condition x_{I0} that results in a trajectory with a minimum $x_I^{min} \in [0, 1]$.

When $x_I = 0$, $Q = M$ and we get $x_{I0} = x_{IC}$, that is, we recover (9.11), the critical value for bistability. When $x_I = (R_1 - 1)/R_1$, $Q = 0$ and $x_{I0} = 1$. This shows that any solution with $x_{IC} < x_{I0} < 1$ achieves a minimum value $x_I^{min} \in [0, 1]$.

Finally, note that the same analysis used to find (9.14) is valid in the bistable case, except that the resulting equation describes the minimum value x_I^{min}. □

Figure 9.2 shows a simulation that exhibits resurgent epidemics with $\beta = 0.5$, $\delta = 1$, and $\hat{\beta} = 1.5$. The initial fraction of infected individuals x_{I0} was set to 0.207. The infection decays at first, reaching a value close to zero after 20 time units. The

infection stays close to zero for over 350 time units before increasing towards an endemic state where $x_I = 0.33$.

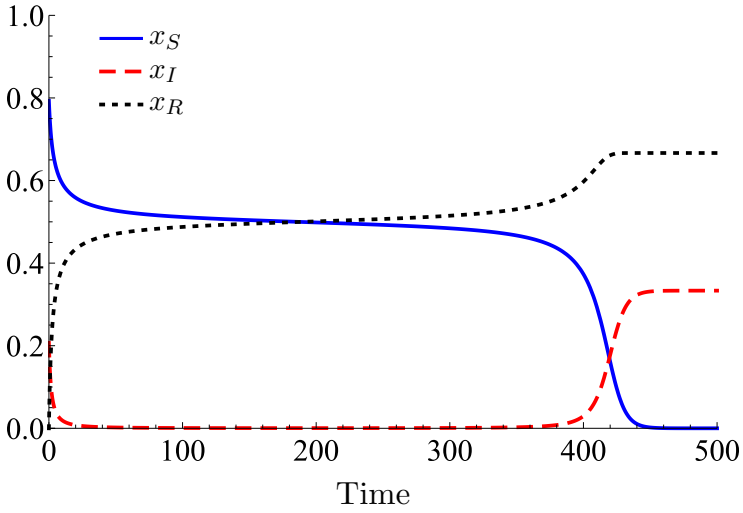

Figure 9.2: Resurgent epidemic for $\beta = 0.5$, $\delta = 1$, $\hat{\beta} = 1.5$, and $x_{I0} = 0.207$.

The time it takes before the resurgent epidemic is observed depends on the difference between the initial condition x_{I0} and the critical value x_{IC}. If x_{I0} is close to x_{IC} then the minimum value x_I^{min} will be close to zero and the rate of growth of x_I will be very slow, leading to long time periods where the infection appears to be under control before the epidemic resurges.

To study this phenomenon in more detail, we define the time to resurgence t_{RS} as the time it takes x_I to decay from x_{I0} to the minimum value x_I^{min}. Figure 9.3 shows t_{RS} versus $x_{I0} > x_{IC}$ for the same parameters as in Figure 9.2. As the difference $x_{I0} - x_{IC}$ goes to zero, t_{RS} goes to infinity.

Theorem 13 (Time to Resurgence)**.** *Consider a solution in the bistable regime that exhibits a resurgent epidemic. The time to resurgence* $t_{RS} \triangleq t_{min} - t_0$ *satisfies the lower bound*

$$t_{RS} \geq \frac{\log x_{I0} - \log x_I^{min}}{\delta - \beta},$$

where t_0 *is the initial time and* t_{min} *is the time at which* $x_I = x_I^{min}$*. Moreover,* $t_{RS} \to \infty$ *as* $x_{I0} - x_{IC} \to 0_+$.

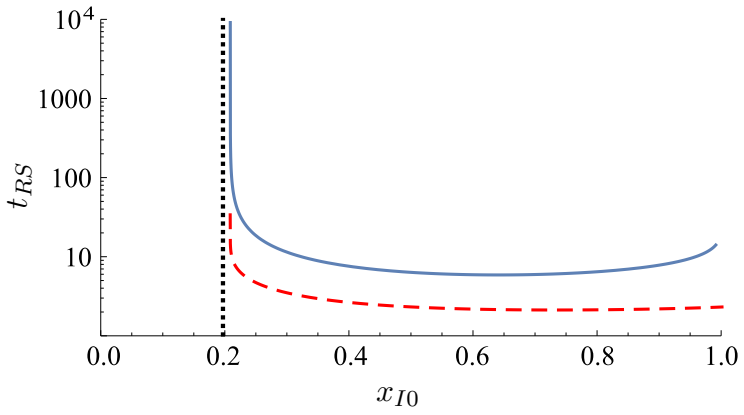

Figure 9.3: Numerical simulations (solid) and lower bound in Theorem 13 (dashed) for time to resurgence t_{RS} versus initial condition x_{I0} for $\beta = 0.5$, $\delta = 1$, $\hat{\beta} = 1.5$. The dotted line shows the critical initial condition $x_{IC} = 0.206$.

Proof. Recall that close to the initial condition, the dynamics of x_I are given by (9.4) with solution $x_I(t) = x_{I0}e^{(\beta-\delta)t}$, where $\beta - \delta < 0$. Setting $x_I = x_I^{min}$ and solving for t, we find the time t_d it takes (9.4) to decay from x_{I0} to x_I^{min}:

$$t_d = \frac{\log x_I^{min} - \log x_{I0}}{\beta - \delta}. \tag{9.16}$$

Along the solution, $x_I(t) \geq x_{I0}e^{-(\delta-\beta)t}$, which implies $t_{RS} \geq t_d$. In the limit $x_{I0} \to x_{IC}$, $\log x_I^{min} \to -\infty$ and $t_d \to \infty$. □

9.5 Conclusions and Future Directions

We have studied the SIRI model for reinfection. We prove that the model has four different behavioral regimes determined by the values R_0 and R_1 that describe the susceptibility of individuals to primary and secondary infections. When both R_0 and R_1 are below or above the critical value of 1, the SIRI model behaves like the SIS model: if $R_0 \leq 1$ and $R_1 \leq 1$ the infection dies out, and if $R_0 > 1$ and $R_1 > 1$ the infection spreads. When $R_0 > 1$ and $R_1 \leq 1$, the SIRI model behaves qualitatively like the SIR model and the infection spreads initially in an epidemic that reaches a

maximum number of infected individuals before dying out. Finally, when $R_0 \leq 1$ and $R_1 > 1$, the model displays bistability in which initial conditions below a critical value lead to an infection-free equilibrium while initial conditions above the critical value lead to the infection spreading through the population. We prove that, in the latter case, solutions exhibit a resurgent epidemic in which the infection decreases at first and reaches a minimum value before rapidly increasing after a long delay.

Possible extensions of the SIRI model include an SIRS-like model in which recovering individuals pass through an additional stage with full immunity before transitioning to the recovered state at a fixed rate.

Our results have implications for spreading processes where individuals adapt after first exposure. Common control strategies focus on preventative measures that seek to minimize the number of exposed individuals. However, the resurgent epidemic phenomenon shows that if reinfections are more likely than primary infections, these control strategies might fail at preventing the spread of the process. More effective control strategies should complement prevention of infection with post-exposure treatment and reinfection prevention.

Our results hold under the restrictive assumption of homogeneous interactions. Although in most cases individuals tend to interact in a non-homogeneous manner with other individuals in the population, our assumption provides invaluable intuition into the dynamics, and the analytical tractability of the model has allowed us to obtain rigorous results. The SIS and SIR models have been successfully adapted to network topologies [69, 74, 79]. In ongoing work we are designing and analyzing SIRI dynamics on networks.

Chapter 10

Adaptive Susceptibility and Heterogeneity in Contagion Models on Networks*

Renato Pagliara and Naomi Ehrich Leonard

Contagion processes, such as the spread of infectious diseases, computer viruses, and social behaviors, form the basis of many biological, societal, and engineering systems. Epidemic models have been successfully used in many cases to understand and control transient and steady-state behaviors in these systems. Prior research has focused mainly on network models with agents who either gain no immunity to the infection after recovery, or gain full immunity after recovery and cannot be reinfected. However, in many systems, agents fall somewhere in between these two extremes; agents adapt after a first exposure to the infection and become either more or less susceptible to reinfection. In this paper, we consider the network SIRI (Susceptible-Infected-Recovered-Infected) model, an epidemic model for the spread of contagious

[1]This chapter is in preparation for submission and appears as Pagliara and Leonard [58].

processes on networks of heterogeneous agents who adapt their susceptibility to infection after first exposure to each of their infected neighbors. We study the network SIRI dynamics in arbitrary strongly connected digraphs and show that the model exhibits four distinct dynamic regimes: infection-free, epidemic, endemic, and bistable. We find necessary and sufficient conditions on model parameters and graph structure for the model to be in each of the regimes. In the bistable regime the infection dies out or spreads depending on initial conditions. Our model presents opportunities for the design and control of multi-agent systems with adaptive susceptibility.

10.1 Introduction

Many engineering, biological, and social systems are the result of contagion processes in which interconnected agents form networks on which information, cultural norms, and/or social behaviors spread. In many cases, "infected" agents, willingly or unwillingly, adapt their susceptibility to reinfection based on their previous experience. For example, in the case of infectious diseases, the susceptibility of individuals to the infection might decrease after a first exposure, resulting in partial immunity as in the case of influenza. Alternatively, the susceptibility of individuals might increase after a first exposure, resulting in a compromised immunity, such as in the case of dengue. Similarly, in the spread of social behaviors, the susceptibility of individuals to the infection might decrease (increase) as a result of a negative (positive) past experience that decreases (increases) the propensity of an individual to engage in the behavior.

Here, we study the role of adaptive susceptibility in the spread of a contagious process in an arbitrary network using the SIRI (Susceptible-Infected-Recovered-Infected) epidemic model for reinfection. The SIRI model describes a group of susceptible agents that become infected by coming in contact with an already infected agent.

Infected agents then recover from the infection at a given rate and become susceptible again. However, unlike the SIS (Susceptible-Infected-Susceptible) model in which there is no difference in susceptibility between a first infection and reinfection, in the SIRI model the susceptibility of agents may increase or decrease irreversibly after a first infection. When every agent in the network drops their susceptibility to reinfection to zero, the SIRI model reduces to the SIR (Susceptible-Infected-Recovered) model. In network settings, the SIRI model describes how agents adapt their susceptibility after a first infection with respect to each of their infected neighbors, leading to heterogeneities in how agents adapt in response to a contagious process.

Epidemic models have been successfully used to study contagious processes in a wide number of systems, ranging from the spread of infectious diseases on populations [139, 182] and memes on social networks [183] to the evolution of riots [52] and power grid failures [19]. The wide applicability of epidemic models has, in turn, lead to an increase in recent years on the number of studies in the controls community focusing on theoretical epidemic models for the propagation of contagious processes in networks [72,74,184,185]. Typically, these models are SIS like models [75,81,135,184,185] in which recovered individuals gain no immunity to the infection, with fewer studies considering SIR like models [78] in which recovered individuals gain full immunity to the infection and cannot become infected again.

Models that consider reinfection usually only consider the case of partial immunity in which the reinfection rate is equal or lower than the infection rate. In the theoretical biology literature, reinfection models have been used to study the role of partial immunity in multi-strain epidemics in very large number of individuals, where every pair of individuals has the same probability of interaction (i.e., under the well-mixed assumption) [134]. In the physics community, reinfection models with partial immunity have been used to study the connection between directed and dynamic percolation in regular lattices [180, 186].

In previous work [57] we analyzed the SIRI model in well-mixed settings and showed that there exist four distinct dynamical regimes characterized by the value of two numbers R_0 and R_1: infection-free, endemic, epidemic, and bistable. R_0 is the ratio of infection rate to recovery rate and R_1 is the ratio of reinfection rate to recovery rate. They are related to the so-called basic reproduction numbers [46]. In the infection-free regime ($R_0, R_1 \leq 1$), all solutions decay monotonically to an infection-free equilibrium. In the endemic regime ($R_0, R_1 > 1$), the infection spreads through the population and all solutions reach an endemic equilibrium corresponding to a nonzero fraction of infected individuals. In the epidemic regime ($R_0 > 1$, $R_1 \leq 1$) the fraction of infected individuals rapidly increases at first before reaching a maximum and then decays towards an infection-free equilibrium. Finally, in the bistable regime ($R_0 \leq 1$, $R_1 > 1$), there exists a critical initial fraction of infected individuals below which the infection dies out and above which the infection spreads through the population and remains endemic. We showed that in the bistable regime there is a resurgent epidemic in which the fraction of infected individuals initially decreases to close to zero before ramping up after an arbitrarily long delay.

In this paper we extend the well-mixed SIRI results to arbitrary strongly connected network topologies with heterogeneous agents. Each agent has its own recovery rate and every ordered pair of agents has its own infection and reinfection rates. Following the individual-based mean-field approach (IBMF) used in the derivation of the N-intertwined SIS model [75,77], we reduce the Markov chain model for the SIRI dynamics on networks to a system of deterministic nonlinear differential equations that can be analytically studied. In [72] the authors studied the Markovian SIRI model on networks with global recovery, infection, and reinfection rates, and showed through numerical simulations that when the reinfection rate is larger than the infection rate, the transition from an infection-free steady-state to an endemic one changes from smooth to abrupt. In the present paper we formalize this observation by proving new

results on the existence of a bistable regime in which a critical manifold of initial conditions separates solutions where the infection dies out from solutions where the infection spreads through the network and remains endemic.

Our contributions in the present paper are as follows. First, we introduce the network SIRI model over strongly connected digraphs and present a rigorous stability analysis of its dynamics under the assumption that the infection can reach every agent in the network. We show that there exists a set of non-isolated infection-free equilibria, and an isolated endemic equilibrium. We find conditions on the graph structure and system parameters for equilibria to be locally stable or unstable. In the case of the set of infection-free equilibria, we find conditions for the set to be split into two or more connected sets of locally stable and unstable points. Second, we study transient and steady-state responses to show that the model exhibits the same four distinct dynamical regimes observed in the well-mixed SIRI model: infection-free, epidemic, endemic, and bistable. We show how the four dynamical regimes are characterized by four numbers that generalize R_0 and R_1 in [57] to network settings and also generalize previous results for the SIS and SIR models in networks [75, 77, 78, 81]. Third, we show that, in the bistable regime, the infection dies out or spreads depending on initial conditions. We show through simulations that when, for every pair of ordered agents in the graph, the reinfection rate is larger or equal than the infection rate, solutions for which the infection spreads exhibit a resurgent epidemic in which the infection initially decreases and reaches a minimum before ramping up and spreading through the network.

The structure of the paper is as follows. In Section 10.2 we present mathematical notation and results that are used throughout the paper. Section 10.3 introduces the network SIRI model and enumerates the different special cases of the model. In Section 10.4 we describe the equilibria of the model and introduce the notion of reproduction numbers. In Section 10.5 we perform a stability analysis under the

assumption that there is always a directed path for the infection to reach all agents in the network, with particular emphasis on the stability properties of the manifold of infection-free equilibria. In Section 10.6 we prove our main result on the existence of the four dynamical regimes. In Section 10.4 we study the geometry of solutions and present analytical and numerical results for the bistable and epidemic regimes. In Section 10.8 we discuss possible control strategies through simulations. We provide closing remarks and future directions in Section 10.9.

10.2 Mathematical Preliminaries

10.2.1 Properties of Gradient Systems

A gradient system on an open set $\Omega \subseteq \mathbb{R}^n$ is a system of the form $\dot{\zeta} = -\nabla V(\zeta)$ where $\zeta(t) \in \Omega$, $V \in C^2(\Omega)$ is the potential function, and $\nabla V = [\partial V/\partial \zeta_1, \ldots, \partial V/\partial \zeta_N]$ is the gradient of V with respect to ζ. The *level surfaces* of V are the subsets $V_c = \{V^{-1}(c) \in \Omega \,|\, c \in \mathbb{R}\}$. A point $\zeta_0 \in \Omega$ is a *regular point* if $\nabla V(\zeta_0) \neq \mathbf{0}$ and a *critical point* if $\nabla V(\zeta_0) = \mathbf{0}$. If $\nabla V(\zeta) \neq \mathbf{0}$ for all $\zeta \in V_c$, then c is a *regular value* for V.

Proposition 6 (Properties of Gradient Systems [187, 188])**.** *Consider the gradient system $\dot{\zeta} = -\nabla V(\zeta)$ where $V \in C^2(\Omega)$, $\zeta(t) \in \Omega \subseteq \mathbb{R}^n$. Then,*

1. *$V(\zeta)$ is a Lyapunov function of the gradient system. Moreover, $\dot{V}(\zeta) = 0$ if and only if ζ is an equilibrium.*

2. *The critical points of V are the equilibria.*

3. *If c is a regular value for V, then the surface set V_c forms an $N-1$ dimensional surface in Ω and the vector field is perpendicular to V_c.*

4. *At every point $\zeta \in \Omega$, the directional derivative along $\mathbf{w} \in \mathbb{R}^N$ is given by $D_{\mathbf{w}}V(\zeta) = \mathbf{w}^T \nabla V(\zeta)$.*

5. *Let ζ_0 be an α-limit point or an ω-limit point of a solution of the gradient system. Then ζ_0 is an equilibrium.*

6. *The linearized system at any equilibrium has only real eigenvalues. No periodic solutions are possible.*

10.3 Network SIRI Model Dynamics

In this section we present the network SIRI model dynamics, which represents a contagious process with reinfection in a population of N agents. Consider a strongly connected digraph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with adjacency matrix A, where each node in $\mathcal{V}$ represents an agent. The state of each agent j is given by the random variable $X_j(t) \in \{S, I, R\}$, where S is "susceptible", I is "infected", and R is "recovered". Let transitions between states for each agent be independent Poisson processes with rates defined as follows. Susceptible agent j becomes infected through contact with infected neighbor k at the rate $\beta_{jk} = f_{jk}(a_{jk}) \geq 0$. We assume $f_{jk}(a_{jk}) = 0 \iff a_{jk} = 0$. Infected agent j recovers from the infection at the rate $\delta_j \geq 0$. Recovered agent j becomes reinfected through contact with infected neighbor k at the rate $\hat{\beta}_{jk} = \hat{f}_{jk}(a_{jk}) \geq 0$, where $\hat{f}_{jk}(0) = 0$. These transitions are summarized as

$$S_j + I_k \xrightarrow{\beta_{jk}} I_j + I_k \xleftarrow{\hat{\beta}_{jk}} R_j + I_k$$

$$I_j \xrightarrow{\delta_j} R_j.$$

The dynamics are described by a continuous-time Markov chain, where the probability that an agent transitions state at time t can depend on the state of its neighbors at time t. Thus, the dimension of the state space can be as large as 3^N.

To reduce the size of the state space, we use an individual mean-field approximation similar to that described in [75] for the SIS model. This approach assumes that the state of every node is statistically independent from the state of its neighbors. The approximation reduces the state of every agent j to the probabilities $p_j^S(t)$, $p_j^I(t)$, and $p_j^R(t)$ of agent j being in state S, I, and R, respectively, at time $t \geq 0$. Since at every time $t \geq 0$, these probabilities sum to 1, the state of every agent j evolves on the 2-simplex $\Delta := \{(p_j^S, p_j^I, p_j^R) \in [0,1]^3 |\ p_j^S + p_j^I + p_j^R = 1\}$. The reduced state space corresponds to N copies of Δ, denoted Δ_N, which has dimension $2N$.

The dynamics retain the full topological structure of the network encoded in the infection and reinfection rates β_{jk} and $\hat{\beta}_{jk}$, which depend on the entries of the adjacency matrix A. We refer the reader to [77] for a detailed derivation of the individual mean-field approximation for the SIS model, and to [135, 136] for a discussion and numerical exploration of the accuracy of mean-field approximations in network dynamics.

Under the individual mean-field approximation, the dynamics of the network SIRI model are given by

$$\begin{aligned}
\dot{p}_j^S &= -\, p_j^S \sum_{k=1}^{N} \beta_{jk} p_k^I \\
\dot{p}_j^I &= -\, \delta_j p_j^I + p_j^S \sum_{k=1}^{N} \beta_{jk} p_k^I + p_j^R \sum_{k=1}^{N} \hat{\beta}_{jk} p_k^I \\
\dot{p}_j^R &= -\, p_j^R \sum_{k=1}^{N} \hat{\beta}_{jk} p_k^I + \delta_j p_k^I.
\end{aligned} \tag{10.1}$$

We can reduce the number of equations from $3N$ to $2N$ by using the substitution $p_j^R = 1 - p_j^S - p_j^I$ in (10.1):

$$\begin{aligned} \dot{p}_j^S &= - p_j^S \sum_{k=1}^{N} \beta_{jk} p_k^I \\ \dot{p}_j^I &= \sum_{k=1}^{N} \big((1 - p_j^S)\hat{\beta}_{jk} + p_j^S \beta_{jk}\big) p_k^I - \delta_j p_j^I - p_j^I \sum_{j=1}^{N} \hat{\beta}_{jk} p_k^I. \end{aligned} \tag{10.2}$$

The dynamics can be written in matrix form where $\mathbf{p}^\Omega = [p_1^\Omega, \cdots, p_N^\Omega]^T$ and $P^\Omega = \text{diag}(\mathbf{p}^\Omega)$ for $\Omega \in \{S, I\}$:

$$\begin{aligned} \dot{\mathbf{p}}^S &= -P^S B \mathbf{p}^I \\ \dot{\mathbf{p}}^I &= \big(B^*(\mathbf{p}^S) - \Gamma\big)\mathbf{p}^I - P^I \hat{B} \mathbf{p}^I, \end{aligned} \tag{10.3}$$

where

$$B^*(\mathbf{p}^S) = (\mathbb{I} - P^S)\hat{B} + P^S B$$

and

$$\begin{aligned} B &= \{\beta_{jk}\} \succ \bar{\mathbf{0}} && \text{(infection matrix)}, \\ \hat{B} &= \{\hat{\beta}_{jk}\} \succeq \bar{\mathbf{0}} && \text{(reinfection matrix)}, \\ \Gamma &= \text{diag}(\delta_1, \ldots, \delta_N) \succeq \bar{\mathbf{0}} && \text{(recovery matrix)}. \end{aligned}$$

Further, we define

$$\bar{B}_{max} = [\max(\beta_{jk}, \hat{\beta}_{jk})], \quad \bar{B}_{min} = [\min(\beta_{jk}, \hat{\beta}_{jk})]. \tag{10.4}$$

The network SIRI model dynamics provide sufficient richness to describe a family of models which can be categorized into six different cases (summarized in Table 10.1):

Table 10.1: Network SIRI model cases.

Case	**Parameter Value**	**Equivalent Model**
1	$\Gamma = \bar{\mathbf{0}}$	SI
2	$\hat{B} = \bar{\mathbf{0}}$	SIR
3	$B = \hat{B}$	SIS
4	$B \succ \hat{B} \succ \bar{\mathbf{0}}$	Partial Immunity
5	$\hat{B} \succ B \succ \bar{\mathbf{0}}$	Compromised Immunity
6	Otherwise	Mixed Immunity

- Case 1 (SI): When $\Gamma = \bar{\mathbf{0}}$ the network SIRI model specializes to the network SI model.
- Case 2 (SIR): When $\hat{B} = \bar{\mathbf{0}}$, the network SIRI model specializes to the network SIR model.
- Case 3 (SIS): When $B = \hat{B}$ the network SIRI model specializes to the network SIS model with $\mathbf{p}^S \mapsto \mathbf{p}^S + \mathbf{p}^R$.
- Case 4 (Partial Immunity): When $B \succ \hat{B} \succ \bar{\mathbf{0}}$, every recovered agent acquires partial (or no) immunity to each of its infected neighbors.
- Case 5 (Compromised Immunity): When $\hat{B} \succ B \succ \bar{\mathbf{0}}$, every recovered agent acquires compromised (or no) immunity to each of its infected neighbors.
- Case 6 (Mixed Immunity): Models not in Cases 1-5. Notably, there is at least one pair of edges (j,k) and (l,m) such that $\beta_{jk} \geq \hat{\beta}_{jk}$ and $\beta_{lm} < \hat{\beta}_{lm}$. We classify mixed immunity into two sub-cases:
  - Case 6a (Weak Mixed Immunity): For every agent j, $\beta_{jk} - \hat{\beta}_{jk} \geq 0$ for all $k \in \mathcal{N}_j$ or $\beta_{jk} - \hat{\beta}_{jk} \leq 0$ for all $k \in \mathcal{N}_j$.
  - Case 6b (Strong Mixed Immunity): Mixed immunity that is not weak.

10.4 Equilibria and Reproduction Numbers

In this section we analyze the equilibria of the network SIRI model dynamics and define the notion of basic and extreme basic reproduction numbers, which we use throughout the rest of the paper. We denote the value of $\mathbf{p}^S$ and $\mathbf{p}^I$ at equilibrium as $\mathbf{p}^{S*}$ and $\mathbf{p}^{I*}$, respectively.

10.4.1 Equilibria

Proposition 7. *The only equilibria of the network SIRI model* (10.3) *are an invariant set of infection-free equilibria (IFE)* $\mathcal{M} = \{(\mathbf{p}^{S*}, \mathbf{0}) \in \Delta_N |\ \mathbf{0} \preceq \mathbf{p}^{S*} \preceq \mathbf{1}\}$ *and one or more isolated endemic equilibria (EE) where* $\mathbf{p}^{I*} \succ \mathbf{0}$ *satisfy*

$$p_j^{I*} = \frac{\sum_{k=1}^{N} \hat{\beta}_{jk} p_k^{I*}}{\delta_j + \sum_{k=1}^{N} \hat{\beta}_{jk} p_k^{I*}}. \tag{10.5}$$

If $\hat{B}$ *is irreducible then* $\mathbf{p}^{S*} = \mathbf{0}$ *and* $\mathbf{p}^{I*} \gg \mathbf{0}$ *for every EE.*

Proof. Setting $\dot{\mathbf{p}}^S = \mathbf{0}$ in (10.3), we get $P^{S*} B \mathbf{p}^{I*} = \mathbf{0}$. Since $\mathcal{G}$ is strongly connected and B preserves the connectivity of A, then for every agent j we must have $\mathbf{p}_j^{S*} = 0$ or $\sum_{\mathcal{N}_j} \mathbf{p}_k^{I*} = 0$. Moreover, since $\hat{B}$ has a zero at every entry where B has a zero, it follows that $P^{S*} \hat{B} \mathbf{p}^{I*} = \mathbf{0}$.

Setting $\dot{\mathbf{p}}^I = \mathbf{0}$ in (10.3) and using $P^{S*} B \mathbf{p}^{I*} = P^{S*} \hat{B} \mathbf{p}^{I*} = \mathbf{0}$ we get

$$\mathbf{0} = (\hat{B} - \Gamma - P^{I*} \hat{B}) \mathbf{p}^{I*} = (\hat{B} - \Gamma - \mathrm{diag}(\hat{B} \mathbf{p}^{I*})) \mathbf{p}^{I*}. \tag{10.6}$$

One solution is the invariant set $\mathcal{M} = \{(\mathbf{p}^{S*}, \mathbf{0}) \in \Delta_N |\ \mathbf{0} \preceq \mathbf{p}^{S*} \preceq \mathbf{1}\}$. The only other solutions are isolated equilibria $\mathbf{p}^{I*} \succ \mathbf{0}$ satisfying (10.5).

If $\hat{B}$ is irreducible then $\beta_{jk} > 0$ for any $(j,k) \in \mathcal{E}$, and if $p_k^{I*} > 0$ for any $k \in \mathcal{V}$, it follows from (10.5) that $p_j^{I*} > 0$ for any j where $k \in \mathcal{N}_j$. This in turn implies that

$p_i^{I*} > 0$ for any i where $j \in \mathcal{N}_i$. This argument can be recursively applied until all nodes in $\mathcal{G}$ are covered. Moreover, since $P^{S*}B\mathbf{p}^{I*} = \mathbf{0}$, $\mathbf{p}^{I*} \gg \mathbf{0}$ implies $\mathbf{p}^{S*} = \mathbf{0}$. $\square$

Definition 10.4.1. *The boundary of* $\mathcal{M}$ *is* $\partial\mathcal{M} = \{\mathbf{x} = (\mathbf{p}^{S*}, \mathbf{0}) \in \mathcal{M} \,|\, \exists j,\, p_j^{S*} \in \{0,1\}\}$. *The corner set of* $\mathcal{M}$ *is* $\hat{\mathcal{M}} = \{\mathbf{x} = (\mathbf{p}^{S*}, \mathbf{0}) \in \partial\mathcal{M} \,|\, p_j^{S*} \in \{0,1\},\, \forall j\}$. *The interior of* $\mathcal{M}$ *is* $\text{int}\,(\mathcal{M}) = \mathcal{M} \setminus \partial\mathcal{M}$.

Remark 15. The equilibria of the network SIRI model are equivalent to the equilibria of the network SIS model (Case 3), where $B = \hat{B}$. This follows since any equilibrium of (10.3) satisfies (10.6), and therefore is also an equilibrium of the network SIS dynamics [77, 81, 139]:

$$\dot{\mathbf{p}}^I = (B - \Gamma)\mathbf{p}^I - P^I B \mathbf{p}^I. \tag{10.7}$$

For the network SI model (Case 1), the only equilibrium is a unique EE with $\mathbf{p}^{I*} = \mathbf{1}$ and $\mathbf{p}^{S*} = \mathbf{0}$. For the network SIR model (Case 2), the only equilibria are the IFE set $\mathcal{M}$.

Remark 16. For initial conditions $\mathbf{p}^S(0) = \mathbf{1} - \mathbf{p}^I(0)$, the network SIRI dynamics (10.3) initially behave as the network SIS model (10.7) with infection matrix B. As agents become exposed to the infection for the first time, the dynamics transition to network SIS dynamics with infection matrix $\hat{B}$. This model of adaptive susceptibility of agents can be used in other contexts, including the spread of behaviors, such as the regulation of foraging by desert harvester ants [43].

In the remainder of this paper, we assume Γ is non-singular and that $\hat{B}$ is irreducible, and thus every EE is *strong* since $\mathbf{p}^{I*} \gg \mathbf{0}$. The generalization to the case of reducible $\hat{B}$ is straightforward.[2]

[2]If $\hat{B}$ is reducible, then the graph $\mathcal{G}_{\hat{B}}$ generated by interpreting $\hat{B}$ as its adjacency matrix is weakly connected or disconnected. If $\mathcal{G}_{\hat{B}}$ is weakly connected, the adjacency matrix of $\mathcal{G}_{\hat{B}}$ can always be written as an upper block triangular matrix with K diagonal irreducible blocks which describe the K strongly connected subgraphs of $\mathcal{G}$ [138]. If $\mathcal{G}_{\hat{B}}$ is disconnected, it is sufficient to study each connected subgraph of $\mathcal{G}_{\hat{B}}$.

10.4.2 Basic Reproduction Numbers

In the network SIS model (10.7), the steady-state behavior of solutions depends on the value of the basic reproduction number $R_0 = \rho(B\Gamma^{-1})$. If $R_0 \leq 1$, solutions reach the IFE set $\mathcal{M}$ as $t \to \infty$ while if $R_0 > 1$, solutions reach the EE (10.5) as $t \to \infty$ [81, 139]. A key concept in epidemiology, the basic reproduction number is the expected number of new cases of infection caused by a typical infected individual in a population of susceptible individuals [46, 68].

In previous work [57] we proved that, in well-mixed settings, the transient and steady-state behavior of solutions in the SIRI model depend on two numbers R_0 and R_1, corresponding to the basic reproduction number for a population of susceptible individuals and for a population of recovered individuals, respectively. Here we extend the definition of R_0 and R_1 in [57] to network topologies and introduce the notion of extreme basic reproduction numbers.

Definition 10.4.2 (Basic Reproduction Numbers)**.** *The basic infection reproduction number is* $R_0 = \rho(B\Gamma^{-1})$ *and the basic reinfection reproduction number is* $R_1 = \rho(\hat{B}\Gamma^{-1})$.

Definition 10.4.3 (Extreme Basic Reproduction Numbers)**.** *Let* $B^*(\mathbf{p}^S) = (\mathbb{I} - P^S)\hat{B} + P^S B$. *Maximum basic reproduction number is* $R_{max} = \max_{\mathbf{p}^S} \rho(B^*(\mathbf{p}^S)\Gamma^{-1})$ *and minimum basic reproduction number is* $R_{min} = \min_{\mathbf{p}^S} \rho(B^*(\mathbf{p}^S)\Gamma^{-1})$.

Remark 17. Each reproduction number corresponds to $K = B^*(\mathbf{p}^S) - \Gamma$ for a particular value of $\mathbf{p}^S$, where K is the linear term of the dynamics of $\mathbf{p}^I$ in (10.3). K is a Metzler matrix and each reproduction number R is the spectral radius associated with this regular splitting, i.e., $R = \rho(B^*(\mathbf{p}^S)\Gamma^{-1})$ as defined in Proposition 3. R_0 is R for $P^S = \mathbb{I}$, R_1 is R for $P^S = \bar{\mathbf{0}}$, R_{max} is R for $\mathbf{p}^S = \text{argmax}_{\mathbf{p}^S} \rho(B^*(\mathbf{p}^S)\Gamma^{-1})$ and R_{min} is R for $\mathbf{p}^S = \text{argmin}_{\mathbf{p}^S} \rho(B^*(\mathbf{p}^S)\Gamma^{-1})$.

Proposition 8. *Let* $\bar{R}_{max} = \rho(\bar{B}_{max}\Gamma^{-1})$ *and* $\bar{R}_{min} = \rho(\bar{B}_{min}\Gamma^{-1})$. *Then,*

$$\bar{R}_{min} \leq R_{min} \leq \lambda_{max}(B^*(\mathbf{p}^S)\Gamma^{-1}) \leq R_{max} \leq \bar{R}_{max} \tag{10.8}$$

for any $\mathbf{0} \preceq \mathbf{p}^S \preceq \mathbf{1}$. *If* $B \succeq \hat{B}$, *then* $\bar{R}_{max} = R_{max} = R_0$ *and* $\bar{R}_{min} = R_{min} = R_1$. *If* $\hat{B} \succeq B$, *then* $\bar{R}_{max} = R_{max} = R_1$, *and* $\bar{R}_{min} = R_{min} = R_0$.

Proof. Any matrix with nonnegative entries is Metzler. Thus, $Y(\mathbf{p}^S) = B^*(\mathbf{p}^S)\Gamma^{-1}$ is an irreducible Metzler matrix since $B \succ \bar{\mathbf{0}}$ and $\hat{B} \succeq \bar{\mathbf{0}}$ are irreducible. By Proposition 1, $\lambda_{max}(Y(\mathbf{p}^S))$ increases (decreases) as any entry in $Y(\mathbf{p}^S)$ increases (decreases). Since every non-zero entry of $Y(\mathbf{p}^S)$ is a scaled convex sum of β_{jk} and $\hat{\beta}_{jk}$, it follows that $\bar{B}_{min}\Gamma^{-1} \preceq Y(\mathbf{p}^S) \preceq \bar{B}_{max}\Gamma^{-1}$ for any $\mathbf{0} \preceq p^S \preceq \mathbf{1}$. Consequently, (10.8) holds for any $\mathbf{0} \preceq \mathbf{p}^S \preceq \mathbf{1}$. If $B \succeq \hat{B}$, then $\bar{B}_{max} = B$ and $\bar{B}_{min} = \hat{B}$. If $\hat{B} \succeq B$, then $\bar{B}_{max} = \hat{B}$ and $\bar{B}_{min} = B$. The stated results then follow from the definitions for the extreme basic and basic reproduction numbers. □

10.5 Stability of equilibria

In this section we prove conditions for the local stability of the EE and of points in the IFE set $\mathcal{M}$.

10.5.1 Stability of the Endemic Equilibria

Proposition 9. *The network SIRI dynamics* (10.3) *have a unique EE if and only if* $R_1 > 1$. *The EE is locally stable.*

Proof. Proof of the existence and uniqueness of the EE if and only if $R_1 > 1$ for (10.3) follows from the proof in Section 2.2 of [81] for the network SIS model (10.7) (see Remark 15).

Recall that B and $\hat{B}$ are irreducible, so $\mathbf{p}^{S*} = \mathbf{0}$. To prove local stability, we compute the Jacobian of (10.3) at the EE (10.5):

$$J_{EE} = \begin{bmatrix} -\text{diag}(B\mathbf{p}^{I*}) & \bar{\mathbf{0}} \\ \text{diag}((B - \hat{B})\mathbf{p}^{I*}) & J_a \end{bmatrix} \tag{10.9}$$

where $J_a = \hat{B} - \Gamma - P^{I*}\hat{B} - \text{diag}(\hat{B}\mathbf{p}^{I*})$. Since $-\text{diag}(B\mathbf{p}^{I*})$ is Hurwitz, showing that the EE is locally stable is equivalent to showing that the Metzler matrix J_a is Hurwitz.

By (10.6), $(\hat{B} - \Gamma - P^{I*}\hat{B})\mathbf{p}^{I*} = 0$, and thus

$$J_a\mathbf{p}^{I*} = -\text{diag}(\hat{B}\mathbf{p}^{I*})\mathbf{p}^{I*} \ll \mathbf{0}. \tag{10.10}$$

where the inequality follows from $\mathbf{p}^{I*} \gg \mathbf{0}$.

By Proposition 2 we conclude that J_a is Hurwitz. □

10.5.2 Stability of Infection-Free Equilibria

In this section we prove results on the stability of the IFE set $\mathcal{M}$. The equilibria in $\mathcal{M}$ are non-hyperbolic: the Jacobian of (10.3) at a point $\mathbf{x} \in \mathcal{M}$ has N zero eigenvalues corresponding to the N-dimensional space tangent to $\mathcal{M}$. The remaining N eigenvalues are called *transverse* as they correspond to the N-dimensional space transverse to $\mathcal{M}$. The Shoshitaishvili Reduction Principle [90], which extends the Hartman-Grobman Theorem to non-hyperbolic equilibria, can be used to study the local stability of points in $\mathcal{M}$ in terms of the transverse eigenvalues of the Jacobian and the dynamics on the center manifold. We show how the irreducibility of B and $\hat{B}$ imply that the behavior of solutions in Δ_N close to a point $\mathbf{x} \in \mathcal{M}$ depends only on the sign of the leading transverse eigenvalue of the Jacobian at $\mathbf{x}$.

Throughout the rest of this paper, we consider the topological space Δ_N as a subspace of $\mathbb{R}^{2N}$. This allows us to study points in $\partial\mathcal{M}$ and in $\text{int}(\mathcal{M})$ simultaneously. In $\mathbb{R}^{2N}$, the invariant set $\mathcal{M}$ of IFE points becomes a subset of the invariant manifold of equilibria $\mathcal{M}' = \{(\mathbf{p}, \mathbf{0}) | \mathbf{p} \in \mathbb{R}^N\}$.

Lemma 13 (Local Stability of Points in the IFE set $\mathcal{M}$)**.** *Let $\mathbf{x} = (\mathbf{p}^{S*}, \mathbf{0}) \in \mathcal{M}$. Let $J_{\mathcal{M}}(\mathbf{x})$ be the Jacobian of (10.3) at $\mathbf{x}$ and $\lambda_{Tmax}(J_{\mathcal{M}}(\mathbf{x}))$ the leading transverse eigenvalue of $J_{\mathcal{M}}(\mathbf{x})$. Then, $\lambda_{Tmax}(J_{\mathcal{M}}(\mathbf{x})) \in \mathbb{R}$ and the following hold.*

- *Suppose $\lambda_{Tmax}(J_{\mathcal{M}}(\mathbf{x})) < 0$. Then, $\mathbf{x}$ is locally stable. I.e., given a neighborhood U of $\mathbf{x}$ on $\mathcal{M}'$ such that $\lambda_{Tmax}(J_{\mathcal{M}}(\mathbf{u})) < 0$ for all $\mathbf{u} \in U$, there exists $V \subset \Delta_N$ and $\mathbf{x} \in V$ such that any solution starting in V converges exponentially to a point in $U \cap \Delta_N$.*
- *Suppose $\lambda_{Tmax}(J_{\mathcal{M}}(\mathbf{x})) > 0$. Then, $\mathbf{x}$ is unstable. I.e., there exists $W \subset \Delta_N$ and $\mathbf{x} \in W$, such that any solution starting in W leaves W.*

Proof. For an arbitrary point $\mathbf{x} = (\mathbf{p}^{S*}, \mathbf{0}) \in \mathcal{M}$,

$$J_{\mathcal{M}}(\mathbf{x}) = \begin{bmatrix} \bar{\mathbf{0}} & -P^{S*}B \\ \bar{\mathbf{0}} & J_T(\mathbf{p}^{S*}) \end{bmatrix} \tag{10.11}$$

where $J_T(\mathbf{p}^{S*}) = B^*(\mathbf{p}^{S*}) - \Gamma$. The N transverse eigenvalues of $J_{\mathcal{M}}(\mathbf{x})$ are the eigenvalues of $J_T(\mathbf{p}^{S*})$ and so $\lambda_{Tmax}(J_{\mathcal{M}}(\mathbf{x})) = \lambda_{max}(J_T(\mathbf{p}^{S*}))$. The matrix $J_T(\mathbf{p}^{S*})$ is Metzler irreducible since B and $\hat{B}$ are Metzler irreducible. By Proposition 1 $\lambda_{max}(J_T(\mathbf{p}^{S*})) \in \mathbb{R}$.

Consider an arbitrary point $\mathbf{x}' = (\mathbf{p}', \mathbf{0}) \in \mathcal{M}' \setminus \mathcal{M}$. The Jacobian of (10.3) at $\mathbf{x}'$ takes on the same form as (10.11), and $J_T(\mathbf{p}')$ is Metzler irreducible if every entry of $\mathbf{p}'$ satisfies

$$\begin{cases} p'_j > -\frac{\sum_{j=1}^N \hat{\beta}_{jk}}{\sum_{j=1}^N (\beta_{jk} - \hat{\beta}_{jk})} & \text{if } \sum_{\mathrm{j}=1}^{\mathrm{N}} (\beta_{\mathrm{jk}} - \hat{\beta}_{\mathrm{jk}}) \geq 0 \\ p'_j < \frac{\sum_{j=1}^N \hat{\beta}_{jk}}{\sum_{j=1}^N (\hat{\beta}_{jk} - \beta_{jk})} & \text{if } \sum_{\mathrm{j}=1}^{\mathrm{N}} (\beta_{\mathrm{jk}} - \hat{\beta}_{\mathrm{jk}}) \leq 0. \end{cases}$$

Since $B, \hat{B}$ are irreducible, $\mid \sum_{j=1}^{N} \hat{\beta}_{jk} / \sum_{j=1}^{N} (\beta_{jk} - \hat{\beta}_{jk}) \mid > 1$ for all j. So, for any $\mathbf{x} \in \partial\mathcal{M}$, there exists a neighborhood $\bar{U}$ of $\mathbf{x}$ on $\mathcal{M}'$ such that $J_T(\bar{\mathbf{u}}')$ is Metzler irreducible for every $\bar{\mathbf{u}} = (\bar{\mathbf{u}}', \mathbf{0}) \in \bar{U}$. By Proposition 1 $\lambda_{max}(J_T(\mathbf{p}')) \in \mathbb{R}$.

Let U be a neighborhood of $\mathbf{x}$ on $\mathcal{M}'$ such that $\lambda_{Tmax}(J_{\mathcal{M}}(\mathbf{u}))$ has the same sign as $\lambda_{Tmax}(J_{\mathcal{M}}(\mathbf{x}))$ for all $\mathbf{u} \in U$. Then, $\lambda_{max}(J_T(\mathbf{u}'))$ has the same sign as $\lambda_{max}(J_T(\mathbf{p}^{S*}))$ for all $\mathbf{u} = (\mathbf{u}', \mathbf{0}) \in U$. By Proposition 1 every left and right eigenvector of every eigenvalue of $J_T(\mathbf{u}')$, other than $\lambda_{max}(J_T(\mathbf{u}'))$, contains at least one negative entry. Thus, for any $\bar{\mathbf{u}} \in U \cap \Delta_N$, the eigenvector corresponding to $\lambda_{Tmax}(J_{\mathcal{M}}(\bar{\mathbf{u}}))$ lies in Δ_N, and the eigenvectors corresponding to the other $N-1$ transverse eigenvalues lie outside Δ_N.

If $\lambda_{max}(J_T(\mathbf{p}^{S*})) < 0$, then every transverse eigenvalue of $J_{\mathcal{M}}(\mathbf{x})$ has negative real part. By the Shoshitaishvili Reduction Principle, there exists a neighborhood $V' \in \mathbb{R}^{2N}$ of $\mathbf{x}$ that is positively invariantly foliated by a family of stable manifolds corresponding to the family of stationary solutions in U (see [86,189,190]), each stable manifold spanned by the (generalized) eigenvectors associated with the N negative transverse eigenvalues of $J_{\mathcal{M}}(\mathbf{u})$. Let $V = V' \cap \Delta_N$. Then $V \subset \Delta_N$ is positively invariantly foliated by a family of stable manifolds. The invariance of Δ_N implies each of these stable manifolds correspond to a point $\bar{\mathbf{u}} \in U \cap \mathcal{M}$. Thus, any solution starting in V converges exponentially along a stable manifold to the corresponding stationary solution in $U \cap \mathcal{M}$.

If $\lambda_{max}(J_T(\mathbf{p}^{S*})) > 0$, then there is at least one transverse eigenvalue of $J_{\mathcal{M}}(\mathbf{x})$ with positive real part. The trace of $J_T(\mathbf{p}^{S*})$ is negative for any $\mathbf{0} \preceq \mathbf{p}^{S*} \preceq \mathbf{1}$, so the sum of the eigenvalues of $J_T(\mathbf{p}^{S*})$ is always negative and $J_{\mathcal{M}}(\mathbf{x})$ has at least one transverse eigenvalue with negative real part. By the Shoshitaishvili Reduction Principle there exists a neighborhood $W' \subset \mathbb{R}^{2N}$ of $\mathbf{x}$ that is positively invariantly foliated by a family of stable, unstable, and possibly center manifolds corresponding to the family of stationary solutions in U. Let $W = W' \cap \Delta_N$. Then, the stable and

center manifolds of each stationary solution $\bar{\mathbf{u}} \in W \cap \mathcal{M}$ lie outside Δ_N. Thus, no solution starting in W can remain in W for all time, i.e., any solution starting in W leaves W. □

Definition 10.5.1 (Stable, unstable, and center IFE subsets)**.** *The stable IFE subset is* $\mathcal{M}_- = \{\mathbf{x} \in \mathcal{M} \,|\, \lambda_{Tmax}(J_{\mathcal{M}}(\mathbf{x})) < 0\}$. *The unstable IFE subset is* $\mathcal{M}_+ = \{\mathbf{x} \in \mathcal{M} \,|\, \lambda_{Tmax}(J_{\mathcal{M}}(\mathbf{x})) > 0\}$. *The center IFE subset is* $\mathcal{M}_0 = \{\mathbf{x} \in \mathcal{M} \,|\, \lambda_{Tmax}(J_{\mathcal{M}}(\mathbf{x})) = 0\}$.

Proposition 10. $\mathcal{M}_- \cup \mathcal{M}_+ \cup \mathcal{M}_0 = \mathcal{M}$. *Every point in* $\mathcal{M}_-$ *is locally stable and every point in* $\mathcal{M}_+$ *is unstable.*

Proof. This follows from Definition 10.5.1 and Lemma 13. □

We now state two main theorems of the paper, which relate the extreme basic reproduction numbers R_{max} and R_{min} to the stable, unstable, and center subsets of the IFE set $\mathcal{M}$.

Theorem 14 (Stability of the IFE set $\mathcal{M}$)**.**

(A) If $R_{max} < 1$, *then* $\mathcal{M}_- = \mathcal{M}$.

(B) If $R_{min} > 1$, *then* $\mathcal{M}_+ = \mathcal{M}$.

(C) If $R_{max} = R_{min} = 1$, *then* $\mathcal{M}_0 = \mathcal{M}$.

(D) If $R_{min} < R_{max} = 1$, *then* $\mathcal{M}_- = \mathcal{M} \setminus \mathcal{M}_0$ *and* $\mathcal{M}_0 \subset \partial\mathcal{M}$.

(E) If $R_{max} > R_{min} = 1$, *then* $\mathcal{M}_+ = \mathcal{M} \setminus \mathcal{M}_0$ *and* $\mathcal{M}_0 \subset \partial\mathcal{M}$.

(F) If $R_{max} > 1$ *and* $R_{min} < 1$, *then* $\mathcal{M}_-, \mathcal{M}_+, \mathcal{M}_0 \neq \emptyset$ *and each subset consists of* n_-, n_+, n_0 *connected sets, respectively. Each of the center connected sets* $\mathcal{M}_0^j$, $j = 1, \ldots, n_0$, *is an* $N-1$*-dimensional smooth hypersurface with boundary* $\partial\mathcal{M}_0^j \subset \partial\mathcal{M}$. *Each* $\mathcal{M}_0^j$ *separates an* N*-dimensional stable connected hypervolume from an* N*-dimensional unstable connected hypervolume.*

Remark 18. Theorem 14 applies to the six different cases of the network SIRI model (see Table 10.1) as follows. (A) applies to Cases 2, 3, 4, 5, and 6. (B) applies to Cases 3, 4, 5, and 6. (C) applies to Case 3. (D) applies to Cases 2, 4, 5, and 6. (E) applies to Cases 4, 5, and 6. (F) applies to Cases 2, 4, 5, and 6. We specialize (F) in Theorem 15 to provide the key to characterizing global behavior in Cases 2, 4, 5, and 6a.

Theorem 15 (Uniqueness of stable, unstable, and center subsets)**.** *If $R_{max} > 1$ and $R_{min} < 1$, then for Case 2 (SIR), Case 4 (partial immunity), Case 5 (compromised immunity), and Case 6a (weak mixed immunity), $\mathcal{M}_0$ consists of a unique $N-1$-dimensional surface with boundary $\partial\mathcal{M}_0 \subset \partial\mathcal{M}$ dividing $\mathcal{M}$ into $\mathcal{M}_-$ and $\mathcal{M}_+$.*

Remark 19. We conjecture that Theorem 15 can be extended to Case 6b. Extensive computations of $\mathcal{M}_0$, $\mathcal{M}_-$, and $\mathcal{M}_+$, for an $N = 3$ agent network, with different network configurations and parameter values, consistently show a unique connected surface $\mathcal{M}_0$ dividing $\mathcal{M}$ into $\mathcal{M}_-$ and $\mathcal{M}_+$.

The proofs of Theorem 14 and 15 make use of the following definitions and lemmas.

Definition 10.5.2 (Neighborhood $E \subset \mathcal{M}'$ of $\mathcal{M}$)**.** *Neighborhood $E \subset \mathcal{M}'$ of $\mathcal{M}$ is the union of $\mathcal{M}$ and the neighborhoods $\bar{U} \subset \mathcal{M}'$ of every $\bar{\mathbf{x}} \in \partial\mathcal{M}$ described in the proof of Lemma 13.*

Definition 10.5.3 (J_T and Λ)**.** *Let $J_T(\mathbf{p}) = B^*(\mathbf{p}) - \Gamma$ for $(\mathbf{p}, \mathbf{0}) \in E$. The λ_{Tmax} function is $\Lambda : E \to \mathbb{R}$, $(\mathbf{p}, \mathbf{0}) \mapsto \lambda_{max}(J_T(\mathbf{p}))$. For ease of notation we use $\Lambda(\mathbf{p})$ for $\Lambda(\mathbf{p}, \mathbf{0})$.*

Definition 10.5.4 (Stubborn agents)**.** *An agent $j \in \mathcal{V}$ is stubborn if $(\beta_{jk} - \hat{\beta}_{jk}) = 0$ for all $k \in \mathcal{N}_j$.*

Lemma 14 (IFE subsets as level surfaces of Λ)**.** *Let $\Lambda_c = \{\Lambda^{-1}(c) \,|\, c \in \mathbb{R}\}$ be the level surface of Λ on $E \subset \mathcal{M}'$ corresponding to $c \in \mathbb{R}$. Then, $\mathcal{M}_0 = \Lambda_0 \cap \mathcal{M}$, $\mathcal{M}_- = \bigcup_{c<0} \Lambda_c \cap \mathcal{M}$ and $\mathcal{M}_+ = \bigcup_{c>0} \Lambda_c \cap \mathcal{M}$.*

Proof. This follows from Definitions 10.5.1 and 10.5.3. □

Lemma 15 (Gradient of Λ)**.** *For $\acute{\mathbf{x}} = (\mathbf{p}, \mathbf{0}) \in E$, let $\mathbf{w}^T, \mathbf{v} \in \mathbb{R}^N$ be left and right eigenvectors of $J_T(\mathbf{p})$ for $\Lambda(\mathbf{p})$. Then, Λ is smooth on E, i.e., $\Lambda(\cdot) \in C^\infty(E)$, with partial derivatives*

$$\frac{\partial \Lambda}{\partial p_j}(\mathbf{p}) = w_j \sum_{k=1}^{N} (\beta_{jk} - \hat{\beta}_{jk}) v_k \tag{10.12}$$

and gradient

$$\nabla \Lambda(\mathbf{p}) = \operatorname{diag}(\mathbf{w})(B - \hat{B})\mathbf{v}. \tag{10.13}$$

In addition, the following hold.

- *If $B = \hat{B}$, all points $\acute{\mathbf{x}} \in E$ are critical points of Λ.*
- *If j is a stubborn agent, $(\partial \Lambda / \partial p_j)(\mathbf{p}) = 0$ for all $\acute{\mathbf{x}} \in E$.*
- *If $B \succ \hat{B}$, and there are no stubborn agents in $\mathcal{G}$, $\nabla \Lambda(\mathbf{p}) \gg \mathbf{0}$ for all $\acute{\mathbf{x}} \in E$, and Λ has no critical points.*
- *If $\hat{B} \succ B$, and there are no stubborn agents in $\mathcal{G}$, $\nabla \Lambda(\mathbf{p}) \ll \mathbf{0}$ for all $\acute{\mathbf{x}} \in E$ and Λ has no critical points.*
- *If $B \neq \hat{B}$ then either Λ has no critical points or all points $\acute{\mathbf{x}} \in E$ are critical points of Λ.*

Proof. Let $\acute{\mathbf{x}} = (\mathbf{p}, \mathbf{0}) \in E$. By the proof of Lemma 13, $J_T(\mathbf{p})$ is Metzler irreducible. By Proposition 1, $\lambda_{max}(J_T(\mathbf{p})) \in \mathbb{R}$ and has multiplicity one. Thus by [191, p. 66-67], $\Lambda(\cdot) \in C^\infty(E)$.

Differentiating the right-eigenvector equation $J_T(\mathbf{p})\mathbf{v} = \Lambda(\mathbf{p})\mathbf{v}$ with respect to p_j, and premultiplying by $\mathbf{w}^T$ we get

$$\mathbf{w}^T \frac{\partial J_T}{\partial p_j}\mathbf{v} + \mathbf{w}^T J_T \frac{\partial \mathbf{v}}{\partial p_j} = \mathbf{w}^T \mathbf{v} \frac{\partial \Lambda}{\partial p_j} + \mathbf{w}^T \Lambda \frac{\partial \mathbf{v}}{\partial p_j},$$

where all terms are evaluated at $\mathbf{p}$, $\partial J_T/\partial p_j(\mathbf{p}) = \{z_{kl}\}$ and

$$z_{kl} = \begin{cases} \beta_{kl} - \hat{\beta}_{kl} & k = j \text{ and } (k,l) \in \mathcal{E}, \\ 0 & \text{otherwise.} \end{cases}$$

Because $\Lambda(\mathbf{p})$ has multiplicity one, we can always pick $\mathbf{w}$ such that $\mathbf{w}^T\mathbf{v} = 1$. Using $\mathbf{w}^T J_T(\mathbf{p}) = \mathbf{w}^T\Lambda(\mathbf{p})$, we obtain

$$\frac{\partial \Lambda}{\partial p_j}(\mathbf{p}) = \mathbf{w}^T \frac{\partial J_T}{\partial p_j}(\mathbf{p})\mathbf{v} = w_j \sum_{k=1}^{N} (\beta_{jk} - \hat{\beta}_{jk}) v_k. \tag{10.14}$$

In vector form, (10.14) becomes $\nabla\Lambda(\mathbf{p}) = \text{diag}(\mathbf{w})(B - \hat{B})\mathbf{v}$.

By Proposition 1 $\mathbf{w}, \mathbf{v} \gg \mathbf{0}$. If $B = \hat{B}$, then $\nabla\Lambda(\mathbf{p}) = \mathbf{0}$ for all $\acute{\mathbf{x}} \in E$. If j is a stubborn agent, then $\beta_{jk} - \hat{\beta}_{jk} = 0$ for all $k \in \mathcal{N}_j$. Therefore, $(\partial\Lambda/\partial p_j)(\mathbf{p}) = 0$ for all $\acute{\mathbf{x}} \in E$. If $B \succ \hat{B}$ and there are no stubborn agents in $\mathcal{G}$, then $\sum_{k=1}^{N}(\beta_{jk} - \hat{\beta}_{jk})v_k > 0$ for any j. Therefore, $\nabla\Lambda(\mathbf{p}) \gg \mathbf{0}$ for all $\acute{\mathbf{x}} \in E$. A similar argument is made if $\hat{B} \succ B$, and there are no stubborn agents in $\mathcal{G}$.

By Proposition 6 and (10.13), $\acute{\mathbf{x}}$ is a critical point of Λ if and only if $(B - \hat{B})\mathbf{v} = \mathbf{0}$. Let $\acute{\mathbf{x}}_0 = (\mathbf{p_0}, \mathbf{0}) \in E$ be a critical point, then $(B - \hat{B})\mathbf{v_0} = \mathbf{0}$. Thus, for any $\mathbf{p}$, we have

$$\begin{aligned} J_T(\mathbf{p})\mathbf{v_0} &= (B^*(\mathbf{p}) - \Gamma)\mathbf{v_0} \\ &= (\hat{B} - \Gamma)\mathbf{v_0} + \text{diag}(\mathbf{p})(B - \hat{B})\mathbf{v_0} \\ &= (\hat{B} - \Gamma)\mathbf{v_0}. \end{aligned} \tag{10.15}$$

In particular, $(\hat{B} - \Gamma)\mathbf{v_0} = \Lambda(\mathbf{p_0})\mathbf{v_0}$. Hence, $(\Lambda(\mathbf{p_0}), \mathbf{v_0})$ is an eigenpair of $\hat{B} - \Gamma$ and by (10.15) also an eigenpair of $J_T(\mathbf{p})$ for any $\mathbf{p}$. Since $\Lambda(\mathbf{p})$ has multiplicity one and $\mathbf{v}$ is the only eigenvector satisfying $\mathbf{v} \gg \mathbf{0}$, then $(\Lambda(\mathbf{p}), \mathbf{v}) = (\Lambda(\mathbf{p_0}), \mathbf{v_0})$ and $\acute{\mathbf{x}} \in E$ is a critical point of Λ. □

Lemma 16 (Λ_c with stubborn agents)**.** *If agents $1, \ldots, k$, $k < N$, are stubborn, then at any $(p, \mathbf{0}) \in E$, the subspace spanned by $\{\mathbf{e}^1, \ldots, \mathbf{e}^k\}$ is tangent to the level surface Λ_c with $\mathbf{p} \in \Lambda_c$.*

Proof. By Lemma 15, for any stubborn agent j, $(\mathbf{e}^j)^T \nabla \Lambda(\mathbf{p}) = 0$ for any $(\mathbf{p}, \mathbf{0}) \in E$. By Proposition 6, $\mathbf{e}^j$ is tangent to the level surface Λ_c with $\mathbf{p} \in \Lambda_c$. □

Lemma 17 (Maximum and minimum values of Λ on $\mathcal{M}$)**.** *Λ achieves its global maximum and minimum c_{max}, c_{min} on $\mathcal{M}$ at one or more points in $\partial\mathcal{M}$. In Cases 2, 4, 5, and 6a, there exist unique corner points $(\mathbf{p}^{max}, \mathbf{0}) \in \hat{\mathcal{M}}$, $(\mathbf{p}^{min}, \mathbf{0}) \in \hat{\mathcal{M}}$ such that $\Lambda(\mathbf{p}^{max}) = c_{max}$, $\Lambda(\mathbf{p}^{min}) = c_{min}$, $R_{max} = \rho(B^*(\mathbf{p}^{max})\Gamma^{-1})$, and $R_{min} = \rho(B^*(\mathbf{p}^{min})\Gamma^{-1})$. Moreover, $\mathbf{p}^{max}$ and $\mathbf{p}^{min}$ are the respective unique global maximum and minimum points of $\Lambda(\mathbf{p})$ if and only if there are no stubborn agents in $\mathcal{G}$.*

Proof. The case in which every point in E is a critical point of Λ is trivial. Assume Λ has no critical points in E. Since $\mathcal{M} \subset \mathcal{M}'$ is a compact set and Λ is continuous, then by the Extreme Value Theorem, Λ achieves c_{max} and c_{min} at one or more points in $\partial\mathcal{M}$. Let $(\mathbf{p}^{max}, \mathbf{0}) \in \partial\mathcal{M}$ and $(\mathbf{p}^{min}, \mathbf{0}) \in \partial\mathcal{M}$ be points such that $\Lambda(\mathbf{p}^{max}) = c_{max}$ and $\Lambda(\mathbf{p}^{min}) = c_{min}$.

Let $(\mathbf{p}^S, \mathbf{0}) \in \mathcal{M}$. Assume there are no stubborn agents. The components $(J_T(\mathbf{p}^S))_{jk} = (1 - p_j^S)\hat{\beta}_{jk} + p_j^S \beta_{jk}$, $j \neq k$, are maximized and minimized at $p_j^S = 0$ or $p_j^S = 1$ for all j. If $B \succ \hat{B}$ (Cases 2 and 4), the entries $(J_T(\mathbf{p}^S))_{jk}$ are maximized at $p_j^S = 1$ and minimized at $p_j^S = 0$ for all j. It follows from Proposition 1 that $\mathbf{p}^{max} = \mathbf{1} \in \hat{\mathcal{M}}$ and $\mathbf{p}^{min} = \mathbf{0} \in \hat{\mathcal{M}}$. Using a similar argument if $\hat{B} \succ B$ (Case 5) $\mathbf{p}^{max} = \mathbf{0} \in \hat{\mathcal{M}}$ and $\mathbf{p}^{min} = \mathbf{1} \in \hat{\mathcal{M}}$. Further, in Case 6a with no stubborn agents, $\beta_{jk} - \hat{\beta}_{jk} > 0$ for all $k \in \mathcal{N}_j$ or $\beta_{jk} - \hat{\beta}_{jk} < 0$ for all $k \in \mathcal{N}_j$. Thus, similarly, $p_j^{max}, p_j^{min} \in \{0, 1\}$ for all j.

For any $\mathbf{p}^S \neq \mathbf{p}^{max}, \mathbf{p}^{min}$, it follows that $B^*(\mathbf{p}^{max})\Gamma^{-1} \succ B^*(\mathbf{p}^S)\Gamma^{-1} \succ B^*(\mathbf{p}^{min})\Gamma^{-1}$. By Proposition 1, $R_{max} = \rho(B^*(\mathbf{p}^{max})\Gamma^{-1})$ and $R_{min} = \rho(B^*(\mathbf{p}^{max})\Gamma^{-1})$.

Now suppose there are stubborn agents $1, \ldots, k$, $k < N$. Let $\mathbf{v_e} \in \text{span}\{\mathbf{e}^1, \ldots, \mathbf{e}^k\}$. Then, by Lemma 16, if $\mathbf{p}^{max} \in \Lambda_{c_{max}}$ then $\mathbf{p}^{max} + \mathbf{v_e} \in \Lambda_{c_{max}}$ as long as $(\mathbf{p}^{max} + \mathbf{v_e}, \mathbf{0}) \in E$. Similarly, if $\mathbf{p}^{min} \in \Lambda_{c_{min}}$ then $\mathbf{p}^{min} + \mathbf{v_e} \in \Lambda_{c_{min}}$ as long as $(\mathbf{p}^{min} + \mathbf{v_e}, \mathbf{0}) \in E$. □

We now present the proof for Theorems 14 and 15.

Proof of Theorem 14. To prove (A), assume $R_{max} < 1$. By Proposition 3, $\lambda_{Tmax}(J_{\mathcal{M}}(\mathbf{x})) < 0$ for all $\mathbf{x} \in \mathcal{M}$. Therefore, $\mathcal{M}_- = \mathcal{M}$, $\mathcal{M}_+ = \emptyset$, and $\mathcal{M}_0 = \emptyset$. The proof of (B) and (C) follow similarly.

To prove (D), assume $R_{min} < R_{max} = 1$. By Proposition 3, $\max_{\mathbf{x} \in \mathcal{M}} \lambda_{Tmax}(J_{\mathcal{M}}(\mathbf{x})) = 0$. Therefore, $\mathcal{M}_0 \neq \emptyset$ and $\mathcal{M}_- = \mathcal{M} \setminus \mathcal{M}_0$. By Lemma 17, $\mathcal{M}_0 \subset \partial\mathcal{M}$. The proof of (E) follows similarly.

To prove (F), assume $R_{max} > 1$ and $R_{min} < 1$. By Proposition 3 and Lemma 17, there exist points $\mathbf{x}_{max} = (\mathbf{p}^S_{max}, \mathbf{0})$ and $\mathbf{x}_{min} = (\mathbf{p}^S_{min}, \mathbf{0})$ in $\partial\mathcal{M}$ such that $\lambda_{Tmax}(J_{\mathcal{M}}(\mathbf{x}_{max})) > 0$ and $\lambda_{Tmax}(J_{\mathcal{M}}(\mathbf{x}_{min})) < 0$. By the continuity of Λ on E and Lemma 14, it follows that $\mathcal{M}_-, \mathcal{M}_+, \mathcal{M}_0 \neq \emptyset$, with each subset consisting of n_-, n_+, and n_0 connected sets, respectively, where each of the n_0 center sets separates n_- stable connected sets from n_+ unstable connected sets.

Since, by Lemma 15, Λ has no critical points in E, it follows that c is a regular value of Λ for any $c \in \mathbb{R}$. Hence, by the Implicit Function Theorem, every center connected set M_0 is an $(N-1)$-dimensional smooth hypersurface, and every stable and unstable connected set is an N-dimensional hypervolume.

Since Λ has no critial points in E, the gradient dynamics of Λ have no equilibria in $\mathcal{M} \subset E$. Thus, no center connected set M_0 in $\mathcal{M}_0$ is compact in $\text{int}(\mathcal{M})$ in any direction, since by Proposition 6 there cannot be any α-limit points or ω-limit points in $\mathcal{M}$. Thus, ∂M_0 in $\mathcal{M}$ must be contained in $\partial\mathcal{M}$. □

Proof of Theorem 15. Let $R_{max} > 1$, $R_{min} < 1$. In Cases 2, 4, 5 and 6a, by Lemma 17, there exists a unique corner point $\mathbf{x}^{min} = (\mathbf{p}^{min}, \mathbf{0}) \in \hat{\mathcal{M}}$ where $\Lambda(\mathbf{p}^{min}) \leq \Lambda(\mathbf{p}^S)$ for any $\mathbf{x} = (\mathbf{p}^S, \mathbf{0}) \in \mathcal{M}$, and $R_{min} = \rho(B^*(\mathbf{p}^{min})\Gamma^{-1})$. By Remark 17 and Proposition 3, it follows that $\Lambda(\mathbf{p}^{min}) < 0$.

Given $\mathbf{p}^S$, there exists a point $(\mathbf{p}^{end}, \mathbf{0}) \in \partial\mathcal{M}$ such that the line connecting $\mathbf{p}^{min}$ to $\mathbf{p}^{end}$ passes through $\mathbf{p}^S$. Each point on the line is parameterized by $r \in [0, 1]$ as follows: $\mathbf{s}(r, \mathbf{p}^{end}) = \{(\mathbf{p}(r), \mathbf{0}) \in \mathcal{M} \,|\, \mathbf{p}(r) = (1 - r)\mathbf{p}^{min} + r\mathbf{p}^{end}\}$.

Let $\bar{\mathbf{p}} = \mathbf{p}^{end} - \mathbf{p}^{min}$. Then, $\bar{\mathbf{p}}$ is tangent to $\mathbf{s}(r, \mathbf{p}^{end})$, and $\bar{\mathbf{p}}_j \geq 0$ if $\mathbf{p}_j^{min} = 0$ and $\bar{\mathbf{p}}_j \leq 0$ if $\mathbf{p}_j^{min} = 1$, for all j. By Proposition 6, the directional derivative of Λ at $\mathbf{s}(r, \mathbf{p}^{end})$ in the direction $\bar{\mathbf{p}}$ is $D_{\bar{\mathbf{p}}}\Lambda(\mathbf{s}(r, \mathbf{p}^{end})) = \bar{\mathbf{p}}^T \nabla\Lambda(\mathbf{s}(r, \mathbf{p}^{end}))$.

By Lemma 17, if $B \succ \hat{B}$ (Cases 2 and 4), then $\mathbf{p}^{min} = \mathbf{0}$ and $\bar{\mathbf{p}} \succ \mathbf{0}$. By Lemma 15, $\nabla\Lambda(\mathbf{s}(r, \mathbf{p}^{end})) \succ 0$ and so $\bar{\mathbf{p}}^T \nabla\Lambda(\mathbf{s}(r, \mathbf{p}^{end})) \geq 0$ for any $r \in [0, 1]$. Similarly, for Cases 5 and 6a, $\bar{\mathbf{p}}^T \nabla\Lambda(\mathbf{s}(r, \mathbf{p}^{end})) \geq 0$ for all $r \in [0, 1]$.

If $\bar{\mathbf{p}}^T \nabla\Lambda(\mathbf{s}(r, \mathbf{p}^{end})) = 0$, Λ is constant along the line $\mathbf{s}(r, \mathbf{p}^{end})$, $r \in [0, 1]$, and the line describes a level-surface Λ_c. Moreover, since $\Lambda(\mathbf{p}^{min}) < 0$, by Lemma 14 $c \neq 0$, i.e., $\mathcal{M}_0$ does not intersect Λ_c. For all other lines, $\bar{\mathbf{p}}^T \nabla\Lambda(\mathbf{s}(r, \mathbf{p}^{end})) > 0$, which implies that Λ is strictly increasing from a negative value at the corner $(\mathbf{p}^{min}, \mathbf{0})$ to the value at $(\mathbf{p}^{end}, \mathbf{0}) \in \partial\mathcal{M}$. By Theorem 14, $\partial\mathcal{M}_0 \subset \partial\mathcal{M}$. Thus, there is only the possibility of a single crossing of Λ_0 on each of these lines. By Lemma 14 there is a unique center connected hypersurface. □

10.6 Reproduction Numbers and Regimes of Dynamical Behavior

In this section we prove our main theorem, which provides conditions that determine whether solutions of (5.6) converge to a point in the IFE set $\mathcal{M}$ or to the EE as $t \to \infty$. We show that the basic and extreme basic reproduction numbers $R_0, R_1, R_{min}, R_{max}$

distinguish four behavioral regimes in the network SIRI model, each characterized by qualitatively different transient and steady-state behaviors.

10.6.1 The ω-limit Set of Solutions

The components of $\mathbf{p}$ decrease monotonically along solutions of (5.6). Here we show that this monotonicity implies that all solutions either converge to a point in $\mathcal{M}$ or to the EE as $t \to \infty$. Moreover, this means that when the EE is not an equilibrium of the dynamics, the infection cannot survive in the network and all solutions reach an IFE point in $\mathcal{M}$. The results in this section are valid even if $\hat{B}$ is not irreducible.

Lemma 18. *Let $\mathbf{y}(t, \mathbf{y}_0) = (\mathbf{p}^S(t), \mathbf{p}^I(t))$ be the solution of (10.3) with initial condition $\mathbf{y}_0 \in \Delta_N$. Then the following hold:*

- *The ω-limit set $\Omega(\mathbf{y}_0)$ of $\mathbf{y}(t, \mathbf{y}_0)$ is either a point in the IFE set $\mathcal{M}$ or the EE.*
- *$\mathbf{y}(t, \mathbf{y}_0)$ cannot exhibit non-trivial periodic orbits.*
- *$\mathbf{y}(t, \mathbf{y}_0)$ converges to a point in $\mathcal{M}$ as $t \to \infty$ if $R_1 \leq 1$.*

Proof. By invariance of Δ_N, any solution $\mathbf{y} = \mathbf{y}(t, \mathbf{y}_0)$ of (5.6) with initial condition $\mathbf{y}_0 \in \Delta_N$ is bounded and stays in Δ_N for $t \geq 0$. Therefore its ω-limit set $\Omega(\mathbf{y}_0)$ is a nonempty, compact, invariant set, and $\mathbf{y}$ approaches $\Omega(\mathbf{y}_0)$ as $t \to \infty$ (see Lemma 4.1 in [192]). Let $V = \mathbf{1}^T \mathbf{p}^S$, then $\dot{V} = -(\mathbf{p}^S)^T B \mathbf{p}^I \leq 0$ in Δ_N. By LaSalle's Invariance Principle [192], $\mathbf{y}$ approaches the largest invariant set W in the set $L = \{(\mathbf{p}^S, \mathbf{p}^I) \in \Delta_N \,|\, \dot{V} = 0\}$. It follows that, on L, $P^S B \mathbf{p}^I = \mathbf{0}$ which implies $\dot{\mathbf{p}}^S = \mathbf{0}$. Moreover, since $\hat{B}$ has a zero at every entry where B has a zero, it follows that $P^S \hat{B} \mathbf{p}^I = \mathbf{0}$ on L. This in turn implies that the $\mathbf{p}^I$ dynamics on L are given by the network SIS dynamics (10.7). Since solutions of (10.7) either converge to the IFE point $\mathbf{p}^{I*} = \mathbf{0}$ or to the EE (10.5) as $t \to \infty$ [81], it follows that every invariant set of L consists only of equilibria of (10.3) (see Remark 15). By Proposition 7, W is the union of the

IFE set $\mathcal{M}$ and the EE (10.5). Furthermore, since all w-limit points are equilibria, $\Omega(\mathbf{y}_0)$ contains a single point corresponding to either a point in the IFE set $\mathcal{M}$ or the EE.

Consequently, $\mathbf{y}$ converges to a point in the IFE or to the EE as $t \to \infty$, and $\mathbf{y}$ cannot exhibit non-trivial periodic orbits. In addition, if $R_1 \leq 1$ it follows from Propositions 7 and 9 that the IFE set $\mathcal{M}$ comprises the only equilibria of (10.3). Therefore, $\mathbf{y}$ converges to a point in $\mathcal{M}$ as $t \to \infty$. □

10.6.2 Behavioral Regimes

We now state the main theorem of the paper. We interpret and illustrate in Figure 10.1. The proof follows.

Theorem 16 (Behavioral Regimes)**.** *Let* $\mathbf{p}^I(0) \succ \mathbf{0}$*, and* $\mathbf{w}_m^T$ *be the leading left-eigenvector of* $B^*(\mathbf{p}^S)\Gamma^{-1}$*, where* $\mathbf{p}^S = \text{argmax}_{\mathbf{p}^S} \rho(B^*(\mathbf{p}^S)\Gamma^{-1})$*. Then the network SIRI model* (10.3) *exhibits four qualitatively distinct behavioral regimes:*

1. *Infection-Free Regime: If* $R_{max} \leq 1$ *the following hold:*
   (a) *All solutions converge to a point in* $\mathcal{M}$ *as* $t \to \infty$*.*
   (b) *If* $B \succeq \hat{B}$ *or* $\hat{B} \succeq B$*, the weighted infected average* $p^I_{\text{avg}}(t) = \mathbf{w}_m^T \Gamma^{-1} \mathbf{p}^I(t)$ *decays monotonically to zero.*
2. *Endemic Regime: If* $R_{min} > 1$*, all solutions converge to the EE as* $t \to \infty$*.*
3. *Epidemic Regime: If* $R_{min} < 1, R_{max} > 1$*, and* $R_1 \leq 1$*, the following hold:*
   (a) *All solutions converge to a point in* $\mathcal{M}$ *as* $t \to \infty$*.*
   (b) *There exists* $H \in \Delta_N$ *and* $H \supset \mathcal{M}_+$ *that is foliated by families of heteroclinic orbits, each orbit connecting two points in* $\mathcal{M}$*.*

Table 10.2: Behavioral Regimes of the Network SIRI Cases

Case	**Model**	**Inf.-Free**	**Endemic**	**Epidemic**	**Bistable**
1	SI		✓		
2	SIR	✓		✓	
3	SIS	✓	✓		
4	Partial	✓	✓	✓	
5	Comprom.	✓	✓		✓
6	Mixed	✓	✓	✓	✓

4. Bistable Regime: If $R_{min} < 1, R_{max} > 1$ and $R_1 > 1$, then, depending on the initial conditions, solutions converge to a point in $\mathcal{M}$ or to the EE as $t \to \infty$.

Table 10.2 summarizes which regimes of Theorem 16 exist for each of the six cases of the network SIRI model. Figure 10.1 illustrates the four regimes of Theorem 16 near the IFE set $\mathcal{M}$ when $B \succeq \hat{B}$ or $\hat{B} \succeq B$.

In Case 2 (SIR), since $\hat{B} = \bar{\mathbf{0}}$, $R_{max} = R_0$ and $R_{min} = R_1 = 0$. So only the infection-free and epidemic regimes are possible. This corresponds to the line $R_1 = 0$ in Figure 10.1.

In Case 3 (SIS), since $B = \hat{B}$, $R_{max} = R_0 = R_1 = R_{min}$. So only the infection-free and endemic regimes are possible. This corresponds to the line $R_1 = R_0$ in Figure 10.1.

In Case 4 (partial immunity), since $B \succ \hat{B}$, by Proposition 8, $R_{max} = R_0$ and $R_{min} = R_1$. So only the infection-free, endemic, and epidemic regimes are possible. This corresponds to the region $R_1 < R_0$ in Figure 10.1.

In Case 5 (compromised immunity), since $\hat{B} \succ B$, by Proposition 8, $R_{max} = R_1$ and $R_{min} = R_0$. So only the infection-free, endemic, and bistable regimes are possible. This corresponds to the region $R_1 > R_0$ in Figure 10.1.

In Case 6 (mixed immunity), all four regimes are possible.

The N-dimensional set $\mathcal{M}$ is illustrated in Figure 10.1 as a plane ($N = 2$) for ease of visualization. The blue region represents $\mathcal{M}_-$ (the set of stable points in $\mathcal{M}$)

and the red region represents $\mathcal{M}_+$ (the set of unstable points in $\mathcal{M}$). Black arrows illustrate the flow of solutions near $\mathcal{M}$. Theorems 11 and 15 prove which regions of $\mathcal{M}$ exist in each of the regimes. In the infection-free regime, $\mathcal{M} = \mathcal{M}_-$ and all solutions converge to the IFE set $\mathcal{M}$. In the endemic regime, $\mathcal{M} = \mathcal{M}_+$ and all solutions converge to the EE. In Cases 2 and 4 in the epidemic regime and in Case 5 in the bistable regime, $\mathcal{M}_-$ and $\mathcal{M}_+$ both exist and there is a unique hypersurface $\mathcal{M}_0$ shown as a black dashed line separating the two.

In Section 10.7 we study the geometry of solutions near $\mathcal{M}$ and the stable manifold (green translucent surface) and unstable manifold (magenta translucent surface) of $\mathcal{M}_0$ in the epidemic regime of Cases 2 and 4 and the bistable regime of Case 5. These manifolds are included in Figure 10.1 and help illustrate how solutions can flow.

proof of Theorem 16. To prove 1, let $R_{max} \leq 1$. Then by definition $R_1 \leq 1$. By Lemma 18, all solutions converge to a point in $\mathcal{M}$ as $t \to \infty$.

By (10.3), the dynamics of p^I_{avg} are

$$\begin{aligned}
\dot{p}^I_{\text{avg}} &= \mathbf{w}_m^T \Gamma^{-1}(B^*(\mathbf{p}) - \Gamma)\mathbf{p}^I - \mathbf{w}_m^T \Gamma^{-1} P^I \hat{B} \mathbf{p}^I \\
&\leq \mathbf{w}_m^T \Gamma^{-1}(\bar{B}_{max} - \Gamma)\mathbf{p}^I - \mathbf{w}_m^T \Gamma^{-1} P^I \hat{B} \mathbf{p}^I \\
&= (\bar{R}_{max} - 1)\mathbf{w}_m^T \mathbf{p}^I - \mathbf{w}_m^T \Gamma^{-1} P^I \hat{B} \mathbf{p}^I. \qquad (10.16)
\end{aligned}$$

The inequality follows from (10.4), and the last equality follows from $\rho(\Gamma^{-1}\bar{B}_{max}) = \rho(\bar{B}_{max}\Gamma^{-1})$. Since $\hat{B}$ is irreducible and by Proposition 1 $\mathbf{w}_m \gg \mathbf{0}$, the nonlinear term $\mathbf{w}_m^T \Gamma^{-1} P^I \hat{B} \mathbf{p}^I$ is nonnegative. And since $B \succeq \hat{B}$ or $\hat{B} \succeq B$, by Proposition 8, $\bar{R}_{max} = R_{max}$. Therefore, if $R_{max} < 1$, then $\dot{p}^I_{\text{avg}} < 0$ and p^I_{avg} decays monotonically to zero as $t \to \infty$.

If $R_{max} = 1$, $\dot{p}^I_{\text{avg}} \leq 0$ with equality holding only at points in $\Sigma = \{\mathbf{0} \preceq \mathbf{p}^I \preceq \mathbf{1} \,|\, p^I_j > 0 \text{ implies } p^I_k = 0, k \in \mathcal{N}_j\}$.

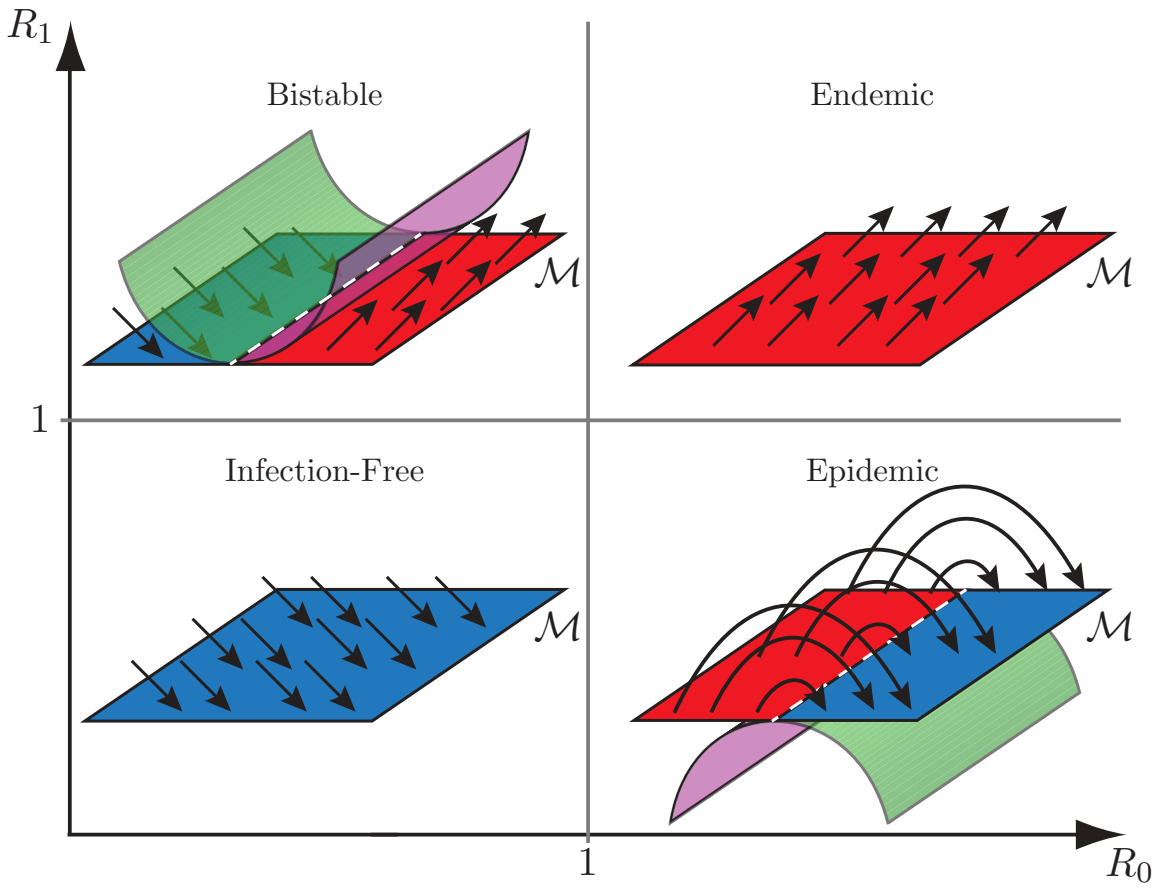

Figure 10.1: Illustration of local dynamics near $\mathcal{M}$ for the four different behavioral regimes of the network SIRI model (10.3) when $B \succeq \hat{B}$ or $\hat{B} \succeq B$. The diagrams are arranged where they exist in the R_0, R_1 parameter space according to Theorem 16. $\mathcal{M}_-$ is blue, $\mathcal{M}_+$ is red, and $\mathcal{M}_0$ is the black dashed line. The stable and unstable manifolds of $\mathcal{M}_0$ are green and magenta, respectively.

At any point in Σ, the dynamics of node j where $p_j^I > 0$ reduce to $\dot{p}_j^I = -\delta_j p_j^I$. Thus, no solution can stay in Σ except for the trivial solution $\mathbf{p}^I = \mathbf{0}$. By LaSalle's Invariance Principle, we conclude that p_{avg}^I decays monotonically to zero as $t \to \infty$.

To prove 2, let $R_{min} > 1$. Then by definition $R_1 > 1$. By Theorem 14, all points in $\mathcal{M}$ are unstable. Therefore, no non-trivial solution can converge to a point in $\mathcal{M}$. By Lemma 18, it follows that the ω-limit set of all solutions with $\mathbf{p}^I(0) \succ \mathbf{0}$ is the EE. Therefore, all solutions converge to the EE as $t \to \infty$.

To prove 3, let $R_{min} < 1, R_{max} > 1$, and $R_1 \leq 1$. By Theorem 14, $\mathcal{M}_-, \mathcal{M}_+, \mathcal{M}_0 \neq \emptyset$. By Lemma 18, all solutions converge to a point in $\mathcal{M}$ as $t \to \infty$. By the proof

of Lemma 13, the unstable manifold of any unstable point $\mathbf{x} \in \mathcal{M}_+$, lies partially or entirely in Δ_N. Let $\mathbf{y} = \mathbf{y}(t, \mathbf{y}_0)$ be a solution of (10.3) with initial condition $\mathbf{y}_0$ on the unstable manifold of $\mathbf{x}$. Then, $\mathbf{y}$ converges to $\mathbf{x}$ as $t \to -\infty$. By Lemma 18, $\mathbf{y}$ converges to a point $\mathbf{x}' \in \mathcal{M}$, $\mathbf{x}' \neq \mathbf{x}$, as $t \to \infty$. Thus, $\mathbf{y}$ forms a heteroclinic orbit. Let $H \subset \Delta_N$ be the union of the unstable manifolds of all points $\mathbf{x} \in \mathcal{M}_+$. Then, every solution in H forms a heteroclinic orbit connecting two points in $\mathcal{M}$.

To prove 4, let $R_1 > 1$. By Proposition 9, the EE exists and it is locally stable. Therefore any solution in the region of attraction of the EE converges to the EE as $t \to \infty$. By Lemma 13, for any locally stable point $\mathbf{x} \in \mathcal{M}_-$, there exists $V \subset \Delta_N$ and $\mathbf{x} \in V$ such that any solution starting in V converges to a point in $V \cap \mathcal{M}_-$ at an exponential rate. □

10.7 Bistable and Epidemic Regimes

10.7.1 Geometry of solutions near $\mathcal{M}$

In this section we examine the geometry of solutions near the IFE set $\mathcal{M}$ in the epidemic regime for Case 2 (SIR) and Case 4 (partial immunity), and the bistable regime for Case 5 (compromised immunity). The bistable regime for the network SIRI model, which doesn't exist for the well-studied SIS and SIR models, generalizes that proved for the well-mixed SIRI model studied in [57].

Definition 10.7.1 (Transversal crossing of Λ_c)**.** *Let $\mathbf{y}(t) = (\mathbf{p}^S(t), \mathbf{p}^I(t)) \in \Delta_N$, $t \geq 0$, be a solution of* (10.3)*. We say that $\mathbf{y}$ crosses Λ_c transversally if $\mathbf{p}^S$, the projection of $\mathbf{y}$ onto $\mathcal{M}$, crosses Λ_c transversally. This holds if there exists a time $t' > 0$ and $\mathbf{m} \in \Lambda_c$ such that $\mathbf{p}^S(t') = \mathbf{m}$ and $\dot{\mathbf{p}}^S(t')^T \nabla \Lambda(\mathbf{m}) \neq 0$.*

Proposition 11 (Transversal crossing direction)**.** *Let $\mathbf{y}(t) = (\mathbf{p}^S(t), \mathbf{p}^I(t)) \in \Delta_N$, $t \geq 0$, be a solution of* (10.3) *that crosses Λ_c transversally at the point $(\mathbf{m}, \mathbf{0}) \in \mathcal{M}$*

and time $t = t'$, where $c = \Lambda(\mathbf{m})$. If $\dot{\mathbf{p}}^S(t')^T \nabla\Lambda(\mathbf{m}) < 0$, then Λ decreases as $\mathbf{p}^S$ crosses Λ_c, and if $\dot{\mathbf{p}}^S(t')^T \nabla\Lambda(\mathbf{m}) > 0$, then Λ increases as $\mathbf{p}^S$ crosses Λ_c. Suppose $(\mathbf{m}, \mathbf{0}) \in \mathcal{M}_0$, then if $\dot{\mathbf{p}}^S(t')^T \nabla\Lambda(\mathbf{m}) < 0$, $\mathbf{p}^S$ crosses $\mathcal{M}_0$ from $\mathcal{M}_+$ to $\mathcal{M}_-$ and if $\dot{\mathbf{p}}^S(t')^T \nabla\Lambda(\mathbf{m}) > 0$, then $\mathbf{p}^S$ crosses $\mathcal{M}_0$ from $\mathcal{M}_-$ to $\mathcal{M}_+$.

Proof. Since $\dot{\mathbf{p}}^S = -P^S B \mathbf{p}^I$, it follows that $\dot{\mathbf{p}}^S(t')^T \nabla\Lambda(\mathbf{m})$ is the derivative of Λ at $\mathbf{m}$ along solutions of (10.3) (see Proposition 6). If $\dot{\mathbf{p}}^S(t')^T \nabla\Lambda(\mathbf{m}) < 0$, Λ decreases as $\mathbf{p}^S(t)$ crosses Λ_c and if $\dot{\mathbf{p}}^S(t')^T \nabla\Lambda(\mathbf{m}) > 0$, Λ increases as $\mathbf{p}^S(t)$ crosses Λ_c. Suppose $(\mathbf{m}, \mathbf{0}) \in \mathcal{M}_0$. By Lemma 14, if $\dot{\mathbf{p}}^S(t')^T \nabla\Lambda(\mathbf{m}) < 0$, $\mathbf{p}^S$ crosses $\mathcal{M}_0$ from $\mathcal{M}_+$ to $\mathcal{M}_-$. Similarly, if $\dot{\mathbf{p}}^S(t')^T \nabla\Lambda(\mathbf{m}) > 0$, $\mathbf{p}^S$ crosses $\mathcal{M}_0$ from $\mathcal{M}_-$ to $\mathcal{M}_+$. □

Theorem 17 (Transversality of solutions)**.** *Consider Cases 2 and 4 in the epidemic regime ($R_0 > 1$, $R_1 < 1$) and Case 5 in the bistable regime ($R_0 < 1$, $R_1 > 1$). Assume no stubborn agents. Let $\mathbf{y}(t) = (\mathbf{p}^S(t), \mathbf{p}^I(t)) \in \Delta_N$, $t \geq 0$, be a solution of* (10.3) *for which there exists a time $t' > 0$ and $(\mathbf{m}, \mathbf{0}) \in$ int$(\mathcal{M})$, such that $\mathbf{p}^S(t') = \mathbf{m}$ and $\mathbf{p}^I(t') \succ \mathbf{0}$. Let $c = \Lambda(\mathbf{m})$ so that $\mathbf{m} \in \Lambda_c$. Then $\mathbf{y}$ crosses Λ_c transversally. Suppose $(\mathbf{m}, \mathbf{0}) \in$ int$(\mathcal{M}_0)$. In the epidemic regime of Cases 2 and 4, $\mathbf{p}^S$ crosses $\mathcal{M}_0$ from $\mathcal{M}_+$ to $\mathcal{M}_-$, and the stable and unstable manifolds of $\mathcal{M}_0$ lie outside Δ_N. In the bistable regime of Case 5, $\mathbf{p}^S$ crosses $\mathcal{M}_0$ from $\mathcal{M}_-$ to $\mathcal{M}_+$, and the stable and unstable manifolds of $\mathcal{M}_0$ lie inside Δ_N.*

Proof. Assume no stubborn agents. Let $t' > 0$ such that $\mathbf{p}^S(t') = \mathbf{m}$, $\mathbf{p}^I(t') \succ \mathbf{0}$ and $(\mathbf{m}, \mathbf{0}) \in \text{int}(\mathcal{M})$. By Definition 10.4.1, $\mathbf{m} \gg \mathbf{0}$. Let $c = \Lambda(\mathbf{m})$. So $\dot{\mathbf{p}}^S(t') = -\text{diag}(\mathbf{m}) B p^I(t') \prec \mathbf{0}$. By Lemma 15, if $B \succ \hat{B}$ (Cases 2 and 4) then $\nabla\Lambda \gg \mathbf{0}$ and if $\hat{B} \succ B$ (Case 5) then $\nabla\Lambda \ll \mathbf{0}$. Thus $\dot{\mathbf{p}}^S(t')^T \nabla\Lambda(\mathbf{m}) \neq 0$, and so, by Proposition 11, $\mathbf{y}$ crosses Λ_c transversally. Suppose $(\mathbf{m}, \mathbf{0}) \in \text{int}(\mathcal{M}_0)$. Then, by Proposition 11, $\mathbf{p}^S$ crosses from $\mathcal{M}_+$ to $\mathcal{M}_-$ in Cases 2 and 4, and from $\mathcal{M}_-$ to $\mathcal{M}_+$ in Case 5. By continuity of solutions with respect to initial conditions, it follows that the stable and unstable manifolds of $\mathcal{M}_0$ must lie outside Δ_N in the epidemic regime of Cases 2 and 4 and inside Δ_N in the bistable regime of Case 5, as illustrated in Figure 10.1. □

Corollary 4. *In the epidemic regime of Cases 2 and 4, every heteroclinic orbit in Δ_N connects a point in $\mathcal{M}_+$ to a point in $\mathcal{M}_-$.*

Proof. Since by Theorem 17 the stable manifold of any point in $\mathcal{M}_0$ lies outside Δ_N, it follows by Theorem 16 that the orbit connects a point in $\mathcal{M}_+$ to a point in $\mathcal{M}_-$. □

Corollary 5. *Consider Case 5 in the bistable regime. Let $\mathbf{y}(t) = (\mathbf{p}^S(t), \mathbf{p}^I(t))$ be a solution of* (10.3)*. Then it holds that*

- *If $\mathbf{y}$ crosses $\mathcal{M}_0$ transversally or $(\mathbf{p}^S(0), \mathbf{0}) \in \mathcal{M}_+$, then $\mathbf{y}$ converges to the EE as $t \to \infty$. Moreover, the EE lies on the unstable manifold of $\mathcal{M}_0$.*
- *The stable manifold of $\mathcal{M}_0$ intersects the boundary of Δ_N where $\mathbf{p}^S = \mathbf{1} - \mathbf{p}^I$.*

Proof. If y crosses $\mathcal{M}_0$ transversally at $t = t'$, then, by Theorem 17, $\mathbf{p}^S$ crosses every level surface Λ_c in $\mathcal{M}$ transversally and crosses $\mathcal{M}_0$ from $\mathcal{M}_-$ to $\mathcal{M}_+$. By Proposition 11, Λ strictly increases along $\mathbf{p}^S$. It follows by Lemma 14 that $(\mathbf{p}^S, \mathbf{0}) \in \mathcal{M}_+$ for all $t > t'$. Since by Lemma 13 $\mathbf{y}$ cannot converge to a point in the IFE subset $\mathcal{M}_+$, it follows by Lemma 18 that $\mathbf{y}$ converges to the EE as $t \to \infty$. The same argument holds if $(\mathbf{p}^S(0), \mathbf{0}) \in \mathcal{M}_+$. It follows that the EE lies on the unstable manifold of $\mathcal{M}_0$.

Consider any point $\mathbf{y} = (\mathbf{p}^S, \mathbf{p}^I) \in \Delta_N$ on the stable manifold of $\mathcal{M}_0$. Then, as $t \to -\infty$ and $\mathbf{y}(t) \in \Delta_N$, $\mathbf{y}$ remains on the stable manifold of $\mathcal{M}_0$ and $(p^S, \mathbf{0})$ remains in $\mathcal{M}_-$. By Lemma 13, points in $\mathcal{M}_-$ have no unstable manifold in Δ_N so the stable manifold of $\mathcal{M}_0$ cannot intersect $\mathcal{M}_-$. Instead, because the components of $\mathbf{p}^S$ increase monotonically as $t \to -\infty$, $\mathbf{y}$ intersects $\partial\Delta_N$ where $\mathbf{p}^S = \mathbf{1} - \mathbf{p}^I$. □

The locations of the stable and unstable manifolds of $\mathcal{M}_0$, as proved in Theorem 17, are illustrated in Figure 10.1: outside Δ_N in the epidemic regime and inside Δ_N in the bistable regime. The figure shows the heteroclinic orbits proved in Corollary 4 for the epidemic regime. The solutions along the heteroclinic orbits cross $\mathcal{M}_0$

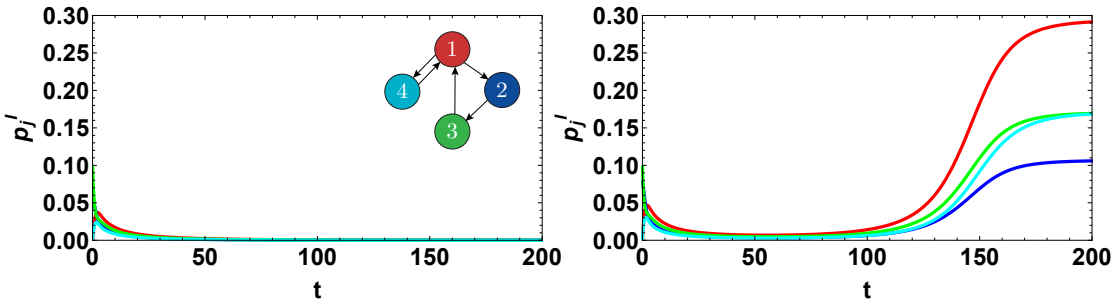

Figure 10.2: Bistability and resurgent epidemic. Simulation of $\mathbf{p}_j^I$ versus time t for network of $N = 4$ agents in bistable regime of Case 5 (compromised immunity) with $\mathbf{p}^S(0) = \mathbf{1} - \mathbf{p}^I(0)$. A is the unweighted adjacency matrix of the digraph shown. $B = 0.7A$, $\hat{B} = \text{diag}([1.5, 0.7, 0.7, 0.7])A$, and $\Gamma = \mathbb{I}$. Left. $\mathbf{p}^I(0) = [0, 0.05, 0.1, 0]^T$. Right. $\mathbf{p}^I(0) = [0, 0.08, 0.1, 0]^T$.

transversally, with their projection onto $\mathcal{M}$ crossing from $\mathcal{M}_+$ to $\mathcal{M}_-$, as proved in Theorem 17.

Figure 10.1 also shows the local flow in the bistable regime, as proved in Corollary 5. These solutions also cross $\mathcal{M}_0$ transversally, but with their projection onto $\mathcal{M}$ crossing from $\mathcal{M}_-$ to $\mathcal{M}_+$, as proved in Theorem 17.

10.7.2 Bistability Conditions for d-Regular Digraphs

In this section we examine the bistable regime of Theorem 16, which exists for Case 5 (compromised immunity) and for Case 6 (mixed immunity). We show how solutions in this regime can exhibit a *resurgent epidemic* in which an initial infection appears to die out for an arbitrarily long period of time, but then abruptly and surprisingly resurges to the EE.

Conditions on the initial state that predict a resurgent epidemic were proved for the well-mixed SIRI model in [57]. Here we compute a critical condition on the initial state in the special case that $\mathcal{G}$ is a d-regular digraph and every node has the same initial state. We then illustrate numerically for a more general digraph with compromised immunity in Figure 10.2.

Consider a d-regular digraph with global recovery, infection, and reinfection rates $\Gamma = \delta\mathbb{I}$, $B = \beta A$, and $\hat{B} = \hat{\beta}A$, where A is the adjacency matrix. Because the graph is d-regular, we have that, for any agent j, $\sum_{k=1}^{N} a_{jk} = d > 0$. The network SIRI agent dynamics (10.2) can be rewritten as

$$\begin{aligned}
\dot{p}_j^S &= -\beta d p_j^S p_j^I - \beta p_j^S \sum_{k=1}^{N} a_{jk}(p_k^I - p_j^I) \\
\dot{p}_j^I &= -\delta p_j^I + (\beta - \hat{\beta}) d p_j^S p_j^I + \hat{\beta} d p_j^I - \hat{\beta} d (p_j^I)^2 \\
&\quad + (\beta - \hat{\beta}) p_j^S \sum_{k=1}^{N} a_{jk}(p_k^I - p_j^I) \\
&\quad + \hat{\beta}(1 - p_j^I) \sum_{k=1}^{N} a_{jk}(p_k^I - p_j^I),
\end{aligned}$$

where we have used the identity $\sum_{k=1}^{N} a_{jk} p_k^I = p_j^I d + \sum_{k=1}^{N} a_{jk}(p_k^I - p_j^I)$. Setting $\mathbf{p}^I(0) = p_{ic}\mathbf{1}$, the dynamics reduce to

$$\begin{aligned}
\dot{p}_j^S &= -\beta d p_j^S p_j^I \\
\dot{p}_j^I &= -\delta p_j^I + (\beta - \hat{\beta}) d p_j^S p_j^I + \hat{\beta} d p_j^I - \hat{\beta} d.
\end{aligned} \tag{10.17}$$

(10.17) describes identical and uncoupled dynamics for every agent j, which are equivalent to the dynamics of the well-mixed SIRI model [57] with infection rate βd and reinfection rate $\hat{\beta} d$. Following [57], we find the critical initial condition $p_{crit} = 1 - \xi(R_0 d \xi)^{-\beta/\hat{\beta}}$, where $\xi = (R_1 - 1/d)/(R_1 - R_0)$. If $p_{ic} < p_{crit}$ solutions converge to a point in the IFE as $t \to \infty$. If $p_{ic} > p_{crit}$ solutions converge to the EE, $p^{I*} = (1 - \delta/(\hat{\beta} d))\mathbf{1}$, as $t \to \infty$. If $p_{ic} = p_{crit}$, the solution flows along the stable manifold of the point $\xi\mathbf{1} \in \mathcal{M}_0$ and converges to $\xi\mathbf{1}$. These results suggest more generally that the stable manifold of $\mathcal{M}_0$ separates solutions that converge to the IFE from those that converge to the EE, as in [57].

Further, as in [57], if $p_{ic} > p_{crit}$, the solution exhibits a *resurgent epidemic* in which p_j^I initially decreases to a minimum value p_{min}^I and then increases to the EE. As $(p_{ic} - p_{crit}) \to 0$, $p_{min}^I \to 0$ and the time it takes for the solution to resurge goes to infinity. That is, the infection may look like it is gone for a long time before resurging without warning. We note that the SIR and SIS models do not admit a bistable regime and therefore fail to account for the possibility of a resurgent epidemic, which could result when there is variability in infection or reinfection rates.

Figure 10.2 illustrates the bistability and resurgent epidemic phenomena for an example network of $N = 4$ agents for Case 5 (compromised immunity). $B = 0.7A$, $\hat{B} = \text{diag}([1.5, 0.7, 0.7, 0.7])A$, and $\Gamma = \mathbb{I}$. That is, agents $2, 3, 4$ are stubborn and agent 1 acquires compromised immunity to all its infected neighbors (agents 2 and 4). So $R_{max} = R_1 = 1.28$ and $R_{min} = R_0 = 0.85$, placing the system in the bistable regime. In both panels of Figure 10.2, $\mathbf{p}^S(0) = \mathbf{1} - \mathbf{p}^I(0)$. In the left panel $\mathbf{p}^I(0) = [0, 0.05, 0.1, 0]^T$ and the solution can be observed to converge to the IFE, i.e., $p_j^I \to 0$ for all j. In the right panel $\mathbf{p}^I(0) = [0, 0.08, 0.1, 0]^T$ and there is a resurgent epidemic: each p_j^I initially decays and then remains close to zero until $t \approx 100$, after which the p_j^I increase rapidly to the EE, which is $\mathbf{p}^{I*} = [0.29, 0.11, 0.17, 0.17]^T$.

10.8 Control Strategies

We apply our theory to design control strategies that guarantee desired steady-state behavior, such as the eradication of an infection. We begin with an example network with mixed immunity in the endemic regime, which has an infected steady state. We show three strategies for changing parameters that modify the reproduction numbers R_0, R_1, R_{min}, and R_{max}, and control the dynamics to a behavioral regime that results in an infection-free steady state, according to Theorem 16. We then consider two example networks with mixed immunity in the bistable regime. We illustrate how

vaccination of well-chosen agents increases the set of initial conditions that yield an infection-free steady state or at least delay a resurgent epidemic so that further control can be introduced.

10.8.1 Control from Endemic to Infection-Free Steady State

Consider the network of four agents shown in the top left panel of Figure 10.3. Let $B = A$, the unweighted adjacency matrix for the digraph shown. Let $\hat{B} = \text{diag}([0.3, 1, 2, 1])A$ and $\Gamma = \mathbb{I}$. The network has weak mixed immunity (Case 6a): agent 1 acquires partial immunity to reinfection, agent 3 acquires compromised immunity to reinfection, and agents 2 and 4 are stubborn. We compute the reproduction numbers: $R_0 = 1.32$, $R_1 = 1.22$, $R_{min} = 1.13$, and $R_{max} = 1.52$, which imply dynamics in the endemic regime and an infected steady state for every initial condition. This is illustrated in the simulation of p_j^I versus time t for initial conditions $\mathbf{p}^S(0) = \mathbf{1} - \mathbf{p}^I(0)$, $\mathbf{p}^I(0) = [0.01, 0.01, 0.01, 0.2]^T$. By (10.5), the solution converges to the EE: $\mathbf{p}^{I*} = [0.07, 0.23, 0.12, 0.19]^T$.

Modification of agent recovery rate

This strategy controls the network behavior by selecting one or more agents for treatment to increase its recovery rate. In the epidemiological setting, this could mean medication. In the behavioral setting, this could mean providing incentives or training. We make just one modification to the example network of four agents: $\delta_2 = 1$ becomes $\delta_2 = 3.5$. The corresponding reproduction numbers are $R_0 = 0.80$, $R_1 = 0.72$, $R_{min} = 0.65$, and $R_{max} = 0.94$, which imply dynamics in the infection-free regime. Using the same initial conditions as in the top left panel, we simulate the modified system in the bottom left panel of Figure 10.3. The solution converges to an infection-free steady state as predicted by Theorem 16.

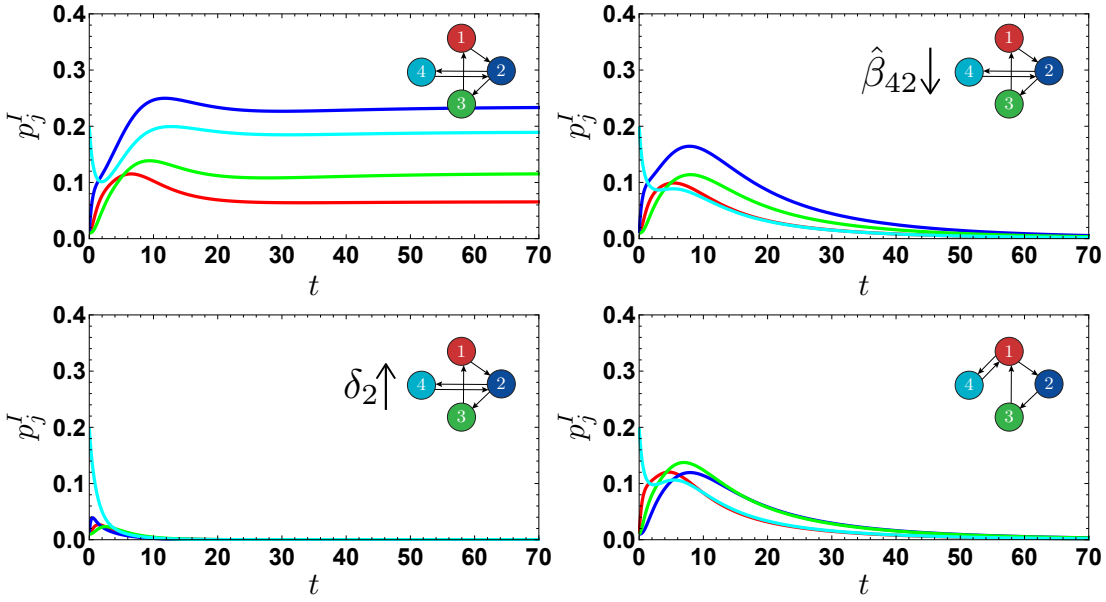

Figure 10.3: Simulations of p_j^I vs. t to illustrate control strategies that eradicate infection: $j = 1$ in red, $j = 2$ in blue, $j = 3$ in green, and $j = 4$ in cyan. Top Left. Example network of 4 agents with weak mixed immunity (Case 6a) in the endemic regime. $B = A$, $\hat{B} = \text{diag}([0.3, 1, 2, 1])A$, and $\Gamma = \mathbb{I}$. Bottom Left. Modification of recovery rate of agent 2 from $\delta_2 = 1$ to $\delta_2 = 3.5$. Top Right. Modification of reinfection rate $\hat{\beta}_{42} = 1$ to $\hat{\beta}_{42} = 0.3$. Bottom Right. Modification of network topology as shown.

Modification of agent reinfection rate

This strategy controls the network behavior by selecting one or more agents for treatment to decrease its reinfection rate. In the epidemiological setting, this could mean innoculation after first exposure. In the behavioral setting, this could mean education that leverages first exposure. We make just one modification to the example network of four agents: $\hat{\beta}_{42} = 1$ becomes $\hat{\beta}_{42} = 0.3$. The corresponding reproduction numbers are $R_0 = 1.32$, $R_1 = 0.96$, $R_{min} = 0.82$, and $R_{max} = 1.52$, which imply dynamics in the epidemic regime. Using the same initial conditions as in the top left panel, we simulate the modified system in the top right panel of Figure 10.3. After a small and short-lived epidemic, the solution converges to an infection-free steady state, as predicted by Theorem 16.

Modification of network topology

This strategy controls the network behavior by selecting one or more edges in the network graph for re-wiring. In the epidemiological and behavioral settings, this could mean affecting who comes in contact with whom. For the example network of four agents, we move the connections between agent 4 and agent 1 to instead be connections between agent 4 and agent 2. The corresponding reproduction numbers are identical to the reproduction numbers in the modification of reinfection rate example and therefore also imply dynamics in the epidemic regime. Using the same initial conditions as in the top left panel, we simulate the modified system in the bottom right panel of Figure 10.3. The solution converges to an infection-free steady state, as predicted by Theorem 16, with an initial epidemic smaller than in the top right panel.

10.8.2 Control in Bistable Regime

Small network

A network of two agents with weak mixed immunity in the bistable regime is shown in Figure 10.4. Agent 1 acquires compromised immunity while agent 2 acquires partial immunity: $\beta_{12} = \hat{\beta}_{21} = 0.8$, $\hat{\beta}_{12} = \beta_{21} = 1.3$, $\delta_1 = \delta_2 = 1$. Suppose initially there are no recovered agents, i.e., $\mathbf{p}^S(0) = \mathbf{1} - \mathbf{p}^I(0)$. Then, it can be shown that the solution will always converge to the EE. Now suppose we apply a control strategy in which we vaccinate the agent who acquires partial immunity, where vaccination is equivalent to exposing the agent to the infection. In the behavioral setting, this could mean preemptive exposure to the infection; see for example [193] where this was explored in the context of people resisting the spread of misinformation. After the vaccination of agent 2 in our example, we have $p_1^S(0) = 1 - p_1^I(0)$ and $p_2^S(0) = 0$.

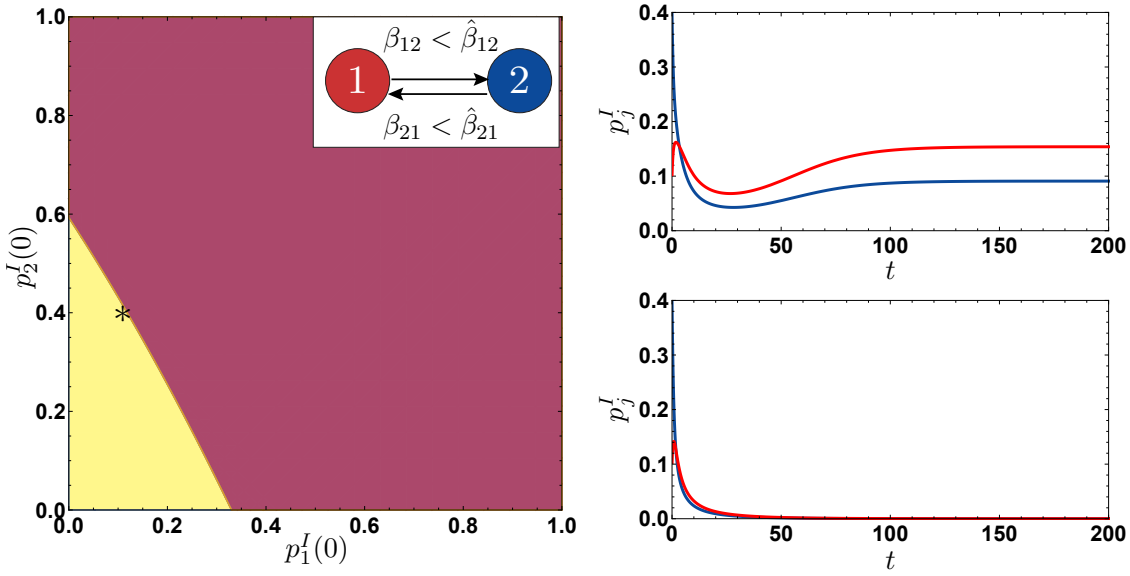

Figure 10.4: Vaccination of agent 2 in network of two agents with mixed immunity in bistable regime. Agent 1 acquires compromised immunity and agent 2 acquires partial immunity: $\beta_{12} = \beta_{21} = 0.8$, $\hat{\beta}_{12} = \hat{\beta}_{21} = 1.3$, $\delta_1 = \delta_2 = 1$. Left. With agent 2 vaccinated, solutions with initial conditions in magenta converge to the EE, and solutions with initial conditions in yellow converge to an IFE. Top Right. Simulation with no vaccination. Bottom Right. Simulation with vaccination. The initial condition is $p_1^I(0) = 0.1$ and $p_2^I(0) = 0.4$, shown as a star in the left panel.

We illustrate the results of the vaccination of agent 2 in the left panel of Figure 10.4. Initial conditions that lead to the EE are shown in magenta and to an infection-free steady state in yellow. We illustrate with simulations using the initial conditions $p_1^I(0)$ and $p_2^I(0)$ denoted by the star in the left panel of Figure 10.4. The top right simulation is of the system with no vaccination: there is an infected steady state as predicted. The bottom right simulation is of the system with agent 2 vaccinated: there is an infection-free steady state as predicted.

Large network

A network of twenty agents with weak mixed immunity in the bistable regime is shown in Figure 10.5. Agents 2, 3, 5, 7, 11, 13, 17, and 19 (dark gray) acquire partial immunity to reinfection: $\beta_{jk} = 0.5$, $\hat{\beta}_{jk} = 0.1$. Agents 1 and 20 (gray) are

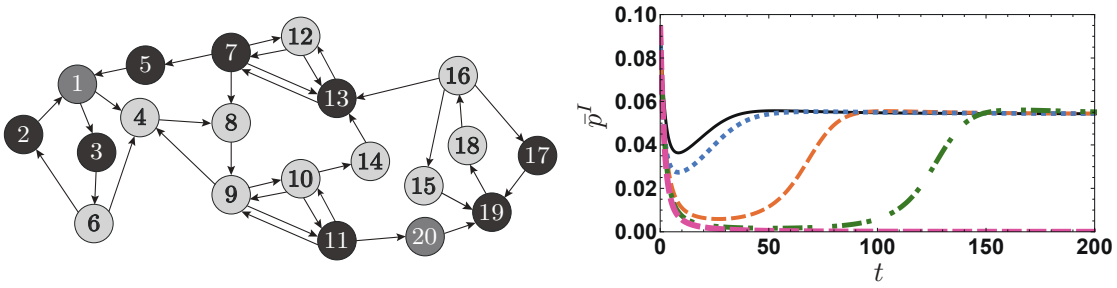

Figure 10.5: Left. Network of twenty agents with weak mixed immunity in bistable regime. Agents in dark gray acquire partial immunity to reinfection: $\beta_{jk} = 0.5$, $\hat{\beta}_{jk} = 0.1$. Agents in gray are stubborn: $\beta_{jk} = \hat{\beta}_{jk} = 0.5$. Agents in light gray acquire compromised immunity to reinfection: $\beta_{jk} = 0.5$, $\hat{\beta}_{jk} = 0.875$. $\Gamma = \mathbb{I}$. Right. Simulations of $\bar{p}^I$ for no vaccination (solid black), and vaccinations of agents 2,3,5,7 (dotted blue), agent 11 (dashed orange), agents 7, 11 (point-dashed green), and agents 7, 11, 13 (violet dashed). Initial conditions are $p_1^I(0) = p_{20}^I(0) = 0.5$ and $p_j^I(0) = 0.05$ for $j \neq 1, 20$.

stubborn: $\beta_{jk} = \hat{\beta}_{jk} = 0.5$. Agents 4, 6, 8, 9, 10, 12, 14, 16, and 18 (light gray) acquire compromised immunity to reinfection: $\beta_{jk} = 0.5$, $\hat{\beta}_{jk} = 0.875$. $\Gamma = \mathbb{I}$.

In Figure 10.5, we plot simulations of the average infected state $\bar{p}^I = \mathbf{1}^T p^I / N$ versus t for no vaccinations (solid black) and for different sets of vaccinated agents: agents 2, 3, 5, and 7 (dotted blue), agent 11 (dashed orange), agents 7 and 11 (point-dashed green), and agents 7, 11, and 13 (violet dashed). Vaccinating agent 11 has the strongest effect. The infection appears to be eradicated by vaccinating agents 7, 11, and 13. The other vaccination cases delay the epidemic, which provides time for treatments or other control interventions.

10.9 Conclusion and Future Directions

Many proposed control methodologies for contagion processes focus on mechanisms that affect R_0 via changes to the network structure and the system parameters [72, 74]. Example mechanisms include optimal node removal, optimal link removal, and budget-constrained allocation [73, 74, 147, 148]. Under the assumption of perfect knowledge of the network structure and the system parameters, these same mech-

anisms could be used to derive control strategies for the network SIRI model that affect the value of the basic and extreme basic reproduction numbers: R_{max}, R_{min}, R_0, and R_1. However, optimal node and link removal have been shown to be NP-complete and NP-hard problems, respectively [74, 149]. Alternatives include the use of centrality measures to select the nodes that should be removed first. Moreover, if knowledge of the network structure or system parameters is uncertain or not available, these control strategies are not viable. Instead, distributed control strategies that affect the system parameters are required. Our network SIRI model results invite an interpretation of the changes in susceptibility by the individuals in the network as a distributed heuristic feedback policy: individuals assess the quality of the process spreading through the network and adjust their susceptibility accordingly, leading to the eradication or spread of the infection. Future studies could exploit our results for the network SIRI model and use centrality measures for contagion processes to derive necessary and sufficient conditions on the reinfection and/or recovery rates of a small number of selected agents that would guarantee that the infection dies out or spreads through the network. Our results where obtained on a reduced model based on the IBMF approach [72, 75]. Previous work has shown that solutions in an IBMF model converge to an infection-free equilibrium, then the stochastic Markov model reaches the infection-free absorbing state in sublinear time with respect to the size of the network in expectation [74]. The bistable regime in the network SIRI model suggests that bistability in stochastic Markov models may be highly sensitive to noise. Future work will focus on extending the results to cases in which the network is weakly connected and designing control strategies that leverage the adaptation of agents to the infection.

Bibliography

[1] J. Krause and G. D. Ruxton, *Living in Groups*. Oxford University Press, 2002.

[2] J. K. Parrish and L. Edelstein-Keshet, "Complexity, pattern, and evolutionary trade-offs in animal aggregation," *Science*, vol. 284, no. 5411, pp. 99–101, 1999.

[3] I. D. Couzin, "Collective cognition in animal groups," *Trends in Cognitive Sciences*, vol. 13, no. 1, pp. 36–43, 2009.

[4] J. A. Marshall and N. R. Franks, "Colony-level cognition," *Current Biology*, vol. 19, no. 10, pp. R395–R396, 2009.

[5] E. Codling, J. Pitchford, and S. Simpson, "Group navigation and the many-wrongs principle in models of animal movement," *Ecology*, vol. 88, no. 7, pp. 1864–1870, 2007.

[6] D. J. Sumpter, J. Krause, R. James, I. D. Couzin, and A. J. Ward, "Consensus decision making by fish," *Current Biology*, vol. 18, no. 22, pp. 1773–1777, 2008.

[7] D. J. Sumpter, *Collective Animal Behavior*. Princeton: Princeton University Press, 2010.

[8] D. M. Gordon, "The organization of work in social insect colonies," *Nature*, vol. 380, pp. 121–124, 1996.

[9] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proceedings of the Natural Academy of Sciences*, vol. 79, no. 8, pp. 2554–2558, 1982.

[10] W. Bialek, A. Cavagna, I. Giardina, T. Mora, E. Silvestri, M. Viale, and A. M. Walczak, "Statistical mechanics for natural flocks of birds," *Proceedings of the National Academy of Sciences*, vol. 109, no. 13, pp. 4786–4791, 2012.

[11] G. F. Young, L. Scardovi, A. Cavagna, I. Giardina, and N. E. Leonard, "Starling flock networks manage uncertainty in consensus at low cost," *PLoS computational biology*, vol. 9, no. 1, p. e1002894, 2013.

[12] C. C. Ioannou, V. Guttal, and I. D. Couzin, "Predatory fish select for coordinated collective motion in virtual prey," *Science*, vol. 337, no. 6099, pp. 1212–1215, 2012.

[13] M. B. Miller and B. L. Bassler, "Quorum sensing in bacteria.," *Annual Review of Microbiology*, vol. 55, pp. 165–99, 2001.

[14] D. M. Gordon, *Ants at Work: How an Insect Society is Organized*. New York: The Free Press, 1999.

[15] G. E. Robinson, "Regulation of division of labor in insect societies," *Annual Review of Entomology*, vol. 37, no. 1, pp. 637–665, 1992.

[16] T. D. Seeley, S. Camazine, and J. Sneyd, "Collective decision-making in honey bees: how colonies choose among nectar sources," *Behavioral Ecology and Sociobiology*, vol. 28, no. 4, pp. 277–290, 1991.

[17] S. Camazine, J.-L. Deneubourg, N. R. Franks, J. Sneyd, E. Bonabeau, and G. Theraula, *Self-Organization in biological systems*, vol. 7. Princeton university press, 2003.

[18] M. Ballerini, N. Cabibbo, R. Candelier, A. Cavagna, E. Cisbani, I. Giardina, V. Lecomte, A. Orlandi, G. Parisi, A. Procaccini, *et al.*, "Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study," *Proceedings of the National Academy of Sciences*, vol. 105, no. 4, pp. 1232–1237, 2008.

[19] A. Bernstein, D. Bienstock, D. Hay, M. Uzunoglu, and G. Zussman, "Power grid vulnerability to geographically correlated failures analysis and control implications," in *IEEE INFOCOM 2014 - IEEE Conference on Computer Communications*, pp. 2634–2642, April 2014.

[20] J. D. Farmer, M. Gallegati, C. Hommes, A. Kirman, P. Ormerod, S. Cincotti, A. Sanchez, and D. Helbing, "A complex systems approach to constructing better models for managing financial markets and the economy," *The European Physical Journal Special Topics*, vol. 214, no. 1, pp. 295–324, 2012.

[21] G. Turrigiano, "Too many cooks? intrinsic and synaptic homeostatic mechanisms in cortical circuit refinement," *Annual Review of Neuroscience*, vol. 34, pp. 89–103, 2011.

[22] T. O'Leary and D. J. Wyllie, "Neuronal homeostasis: time for a change?," *The Journal of Physiology*, vol. 589, no. 20, pp. 4811–4826, 2011.

[23] N. J. Mlot, C. A. Tovey, and D. L. Hu, "Fire ants self-assemble into waterproof rafts to survive floods," *Proceedings of the National Academy of Sciences*, vol. 108, no. 19, pp. 7669–7673, 2011.

[24] N. E. Leonard, "Multi-agent system dynamics: Bifurcation and behavior of animal groups," *Annual Reviews in Control*, vol. 38, no. 2, pp. 171–183, 2014.

[25] M. J. Mataric, "Designing emergent behaviors: From local interactions to collective intelligence," in *Proceedings of the Second International Conference on Simulation of Adaptive Behavior*, pp. 432–441, 1993.

[26] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, 2007.

[27] B. Nabet, N. E. Leonard, I. D. Couzin, and S. A. Levin, "Dynamics of decision making in animal group motion," *Journal of Nonlinear Science*, vol. 19, no. 4, pp. 399–435, 2009.

[28] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *Departmental Papers (ESE)*, p. 29, 2003.

[29] M. Newman, *Networks: an Introduction*. Oxford University Press, 2010.

[30] J. Guckenheimer and P. J. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, vol. 42. Springer Science & Business Media, 2013.

[31] M. Dorigo, G. D. Caro, and L. M. Gambardella, "Ant algorithms for discrete optimization," *Artificial Life*, vol. 5, no. 2, pp. 137–172, 1999.

[32] J. Lighton and G. A. Bartholomew, "Standard energy metabolism of a desert harvester ant, pogonomyrmex rugosus: effects of temperature, body mass, group size, and humidity," *Proceedings of the Naturall Academy of Sciences*, vol. 85, no. 13, pp. 4765–4769, 1988.

[33] J. R. Lighton and D. H. Feener Jr, "Water-loss rate and cuticular permeability in foragers of the desert ant pogonomyrmex rugosus," *Physiological Zoology*, vol. 62, no. 6, pp. 1232–1256, 1989.

[34] M. J. Greene and D. M. Gordon, "Social insects: cuticular hydrocarbons inform task decisions," *Nature*, vol. 423, no. 6935, pp. 32–32, 2003.

[35] M. J. Greene and D. M. Gordon, "Interaction rate informs harvester ant task decisions," *Behavioral Ecology*, vol. 18, no. 2, pp. 451–455, 2007.

[36] M. J. Greene, N. Pinter-Wollman, and D. M. Gordon, "Interactions with combined chemical cues inform harvester ant foragers' decisions to leave the nest in search of food," *PLoS One*, vol. 8, p. e52219, 01 2013.

[37] N. Pinter-Wollman, A. Bala, A. Merrell, J. Queirolo, M. C. Stumpe, S. Holmes, and D. M. Gordon, "Harvester ants use interactions to regulate forager activation and availability," *Animal Behavior*, vol. 86, no. 1, pp. 197–207, 2013.

[38] J. D. Davidson, R. P. Arauco-Aliaga, S. Crow, D. M. Gordon, and M. S. Goldman, "Effect of interactions between harvester ants on forager decisions," *Frontiers in Ecology and Evolution*, vol. 4, p. 115, 2016.

[39] B. D. Beverly, H. McLendon, S. Nacu, S. Holmes, and D. M. Gordon, "How site fidelity leads to individual differences in the foraging activity of harvester ants," *Behavioral Ecology*, vol. 20, no. 3, pp. 633–638, 2009.

[40] B. Prabhakar, K. N. Dektar, and D. M. Gordon, "The regulation of ant colony foraging activity without spatial information," *PLoS Comput Biol*, vol. 8, no. 8, p. e1002670, 2012.

[41] D. M. Gordon, "The rewards of restraint in the collective regulation of foraging by harvester ant colonies," *Nature*, vol. 498, no. 7452, pp. 91–93, 2013.

[42] D. M. Gordon, "The of an ant colony's foraging range," *Animal Behavior*, vol. 49, no. 3, pp. 649 – 659, 1995.

[43] R. Pagliara, D. M. Gordon, and N. E. Leonard, "Regulation of harvester ant foraging as a closed-loop excitable system," *PLoS Computational Biology*, vol. 14, pp. 1–25, 12 2018.

[44] D. H. Anderson, *Compartmental Modeling and Tracer Kinetics*, vol. 50. Springer Science & Business Media, 2013.

[45] E. D. Sontag, "Structure and stability of certain chemical networks and applications to the kinetic proofreading model of t-cell receptor signal transduction," *IEEE Transactions on Automatic Control*, vol. 46, no. 7, pp. 1028–1047, 2001.

[46] R. M. Anderson and R. M. May, *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press, 1992.

[47] H. W. Hethcote, "The mathematics of infectious diseases," *SIAM Review*, vol. 42, no. 4, pp. 599–653, 2000.

[48] L. Weng, A. Flammini, A. Vespignani, and F. Menczer, "Competition among memes in a world with limited attention," *Scientific Reports*, vol. 2, p. 335, 2012.

[49] L. Weng, F. Menczer, and Y.-Y. Ahn, "Virality prediction and community structure in social networks," *Scientific Reports*, vol. 3, p. 2522, 2013.

[50] D. Centola, "The spread of behavior in an online social network experiment," *Science*, vol. 329, no. 5996, pp. 1194–1197, 2010.

[51] D. J. Daley and D. G. Kendall, "Stochastic rumours," *IMA Journal of Applied Mathematics*, vol. 1, no. 1, pp. 42–55, 1965.

[52] L. Bonnasse-Gahot, H. Berestycki, M.-A. Depuiset, M. B. Gordon, S. Roché, N. Rodriguez, and J.-P. Nadal, "Epidemiological modelling of the 2005 French riots: a spreading wave and the role of contagion," *Scientific Reports*, vol. 8, no. 1, p. 107, 2018.

[53] R. Sachak-Patwa, N. T. Fadai, and R. A. VanGorder, "Understanding viral video dynamics through an epidemic modelling approach," *Physica A: Statistical Mechanics and its Applications*, vol. 502, pp. 416 – 435, 2018.

[54] W. Goffman and V. Newill, "Generalization of epidemic theory," *Nature*, vol. 204, no. 4955, pp. 225–228, 1964.

[55] B. Min and M. San Miguel, "Competing contagion processes: Complex contagion triggered by simple contagion," *Scientific Reports*, vol. 8, no. 1, p. 10422, 2018.

[56] L. Adamic *et al.*, "The diffusion of support in an online social movement: Evidence from the adoption of equal-sign profile pictures," in *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, pp. 1741–1750, ACM, 2015.

[57] R. Pagliara, B. Dey, and N. E. Leonard, "Bistability and resurgent epidemics in reinfection models," *IEEE Control Systems Letters*, vol. 2, pp. 290–295, April 2018.

[58] R. Pagliara and N. E. Leonard, "Adaptive susceptibility and heterogeneity in contagion models on networks," In Preparation.

[59] D. Bernoulli, "Essai d'une nouvelle analyse de la mortalité causée par la petite vérole, et des avantages de l'inoculation pour la prévenir," *Histoire de l'Academie Royale des Sciences. Avec les Mémoires de Mathematique & de Physique. Paris.*, pp. 1–45, 1760.

[60] W. O. Kermack and A. G. McKendrick, "A contribution to the mathematical theory of epidemics," *Proceedings of the Royal Society of London. Series A, Containing papers of a mathematical and physical character*, vol. 115, no. 772, pp. 700–721, 1927.

[61] F. Brauer, C. Castillo-Chavez, and C. Castillo-Chavez, *Mathematical Models in Population Biology and Epidemiology*, vol. 40. Springer, 2001.

[62] K. Dietz, "The estimation of the basic reproduction number for infectious diseases," *Statistical Methods in Medical Research*, vol. 2, no. 1, pp. 23–41, 1993.

[63] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry*, vol. 1. Elsevier, 1992.

[64] S. Ross, *Stochastic Processes*. Wiley series in mathematical statistics. Probability and mathematical statistics, Wiley, 2nd ed., 1996.

[65] T. D. Seeley, P. K. Visscher, T. Schlegel, P. M. Hogan, N. R. Franks, and J. A. Marshall, "Stop signals provide cross inhibition in collective decision-making by honeybee swarms," *Science*, vol. 335, no. 6064, pp. 108–111, 2012.

[66] P. Érdi and J. Tóth, *Mathematical Models of Chemical Reactions: Theory and Applications of Deterministic and Stochastic models*. Manchester University Press, 1989.

[67] D. Sumpter and S. Pratt, "A modelling framework for understanding social insect foraging," *Behavioral Ecology and Sociobiology*, vol. 53, pp. 131–144, Feb 2003.

[68] O. Diekmann, J. A. P. Heesterbeek, and J. A. Metz, "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*, vol. 28, no. 4, pp. 365–382, 1990.

[69] W. Mei, S. Mohagheghi, S. Zampieri, and F. Bullo, "On the dynamics of deterministic epidemic propagation over networks," *Annual Reviews in Control*, vol. 44, pp. 116 – 128, 2017.

[70] E. Lieberman, C. Hauert, and M. A. Nowak, "Evolutionary dynamics on graphs," *Nature*, vol. 433, no. 7023, p. 312, 2005.

[71] M. Mesbahi and M. Egerstedt, *Graph Theoretic Methods in Multiagent Networks*, vol. 33. Princeton University Press, 2010.

[72] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, "Epidemic processes in complex networks," *Reviews of Modern Physics*, vol. 87, no. 3, p. 925, 2015.

[73] C. Nowzari, V. M. Preciado, and G. J. Pappas, "Optimal resource allocation for control of networked epidemic models," *IEEE Transactions on Control of Network Systems*, vol. 4, no. 2, pp. 159–169, 2017.

[74] C. Nowzari, V. M. Preciado, and G. J. Pappas, "Analysis and control of epidemics: A survey of spreading processes on complex networks," *IEEE Control Systems*, vol. 36, no. 1, pp. 26–46, 2016.

[75] P. Van Mieghem, J. Omic, and R. Kooij, "Virus spread in networks," *IEEE/ACM Transactions on Networking (TON)*, vol. 17, no. 1, pp. 1–14, 2009.

[76] Y. Wang, D. Chakrabarti, C. Wang, and C. Faloutsos, "Epidemic spreading in real networks: An eigenvalue viewpoint," in *22nd International Symposium on Reliable Distributed Systems, 2003. Proceedings.*, pp. 25–34, IEEE, 2003.

[77] P. Van Mieghem and J. Omic, "In-homogeneous virus spread in networks," *arXiv preprint arXiv:1306.2588*, 2013.

[78] M. Youssef and C. Scoglio, "An individual-based approach to SIR epidemics in contact networks," *Journal of Theoretical Biology*, vol. 283, no. 1, pp. 136–144, 2011.

[79] A. Khanafer, T. Başar, and B. Gharesifard, "Stability of epidemic models over directed graphs: A positive systems approach," *Automatica*, vol. 74, pp. 126–134, 2016.

[80] S. Saha, A. Adiga, B. A. Prakash, and A. K. S. Vullikanti, "Approximation algorithms for reducing the spectral radius to control epidemic spread," in *Proceedings of the 2015 SIAM International Conference on Data Mining*, pp. 568–576, SIAM, 2015.

[81] A. Fall, A. Iggidr, G. Sallet, and J.-J. Tewa, "Epidemiological models and Lyapunov functions," *Mathematical Modelling of Natural Phenomena*, vol. 2, no. 1, pp. 62–83, 2007.

[82] O. Diekmann, J. Heesterbeek, and M. G. Roberts, "The construction of next-generation matrices for compartmental epidemic models," *Journal of the Royal Society Interface*, vol. 7, no. 47, pp. 873–885, 2009.

[83] R. Pastor-Satorras and A. Vespignani, "Epidemic spreading in scale-free networks," *Physical Review Letters*, vol. 86, no. 14, p. 3200, 2001.

[84] M. Boguná, R. Pastor-Satorras, and A. Vespignani, "Cut-offs and finite size effects in scale-free networks," *The European Physical Journal B*, vol. 38, no. 2, pp. 205–209, 2004.

[85] J. P. Hespanha, *Linear Systems Theory*. Princeton university press, 2018.

[86] S. Liebscher, *Bifurcation Without Parameters*. Springer, 2015.

[87] H. K. Khalil and J. Grizzle, *Nonlinear Systems*, vol. 3. Prentice hall New Jersey, 1996.

[88] Y. A. Kuznetsov, *Elements of Applied Bifurcation Theory*, vol. 112. Springer Science & Business Media, 2 ed., 1998.

[89] M. W. Hirsch, C. C. Pugh, and M. Shub, *Invariant Manifolds*, vol. 583. Springer, 1977.

[90] A. Shoshitaishvili, "Bifurcations of topological type of a vector field near a singular point," *Trudy Seminarov IG Petrovskogo*, vol. 1, pp. 279–309, 1975.

[91] R. Gray, A. Franci, V. Srivastava, and N. E. Leonard, "Multiagent decision-making dynamics inspired by honeybees," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 2, pp. 793–806, 2018.

[92] A. L. Hodgkin and A. F. Huxley, "A quantitative description of membrane current and its application to conduction and excitation in nerve," *The Journal of Physiology*, vol. 117, no. 4, p. 500, 1952.

[93] C. O. Weiss and R. Vilaseca, "Dynamics of lasers," *NASA STI/Recon Technical Report A*, vol. 92, p. 39875, 1991.

[94] J. D. Murray, *Mathematical Biology. II Spatial Models and Biomedical Applications {Interdisciplinary Applied Mathematics V. 18}*. Springer-Verlag New York Incorporated, 2001.

[95] L. Glass, "Synchronization and rhythmic processes in physiology," *Nature*, vol. 410, no. 6825, pp. 277–284, 2001.

[96] E. Meron, "Pattern formation in excitable media," *Physics Reports*, vol. 218, no. 1, pp. 1–66, 1992.

[97] R. Sepulchre, G. Drion, and A. Franci, "Excitable behaviors," in *Emerging Applications of Control and Systems Theory*, pp. 269–280, Springer, 2018.

[98] E. M. Izhikevich, *Dynamical Systems in Neuroscience*. Cambridge: The MIT press, 1st ed., 2007.

[99] A. L. Hodgkin, "The local electric changes associated with repetitive action in a non-medullated axon," *The Journal of Physiology*, vol. 107, no. 2, p. 165, 1948.

[100] J. Rinzel, "Excitation dynamics: insights from simplified membrane models," in *Federation Proceedings*, vol. 44, pp. 2944–2946, 1985.

[101] J. Rinzel and G. B. Ermentrout, "Analysis of neural excitability and oscillations," in *Methods in Neuronal Modeling*, pp. 135–169, MIT press, 1989.

[102] E. M. Izhikevich, "Neural excitability, spiking and bursting," *International Journal of Bifurcation and Chaos*, vol. 10, no. 06, pp. 1171–1266, 2000.

[103] R. FitzHugh, "Impulses and physiological states in theoretical models of nerve membrane," *Biophysical Journal*, vol. 1, no. 6, p. 445, 1961.

[104] J. Nagumo, S. Arimoto, and S. Yoshizawa, "An active pulse transmission line simulating nerve axon," *Proceedings of the IEEE Institute of Electrical and Electronics Engineers*, vol. 50, no. 10, pp. 2061–2070, 1962.

[105] H. R. Wilson, *Spikes, Decisions, and Actions: the Dynamical Foundations of Neurosciences*. Oxford University Press, 1999.

[106] E. N. Davison, Z. Aminzare, B. Dey, and N. E. Leonard, "Mixed mode oscillations and phase locking in coupled FitzHugh-Nagumo model neurons," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 29, no. 3, p. 033105, 2019.

[107] É. Benoît, "Systemes lents-rapides dans $\mathbb{R}^3$ et leurs canards," *Astérisque*, vol. 109, no. 110, pp. 159–191, 1983.

[108] D. Gross, *Fundamentals of Queueing Theory*. John Wiley & Sons, 2008.

[109] D. G. Kendall, "Stochastic processes occurring in the theory of queues and their analysis by the method of the imbedded Markov chain," *The Annals of Mathematical Statistics*, pp. 338–354, 1953.

[110] S. G. Eick, W. A. Massey, and W. Whitt, "The physics of the $M(t)/G/\infty$ queue," *Operations Research*, vol. 41, no. 4, pp. 731–742, 1993.

[111] A. Berman and R. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*. Society for Industrial and Applied Mathematics, 1994.

[112] L. Farina and S. Rinaldi, *Positive Linear Systems: Theory and Applications*, vol. 50. John Wiley & Sons, 2011.

[113] O. Perron, "Zur theorie der matrices," *Mathematische Annalen*, vol. 64, no. 2, pp. 248–263, 1907.

[114] O. Perron, "Grundlagen für eine theorie des Jacobischen kettenbruchalgorithmus," *Mathematische Annalen*, vol. 64, no. 1, pp. 1–76, 1907.

[115] G. Frobenius, F. G. Frobenius, F. G. Frobenius, F. G. Frobenius, and G. Mathematician, "Über matrizen aus nicht negativen elementen," *Sitzung der Physikalisch-Mathematischen*, vol. 12, 1912.

[116] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*, vol. 71. Siam, 2000.

[117] H. Wielandt, "Unzerlegbare, nicht negative matrizen," *Mathematische Zeitschrift*, vol. 52, no. 1, pp. 642–648, 1950.

[118] D. M. Gordon, *Ant Encounters: Interaction Networks and Colony Behavior*. Princeton University Press, 2010.

[119] D. M. Gordon, "The evolution of the algorithms for collective behavior," *Cell Systems*, vol. 3, no. 6, pp. 514–520, 2016.

[120] D. J. Sumpter and M. Beekman, "From nonlinearity to optimality: pheromone trail foraging by ants," *Animal Behaviour*, vol. 66, no. 2, pp. 273–280, 2003.

[121] R. Beckers, J.-L. Deneubourg, S. Goss, and J. M. Pasteels, "Collective decision making through food recruitment," *Insectes sociaux*, vol. 37, no. 3, pp. 258–267, 1990.

[122] M. J. Greene and D. M. Gordon, "Cuticular hydrocarbons inform task decisions," *Nature*, vol. 423, pp. 32 EP –, 05 2003.

[123] D. M. Gordon, "Behavioral flexibility and the foraging ecology of seed-eating ants," *The American Naturalist*, pp. 379–411, 1991.

[124] D. M. Gordon, A. Guetz, M. J. Greene, and S. Holmes, "Colony variation in the collective regulation of foraging by harvester ants," *Behavioral Ecology*, vol. 22, no. 2, pp. 429–435, 2011.

[125] D. M. Gordon, K. N. Dektar, and N. Pinter-Wollman, "Harvester ant colony variation in foraging activity and response to humidity," *PLoS One*, vol. 8, no. 5, p. e63363, 2013.

[126] D. M. Gordon, "How colony growth affects forager intrusion between neighboring harvester ant colonies," *Behavioral Ecology and Sociobiology*, vol. 31, no. 6, pp. 417–427, 1992.

[127] D. M. Lane, D. Scott, M. Hebl, R. Guerra, D. Osherson, and H. Zimmer, "Introduction to statistics, online edition," 2017.

[128] K. M. Passino, T. D. Seeley, and P. K. Visscher, "Swarm cognition in honey bees," *Behavioral Ecology and Sociobiology*, vol. 62, no. 3, pp. 401–414, 2008.

[129] J. A. Marshall, R. Bogacz, A. Dornhaus, R. Planqué, T. Kovacs, and N. R. Franks, "On optimal decision-making in brains and social insect colonies," *Journal of the Royal Society Interface*, vol. 6, no. 40, pp. 1065–1074, 2009.

[130] C. Curto, S. Sakata, S. Marguet, V. Itskov, and K. D. Harris, "A simple model of cortical dynamics explains variability and state dependence of sensory responses in urethane-anesthetized auditory cortex," *Journal of Neuroscience*, vol. 29, no. 34, pp. 10600–10612, 2009.

[131] H. Safaai, R. Neves, O. Eschenko, N. K. Logothetis, and S. Panzeri, "Modeling the effect of locus coeruleus firing on cortical state dynamics and single-trial sensory processing," *Proceedings of the National Academy of Sciences*, vol. 112, no. 41, pp. 12834–12839, 2015.

[132] E. Pless, J. Queirolo, N. Pinter-Wollman, S. Crow, K. Allen, M. B. Mathur, and D. M. Gordon, "Interactions increase forager availability and activity in harvester ants," *PloS One*, vol. 10, no. 11, p. e0141971, 2015.

[133] D. Tudor, "A deterministic model for herpes infections in human and animal populations," *SIAM Review*, vol. 32, no. 1, pp. 136–139, 1990.

[134] M. G. M. Gomes, L. J. White, and G. F. Medley, "Infection, reinfection, and vaccination under suboptimal immune protection: epidemiological perspectives," *Journal of Theoretical Biology*, vol. 228, no. 4, pp. 539–549, 2004.

[135] P. Van Mieghem and R. van de Bovenkamp, "Accuracy criterion for the mean-field approximation in susceptible-infected-susceptible epidemics on networks," *Physical Review E*, vol. 91, no. 3, p. 032812, 2015.

[136] J. P. Gleeson, S. Melnik, J. A. Ward, M. A. Porter, and P. J. Mucha, "Accuracy of mean-field theory for dynamics on real-world networks," *Physical Review E*, vol. 85, no. 2, p. 026106, 2012.

[137] M. Pinsky and S. Karlin, *An Introduction to Stochastic Modeling*. Academic press, 2010.

[138] A. Khanafer, T. Başar, and B. Gharesifard, "Stability of epidemic models over directed graphs: A positive systems approach," *Automatica*, vol. 74, pp. 126 – 134, 2016.

[139] A. Lajmanovich and J. A. Yorke, "A deterministic model for gonorrhea in a nonhomogeneous population," *Mathematical Biosciences*, vol. 28, no. 3, pp. 221 – 236, 1976.

[140] M. Alizadeh, B. Atikoglu, A. Kabbani, A. Lakshmikantha, R. Pan, B. Prabhakar, and M. Seaman, "Data center transport mechanisms: Congestion control theory and IEEE standardization," in *2008 46th Annual Allerton Conference on Communication, Control, and Computing*, pp. 1270–1277, IEEE, 2008.

[141] K. Klemm, M. Á. Serrano, V. M. Eguíluz, and M. San Miguel, "A measure of individual role in collective dynamics," *Scientific Reports*, vol. 2, p. 292, 2012.

[142] G. Lawyer, "Understanding the influence of all nodes in a network," *Scientific Reports*, vol. 5, p. 8665, 2015.

[143] J.-G. Liu, J.-H. Lin, Q. Guo, and T. Zhou, "Locating influential nodes via dynamics-sensitive centrality," *Scientific Reports*, vol. 6, p. 21380, 2016.

[144] M. Šikić, A. Lančić, N. Antulov-Fantulin, and H. Štefančić, "Epidemic centralityis there an underestimated epidemic impact of network peripheral nodes?," *The European Physical Journal B*, vol. 86, no. 10, p. 440, 2013.

[145] S. Tang, X. Teng, S. Pei, S. Yan, and Z. Zheng, "Identification of highly susceptible individuals in complex networks," *Physica A: Statistical Mechanics and its Applications*, vol. 432, pp. 363–372, 2015.

[146] N. J. Watkins, C. Nowzari, and G. J. Pappas, "Inference, prediction and control of networked epidemics," in *2017 American Control Conference (ACC)*, pp. 5611–5616, May 2017.

[147] A. Khanafer and T. Basar, "An optimal control problem over infected networks," in *Proceedings of the International Conference of Control, Dynamic Systems, and Robotics, Ottawa, Ontario, Canada*, 2014.

[148] V. M. Preciado and A. Jadbabaie, "Spectral analysis of virus spreading in random geometric networks," in *Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference*, pp. 4802–4807, IEEE, 2009.

[149] P. Van Mieghem, D. Stevanović, F. Kuipers, C. Li, R. Van De Bovenkamp, D. Liu, and H. Wang, "Decreasing the spectral radius of a graph by link removals," *Physical Review E*, vol. 84, no. 1, p. 016101, 2011.

[150] R. Thenius, T. Schmickl, and K. Crailsheim, "Optimisation of a honeybee-colony's energetics via social learning based on queuing delays," *Connection Science*, vol. 20, no. 2-3, pp. 193–210, 2008.

[151] M. D. Cox and M. R. Myerscough, "A flexible model of foraging by a honey bee colony: the effects of individual behaviour on foraging success," *Journal of Theoretical Biology*, vol. 223, no. 2, pp. 179 – 197, 2003.

[152] T. Schmickl and K. Crailsheim, "Costs of environmental fluctuations and benefits of dynamic decentralized foraging decisions in honey bees," *Adaptive Behavior*, vol. 12, no. 3-4, pp. 263–277, 2004.

[153] B. Granovskiy, T. Latty, M. Duncan, D. J. T. Sumpter, and M. Beekman, "How dancing honey bees keep track of changes: the role of inspector bees," *Behavioral Ecology*, vol. 23, no. 3, pp. 588–596, 2012.

[154] D. M. Gordon, "Dynamics of task switching in harvester ants," *Animal Behaviour*, vol. 38, no. 2, pp. 194 – 204, 1989.

[155] D. M. Gordon, "Behavioral flexibility and the foraging ecology of seed-eating ants," *The American Naturalist*, vol. 138, no. 2, pp. 379–411, 1991.

[156] D. Wagner, M. Tissot, and D. Gordon, "Task-related environment alters the cuticular hydrocarbon composition of harvester ants," *Journal of Chemical Ecology*, vol. 27, no. 9, pp. 1805–1819, 2001.

[157] O. Feinerman and A. Korman, "Individual versus collective cognition in social insects," *Journal of Experimental Biology*, vol. 220, no. 1, pp. 73–82, 2017.

[158] D. A. Friedman, A. Pilko, D. Skowronska-Krawczyk, K. Krasinska, J. W. Parker, J. Hirsh, and D. M. Gordon, "The role of dopamine in the collective regulation of foraging in harvester ants," *iScience*, vol. 8, pp. 283 – 294, 2018.

[159] D. M. Gordon, "The development of an ant colony's foraging range," *Animal Behavior*, vol. 49, no. 3, pp. 649–659, 1995.

[160] D. M. Gordon, S. Holmes, and S. Nacu, "The short-term regulation of foraging in harvester ants," *Behavioral Ecology*, vol. 19, no. 1, pp. 217–222, 2008.

[161] A. Roth and M. C. van Rossum, "Modeling synapses," in *Computational Modeling Methods for Neuroscientists* (E. D. Schutter, ed.), ch. 8, pp. 139–160, Cambridge: The MIT press, 2009.

[162] L. Lapicque, "Recherches quantitatives sur lexcitation électrique des nerfs traitée comme une polarisation," *Journal de Physiologie et de Pathologie Generalej*, vol. 9, no. 1, pp. 620–635, 1907.

[163] H. Tuckwell, *Introduction to Theoretical Neurobiology. Vol. 1, Linear Cable Theory and Dendritic Structure and Stochastic Theories*. Cambridge Studies in Mathematical Biology, Cambridge: Cambridge University Press, 1988.

[164] R. B. Stein, "A theoretical analysis of neuronal variability," *Biophysical Journal*, vol. 5, no. 2, p. 173, 1965.

[165] R. B. Cooper, *Introduction to queueing theory*. Washington, DC: North Holland, 2nd ed., 1981.

[166] P. A. Lewis and G. S. Shedler, "Simulation of nonhomogeneous Poisson processes by thinning," *Naval Research Logistics*, vol. 26, no. 3, pp. 403–413, 1979.

[167] D. M. Gordon, S. Holmes, and S. Nacu, "The short-term regulation of foraging in harvester ants," *Behavioral Ecology*, vol. 19, no. 1, pp. 217–222, 2008.

[168] "The Weather Company LLC, weather underground - ramuda drive knmrodeo3." `https://www.wunderground.com/personal-weather-station/dashboard?ID=KNMRODEO3#history`.

[169] E. Mishchenko and K. N. Rozov, *Differential equations with Small Parameters and Relaxation Oscillations*. Boston: Springer, 1980.

[170] K. Borovkov and A. Novikov, "On a piece-wise deterministic markov process model," *Statistics & Probability Letters*, vol. 53, no. 4, pp. 421–428, 2001.

[171] E. Gilbert and H. Pollak, "Amplitude distribution of shot noise," *Bell Systems Technical Journal*, vol. 39, no. 2, pp. 333–350, 1960.

[172] A. Khelil, C. Becker, J. Tian, and K. Rothermel, "An epidemic model for information diffusion in MANETs," in *Proc. 5th ACM International Workshop on Modeling Analysis and Simulation of Wireless and Mobile Systems*, pp. 54–60, 2002.

[173] P. E. Paré, C. L. Beck, and A. Nedich, "Epidemic processes over time-varying networks," *IEEE Transactions on Control of Network Systems*, 2017.

[174] J. Liu, P. E. Paré, A. Nedić, C. Y. Tang, C. L. Beck, and T. Başar, "On the analysis of a continuous-time bi-virus model," in *Proc. 55th IEEE Conference on Decision and Control (CDC)*, pp. 290–295, 2016.

[175] V. M. Preciado, M. Zargham, C. Enyioha, A. Jadbabaie, and G. J. Pappas, "Optimal resource allocation for network protection against spreading processes," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 1, pp. 99–108, 2014.

[176] N. J. Watkins, C. Nowzari, V. M. Preciado, and G. J. Pappas, "Optimal resource allocation for competitive spreading processes on bilayer networks," *IEEE Transactions on Control of Network Systems*, 2016.

[177] K. Kandhway and J. Kuri, "How to run a campaign: Optimal control of SIS and SIR information epidemics," *Applied Mathematics and Computation*, vol. 231, pp. 79–92, 2014.

[178] M. Clements, R. Betts, E. Tierney, and B. Murphy, "Serum and nasal wash antibodies associated with resistance to experimental challenge with influenza A wild-type virus," *Journal of Clinical Microbiology*, vol. 24, no. 1, pp. 157–160, 1986.

[179] S. Verver, R. M. Warren, N. Beyers, M. Richardson, G. D. van der Spuy, M. W. Borgdorff, D. A. Enarson, M. A. Behr, and P. D. van Helden, "Rate of reinfection tuberculosis after successful treatment is higher than rate of new tuberculosis," *American Journal of Respiratory and Critical Care Medicine*, vol. 171, no. 12, pp. 1430–1435, 2005.

[180] N. Stollenwerk, J. Martins, and A. Pinto, "The phase transition lines in pair approximation for the basic reinfection model SIRI," *Physics Letters A*, vol. 371, no. 5-6, pp. 379–388, 2007.

[181] J. Gómez-Gardeñes, A. S. de Barros, S. T. Pinho, and R. F. S. Andrade, "Abrupt transitions from reinfections in social contagions," *Europhysics Letters*, vol. 110, no. 5, p. 58006, 2015.

[182] T. Kuniya, "Global behavior of a multi-group SIR epidemic model with age structure and an application to the chlamydia epidemic in japan," *SIAM Journal on Applied Mathematics*, vol. 79, no. 1, pp. 321–340, 2019.

[183] F. Jin, E. Dougherty, P. Saraf, Y. Cao, and N. Ramakrishnan, "Epidemiological modeling of news and rumors on Twitter," in *Proceedings of the 7th Workshop on Social Network Mining and Analysis*, SNAKDD '13, (New York, NY, USA), pp. 8:1–8:9, ACM, 2013.

[184] P. E. Paré, C. L. Beck, and A. Nedić, "Epidemic processes over time-varying networks," *IEEE Transactions on Control of Network Systems*, vol. 5, pp. 1322–1334, Sep. 2018.

[185] L. Yang, X. Yang, and Y. Y. Tang, "A bi-virus competing spreading model with generic infection rates," *IEEE Transactions on Network Science and Engineering*, vol. 5, pp. 2–13, Jan 2018.

[186] N. Stollenwerk, S. van Noort, J. Martins, M. Aguiar, F. Hilker, A. Pinto, and G. Gomes, "A spatially stochastic epidemic model with partial immunization shows in mean field approximation the reinfection threshold," *Journal of Biological Dynamics*, vol. 4, no. 6, pp. 634–649, 2010.

[187] M. W. Hirsch, S. Smale, and R. L. Devaney, *Differential Equations, Dynamical Systems, and an Introduction to Chaos*. Academic Press, 2012.

[188] P. N. Tu, *Dynamical Systems: an Introduction with Applications in Economics and Biology*. Springer Science & Business Media, 2012.

[189] D. Henry, *Geometric Theory of Semilinear Parabolic Equations*, vol. 840. Springer, 1981.

[190] B. Aulbach, "Continuous and discrete dynamics near manifolds of equilibria.," *Lecture Notes in Mathematics*, 1984.

[191] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, vol. 662. Oxford Clarendon, 1965.

[192] H. K. Khalil and J. Grizzle, *Nonlinear Systems*, vol. 3. Prentice hall Upper Saddle River, NJ, 2002.

[193] J. Roozenbeek and S. van der Linden, "Fake news game confers psychological resistance against online misinformation," *Palgrave Communications*, vol. 5, no. 1, p. 65, 2019.