Montclair State University
Montclair State University Digital Commons

5-2018

Outbreak and Extinction Dynamics in Stochastic Populations

Garrett T. Nieddu

Montclair State University

OUTBREAK AND EXTINCTION DYNAMICS

IN STOCHASTIC POPULATIONS

A DISSERTATION

Submitted to the Faculty of

Montclair State University in partial fulfillment

of the requirements

for the degree of Doctor of Philosophy

by

GARRETT T. NIEDDU

Montclair State University

Upper Montclair, NJ

2018

Dissertation Chair: Dr. Eric Forgoston

MONTCLAIR STATE UNIVERSITY

THE GRADUATE SCHOOL

DISSERTATION APPROVAL

We hereby approve the Dissertation

OUTBREAK AND EXTINCTION DYNAMICS

IN STOCHASTIC POPULATIONS

of

Garrett T. Nieddu

Candidate for the Degree:

Doctor of Philosophy

Department of Earth and
Environmental Studies

Dissertation Committee:

Certified by:

Dr. M. Scott Herness
Vice Provost for Research and
Dean of The Graduate School

Dr. Eric Forgoston
Dissertation Chair

Dr. Lora Billings

April 12, 2018
Date

Dr. Jennifer Krumins

Dr. Michael Khasin

Abstract

OUTBREAK AND EXTINCTION DYNAMICS

IN STOCHASTIC POPULATIONS

by Garrett T. Nieddu

Each piece of work herein examines nonlinear population dynamics using methods from deterministic dynamical systems, stochastic processes and statistical mechanics. This dissertation is the compilation of three independent – but related – pieces of work: the first investigates an isolated population that is capable of maintaining multiple carrying capacities; the second project examines a stochastic Ebola model with a zoonotic disease reservoir; and the third looks at a basic disease-invasion model to characterize outbreak vulnerability and the connectedness of supposedly separate populations.

Each of these three chapters explore the interplay between interconnected systems, without explicitly modeling the elements that are external to the system of interest. The goal is to take a very large and complex lattice of interconnected biological systems and isolate the necessary components, so that modeling is both practical and utilitarian. These works are done in either an ecological or an epidemiological context, but the results in each chapter can be broadly applied to outbreak, invasion, extinction, and connectedness in stochastic population modeling.

Acknowledgements

I would first like to thank my advisor Eric Forgoston, whose mentorship has been invaluable. I have gained as much from my exposure to his personal interests and curious nature as I have from his professional expertise, and I couldn't say which I appreciate the most. He has devoted quite a lot of time to my training as a scientist. I hope that one day I can pay it forward.

I would also like to thank Lora Billings for her mentorship. One of her first bits of advice to me when we started working together was to keep moving forward, which I keep trying to do. She has kept me honest throughout, and my work is better for her input and her contributions.

Additionally I would like to thank my other committee members. Thanks to Jennifer Krumins for her patience and her guidance, which has given me insight into experimental and theoretical ecology that I would otherwise lack. Thanks to Michael Khasin for very useful discussions, particularly related to the practical limitations of the stochastic methods being applied. Both Jenn and Michael have helped to make my work more practical.

Thanks to Mary Lou West, my first research advisor. She is a great teacher and a great purveyor of science.

Thanks to David Trubatch, a good guy and a good mentor. He taught me a lot, especially in mathematical computing, but far beyond as well.

Thanks to Pankaj Lal for good conversation, good collaboration, and his all around good nature. Thanks to Eric Stern for interesting and engaging conversations that encouraged me to keep my mind and eyes open. Thanks to all of my friends and officemates and co-workers over these years. Interesting discussions and companionship has made the last five years enjoyable, and has been personally enriching. Thanks to my partners at Hilltop Tutoring Center.

I owe the most to my parents, who were the ones that initially equipped me for problem solving, mostly through love and the playing of strategy games.
Go figure.

Thanks to Diane Nieddu. Her keen analytical mind and discerning nature have been challenging me for as long as I've been alive. The traits seem to have rubbed off a bit, and without them I wouldn't make much of a scientist. *Thanks Ma.*

Thanks to Phillip Nieddu. He has encouraged me to push myself in all aspects of my life, and has continuously expressed faith in my ability to do better. By trying to emulate his example of hard work and sacrifice, I have managed to achieve far above my apparent potential. *Thanks Pops.*

To Meegs,

Without your support and companionship, this task would have been insurmountable.

Contents

List of Tables

List of Figures

1 Introduction

1.1 Brief Historical Review

The use of data to study infectious disease in humans is at least as old as the work ti-
tled 'Natural and Political Observations Made upon the Bills of Mortality', originally
published in 1662 by John Graunt [Graunt, 1663]. These observations were made upon
bubonic plague data with the hope of predicting new outbreaks. In 1760 Daniel Bernoulli
presented his work, 'An Attempt at a New Analysis of the Mortality Caused by Smallpox
and of the Advantages of Inoculation to Prevent it', to the Royal Academy of Sciences in
Paris [Bernoulli and Blower, 2004]. In that article he used disease infection and mortality
data to argue in favor of universal smallpox inoculation.

In 1798 a basic but insightful observation was made by Robert Malthus in his work
titled 'An Essay on the Principle of Population'. Therein he states that "Population, when
unchecked, increases in a geometrical ratio." [Malthus, 1888](pg. 4), which suggests that
unchecked population-growth should be modeled as an exponential function. This same
result can be derived from a linear ordinary differential equation (ODE) in which the
growth rate is modeled proportionately to population size as

$$\frac{dP}{dt} = rP, \tag{1}$$

where P is the population size and r is the growth rate constant.

In 1838 Pierre-François Verhulst was the first mathematician to apply a nonlinear ODE to the study of population dynamics [Verhulst, 1838]. The Verhulst model

$$\frac{dP}{dt} = rP\left(1 - \frac{P}{k}\right), \tag{2}$$

where P and r represent the population size and the growth rate respectively, does not show unrestricted growth. Instead, the population tends toward the carrying capacity at $P = k$. When $P > k$ the population will decline to the carrying capacity. When $P < k$ the population will initially experience exponential growth, but after some time the growth slows down and the population approaches the carrying capacity. When $P = k$ one sees that $\frac{dP}{dt} = 0$ so that the population will persist unchanged at the carrying capacity. The solution to the Verhulst model is often called the logistic equation. The nonlinear P^2 term in Eq. 2 involves the population P with itself. This can be thought of as an overcrowding or competition term.

Almost two hundred years has passed since the publication of the Verhulst model. During that time many interesting and complex dynamics have been investigated using the basic premise that a two-way interaction among discrete biological population members can be appropriately modeled using a simple cross term $c_1 XY$ where c_1 is a constant, and X and Y represent the size of two interacting populations. These are often called "mass-action" terms because of their relationship to the law of mass-action for chemical reaction equations, where such terms have been shown to correctly predict the dynamic

behavior of chemical reactions, including reaction rates and dynamic equilibrium.

This is the revelation that can be primarily credited as the predecessor to the modern approach to both ecological population modeling and population level infectious disease modeling. This premise is used in seminal work by Robert May in both ecological stability and infectious disease modeling [Anderson et al., 1992, May, 2001, May, 1971, May, 1972].

The modern study of infectious disease dynamics, as well as the more general study of population dynamics, is largely devoted to stability-related investigations. There is quite a lot of work devoted to stability in dynamical population models – the reader can reference [May, 2001, Caswell, 2001, May, 1971, May, 1972, Pimm, 1984] along with the citations therein for expansive reading. In general, a deterministic model describing infectious disease dynamics will consist of a nonlinear system of equations with the form,

$$\dot{\mathbf{X}} = \mathbf{M}\mathbf{X}, \tag{3}$$

where $\mathbf{M}$ is a matrix with entries dependent on the state vector $\mathbf{X}$. Dot notation is used to indicate the time derivative. An equilibrium state is defined as any point $\mathbf{X}^* = [X_1, X_2, \ldots, X_n]$ where $\dot{\mathbf{X}}$ evaluated at $\mathbf{X}^*$ is equal to the zero vector. After a small perturbation, a solution trajectory will tend to return to a stable equilibrium point, and will tend to move away from an unstable equilibrium point. The stability of an equilibrium point is indicated by the eigenvalues of the linearized system's Jacobian matrix. When all of the eigenvalues have negative real components, the equilibrium point at which the Jacobian was evalu-

ated is locally stable. The stability analysis of an endemic disease model is shown as an example.

Example of Stability Analysis: One of the most simple disease models is the Susceptible-Infectious-Susceptible (*SIS*) model. When modeling disease dynamics in an *SIS* model, the population is separated into two classes: the Susceptible *S* class and the Infectious *I* class. Individuals in the Susceptible class may become infected through contact with an infectious individual and move to the Infectious class. A disease specific recovery rate γ is defined as the reciprocal of the average time spent in the Infectious class, and individuals move from the Infectious class back to the Susceptible class with that rate. All children are born into the Susceptible class with rate μ, and natural death is equally likely from both the Susceptible and Infectious classes with rate μ. The deterministic mean-field dynamical model is given as

$$\frac{dS}{dt} = \mu N - \mu S - \beta SI/N + \gamma I = f(S,I), \tag{4}$$

$$\frac{dI}{dt} = \beta SI/N - \mu I - \gamma I = g(S,I), \tag{5}$$

where β is the contact rate, and *N* is the constant population size. The equilibrium points, or steady states, are found by setting the right-hand side of Eq. 4 and Eq. 5 to zero so that

$$\mu N - \mu S - \beta SI/N + \gamma I = 0, \tag{6}$$

$$\beta SI/N - \mu I - \gamma I = 0. \tag{7}$$

There are two equilibrium points that satisfy Eqs. 6 and 7. The first, $(S_0, I_0) = (N, 0)$

is the extinct state where the population consists entirely of susceptible individuals with

zero infectious individuals; this is called the disease-free equilibrium (DFE). The second,

$(S^*, I^*) = \left(\frac{(\mu+\gamma)N}{\beta}, \left[1 - \frac{\mu+\gamma}{\beta}\right]N \right)$ is the endemic state, for which there is a non-zero num-

ber of infectious individuals.

The Jacobian matrix for the linearized SIS system is

$$J = \begin{bmatrix} \frac{\partial}{\partial S}f(S,I) & \frac{\partial}{\partial I}f(S,I) \\ \frac{\partial}{\partial S}g(S,I) & \frac{\partial}{\partial I}g(S,I) \end{bmatrix} = \begin{bmatrix} -\mu - \beta I/N & -\beta S/N + \gamma \\ \beta I/N & \beta S/N - \mu - \gamma \end{bmatrix}. \tag{8}$$

The Jacobian evaluated at the extinct state is

$$J\Big|_{(N,0)} = \begin{bmatrix} -\mu & -\beta + \gamma \\ 0 & \beta - \mu - \gamma \end{bmatrix}. \tag{9}$$

The eigenvalues of J are given as

$$\lambda_1 = -\mu, \tag{10}$$

$$\lambda_2 = \beta - (\mu + \gamma). \tag{11}$$

The spectral radius is defined as the largest eigenvalue. Since μ, γ, and β are all positive,

the spectral radius is λ_2 because λ_1 is always negative. In order to make the DFE stable

the real part of the spectral radius $\lambda_2 < 0$. Therefore, the DFE is stable when $\frac{\beta}{\mu+\gamma} < 1$.

The non-dimensional quantity $\frac{\beta}{\mu+\gamma}$ is known as the basic reproduction number and is denoted R_0. If $R_0 < 1$ then the DFE is stable. If $R_0 > 1$ then the DFE is unstable. By computing the eigenvalues of the Jacobian matrix at the endemic equilibrium point, it is easy to show that when the DFE is stable, the endemic state is unstable, and when the DFE is unstable, the endemic state is stable. In fact, one can show that there is a transcritical bifurcation at the point when $R_0 = 1$ [Strogatz, 2014].

The formal definition of the basic reproduction number of a system is the number of secondary infections that will be caused by a single infectious individual in an otherwise entirely susceptible population. Based on that definition, it follows that for $R_0 > 1$ a disease outbreak will occur, while for $0 < R_0 < 1$ disease extinction will occur. For this reason the reproductive number is often used to indicate the devastation potential of a particular disease in a particular population. Stability analysis and the deterministic modeling approach, however, is only part of the story.

1.2   Stochastic Modeling

The definition for stability that has been presented is only relevant for deterministic studies, and much work pre-supposes that stability is a good analog for the state of natural population systems. Contrary to that supposition, population systems are in a constant state of change, and at best can be called semi-stable or metastable. A state will be called metastable if it is maintained for extended periods of time in the stochastic system, and corresponds to an equilibrium point in the deterministic system. While 'semi-stable'

seems to better describe the behavior of natural population systems, deterministic systems are generally not capable of displaying these nuanced stabilities. Deterministic systems have equilibrium points that are either stable or unstable. Stochastic models, however, are fundamentally uncertain. Even for systems which strongly favor a single state, there will be some probability density function (PDF) that describes the probability of being found at a myriad of different states. That PDF may or may not change in time; a PDF that does not change in time is called stationary.

The range of stability regimes seen in the stochastic modeling approach is qualitatively similar to what is seen in data, while the deterministic description of stability is not. This is not such a surprise, since the primary interactions that define population systems are inherently random. Consider sexual reproduction, natural death, and predation; all of these are fundamentally random events. Sexual reproduction relies on the happenstance coupling of fertile mating partners. Natural death and predation are both heavily influenced by genetics and behavior. A slow or distracted member of a herd is more likely to be eaten than a fast and attentive member of the same herd. Disease spread is similarly reliant on random interactions between infectious and non-infectious individuals. These observations suggest that population level infectious disease modeling should not only be considered from a deterministic perspective, but should also be investigated using the approaches from statistical physics and stochastic processes.

The random events that comprise a stochastic process can have numerous sources. These can be as personal as the unpredictability of decision making in intelligent popula-

tion members, or the randomness can be imposed on the system, such as the random effect of the wind on a pollination process. The randomness that occurs from the interaction of population members is generally considered to be internal stochastic noise. Randomness that is imposed on a system from an external source is generally considered to be external stochastic noise. This dissertation will only be directly concerned with the effects of internal noise.

Stochastic models and statistical methods have a long history of use in the study of disease and population dynamics. Reviews from the early 1950s seem to indicate that the first such uses of stochastic models can be dated between the late 1930s and the late 1940s [Reid, 1953]. In 1948 David Kendall published a paper, 'On The Generalized 'Birth-and-Death' Process', in which the principles of stochastic processes are applied to population dynamics [Kendall, 1948]. In that paper he indicates that the concept of applying methods from stochastic processes to the study of population dynamics can be attributed to William Feller, almost a decade before his own paper [Feller, 1939]. In his 1948 article, 'Some Evolutionary Stochastic Processes', Maurice Bartlett dedicated a section to population growth [Bartlett, 1949]. It is implied through citation that Bartlett covered this topic in his 1946 Fall semester course at North Carolina State University titled "Stochastic Processes".

In that same paper Bartlett had another section titled 'Mixed Processes in Epidemiology' where he presents a stochastic model of disease spread [Bartlett, 1949]. Found in that section is an analogy between the "molecular association of two atoms in statistical

mechanics" and the "infection by 'contact' between infected and uninfected persons" (pg. 225). This analogy was used throughout the twentieth century and continues to be used for modern stochastic disease models. In Peter Whittle's 1956 paper titled, 'On The Use of the Normal Approximation in the Treatment of Stochastic Processes', disease spread is treated as a Markov process, and population sizes are treated as normally distributed random variables [Whittle, 1957].

This approach to the study of population and disease dynamics was expanded upon and refined during the second half of the twentieth century. In 1992 the comprehensive 'Stochastic Processes of Physics and Chemistry' by Nicolaas van Kampen was published, in which the application of stochastic processes to population and disease dynamics is well represented [van Kampen, 1992]. As an additional resource for expanded application to social sciences and population dynamics one can refer to the 'Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences' by Crispin Gardiner [Gardiner, 2004].

Of particular relevance to this dissertation is the study of population and disease extinction; a topic that the deterministic approach to population and disease modeling is ill-equipped to address. Disease extinction will be studied here as a stochastic transition between the metastable endemic equilibrium and the disease-free state. Early investigations into these stochastic transitions used a Gaussian form for the probability density function. Through use of a van Kampen large system-size expansion, the stochastic master equation can be transformed into a Fokker-Planck equation, which results in a normally distributed

PDF. It has been shown that, when determined through use of the Fokker-Planck equation, extinction times will have exponential error with increases in the population size [Gaveau and Schulman, 1996, Doering et al., 2005].

The use of a Wentzel-Kramers-Brillouin (WKB) approximation for the PDF in the master equation was originally reported in 'Fluctuation and Relaxation of Macrovariables' [Kubo et al., 1973]. That paper describes an approximation of the full stochastic system as a Hamiltonian system, an idea which was formalized for discrete systems by Hu Gang in the 1980s (see Ref. [Gang, 1987] and the references therein). These approaches have been refined for use in the context of discrete population systems, and it has been shown repeatedly that the WKB approximation gives accurate results [Elgart and Kamenev, 2004, Kessler and Shnerb, 2007, Forgoston et al., 2011, Assaf and Meerson, 2010]. This approach is used repeatedly in this dissertation, and so the general method will be briefly outlined here, and then will appear in less general forms in the subsequent chapters.

The WKB Approximation in the Stochastic Master Equation: If the discrete transitions in a large-population system are short and uncorrelated in time, then the system is a Markov process and the evolution of the PDF is described by a master equation. Given a state vector $\mathbf{X} \in \mathbb{R}^n$ where n is the number of subpopulations or unique classes in a population model, and each $X_i \in [X_1, X_2, \ldots, X_n]$ describes the size of each of these subpopula-

tions, then the stochastic master equation is given by

$$\frac{d\rho(\mathbf{X},t)}{dt} = \sum_{\mathbf{r}} \left[W(\mathbf{X}\text{-}\mathbf{r};\mathbf{r})\rho(\mathbf{X}\text{-}\mathbf{r},t) - W(\mathbf{X};\mathbf{r})\rho(\mathbf{X},t) \right]. \tag{12}$$

The transition rate from a state $\mathbf{X}$ to $\mathbf{X}+\mathbf{r}$ is given by $W(\mathbf{X};\mathbf{r})$, where $\mathbf{r} \in \mathbb{R}^n$ is the transition's increment in $\mathbf{X}$. A scaled master equation can be determined using the substitutions $\mathbf{x} = \mathbf{X}/N$ and $W(\mathbf{X};\mathbf{r}) \equiv W(N\mathbf{x};\mathbf{r}) = Nw_{\mathbf{r}}(\mathbf{x}) + u_{\mathbf{r}}(\mathbf{x}) + \mathcal{O}(1/N)$, where N is a large system size parameter such as population size.

For N much larger than one, the WKB approximation for the scaled master equation can be used [Kubo et al., 1973, Gang, 1987, Dykman et al., 1994, Elgart and Kamenev, 2004, Kessler and Shnerb, 2007, Forgoston et al., 2011, Schwartz et al., 2011]. The WKB approximation assumes that the PDF can be approximated as the following exponential function,

$$\rho(\mathbf{x},t) = \pi_0 \exp(-NS(\mathbf{x},t)), \tag{13}$$

where $S(\mathbf{x},t)$ is a function known as the action and π_0 is the prefactor. After Eq. (13) is substituted into the scaled master equation, a Taylor series expansion reveals the leading order Hamilton-Jacobi equation, $\mathscr{H}(\mathbf{x},\mathbf{p}) = -\frac{dS}{dt}$, where

$$\mathscr{H}(\mathbf{x},\mathbf{p}) = \sum_{\mathbf{r}} w_{\mathbf{r}}(\mathbf{x}) \left[\exp(\mathbf{p} \cdot \mathbf{r}) - 1 \right] \tag{14}$$

is the effective Hamiltonian. The new state vector $\mathbf{p}$ is the conjugate momentum and is defined as $\mathbf{p} = dS/d\mathbf{x}$. From the Hamiltonian in Eq. (14), one can easily derive Hamil-

ton's equations

$$\dot{\mathbf{x}} = \frac{\partial \mathcal{H}(\mathbf{x}, \mathbf{p})}{\partial \mathbf{p}}, \qquad \dot{\mathbf{p}} = -\frac{\partial \mathcal{H}(\mathbf{x}, \mathbf{p})}{\partial \mathbf{x}}. \tag{15}$$

In the 1992 paper 'Optimal Paths and Prehistory Problem for Large Fluctuations in Noise-Driven Systems', these methods, along with the concept of the optimal path, was applied to a noise driven analog electronic circuit [Dykman et al., 1992]. In that paper, a theoretical most likely path is defined from one stable state to another, and compared to an ensemble of paths taken in the physical system. In his 1994 paper, 'Large Fluctuations and Optimal Paths in Chemical-kinetics', Mark Dykman extends these methods to discrete populations [Dykman et al., 1994]. The optimal path is the path of least action between the two points. In practice this means solving Eq. (15) with the two stability points as boundary conditions and an action minimizing constraint. Often this is a non-trivial task in systems of one or two dimensions, but higher dimensional systems are usually not analytically tractable. Numerical techniques including shooting methods, action minimizing methods, and application of finite-time Lyapunov exponents can be used to solve for the optimal path in higher dimensional systems [Schwartz et al., 2011, Bauver et al., 2016, Forgoston et al., 2011, Lindley and Schwartz, 2013]. This path is particularly important because it has been shown that the mean time to extinction (MTE) is exponentially dependent on the optimal action,

$$\tau \propto \exp\left(N\mathcal{S}_{\text{opt}}\right).$$

The approximation of stochastic systems as Hamiltonian systems is now widely used to study stochastic transitions in population models, and it has been shown to be particularly useful in the context of vaccination. Using these methods it has been shown that under the presence of non-systematic random vaccination the mean time to disease extinction can be improved exponentially [Dykman et al., 2008, Khasin and Dykman, 2011], although it has been shown that there is a benefit to Poissonian pulse vaccination over vaccination administered normally in time [Schwartz et al., 2009]. It then becomes unsurprising that group vaccination leads to even shorter disease extinction times then when population members act independently [Landsman and Schwartz, 2007]. Optimal periodic vaccination schedules can also result in exponential improvement to disease extinction times, even when the vaccination is scarce and large batch vaccination is not possible [Khasin et al., 2010]. In multiple serotype studies, such as are common for dengue virus, much the opposite has been observed. In these studies it is possible for a secondary infection's viral growth rate to be greater than the primary infections [Billings et al., 2008b, Billings et al., 2008a].

A related problem to disease and species extinction is that of disease outbreak and foreign species invasion. As with extinction, outbreak and invasion are fundamentally stochastic processes. Although stochastic disease modeling has been used in the context of disease outbreak (see for example Ref. [Schwartz et al., 2004]), it is still an underattended topic of study. The study of outbreak dynamics offers a new challenge; the random interactions that drive outbreak are between population members and elements exter-

nal to the population being studied. The difficulty is two-fold; first, the external sources of disease infection can be numerous. Secondly, we are likely to have incomplete knowledge of the external disease sources. In this dissertation, in order to ameliorate these obstacles, disease invasion from an external source will be considered using a generic stochastic introduction term, with introduction rate proportional to the number of susceptible individuals in the system of study.

1.3   Context of This Work

Any intelligent fool can make things bigger, more complex, and more violent. It takes a touch of genius – and a lot of courage to move in the opposite direction.

-Ernst F. Schumacher, 'Small is Beautiful', The Radical Humanist Vol. 37

Biological systems are naturally big, complex, and violent; and so, you can forgive the mathematician who chooses to model them as such. Great insight, however, can often be gained through relatively simple models, as is the case while trying to understand shifts in population dynamics. The system-shifts that are being discussed in this work are major deviations of the system state from the average behavior of the system. This includes disease outbreak, population extinction, ecological invasion of foreign biota, and explosions of abundance in both ecological and epidemiological contexts.

Most systems in biology and population ecology (as well as in physics, chemistry, and economics) spend most of their time fluctuating about some average system state.

The stronger the system's attraction to the average, the more stable the average is said to be. Given time, large fluctuations will be observed. Depending on how stable the state is, these large fluctuations may not have long lasting effects, or they may send the system spiraling out of control, or they may affect a transition into a different stability regime. Depending on the context this can be good, bad, or indifferent. For instance, if a diseased population is effectively driven to extinction, then a disease-free state is realized, which is widely accepted as a favorable transition. On the other hand, if it is a keystone species that is driven to extinction, then the the transition will result in more extinction events. Since it is widely accepted among ecologists that reduction in biodiversity is intrinsically undesirable, this would be considered a catastrophic transition.

The interconnectedness of the biological systems of the Earth make a holistic approach very difficult, and maybe impossible. Seemingly separate populations and systems are heavily coupled, and so it is unclear to what degree they are distinct from one another. It is useful to visualize the interconnected systems as an immense lattice of interconnected nodes, like a large net or chain linked fence. This lattice has nodes which have overlap in membership, and between the nodes and within the nodes there are nonlinear interactions that result in complex feedback loops. Although metastability is realized in this huge and complex system, it remains a great wonder and topic of study how this can possibly be the case.

The inhomogeneity of the interactions makes modeling difficult and parameterization impractical. In his essay, 'Perspectives on Theory at the Interface of Physics and

Biology', William Bialek – the John Archibald Wheeler/Battelle Professor in Physics at Princeton University – discusses the importance of a mathematical model's qualitative behavior to real world biological insight, and the comparative value of full parameterization [Bialek, 2017]. The necessary conclusion is that one must be selective about what is represented explicitly in a model, what will be accounted for implicitly, and what need not be included at all.

In this dissertation, although only a single population will be explicitly modeled, each system sits within a larger framework of systems. The interactions with elements external to the system will be considered in as simple a way as possible, while maintaining the complexity and richness of the biological system.

For the population level dynamical systems that are being discussed here, the external interactions being alluded to include human-to-human interactions as well as human-to-animal interactions and human-to-plant interactions. Increases in globalization and development have fundamentally changed the way that humans interact with the natural world, and with one another. Complex and irregular interactions between man and nature keeps zoonotic diseases (diseases that come from non-human disease vectors) including malaria, dengue, zika, and the Ebola virus disease to name just a few, a prime human health concern. As human development continues to encroach on the natural world, vulnerability to zoonotic disease outbreak will also continue to increase. Similarly, there is a constant threat from other human populations; as human populations continue to become more interconnected, we are perpetually increasing our vulnerability to disease introduction from

those populations. This means that the interconnectedness of human populations is also increasingly important to the understanding of disease spread, and that a population's vulnerability to disease outbreak increases as it becomes more tightly coupled with peripheral human and animal populations.

1.4 Summary

Each piece of work herein examines nonlinear population dynamics using methods from both deterministic dynamical systems, as well as stochastic processes and statistical mechanics. This dissertation is the compilation of three independent – but related – pieces of work: the first investigates an isolated population that is capable of maintaining multiple carrying capacities; the second project examines a stochastic Ebola model with a zoonotic disease reservoir; and the third looks at a basic disease-invasion model to characterize outbreak vulnerability and the connectedness of supposedly separate populations.

Chapter 2, 'Analysis and Control of Pre-extinction Dynamics in Stochastic Populations', contains an analytical and numerical study of extinction dynamics in a stochastic population model with multiple non-zero metastable equilibria. The model presented is an extension of an Allee model. In ecological populations that display the Allee effect, persistence of the population will depend on some minimum population size. This population size is called the Allee threshold, and if a population displaying the Allee effect realizes a population size below the threshold at any time, then the whole population will run towards extinction. If the population is above the threshold then the carrying capacity

will be realized and maintained. In the more complex model presented here, there are two additional states, including a second threshold and a second possible carrying capacity. With two realizable, non-zero carrying capacities, the system will be capable of switching between the two viable carrying capacities prior to extinction. A framework for understanding how switching in a one-dimensional population contributes to mean extinction time is presented. Control is applied, and population control efficacy is quantified using the mean time to extinction.

Chapter 3, 'Outbreak and Extinction Dynamics in a Stochastic Ebola Model', presents a stochastic Ebola virus disease (EVD) model that incorporates hospitalization, a deceased class, and a connection with the disease reservoir. The EVD's ability to remain latent in non-human hosts, along with its outbreak cycles into human populations, makes EVD a zoonotic disease. The fact that the natural reservoir is poorly understood makes the reservoir population impossible to dynamically model. In this model the connection (or coupling) between the human population and the reservoir is modeled as a constant threat, and outbreak (or spillover) events occur randomly. Since infection causing interactions between humans and the reservoir are rare and non-systematic, a stochastic approach is necessary. The results give insight into how outbreak and extinction cycles relate to disease control and prevention. Additionally, the concept of outbreak vulnerability is explored from a numerical perspective. It is shown that for an EVD population that is weakly coupled to the disease reservoir, the basic reproduction number from the deterministic model is a good indicator of outbreak vulnerability. The idea of outbreak

vulnerability is further explored in chapter 4.

In Chapter 4, 'Outbreak Vulnerability and Connectedness in a Stochastic $SIS\kappa$ Endemic Disease Model', outbreak vulnerability and and connectedness are quantified. In practical application of modeling results to real biological systems, it is necessary that a population be well-defined in order to ensure that a system and its parameters are relevant. The interconnectedness of physical and biological systems makes inclusion in a population fuzzy, in the sense that membership is non-binary, and non-absolute. The qualification of a group as a population is contextual, but largely relies on the extent to which a group of individuals can be said to be disjoint from a larger population, to which they are a subset. For instance, if one were to study the proliferation of the British royal family, it would be unnecessary to model all the people of Britain. Although the British royalty are part of the British population, for the sake of marriage they belong to a relatively small subset of the population. In this context, British royalty is sufficiently disjoint from the general public so as to constitute a sub-population.

Inversely, it is clear that when two populations are too strongly coupled they become indistinguishable from one another, but how should one decide the threshold for when a group qualifies as a sub-population? This chapter considers disease outbreak in a stochastic SIS invasion model (termed an $SIS\kappa$ model). Disease invasion will originate from a source external to the $SIS\kappa$ population, for which the coupling strength will be measured by the parameter κ. Boundaries in parameter space are identified for when outbreak is very rare, when it is relatively frequent, and when it is overwhelming. As in the EVD

model, the external source is not modeled explicitly. The results generalize and expand upon the numerical results from the Ebola model.

Each of these three chapters explore the interplay between interconnected systems, without explicitly modeling the elements that are external to the system of interest. The goal is to take a very large and complex lattice of interconnected biological systems and isolate the necessary components, so that modeling is both practical and utilitarian. These works are done in either an ecological or an epidemiological context, but the results in each chapter can be broadly applied to outbreak, invasion, extinction, and connectedness in stochastic population modeling.

The final chapter contains the conclusions, in which the work will be tied together, and placed in a larger context of dynamical population models. This will include a discussion of policy ramifications, and future work.

2   Analysis and Control of Pre-extinction Dynamics in Stochastic Populations

2.1   Introduction

It has long been known that noise can significantly affect physical and biological dynamical systems at a wide variety of levels. For example, in biology, noise can play a role in sub-cellular processes, tissue dynamics, and large-scale population dynamics [Tsimring, 2014]. Stochasticity can arise in a number of ways. For example, in epidemiological models, noise is due to the random interactions of individuals in a population as well as uncertainty in epidemic parameter values [Rand and Wilson, 1991, Billings et al., 2002]. In population ecology, noise may be the result of environmental factors including climatic effects, natural enemies, or inter-specific competition, or may be due to demography [Coulson et al., 2004].

Stochasticity manifests itself as either external or internal noise. External noise comes from a source outside the system being considered (e.g the growth of a species under climatic effects), and often is modeled by replacing an external parameter with a random process. Internal noise is inherent in the system itself and is caused by the random interactions of discrete particles (e.g. individuals in a population). This chapter will explore the dynamics of an isolated single-species population undergoing a set of random gain-loss processes that simulate births and deaths. Thus, in this particular case, the internal noise of the population model is demographic noise. Mathematically, the effects of these
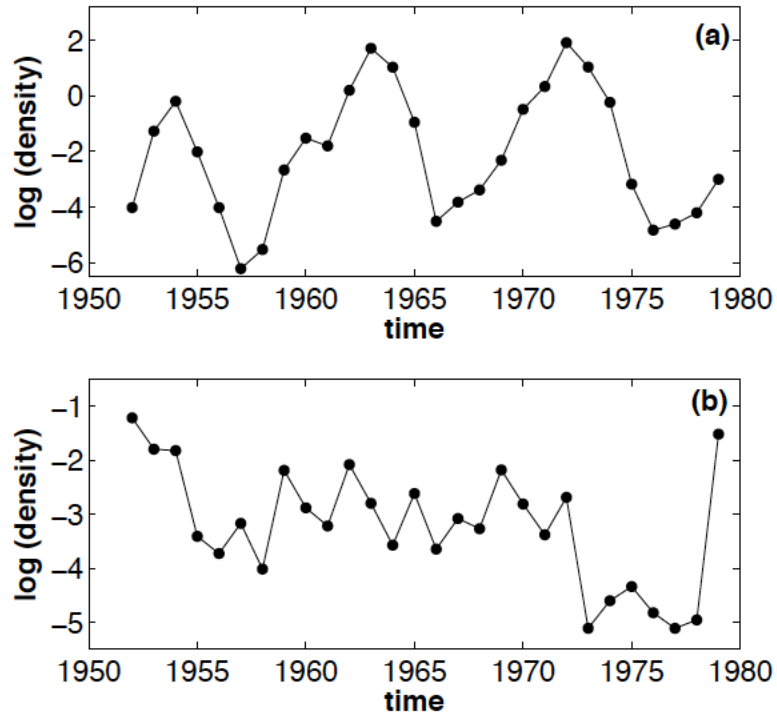
Figure 1: Population density fluctuations of Lepidoptera feeding on larch foliage in the Oberengadin Valley of Switzerland. Data from [Baltensweiler, 1991] reported as the natural logarithm of numbers per 1000 kg of larch foliage: (a) Exapate duratella (Tortricidae) and (b) Teleia saltuum (Gelechidae).

random interactions are described using a master equation [van Kampen, 1992]. Small

fluctuations captured in this modeling approach may act as an effective force that drives

the population to extinction [Assaf and Meerson, 2010]. While population extinction is

assumed to be a rare event, we can study these models to theoretically understand pre-

extinction dynamics and extinction risk.

Extinction risk is an important question in both population dynamics and ecological

community dynamics [Berg et al., 2011]. For single populations, it has long been rec-

ognized that extinction risk is dependent on the carrying capacity of the model [Leigh,

1981, Lande, 1993]. The carrying capacity is a positive stable equilibrium in the deter-

ministic model. In this chapter, a population with multiple deterministically stable equi-

libria is considered. The population size can stochastically fluctuate between these states.

One such example can be found in the data of forest Lepidoptera [Baltensweiler, 1991]

as shown in Fig. 1. The drastic population shifts are attributed to a mix of parasitoids, vi-

ral outbreak among the moths, and the quality of the available foliage [Berryman, 1996].

Figure 1(a) shows that the fluctuations may be regular, although they are not seasonal,

while Fig. 1(b) shows a switch between two states, with vastly different residence times

in each state. Similar fluctuations are observed in the context of human physiology. For

example, neural switching and enzyme level cycling are explored in [Elf and Ehrenberg,

2004, Berndt et al., 2009, Samoilov et al., 2005]. Regardless of the specific context, it is

useful to increase our understanding of pre-extinction dynamics, as well as the mean time

to extinction and the path that optimizes the probability of extinction.

In this chapter, these features are explored for a stochastic model that exhibits extinc-

tion, and for a stochastic model that exhibits pre-extinction cycling that serves to delay the

extinction event. The layout for the chapter is as follows. Section 2.2 presents the master

equation formalism needed to investigate demographic noise in population models and the

general method to find the mean time to extinction. In Section 2.3 the methods for find-

ing the optimal path and calculating the mean time to extinction (MTE) are presented. A

simple population model exhibiting extinction is presented in Section 2.4, and the analyt-

ical and numerical methods used to find the mean time to extinction are presented. The

phenomenon of pre-extinction cycling is described in Section 2.5. Section 2.6 provides

an example of stochastic population that displays cycling dynamics prior to extinction. In

Section 2.7 a general method for predicting mean time to extinction in one-dimensional

cycling models is presented. A control term is introduced in Section 2.8 to increase the

rate of extinction in the model. The methods from the previous sections are used to quan-

tify the effects of the control. In Section 2.9, the results are generalized and a brief discus-

sion is presented.

2.2   Master Equation Formalism

As mentioned in Chapter 1, to study the effects of internal noise on the dynamics of a

population, a stochastic model must be considered. While in Chapter 1.2 a general master

equation formulation was presented, in this chapter it will be recast for a one-dimensional

state variable.

In a population system, if the transitions between states are short and uncorrelated, the

system is a Markov process and the evolution of the probability is described by a mas-

ter equation. In the master equation formulation, the probability of the system taking

on a particular state X (number of agents), at a given time t, is described by $\rho(X,t)$. Let

$W(X;r)$ represent the transition rate from a state X to $X + r$, where r can be a positive or

negative integer. In this case the time evolution of $\rho(X,t)$ can be written as [van Kampen,

1992]:

$$\frac{d\rho(X,t)}{dt} = \sum_{r} \left[W(X-r;r)\rho(X-r,t) - W(X;r)\rho(X,t) \right]. \tag{16}$$

A rescaled coordinate $x = X/K$ is presented, where K is the large parameter of the problem. The transition rates are represented as the following expansion in K:

$$W(X;r) \equiv W(Kx;r) = Kw_r(x) + u_r(x) + \mathcal{O}(1/K), \qquad (17)$$

where $x = \mathcal{O}(1)$, and $w_r(x)$ and $u_r(x)$ also are $\mathcal{O}(1)$.

For $K \gg 1$ the WKB (Wentzel-Kramers-Brillouin) approximation for the scaled master equation can be used [Kubo et al., 1973, Gang, 1987, Dykman et al., 1994, Elgart and Kamenev, 2004, Kessler and Shnerb, 2007, Forgoston et al., 2011, Schwartz et al., 2011]. Accordingly, we look for the probability distribution in the form of the WKB ansatz

$$\rho(x,t) = \exp(-KS(x,t)), \qquad (18)$$

where $S(x,t)$ is a function known as the action.

The approximation Eq. (18) is substituted into the scaled master equation which contains terms with the form $w_r(x - r/K)$ and $S(x - r/K, t)$, where r/K is small. By performing a Taylor series expansion of these functions of $x - r/K$, one arrives at the leading order Hamilton-Jacobi equation $\mathcal{H}(x,p) = -\frac{dS}{dt}$, where

$$\mathcal{H}(x,p) = \sum_r w_r(x)(\exp(pr) - 1) \qquad (19)$$

is the effective Hamiltonian, such that p is the conjugate momentum and is defined as

$p = dS/dx$. In this chapter, we are interested in the special case of a single step process,

for which the only values of r are $+1$ and -1. The Hamiltonian for a single step process

will have the general form

$$\mathscr{H}(x,p) = w_1(x)\left(\exp(p) - 1\right) + w_{-1}(x)\left(\exp(-p) - 1\right). \tag{20}$$

From the Hamiltonian in Eq. (19), one can easily derive Hamilton's equations

$$\dot{x} = \frac{\partial \mathscr{H}(x,p)}{\partial p}, \qquad \dot{p} = -\frac{\partial \mathscr{H}(x,p)}{\partial x}. \tag{21}$$

The x dynamics along the $p = 0$ deterministic line can be described by the equation

$$\dot{x} = \left.\frac{\partial \mathscr{H}(x,p)}{\partial p}\right|_{p=0} = \sum_r r w_r(x), \tag{22}$$

which is simply the rescaled mean-field rate equation associated with the deterministic

problem. For a single step process, this simplifies to $\dot{x} = w_1(x) - w_{-1}(x)$.

2.3   Mean Time to Extinction

We are interested in how intrinsic noise can cause extinction events of long-lived stochas-

tic populations. In this chapter the extinct state x_0 is an attracting point of the determin-

istic mean-field equation. Furthermore, there is an intermediate repelling point $x = x_1$

between the attracting extinct state and another attracting point $x = x_2$. This scenario can

be visualized in Fig. 2 and corresponds to a scenario B extinction as explored in [Assaf

and Meerson, 2010].

In this extinction scenario, the most probable path to extinction, or optimal path to

extinction, is composed of two segments. The first segment is a heteroclinic trajectory

with non-zero momentum that connects the equilibrium point $(x, p) = (x_2, 0)$, where x_2

is an attracting fixed point of the deterministic mean-field equation, with an intermediate

equilibrium point $(x, p) = (x_1, 0)$, where x_1 is a repelling fixed point of the deterministic

mean-field equation. The second segment consists of the segment along $p = 0$ from x_1 to

the extinct state x_0.

The optimal path to extinction $p_{opt}(x)$ between $(x_2, 0)$ and $(x_1, 0)$ is a zero energy

phase trajectory of the Hamiltonian given by Eq. (19). In a single step process, the opti-

mal path will always have the general form

$$p_{opt}(x) = \ln\left(w_{-1}(x)/w_1(x)\right). \tag{23}$$

Using the definition of the conjugate momentum $p = dS/dx$, the action S_{opt} along the

optimal path $p_{opt}(x)$ is given by

$$S_{opt} = \int_{x_2}^{x_1} p_{opt}(x)dx. \tag{24}$$

Therefore, the mean time to extinction (MTE) to escape from $(x_2, 0)$ and arrive at $(x_1, 0)$

can be approximated by

$$\tau = B\exp(KS_{opt}),\tag{25}$$

where B is a prefactor that depends on the system parameters and on the population size.

An accurate approximation of the MTE depends on obtaining B.

To capture the deterministic contribution from x_1 to x_0 in the MTE approximation, the

prefactor derived in [Assaf and Meerson, 2010] is included. Specifically, the following

equation is the general form of the MTE for a single-step scenario B extinction from x_2 to

x_0:

$$\tau_{20} = \frac{2\pi \exp\left(\int_{x_1}^{x_2}\left(\frac{u_1(x)}{w_1(x)} - \frac{u_{-1}(x)}{w_{-1}(x)}\right)dx\right)}{w_1(x_2)\sqrt{|p'_{opt}(x_1)||p'_{opt}(x_2)|}} \exp\left(K\int_{x_2}^{x_1}\ln\left(\frac{w_{-1}(x)}{w_1(x)}\right)dx\right).\tag{26}$$

Note that the general notation τ_{ij} is meant to identify the function $\tau_{ij}(x_i, x_{(i+j)/2})$ that pro-

vides the escape time from state x_i to state x_j. In the case that $i = 2$ and $j = 0$, then one

recovers Eq. (26). It is worth noting that the derivation of Eq. (26) involves matching

the solution from x_2 to x_1 asymptotically with the deterministic solution from x_1 to x_0.

Because this latter solution is associated with $p = 0$, its final form does not involve an

integral from x_1 to x_0. Nevertheless, the deterministic contribution is in fact included in

Eq. (26).

2.4   An Example of Extinction

A model exhibiting the Allee effect is now presented to illustrate the analytical methods

described in Sec. 2.2. The Allee effect is seen in animal populations that benefit from
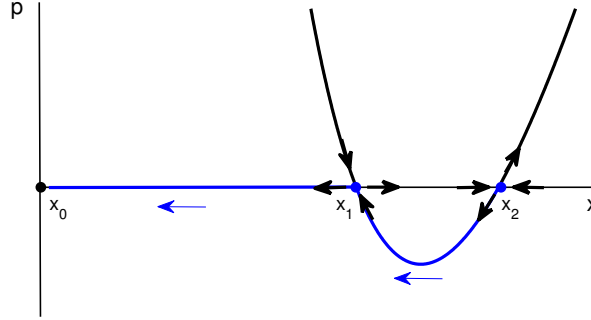
Figure 2: Zero-energy trajectories $p = 0, x = 0$, and $p_{opt}(x)$ of the Hamiltonian for the stochastic Allee population model given by Eq. (28). The optimal path to extinction (blue curve) consists of the heteroclinic trajectory $p_{opt}(x)$ (Eq. (31)) connecting x_2 to x_1, and the $p = 0$ line from x_1 to the extinct state x_0.

conspecific cooperation. These populations tend to perform better in larger numbers. In fact, there is evidence that larger populations are more capable of avoiding predation, can reproduce faster, and are better able to resist toxic environmental conditions [Allee, 1931, Lidicker Jr, 2010]. On the other hand, the growth rate is negative for low densities. Therefore the dynamics are bistable and the population will tend towards a positive state, referred to as the carrying capacity, or an extinct state depending on the initial population. A simple deterministic mathematical model demonstrating the Allee effect can be written as $\dot{x} = f(x)$, where $f(x)$ is a cubic polynomial. Using the notation used in Fig. 2, we could write $f(x) = x(x - x_1)(x - x_2)$. The corresponding stochastic population model is represented by the transition processes and associated rates $W(X; r)$ shown in Table 1.

The death rate of a low-density population is given by μ, and the growth rate of the population when the density is large enough is given by λ. The negative growth rate for

Transition	W(X;r)
$X \xrightarrow{\mu} \varnothing$	$\mu X,$
$2X \xrightarrow{\lambda/K} 3X$	$\lambda \frac{X(X-1)}{2K},$
$3X \xrightarrow{\sigma/K^2} 2X$	$\sigma \frac{X(X-1)(X-2)}{6K^2}.$

Table 1: The transition processes and associated rates for a stochastic Allee model.

an overcrowded population is provided by σ, and K is the carrying capacity of the popula-

tion.

As described in Sec. 2.2, the transition processes and their associated rates are used

to formulate the master equation given by Eq. (16). In this particular example, all of the

transitions are single-step transitions. Therefore, the increment r only takes on the values

of ± 1. The scaled transition rates in Eq. (17) are given as

$$w_1(x) = \frac{\lambda x^2}{2}, \qquad w_{-1}(x) = \mu x + \frac{\sigma x^3}{6},$$

$$\tag{27}$$

$$u_1(x) = -\frac{\lambda x}{2}, \qquad u_{-1}(x) = -\frac{\sigma x^2}{2}.$$

Substitution of Eq. (27) into Eq. (19) leads to the following Hamiltonian:

$$\mathcal{H}(x,p) = \frac{\lambda x^2}{2}(e^p - 1) + \left(\mu x + \frac{\sigma x^3}{6}\right)(e^{-p} - 1). \tag{28}$$

Taking derivatives of Eq. (28) with respect to p and x (Eq. (21)) lead to the following sys-
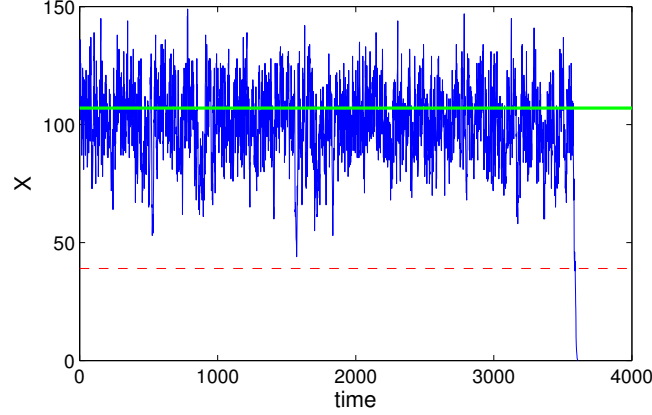
Figure 3: A single realization exhibiting extinction in the stochastic Allee population model. The non-zero deterministic stable state is shown by the green line, while the deterministic unstable state is shown by the red dashed line. The parameter values are $\mu = 0.2$, $\sigma = 3.0$, $\lambda = 1.425$, and $K = 100$.

tem of Hamilton's equations:

$$\dot{x} = \frac{\lambda x^2}{2} e^p - \left(\mu x + \frac{\sigma x^3}{6} \right) e^{-p} \tag{29}$$

$$\dot{p} = -\lambda x (e^p - 1) - \left(\mu + \frac{\sigma x^2}{2} \right) (e^{-p} - 1). \tag{30}$$

By setting the Hamiltonian in Eq. (28) equal to zero and solving for p and x it is possible to find three zero-energy phase trajectories. The solutions are $x = 0$, the extinction line; $p = 0$, the deterministic line and

$$p_{opt}(x) = \ln \left(\frac{6\mu + \sigma x^2}{3\lambda x} \right), \tag{31}$$

the optimal path to extinction. These solutions are shown in Fig. 2.

Using Eq. (22) we can recover the deterministic mean-field equation by substituting

$p = 0$ into Eq. (29) to obtain

$$\dot{x} = -\frac{\sigma}{6}x^3 + \frac{\lambda}{2}x^2 - \mu x. \tag{32}$$

Equation (32) has three steady states: the extinct state $x_0 = 0$, and two non-zero states

$$x_{1,2} = \frac{3\lambda \mp \sqrt{9\lambda^2 - 24\sigma\mu}}{2\sigma}. \tag{33}$$

In the deterministic model exhibiting the Allee effect (Eq. (32)), x_1 is an unstable steady

state that functions as a threshold. For initial conditions whose value lies between x_1 and

x_2, the deterministic solution will increase to x_2, which is a stable steady state. For ini-

tial conditions whose value is less than x_1, the deterministic solution will decrease to the

stable extinct steady state x_0.

However, when intrinsic noise is considered and one performs the analysis described

in Sec. 2.2, then the steady states of Hamilton's equations will be

two-dimensional with both p and x components. Furthermore, it is easy to show that each

of the steady states of the stochastic Allee model will be saddle points, as shown in Fig. 2.

Figure 2 shows that starting from x_2, the optimal path to extinction consists of first travel-

ing along the blue heteroclinic trajectory $p_{opt}(x)$ connecting x_2 to x_1, followed by travel-

ing along the $p = 0$ line from x_1 to the extinct state x_0.

The analytical MTE is found using Eq. (26) and is confirmed using numerical simula-

tions. A Monte Carlo algorithm [Gillespie, 1976] is used to evolve the population in time,
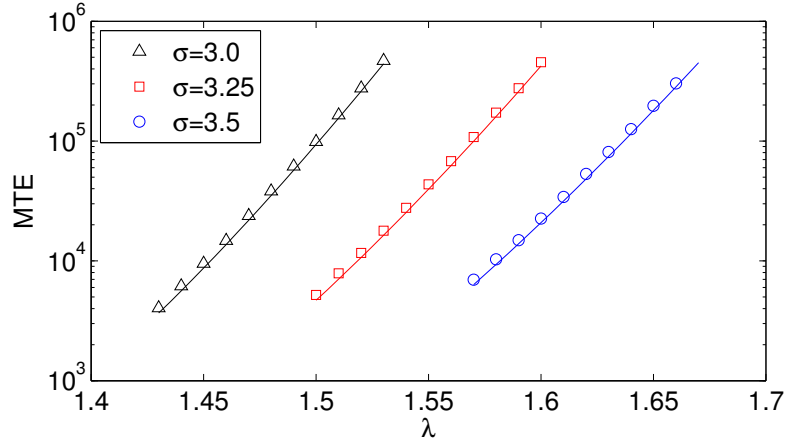
Figure 4: Mean time to extinction for the stochastic Allee population model with an initial population given by X_2. The curves are found using the analytical approximation given by Eq. (26), and the symbols represent the corresponding numerical simulation results. The numerical results are based on 10,000 realizations with $\mu = 0.2$ and $K = 100$ as σ and λ are varied.

and Fig. 3 shows an example of one stochastic realization. Figure 3 shows that the population persists for a very long time near the x_2 state (deterministically stable) but eventually the noise causes the population to go extinct. By numerically computing thousands of stochastic realizations and the associated extinction times, one can calculate the MTE. Figure 4 shows the comparison between the analytical and the numerical mean time to extinction as a function of λ for various choices of σ. Each numerical result is based on 10,000 Monte Carlo simulations, and the agreement is excellent.

2.5 Extinction with Cycling

In the previous section there were three steady states associated with the deterministic Allee model, two of which were stable (a non-zero carrying capacity and the extinct state)

and one of which was unstable (a threshold state). Additionally, for the stochastic Allee

model, the population fluctuated for a long period of time about one of the deterministic-

tically stable states before stochastically switching to the extinct state. There are many

models whose deterministic mean-field equation has multiple non-zero stable steady

states. In these cases, the population can switch between these different population lev-

els repeatedly before going extinct.

To assist in visualization of switching and cycling events consider the diagram in

Fig. 5. Each black ellipse represents a steady state in a population model. Such a system

is capable of switching between three different stable population levels, 'Stable Popula-

tion (C)', 'Stable Population (E)', and 'Extinction (A)'. Each arrow represents a possible

path between the respective states. A red arrow indicates that a transition is only possible

via a large stochastic fluctuation, while a green arrow indicates that a deterministic cur-

rent will force the transition in the direction of the arrow. Points 'D' and 'B' both have

two green arrows heading away from them, so they are deterministically unstable. All of

the green paths associated with points 'A', 'C', and 'E' head towards them, so they are de-

terministically stable. Once one of the threshold points are reached by a large stochastic

fluctuation the deterministic currents take over and force the system state towards a sta-

ble equilibrium point. Notice that state 'A' is the extinct state, and it is an absorbing stable

steady state with no possible escape. In Chapters 3 and 4 we will investigate populations

with migration, where the extinct state is not absorbing.

In one-dimensional systems with either two or three equilibrium points, approxima-
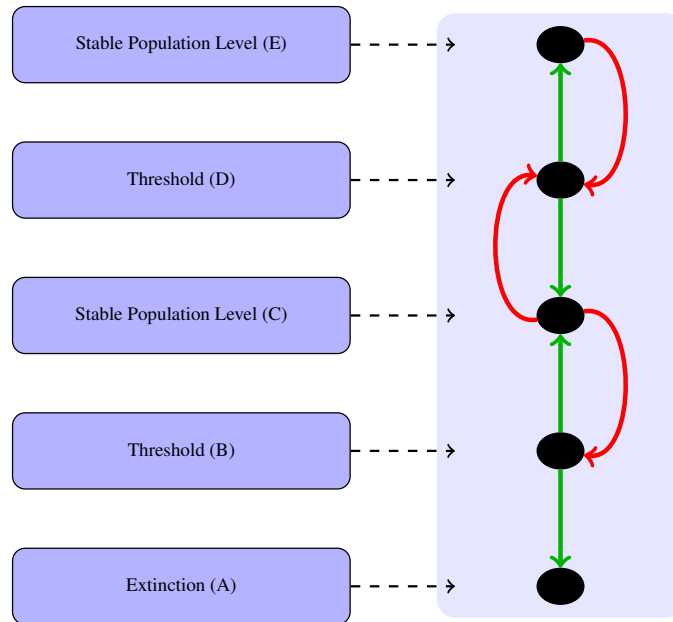
Figure 5: A depiction of a one-dimensional system with five equilibria (black ovals). Potential stochastic transitions are depicted by red arrows, while deterministic transitions are depicted using green arrows.

tion of the MTE can be done analytically [Assaf and Meerson, 2010]. In larger systems,

or systems with more than three steady states, it is generally unknown how successive

switching events contribute to the MTE. Consider the scenario in Fig. 5 where there are

two non-zero stable equilibria located at points 'C' and 'E'. An example of a cycling

event would be a transition from 'C' to 'E' and then back to 'C'. Since the system is one-

dimensional, the system must first move from 'C' to 'D' via a large stochastic fluctuation.

Although the system is capable of transitioning back to 'C', in a cycling event the system

will then travel from 'D' to 'E', forced by the underlying deterministic current. Once at

point 'E' there will be some residence time before a large stochastic fluctuation sends the

system state back down to point 'D', where the deterministic current will carry us back to
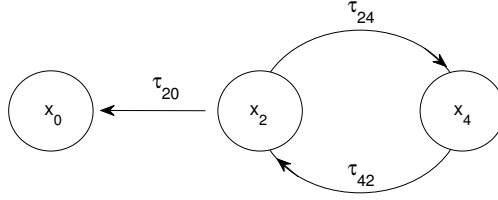
Figure 6: Flowchart for a model whose deterministic mean-field equation has two non-zero stable steady states. Given a state variable x, the stable states are located at x_0 (the extinct state), x_2, and x_4. There are two unstable states at x_1 and x_3 (not shown). The population may stochastically cycle multiple times from x_2 to x_4 and back to x_2 before eventually transitioning to the x_0 extinct state.

the initial position, state 'C'. The 'C-E-C'-transition is called a steady state cycling event, or just cycling. It is generally unknown how these steady state cycling events contribute to the MTE.

In the next section, we investigate a population's MTE when cycling occurs as a pre-extinction event. Furthermore, we derive a new analytical result for the mean time to extinction by considering the probability of stochastic switching events and their associated transition times. This is one of the main results of this chapter.

2.6   An Example of Population Cycling

It is straightforward to extend the Allee model of Sec. 2.4 to a model whose deterministic mean-field equation is a quintic polynomial with five steady states, three of which are stable. The stochastic version of this new model will exhibit pre-extinction cycling as previously discussed. This stochastic population model is represented by the transition processes and associated rates $W(X;r)$ shown in Table 2.

Transition	$W(X;r)$
$X \xrightarrow{\mu} \varnothing$	$\mu X,$
$2X \xrightarrow{\lambda/K} 3X$	$\lambda \frac{X(X-1)}{2K},$
$3X \xrightarrow{\sigma/K^2} 2X$	$\sigma \frac{X(X-1)(X-2)}{6K^2},$
$4X \xrightarrow{\alpha/K^3} 5X$	$\alpha \frac{X(X-1)(X-2)(X-3)}{24K^3},$
$5X \xrightarrow{\beta/K^4} 4X$	$\beta \frac{X(X-1)(X-2)(X-3)(X-4)}{120K^4}.$

Table 2: The transition processes and associated rates for the stochastic cycling model.

The first three events are the same as found in the stochastic Allee model in Sec. 2.4
and allow for fluctuations around a population level – previously called the carrying ca-
pacity – before going extinct. The two additional transition events create two additional
deterministic equilibrium points; one deterministically unstable equilibrium point that
acts as a second Allee-like threshold, and another deterministically stable equilibrium
point. The deterministically stable equilibrium point is metastable in our stochastic model.
Now the stochastic system has two metastable states, and is capable of cycling between
the two corresponding population levels prior to extinction.

As described in Sec. 2.2, the transition processes and their associated rates are used to
formulate the master equation given by Eq. (16). Note that the transitions are single-step
because the increment *r* only takes on the values of ± 1. Therefore, the scaled transition

rates $w_r(x)$ and $u_r(x)$ in Eq. (17) are given as

$$w_1(x) = \frac{\lambda x^2}{2} + \frac{\alpha x^4}{24}, \qquad w_{-1}(x) = \mu x + \frac{\sigma x^3}{6} + \frac{\beta x^5}{120},$$

$$(34)$$

$$u_1(x) = -\frac{\lambda x}{2} - \frac{\alpha x^3}{4}, \qquad u_{-1}(x) = -\frac{\sigma x^2}{2} - \frac{\beta x^4}{12}.$$

Substitution of Eq. (34) into Eq. (19) leads to the following Hamiltonian:

$$\mathscr{H}(x,p) = \left(\frac{\lambda x^2}{2} + \frac{\alpha x^4}{24} \right) (e^p - 1) + \left(\mu x + \frac{\sigma x^3}{6} + \frac{\beta x^5}{120} \right) (e^{-p} - 1). \qquad (35)$$

Taking derivatives of Eq. (35) with respect to p and x (Eq. (21)) lead to the following system of Hamilton's equations:

$$\dot{x} = \left(\frac{\lambda x^2}{2} + \frac{\alpha x^4}{24} \right) e^p - \left(\mu x + \frac{\sigma x^3}{6} + \frac{\beta x^5}{120} \right) e^{-p}, \qquad (36)$$

$$\dot{p} = -\left(\lambda x + \frac{\alpha x^3}{6} \right) (e^p - 1) - \left(\mu + \frac{\sigma x^2}{2} + \frac{\beta x^4}{24} \right) (e^{-p} - 1). \qquad (37)$$

Once again, by setting the Hamiltonian in Eq. (35) equal to zero and solving for p and x it is possible to find the zero-energy phase trajectories to be $x = 0$, the extinction line; $p = 0$, the deterministic line and

$$p_{opt}(x) = \ln \left(\frac{120\mu + 20\sigma x^2 + \beta x^4}{5x(\alpha x^2 + 12\lambda)} \right), \qquad (38)$$

the optimal path to extinction. The $p = 0$ and $p_{opt}(x)$ solutions found using Eq. (35) are
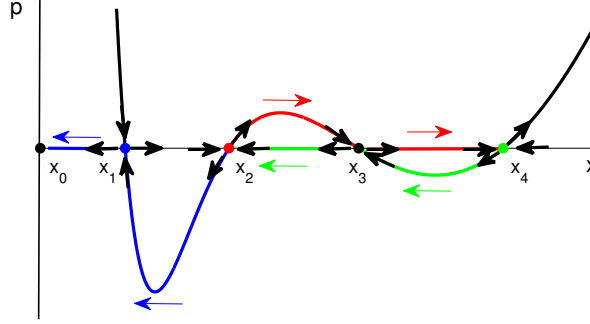
Figure 7: Zero-energy trajectories of the Hamiltonian for the stochastic Allee population model given by Eq. (35). The optimal path of transitioning from one state to another is given by $p = 0$ or $p_{opt}(x)$ (Eq. (38)). A cycling path (red and green) consists of the heteroclinic trajectory connecting x_2 to x_3 (red) and the $p = 0$ line from x_3 to x_4 (red), followed by the heteroclinic trajectory connecting x_4 to x_3 (green) and the $p = 0$ line from x_3 to x_2 (green). The optimal path to extinction consists of the heteroclinic trajectory from x_2 to x_1 (blue) and the $p = 0$ line from x_1 to x_0 (blue).

shown in Fig. 7. Using Eq. (22), the deterministic mean-field equation is recovered by

substituting $p = 0$ into Eq. (36) to obtain

$$\dot{x} = -\frac{\beta}{120}x^5 + \frac{\alpha}{24}x^4 - \frac{\sigma}{6}x^3 + \frac{\lambda}{2}x^2 - \mu x. \tag{39}$$

Equation (39) has five steady states: the extinct state $x_0 = 0$, and four non-zero states,

two of which are stable and two of which are unstable. In the deterministic model given

by Eq. (39), x_1 and x_3 are unstable states. For initial conditions whose value lies between

x_1 and x_2, the deterministic solution will increase to x_2, which is a stable steady state. For

initial conditions whose value is less than x_1, the deterministic solution will decrease to

the stable extinct steady state x_0. Similarly, when initial conditions have a value between

x_3 and x_4, the deterministic solution will increase to x_4, which is a stable steady state.

When initial conditions have a value between x_3 and x_2, the deterministic solution will

decrease to the stable steady state x_2.

As we have already seen in the previous section, the inclusion of intrinsic noise in the

model leads to the steady states of Hamilton's equations being two-dimensional with both

p and x components. Furthermore, the steady states of the stochastic cycling model will

be saddle points, as seen in Fig. 7. Figure 7 shows that starting from x_2 there is a choice

to be made: 1) the population could go extinct by traveling along the blue path, which is

the heteroclinic trajectory connecting x_2 to x_1, followed by traveling along the $p = 0$ line

from x_1 to the extinct state x_0, much like what happens in the stochastic Allee model; or

2) the population could cycle to x_4 and back by traveling along the red path and then the

green path, which includes two stochastic escapes. First, the population travels along the

heteroclinic trajectory connecting x_2 to x_3, followed by traveling along the $p = 0$ line from

x_3 to x_4. After fluctuating for some time about x_4, the population returns to x_2 by traveling

along the heteroclinic trajectory from x_4 to x_3, followed by traveling along the $p = 0$ line

from x_3 to x_2.

A Monte Carlo algorithm [Gillespie, 1976] is used to evolve the population in time,

and Fig. 8 shows one stochastic realization. Figure 8 shows multiple cycling events be-

tween the x_2 and x_4 states (deterministically stable) before the population eventually goes

extinct. By numerically simulating thousands of stochastic realizations, we can compute

the mean time to extinction and compare the numerical result with an analytical result.
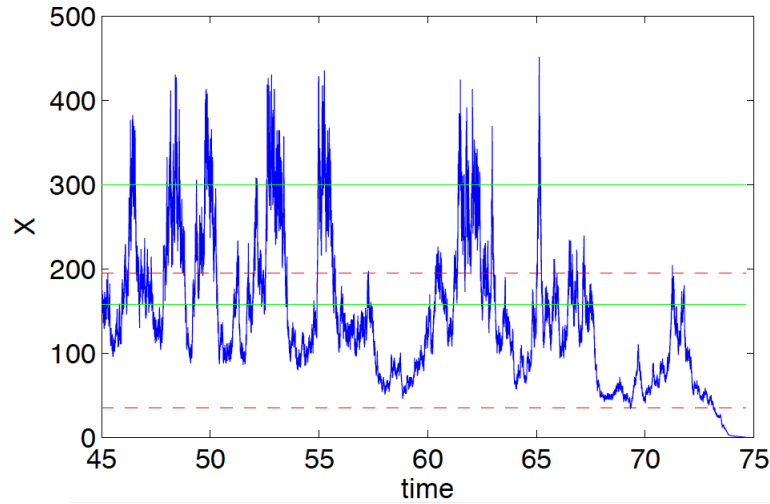
Figure 8: A single realization exhibiting cycling and extinction in the stochastic cycling population model. The non-zero deterministic stable states are shown by the green lines, while the deterministic unstable states are shown by the red dashed lines. The parameter values are $\mu = 3.25$, $\alpha = 0.465$, $\beta = 0.048$, $\lambda = 3.96$, $\sigma = 1.905$, and $K = 14$.

We will now derive a novel analytical form for the mean time to extinction that will account for the additional pre-extinction cycling time that delays the actual extinction event.

2.7   Approximating the Extinction Time

Consider the model whose flowchart is given in Fig. 6. The deterministic mean-field equation for this model has multiple non-zero stable steady states located at x_2 and x_4 along with the extinct state located at x_0. In the corresponding stochastic model, the population may stochastically cycle multiple times from x_2 to x_4 and back to x_2 before eventually transitioning to the extinct state x_0. It is important to note that in this example it is not possible to experience unlimited population growth [Meerson and Sasorov, 2008], and eventually the population will go extinct.

If the system is located at x_2, there are only two options for a stochastic switch: 1)

the population can go to the extinct state x_0, or 2) the population can switch to x_4. Since

the population will eventually go extinct, it follows that any stochastic switch from x_2 to

x_4 must result in a following switch from x_4 back to x_2 at some later time. Furthermore,

the population may cycle from x_2 to x_4 and back to x_2 any number of times before the

population eventually goes extinct by switching from x_2 to the absorbing extinct state x_0.

In isolation, the probability of the population switching from x_2 to x_0 can be approxi-

mated as $1/\tau_{20}$. Similarly, the probability of the population switching from x_2 to x_4 can be

approximated by $1/\tau_{24}$. Recall that both τ_{20} and τ_{24} can be approximated using Eq. (26).

However, in the cycling model, these switches do not occur in isolation. Rather, there is

a "competition" as to which switch will happen first. Therefore, we must compute the

probability of one switch occurring before the other. The probability of the population

switching from x_2 to x_0 before switching from x_2 to x_4 is

$$\mathscr{P}_{20} = \frac{\frac{1}{\tau_{20}}}{\frac{1}{\tau_{20}} + \frac{1}{\tau_{24}}} = \frac{\tau_{24}}{\tau_{20} + \tau_{24}}. \tag{40}$$

Note that we will use the general notation $\mathscr{P}_{ij}$ to denote the probability that an escape

from x_i to x_j happens first. Also, $\mathscr{P}_{24} = 1 - \mathscr{P}_{20}$ because there are only two switching

options.

To find the MTE, we use a probabilistic argument whereby the probability of a given

event (immediate extinction, one cycle followed by extinction, two cycles followed by ex-

tinction, etc.) is weighted by the approximate time of each event. Each transition time is

found using Eq. (26), and it should be noted that each probability term is written in terms

of these approximate transition times (e.g. Eq. (40)). The MTE thus becomes the sum of

the expected times for all possible number of cycles to occur and the final escape from x_2

to x_0:

$$\text{MTE} = \tau_{20}\mathscr{P}_{20} + \sum_{i=0}^{\infty} i(\tau_{24} + \tau_{42})(\mathscr{P}_{24})^i \mathscr{P}_{20} \tag{41a}$$

$$= \tau_{20}\mathscr{P}_{20} + \frac{(\tau_{24} + \tau_{42})\mathscr{P}_{24}}{\mathscr{P}_{20}} \tag{41b}$$

$$= \frac{\tau_{20}\tau_{24}}{\tau_{20} + \tau_{24}} + (\tau_{24} + \tau_{42})\frac{\tau_{20}}{\tau_{24}}. \tag{41c}$$

As in the stochastic Allee model, the analytical mean time to extinction for the stochas-

tic cycling model can be confirmed using Monte Carlo simulations [Gillespie, 1976]. Fig-

ure 9 shows the comparison between the analytical and numerical mean time to extinction

as a function of λ for various choices of carrying capacity K. Each numerical result is

based on 5,000 Monte Carlo simulations, and the agreement is excellent. Note that the

choice of λ values for this example is limited by the quasi-stationarity requirement. How-

ever, in the following section, we continue the exploration of this example using control

and show there is excellent agreement for MTE over several orders of magnitude of MTE.

2.8   Speeding Up Extinction

The previous section presents a way to find the MTE in a population model with cycling

so that extinction is delayed. Often the MTE is of interest in the study of population dy-
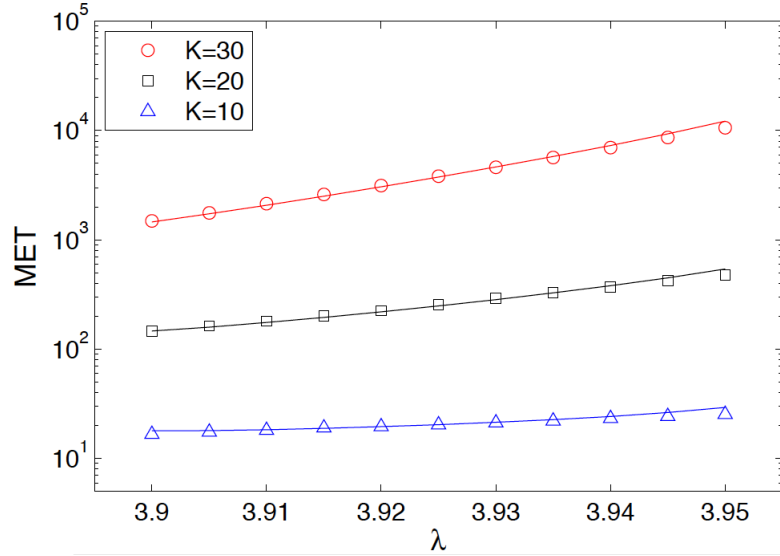
Figure 9: Mean time to extinction for the stochastic cycling population model with an initial population given by X_2. The solid curves are found using the analytical approximation given by Eq. (41c), and the symbols represent the corresponding numerical simulation results. The numerical results are based on 5,000 simulations with $\mu = 3.307$, $\alpha = 0.458$, $\beta = 0.047$, and $\sigma = 1.8874$ as K and λ are varied.

Transition	W(X;r)
$X \xrightarrow{\nu} \varnothing$	ν.

Table 3: The transition process and associated rate for population control. It is effectively a culling or quarantining rate.

namics because either longevity or quick extinction has value. The population studied in

this section should be thought of as pests, and a short MTE should be considered ideal.

The control method we model removes individuals at a particular frequency ν. This popu-

lation will have all the same demographic events that were seen in the cycling population

model, and will have in addition the event shown in Table 3. In an ecological context one

might think of the control term as culling or quarantining.

The only change from the cycling model transition rates given by Eq. (34) is in the
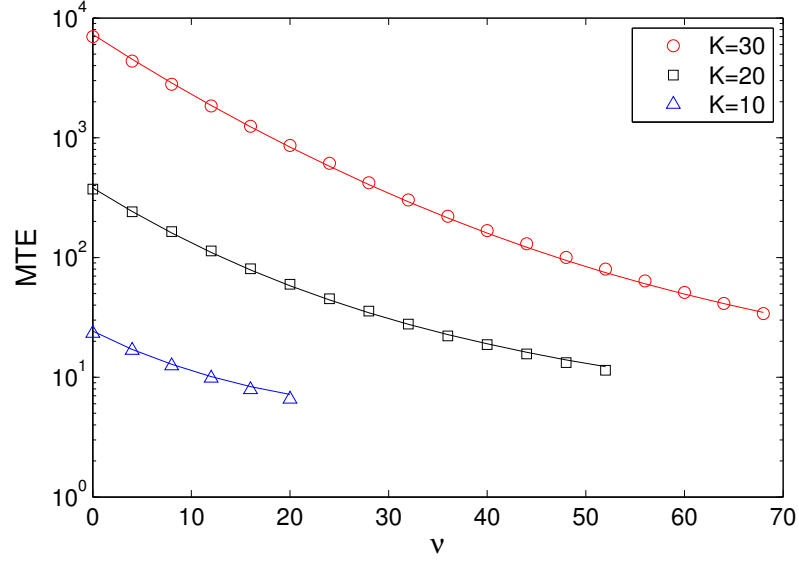
Figure 10: Mean time to extinction for the stochastic cycling population model using control with an initial population given by X_2. The solid curves are found using the analytical approximation. The symbols represent the corresponding numerical simulation results. The numerical results are based on 5,000 simulations with $\mu = 3.307$, $\alpha = 0.458$, $\beta = 0.047$, $\lambda = 3.94$, and $\sigma = 1.8874$ as K and ν are varied.

rate w_{-1} which now has the form

$$w_{-1}(x) = \mu x + \frac{\sigma x^3}{6} + \frac{\beta x^5}{120} + \frac{\nu}{K}. \tag{42}$$

Using Eq. (19), the modified Hamiltonian will be

$$\mathscr{H}(x,p) = \left(\frac{\alpha x^4}{24} + \frac{\lambda x^2}{2}\right)(e^p - 1) + \left(\mu x + \frac{\sigma x^3}{6} + \frac{\beta x^5}{120} + \frac{\nu}{K}\right)(e^{-p} - 1). \tag{43}$$

To quantify the change in the MTE as a function of ν, we use Eq. (41c) and the modified Hamiltonian given by Eq. (43).

The analytical mean time to extinction for the stochastic cycling model with control

can be confirmed using Monte Carlo simulations [Gillespie, 1976]. Figure 10 extends the

examples from Fig. 9 by comparing the analytical and numerical mean time to extinction

as a function of ν. Each numerical result is based on 5,000 Monte Carlo simulations, and

the agreement is excellent over several orders of magnitude of MTE. As expected, when

more individuals are removed from the population the MTE decreases.

2.9   Conclusions

In this chapter we have considered stochastic population models where the intrinsic or

demographic noise eventually causes the population to go extinct. For models that exhibit

stochastic cycling between two states, we described the optimal path to extinction and an

analytical method to approximate the mean time to extinction. We used a probabilistic

argument to understand the pre-extinction dynamics that delay the extinction event.

These results for the MTE can be extended to the general case of a system with $2n + 1$

steady states $((n \in \mathbb{N}), n > 1)$, with the possibility of $n - 1$ cycles. We assume there are

$n + 1$ deterministically stable steady states $\{X_0, X_2, X_4, \dots, X_{2n}\}$ alternating with deter-

ministically unstable steady states, and that X_0 is an absorbing extinct state. For a system

starting at X_2,

$$MTE = \tau_{20}\mathscr{P}_{20} + \sum_{i=1}^{n-1}\left[\left(\tau_{2i,2i+2} + \tau_{2i+2,2i}\mathscr{P}_{2i+2,2i}\right)\prod_{k=1}^{i}\frac{\mathscr{P}_{2k,2k+2}}{\mathscr{P}_{2k,2k-2}}\right]. \qquad (44)$$

Note that when situated at X_{2n}, the stable steady state furthest away from the extinct state,

then there is no choice of switching. The only possible switch is from X_{2n} to X_{2n-2}, and therefore $\mathscr{P}_{2n,2n-2} = 1$. This result can also be extended to find the MTE when the system starts at any of the other deterministically stable steady states. Consider an initial condition X_{2k} for $k < n$. One would need to find the mean time for each escape in the sequence $X_{2k}, X_{2k-2}, \ldots, X_0$. For example, starting at X_{2k} would reduce the problem to the subsystem of deterministically stable steady states $\{X_{2k-2}, X_{2k}, \ldots, X_{2n}\}$ for which Eq. (44) would approximate the mean escape time to X_{k-2}. Repeating this procedure leftward and taking the sum of these mean escape times would result in the total MTE.

Lastly, a control method was introduced to the stochastic cycling population model. The mean time to extinction was calculated analytically and was shown to agree well with numerical Monte Carlo simulations. It was shown that the mean time to extinction decreases monotonically with an increased removal program. From an ecological perspective it is important to work towards a quantitative understanding of how control methods (e.g. bio-control agents, culling programs, quarantine programs, or hunting allowances) may affect the longevity of a population.

3   Extinction Pathways and Outbreak Vulnerability in a Stochastic Ebola Model

3.1   Introduction

The Ebola virus disease (EVD) is an infectious zoonosis found in several mammals including humans, bats, and apes. A zoonosis is a disease that can be naturally transmitted from animals to humans. If there is a non-human disease carrier, we call that species the disease reservoir. Strong evidence for an Ebola virus reservoir can be seen in the invasion/extinction cycles that have taken place over the forty years of recorded EVD history in humans. The first known spillover of EVD into the human population took place in Zaire (now the Democratic Republic of the Congo) near the Ebola River from which the disease took its name. Major EVD epidemics have taken place in the Democratic Republic of the Congo, Gabon, Sudan, Uganda, and most recently in Guinea, Sierra Leone, and Liberia. Although the disease was first recognized and has been primarily located in Central Africa, the most recent epidemic took place along the continent's western coast.

Unlike non-zoonotic diseases, EVD goes through long periods of global extinction in the human population. These EVD-free periods are punctuated by seemingly spontaneous disease reintroduction, which suggests infection from a non-human source. Additionally, there is a large body of work in the biological and ecological sciences providing evidence that the virus is maintained in animal populations [Leroy et al., 2005, Pourrut et al., 2005]. Thus, the seemingly spontaneous reappearance of the disease in the human population

must be understood in the context of unpredictable interactions between humans and an-
imal carriers of the disease. Although they are random in nature, researchers are working
to improve our understanding of these interactions [P. D. Walsh, 2005, Pigott et al., 2014].

The EVD is transmitted via bodily fluids such as blood, saliva, semen, and breast
milk, and is very deadly in both humans and apes, with an average mortality rate of 50%
in humans. Depending on the viral strain the disease may kill as much as 90% of an af-
fected human population [Centers for Disease Control and Prevention, 2016]. When the
disease is transferred from the animal reservoir into the human population it is known as
a spillover event. Although EVD has a relatively difficult time invading and persisting in
a human population, there have been over half a dozen spillover events with more than
100 infected individuals since 1976, when the disease was initially recognized in Zaire
(now the Democratic Republic of Congo). The Centers for Disease Control and Preven-
tion (CDC) estimates over 28,000 infections and over 11,000 deaths in the most recent
West African epidemic [World Health Organization, 2016].

From a modeling perspective, considering an estimated disease incubation time of 7-
21 days, a deterministic Susceptible-Exposed-Infectious-Recovered (*SEIR*) compartmen-
tal model is appropriate to investigate the dynamics [Anderson and May, 1991, Chowell
et al., 2004, Legrand et al., 2007]. Extensions have been proposed to account for various
kinds of intervention [Hu et al., 2015] or, as in the case of gravity models, to account for
the spread of EVD over an explicit spatial region. In a gravity model the force of infec-
tion from a non-contiguous population will be proportional to the size of the respective

populations, and inversely proportional to the square of the distance between the two pop-

ulations [D'Silva and Eisenberg, 2015, Yang et al., 2015].

Studies that explicitly consider the dynamics of EVD while in the presence of a zoonotic

reservoir are less common. Although the zoonotic reservoir for Ebola is still a matter of

contention among researchers, bats make for a likely suspect [Leroy et al., 2005, Pourrut

et al., 2005]. It is known that non-human primates are sometimes involved in the spread

of EVD to humans as intermediate susceptible hosts [P. D. Walsh, 2005]. Accounting

for human exposure to the reservoir is paramount for the study of disease introduction.

However, disease extinction and reintroduction are rare stochastic events that cannot be

captured by deterministic models.

Therefore, the goal of this chapter is to present a stochastic model that explicitly ac-

counts for both the introduction of EVD from the zoonotic reservoir, as well as the fade-

out periods. Perturbations of this work can be used to assess the efficacy of disease con-

trol strategies such as vaccination and quarantine. Additionally, the work supports the use

of the basic reproduction number while quantifying outbreak vulnerability in populations

weakly coupled to a disease reservoir. Although this chapter focuses on EVD, the results

are broadly relevant to outbreak and extinction in zoonotic diseases.

3.2   EVD Model and Methodology

In this chapter, we extend a previously adopted compartmental model for EVD with inter-

vention [Hu et al., 2015] to a stochastic model that captures the internal random dynam-

ics of a population. Figure 11 shows the division of the population into the following six

classes:

1. *Susceptible* class S consists of individuals who may become infected with EVD

   through contact with an infectious individual, a hospitalized individual, or a de-

   ceased but unburied individual.

2. *Exposed* class E consists of individuals who are infected with EVD but are not yet

   infectious.

3. *Infectious* class I consists of individuals who are capable of transmitting EVD to a

   susceptible individual.

4. *Recovered* class R consists of individuals who have recovered from EVD.

5. *Deceased* class D consists of deceased and unburied individuals who are capable of

   transmitting EVD to a susceptible individual.

6. *Hospitalized* class H consists of individuals who have been hospitalized and are

   capable of transmitting EVD to a susceptible individual. Individuals who die while

   in the hospitalized class are immediately buried.

Figure 11 also shows the interactions among these population classes. Note that EVD

transmission can occur though both infectious human contact and the animal reservoir.

We assume that contact with the animal reservoir is always possible, and independent of

the ratio of animal carriers to humans.
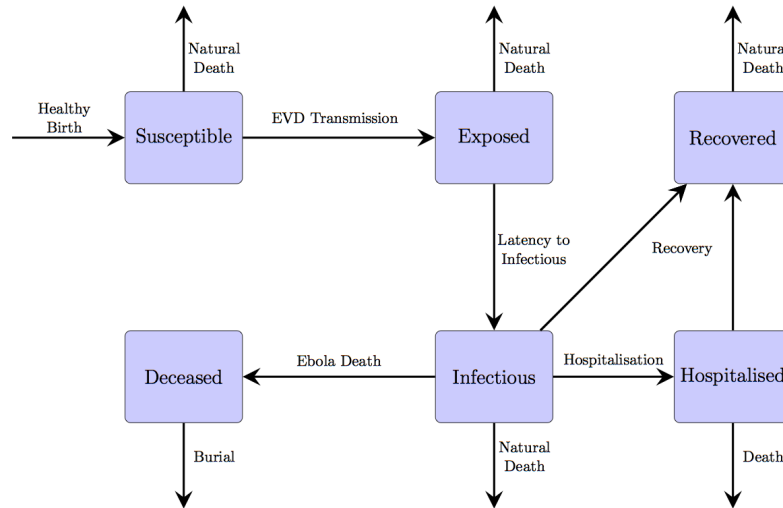
Figure 11: Flow diagram for the EVD model.

The model we study is an extension of an *SEIR* model. Besides explicit consideration of the reservoir, the model includes two additional classes: Hospitalized and Deceased. We note that this model (without a reservoir) has been used in previous studies [Legrand et al., 2007, Hu et al., 2015]. The Hospitalized and Deceased compartments account for important routes of disease transmission which are specific to EVD. Additionally, these routes provide a way to investigate the usefulness of practical intervention through the rate of hospitalization and fast access to a safe burial. The effectiveness of practical intervention strategies can be monitored through these parameters, together with variations in contact rates.

The model captures movement between the classes as stochastic transitions that occur at specified transition rates. Each of these transitions represents a random event that can occur in a population. Table 4 quantifies the possible transition events and associated

rates, which follow the Fig. 11 flow diagram. In the master equation approach to stochastic modeling, these transition rates are used as coefficients that define the probability of each respective transition event [van Kampen, 1992, Gardiner, 2004]. Assuming some discrete period of time during which exactly one of the aforementioned events takes place, the probability that a *particular* event will be the one that *does* take place is equal to its own transition rate divided by the sum of all transition rates. The stochastic simulations reported on in this dissertation use these transitions in a Monte Carlo algorithm as described in [Gillespie, 1976].

If the transitions between states are short and uncorrelated and we assume the population is well mixed, the system is a Markov process and the evolution of the probability is described by a master equation [van Kampen, 1992, Gardiner, 2004].

3.2.1   Master Equation and the Hamiltonian

The general form of the master equation as described by the transitions in Table 4 is

$$
\begin{aligned}
\frac{d\rho(\mathbf{X})}{dt} =\ & \mu N \left(\rho(S-1,E,I,D,H,R) - \rho(S,E,I,D,H,R) \right) \\
& + \tfrac{\beta_i I}{N} \left((S+1)\rho(S+1,E-1,I,D,H,R) - S\rho(S,E,I,D,H,R) \right) \\
& + \tfrac{\beta_d D}{N} \left((S+1)\rho(S+1,E-1,I,D,H,R) - S\rho(S,E,I,D,H,R) \right) \\
& + \tfrac{\beta_h H}{N} \left((S+1)\rho(S+1,E-1,I,D,H,R) - S\rho(S,E,I,D,H,R) \right) \\
& + \kappa \left((S+1)\rho(S+1,E-1,I,D,H,R) - S\rho(S,E,I,D,H,R) \right) \\
& + \mu \left((S+1)\rho(S+1,E,I,D,H,R) - S\rho(S,E,I,D,H,R) \right) \\
& + \mu \left((E+1)\rho(S,E+1,I,D,H,R) - E\rho(S,E,I,D,H,R) \right) \\
& + \sigma \left((E+1)\rho(S,E+1,I-1,D,H,R) - E\rho(S,E,I,D,H,R) \right) \\
& + \mu \left((I+1)\rho(S,E,I+1,D,H,R) - I\rho(S,E,I,D,H,R) \right) \\
& + \tau \left((I+1)\rho(S,E,I+1,D,H-1,R) - I\rho(S,E,I,D,H,R) \right) \\
& + \mu_e \left((I+1)\rho(S,E,I+1,D-1,H,R) - I\rho(S,E,I,D,H,R) \right) \\
& + \gamma_{ir} \left((I+1)\rho(S,E,I+1,D,H,R-1) - I\rho(S,E,I,D,H,R) \right) \\
& + (\mu+\delta) \left((D+1)\rho(S,E,I,D+1,H,R) - D\rho(S,E,I,D,H,R) \right) \\
& + (\mu+\mu_e) \left((H+1)\rho(S,E,I,D,H+1,R) - H\rho(S,E,I,D,H,R) \right) \\
& + \gamma_{hr} \left((H+1)\rho(S,E,I,D,H+1,R-1) - H\rho(S,E,I,D,H,R) \right) \\
& + \mu \left((R+1)\rho(S,E,I,D,H,R+1) - R\rho(S,E,I,D,H,R) \right),
\end{aligned}
\tag{45}
$$

where $\mathbf{X} = [S,E,I,D,H,R]^T$.

   We scale the variables by the average population size as follows: $x_S = S/N$, $x_E = E/N$, $x_I = I/N$, $x_D = D/N$, $x_H = H/N$, and $x_R = R/N$. In the case of a large population,

Event	Transition	Rate
Healthy Birth	$\varnothing \rightarrow S$	$\mu\, N$
EVD Transmission (human)	$S \rightarrow E$	$(\beta_i\, I + \beta_d\, D + \beta_h\, H)\frac{S}{N}$
EVD Transmission (animal)	$S \rightarrow E$	$\kappa\, S$
Latency to Infectious	$E \rightarrow I$	$\sigma\, E$
Recovery	$I \rightarrow R$	$\gamma_{ir}\, I$
EVD Death	$I \rightarrow D$	$\mu_e\, I$
Hospitalization	$I \rightarrow H$	$\tau\, I$
Burial	$D \rightarrow \varnothing$	$\delta\, D$
Death from Hospital	$H \rightarrow \varnothing$	$\mu_e\, H$
Recovery from Hospital	$H \rightarrow R$	$\gamma_{hr}\, H$
Natural Death	$\{S,E,I,D,H,R\} \rightarrow \varnothing$	$\mu\, \{S,E,I,D,H,R\}$

Table 4: The transition events and their associated transition rates for the stochastic EVD model. Each transition involves the movement of a single individual between classes. The classes are represented by the following variables: S = Susceptible, E = Exposed, I = Infectious, R = Recovered, H = hospitalized, and D = Deceased. The average population size is N.

the WKB (Wentzel-Kramers-Brillouin) approximation for the probability distribution in

the scaled master equation can be used [Gang, 1987, Kubo et al., 1973, Dykman et al.,

1994, Elgart and Kamenev, 2004, Kessler and Shnerb, 2007]. By performing a Taylor

series expansion on the equation (45), one arrives at the leading order Hamilton-Jacobi

equation, as seen in Chapter 1.2. In this formulation the stochastic dynamics are captured

in the conjugate momentum variables. The Hamiltonian for the stochastic EVD model

using the WKB approximation is

$$\mathcal{H} = \mu(e^{p_S} - 1) + (\beta_d x_D x_S + \beta_h x_H x_S + \beta_i x_I x_S + \kappa x_S)\left(e^{(-p_S+p_E)} - 1\right)$$

$$+ \mu x_S(e^{-p_S} - 1) + \mu x_E(e^{-p_E} - 1) + \sigma x_E\left(e^{(-p_E+p_I)} - 1\right)$$

$$+ \mu x_I(e^{-p_I} - 1) + \mu_e x_I\left(e^{(-p_I+p_D)} - 1\right) + \gamma_{ir} x_I\left(e^{(-p_I+p_R)} - 1\right) \tag{46}$$

$$+ \tau x_I\left(e^{(-p_I+p_H)} - 1\right) + (\mu + \delta)x_D(e^{-p_D} - 1) + \mu x_R(e^{-p_R} - 1)$$

$$+ (\mu + \mu_e)x_H(e^{-p_H} - 1) + \gamma_{hr} x_H\left(e^{(-p_H+p_R)} - 1\right),$$

and the associated Hamilton's equations are given by the following relations:

$$\begin{aligned}
\frac{dx_S}{dt} &= \frac{\partial \mathcal{H}}{\partial p_S}, & \frac{dx_E}{dt} &= \frac{\partial \mathcal{H}}{\partial p_E}, & \frac{dx_I}{dt} &= \frac{\partial \mathcal{H}}{\partial p_I}, \\
\frac{dx_D}{dt} &= \frac{\partial \mathcal{H}}{\partial p_D}, & \frac{dx_H}{dt} &= \frac{\partial \mathcal{H}}{\partial p_H}, & \frac{dx_R}{dt} &= \frac{\partial \mathcal{H}}{\partial p_R}, \\
\frac{dp_S}{dt} &= -\frac{\partial \mathcal{H}}{\partial x_S}, & \frac{dp_E}{dt} &= -\frac{\partial \mathcal{H}}{\partial x_E}, & \frac{dp_I}{dt} &= -\frac{\partial \mathcal{H}}{\partial x_I}, \\
\frac{dp_D}{dt} &= -\frac{\partial \mathcal{H}}{\partial x_D}, & \frac{dp_H}{dt} &= -\frac{\partial \mathcal{H}}{\partial x_H}, & \frac{dp_R}{dt} &= -\frac{\partial \mathcal{H}}{\partial x_R}.
\end{aligned} \tag{47}$$

As was seen previously in Chapter 2 the master equation/WKB formalism doubles

the dimension of the problem but recasts it as a deterministic system of nonlinear differ-

ential equations. These equations are useful for analysis of the stochastic system's dy-

namics [Assaf and Meerson, 2010, Schwartz et al., 2011, Forgoston et al., 2011]. The full

equations for the EVD model are

$$\frac{dx_S}{dt} = \mu e^{p_S} - (\beta_d x_D + \beta_h x_H + \beta_i x_I + \kappa) x_S e^{-p_S + p_E} - \mu x_S e^{-p_S} \tag{48}$$

$$\frac{dx_E}{dt} = (\beta_d x_D + \beta_h x_H + \beta_i x_I + \kappa) x_S e^{-p_S + p_E} - \sigma x_E e^{-p_E + p_I} - \mu x_E e^{-p_E} \tag{49}$$

$$\frac{dx_I}{dt} = \sigma x_E e^{-p_E + p_I} - \mu_e x_I e^{-p_I + p_D} - \gamma_{ir} x_I e^{-p_I + p_R} - \tau x_I e^{-p_I + p_H}$$

$$- \mu x_I e^{-p_I} \tag{50}$$

$$\frac{dx_D}{dt} = \mu_e x_I e^{-p_I + p_D} - (\mu + \delta) x_D e^{-p_D} \tag{51}$$

$$\frac{dx_H}{dt} = \tau x_I e^{-p_I + p_H} - \gamma_{hr} x_H e^{-p_H + p_R} - (\mu + \mu_e) x_H e^{-p_H} \tag{52}$$

$$\frac{dx_R}{dt} = \gamma_{ir} x_I e^{-p_I + p_R} + \gamma_{hr} x_H e^{-p_H + p_R} - \mu x_R e^{-p_R} \tag{53}$$

$$\frac{dp_S}{dt} = -(\beta_d x_D + \beta_h x_H + \beta_i x_I + \kappa)\left(e^{-p_S + p_E} - 1\right) - \mu\left(e^{-p_S} - 1\right) \tag{54}$$

$$\frac{dp_E}{dt} = -\sigma\left(e^{-p_E + p_I} - 1\right) - \mu\left(e^{-p_E} - 1\right) \tag{55}$$

$$\frac{dp_I}{dt} = -\beta_i x_S\left(e^{-p_S + p_E} - 1\right) - \mu_e\left(e^{-p_I + p_D} - 1\right) - \gamma_{ir}\left(e^{-p_I + p_R} - 1\right)$$

$$- \tau\left(e^{-p_I + p_H} - 1\right) - \mu\left(e^{-p_I} - 1\right) \tag{56}$$

$$\frac{dp_D}{dt} = -\beta_d x_S\left(e^{-p_S + p_E} - 1\right) - (\mu + \delta)\left(e^{-p_D} - 1\right) \tag{57}$$

$$\frac{dp_H}{dt} = -\beta_h x_S\left(e^{-p_S + p_E} - 1\right) - \gamma_{hr}\left(e^{-p_H + p_R} - 1\right) - (\mu + \mu_e)\left(e^{-p_H} - 1\right) \tag{58}$$

$$\frac{dp_R}{dt} = -\mu\left(e^{-p_R} - 1\right) \tag{59}$$

3.2.2   Deterministic Mean Field Equations

In the infinite population size limit, stochastic systems tend towards a zero-variance system; this is by definition the associated mean-field or deterministic system. The associ-

ated mean-field system can be used to determine useful quantities such as the deterministically stable equilibria, and the basic reproduction number.

To find the mean-field equations for the stochastic EVD model, one must evaluate the following partial derivatives of the Hamiltonian with the conjugate momentum variables set to zero (noted by $\mathbf{p} = \mathbf{0}$):

$$\frac{dx_S}{dt} = \frac{\partial \mathcal{H}}{\partial p_S}\Big|_{\mathbf{p}=\mathbf{0}} \qquad \frac{dx_E}{dt} = \frac{\partial \mathcal{H}}{\partial p_E}\Big|_{\mathbf{p}=\mathbf{0}} \qquad \frac{dx_I}{dt} = \frac{\partial \mathcal{H}}{\partial p_I}\Big|_{\mathbf{p}=\mathbf{0}}$$
$$\frac{dx_D}{dt} = \frac{\partial \mathcal{H}}{\partial p_D}\Big|_{\mathbf{p}=\mathbf{0}} \qquad \frac{dx_H}{dt} = \frac{\partial \mathcal{H}}{\partial p_H}\Big|_{\mathbf{p}=\mathbf{0}} \qquad \frac{dx_R}{dt} = \frac{\partial \mathcal{H}}{\partial p_R}\Big|_{\mathbf{p}=\mathbf{0}}. \tag{60}$$

The deterministic mean-field dynamics are given by the following equations:

$$\frac{dx_S}{dt} = \mu - \beta_i x_I x_S - \beta_d x_D x_S - \beta_h x_H x_S - \mu x_S - \kappa x_S \tag{61a}$$

$$\frac{dx_E}{dt} = \beta_i x_I x_S + \beta_d x_D x_S + \beta_h x_H x_S - (\mu + \sigma)x_E + \kappa x_S \tag{61b}$$

$$\frac{dx_I}{dt} = \sigma x_E - (\gamma_{ir} + \mu_e + \tau + \mu)x_I \tag{61c}$$

$$\frac{dx_D}{dt} = \mu_e x_I - (\delta + \mu)x_D \tag{61d}$$

$$\frac{dx_H}{dt} = \tau x_I - (\gamma_{hr} + \mu_e + \mu)x_H \tag{61e}$$

$$\frac{dx_R}{dt} = \gamma_{ir} x_I + \gamma_{hr} x_H - \mu x_R. \tag{61f}$$

Setting the deterministic mean-field equations to zero allows one to find the two steady states of the system in terms of the exposed variable. The exposed class value for the

equilibrium points are

$$x_E^{(e)} = \frac{\mu}{2R_0\,(\mu+\sigma)}\left(R_0 - 1 - \frac{\kappa}{\mu} + \sqrt{\left(R_0 - 1 - \frac{\kappa}{\mu}\right)^2 + \frac{4R_0\kappa}{\mu}}\right) \qquad (62)$$

and

$$x_E^{(i)} = \frac{\mu}{2R_0\,(\mu+\sigma)}\left(R_0 - 1 - \frac{\kappa}{\mu} - \sqrt{\left(R_0 - 1 - \frac{\kappa}{\mu}\right)^2 + \frac{4R_0\kappa}{\mu}}\right). \qquad (63)$$

Note that R_0 is the basic reproduction number whose definition and derivation is provided

in 3.2.3. The values for the remaining variables of the steady states can be determined

using the following relationships:

$$x_S^{(e,i)} = 1 - \left(1 + \frac{\sigma}{\mu}\right) x_E^{(e,i)} \qquad (64)$$

$$x_I^{(e,i)} = \frac{\sigma}{(\gamma_{ir} + \tau + \mu_e + \mu)} x_E^{(e,i)}, \qquad (65)$$

$$x_D^{(e,i)} = \frac{\sigma\mu_e}{(\gamma_{ir} + \tau + \mu_e + \mu)(\delta + \mu)} x_E^{(e,i)}, \qquad (66)$$

$$x_H^{(e,i)} = \frac{\sigma\tau}{(\gamma_{ir} + \tau + \mu_e + \mu)(\gamma_{hr} + \mu_e + \mu)} x_E^{(e,i)}, \qquad (67)$$

$$x_R^{(e,i)} = \frac{\sigma}{\mu(\gamma_{ir} + \tau + \mu_e + \mu)}\left(\gamma_{ir} + \frac{\gamma_{hr}\tau}{(\gamma_{hr} + \mu_e + \mu)}\right) x_E^{(e,i)}. \qquad (68)$$

It is worth noting that if there is no transmission from the reservoir, or $\kappa = 0$, the form

of the steady states for the exposed class simplifies to

$$x_E{}^{(e)} = \frac{(1 - 1/R_0)}{1 + \sigma/\mu} \quad \text{and} \quad x_E{}^{(i)} = 0. \tag{69}$$

We refer to $x_E{}^{(i)}$ as the disease-free equilibrium (DFE):

$$\left(x_S{}^{(i)}, x_E{}^{(i)}, x_I{}^{(i)}, x_D{}^{(i)}, x_H{}^{(i)}, x_R{}^{(i)}\right) = (1, 0, 0, 0, 0, 0). \tag{70}$$

3.2.3   Deterministic Basic Reproduction Number

In both the earlier and the more recent EVD outbreaks in Africa, the virus has been re-

ported to have relatively low infectivity [Chowell et al., 2004, Althaus, 2014]. The R_0 is

a common metric for infectivity of a disease which is defined as the average number of

infections that a single infectious individual will trigger in a fully susceptible population.

The R_0 of an infectious disease model can be found explicitly using the next generation

matrix [van den Driessche and Watmough, 2002]. When $\kappa > 0$ the DFE does not exist.

Since our EVD susceptible population is very weakly coupled to the disease reservoir, the

approximation $\kappa = 0$ will be used. The mean-field equations (Eq. (61)) are rewritten using

the vector

$\mathbf{X} = [x_S, x_E, x_I, x_D, x_H, x_R]^T$. The new infections are separated from the other changes in

the population using the form $\dot{\mathbf{X}} = \hat{\mathscr{F}}(\mathbf{X}) - \hat{\mathscr{V}}(\mathbf{X})$, with

$$
\hat{\mathscr{F}} = \begin{bmatrix} 0 \\ \beta_i x_I x_S + \beta_d x_D x_S + \beta_h x_H x_S \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}
\tag{71}
$$

and

$$
-\hat{\mathscr{V}} = \begin{bmatrix} \mu - \beta_i x_I x_S - \beta_d x_D x_S - \beta_h x_H x_S - \mu x_S \\ -(\mu + \sigma) x_E \\ \sigma x_E - (\gamma_{ir} + \mu_e + \tau + \mu) x_I \\ \mu_e x_I - (\delta + \mu) x_D \\ \tau x_I - (\gamma_{hr} + \mu_e + \mu) x_H \\ \gamma_{ir} x_I + \gamma_{hr} x_H - \mu x_R \end{bmatrix} .
\tag{72}
$$

It is only necessary to consider the classes that have disease transmission. Therefore we

can consider $\mathbf{Y} = (x_E, x_I, x_D, x_H)$ and define $\dot{\mathbf{Y}} = \mathscr{F}(\mathbf{X}) - \mathscr{V}(\mathbf{X})$, with

$$
\mathscr{F} = \begin{bmatrix} \beta_i x_I x_S + \beta_d x_D x_S + \beta_h x_H x_S \\ 0 \\ 0 \\ 0 \end{bmatrix}
\tag{73}
$$

and

$$-\mathscr{V} = \begin{bmatrix} -(\mu + \sigma)x_E \\\\ \sigma x_E - (\gamma_{ir} + \mu_e + \tau + \mu)x_I \\\\ \mu_e x_I - (\delta + \mu)x_D \\\\ \tau x_I - (\gamma_{hr} + \mu_e + \mu)x_H \end{bmatrix}. \tag{74}$$

The vector $\mathscr{F}$ represents all new infectious individuals, and the vector $\mathscr{V}$ is the negative

of the remaining terms in the system.

The evaluation of the Jacobian matrices $\mathscr{F}$ and $\mathscr{V}$ at the DFE will be called $\mathbf{F}$ and $\mathbf{V}$

respectively. The inverse $\mathbf{F}\mathbf{V}^{-1}$ is known as the next generation matrix. Notice that this

expression captures the ratio between the inflow and an outflow from the infected classes

in terms of matrix operations. The basic reproduction number is defined as the spectral

radius (largest eigenvalue) of $\mathbf{F}\mathbf{V}^{-1}$:

$$R_0 = \frac{\sigma \left(\beta_i + \frac{\beta_d \mu_e}{\delta + \mu} + \frac{\beta_h \tau}{\gamma_{hr} + \mu_e + \mu} \right)}{(\gamma_{ir} + \tau + \mu_e + \mu)(\mu + \sigma)}. \tag{75}$$

The basic reproduction number R_0 can be roughly described as the ratio between the

inflow of infected individuals and the outflow of dead or recovered individuals in a fully

susceptible population. If the inflow is greater than the outflow, then $R_0 > 1$, and the dis-

ease will spread. If the outflow is greater than the inflow, then $R_0 < 1$, and the disease will

go extinct. For context, the reproductive number for measles may be as high as $R_0 = 18$

and the reproductive number for influenza is approximately $R_0 \approx 4$. The reproductive

number for EVD is estimated to be approximately $R_0 \leq 2$ [Chowell et al., 2004, Althaus,

Description	Parameter	Value
1/host life span (& birth rate)	μ	0.00005 day^{-1}
contact rate for infectious	β_i	0.5 day^{-1}
contact rate for deceased	β_d	0.6 day^{-1}
contact rate for hospitalized	β_h	0.00016 day^{-1}
1/latency period	σ	0.1 day^{-1}
1/recovery period (no hospital)	γ_{ir}	0.07 day^{-1}
death rate from EVD	μ_e	0.12 day^{-1}
1/mean time to hospitalization	τ	0.2 day^{-1}
1/burial time	δ	0.33 day^{-1}
1/recovery period (hospital)	γ_{hr}	0.10 day^{-1}
reservoir transmission	κ	$5\text{E-}9 \text{ day}^{-1}$

Table 5: The parameter values used in the stochastic EVD model, as reported in Ref. [Hu et al., 2015].

2014]. The parameters in Table 5 agree with this estimation.

3.2.4   Model Parameters

The parameter values we use for EVD are given in Table 5. With the exception of κ, the

parameters are as reported in Ref. [Hu et al., 2015]. The value of κ defines the probability

of an infection from the animal reservoir. The approximate range for κ can be calculated

by dividing the number of outbreak events by the time over which those outbreaks took

place and then dividing by the population size. This gives an approximate range for κ

between 10^{-9} and 10^{-6}.

3.3   Results

The standard approach to analyzing the dynamics of a stochastic system such as this is to

start with the mean-field equations (provided in 3.2.2). This is done by assuming a value

of zero for the conjugate momentum variables in Hamilton's equations. This reduced system captures the deterministic dynamics, which is the limit of the stochastic dynamics in the large population limit. Full analysis of the mean-field equations for the EVD model, shown in Section 3.2.2, reveals dynamics similar to the lower dimensional *SEIR* systems in the literature [Anderson and May, 1991].

By introducing reservoir transmission ($\kappa > 0$), the traditional DFE no longer exists as a small number of exposed individuals are consistently introduced to the system. Therefore, it can be said that the system is either at a true endemic state when $R_0 > 1$, or the system can persist with a very low number of infectious individuals when $R_0 < 1$. We call this the invasion state. While this deterministic system cannot exhibit periods of disease fadeout and re-invasion, its steady states provide guidance in understanding the dynamics of the full stochastic system. The existence of the invasion state depends on the force of infection from the animal reservoir and we will see that it plays an important role in outbreak vulnerability.

Another advantage of analyzing the mean-field equations is the ability to approximate the basic reproduction number, as was shown in Section 3.2.3. Using the next generation matrix as described in [van den Driessche and Watmough, 2002] and assuming no reservoir transmission ($\kappa = 0$), the DFE was used to analytically determine

$$R_0 = \frac{\sigma \left(\beta_i + \frac{\beta_d \mu_e}{\delta + \mu} + \frac{\beta_h \tau}{\gamma_{hr} + \mu_e + \mu} \right)}{\left(\gamma_{ir} + \tau + \mu_e + \mu \right) \left(\mu + \sigma \right)}, \tag{76}$$

which quantifies if EVD can persist in the population. The derivation for the basic repro-

duction number is provided in 3.2.3. Low frequency reservoir transmissions and weak

coupling between the population of study and the external disease reservoir make this a

good approximation of the ability for the disease to persist when randomly introduced.

Intervention methods can also be evaluated by how they decrease the basic reproduction

number from Eq. 76, with an aim of $R_0 < 1$.

Real world data for EVD suggests that the basic reproduction rate is greater than one,

but the outbreak and extinction cycles of EVD are not well explained by deterministic

models with solutions that simply limit to a steady state. The stochastic EVD model al-

lows behavior beyond the dynamics predicted by the mean-field equations. Examples

include solutions that switch steady states, resembling disease outbreak (invasion) and

fadeout (extinction) events. Of particular interest is the effect of a transmission reservoir

in these dynamics.

3.3.1   Invasion

Outbreak vulnerability is a measure of how likely an outbreak is, and how devastating that

outbreak can be. A population with high outbreak vulnerability incurs frequent or large

outbreaks, while a population with low outbreak vulnerability will have infrequent and

relatively small outbreaks. The random disease-introduction occurs through the transmis-

sion reservoir, quantified by the parameter κ. Although large infrequent outbreaks result

in an under-prepared and overwhelmed healthcare system, long and sustained outbreaks

are more taxing on a population. Assessing the outbreak vulnerability of a specific popu-
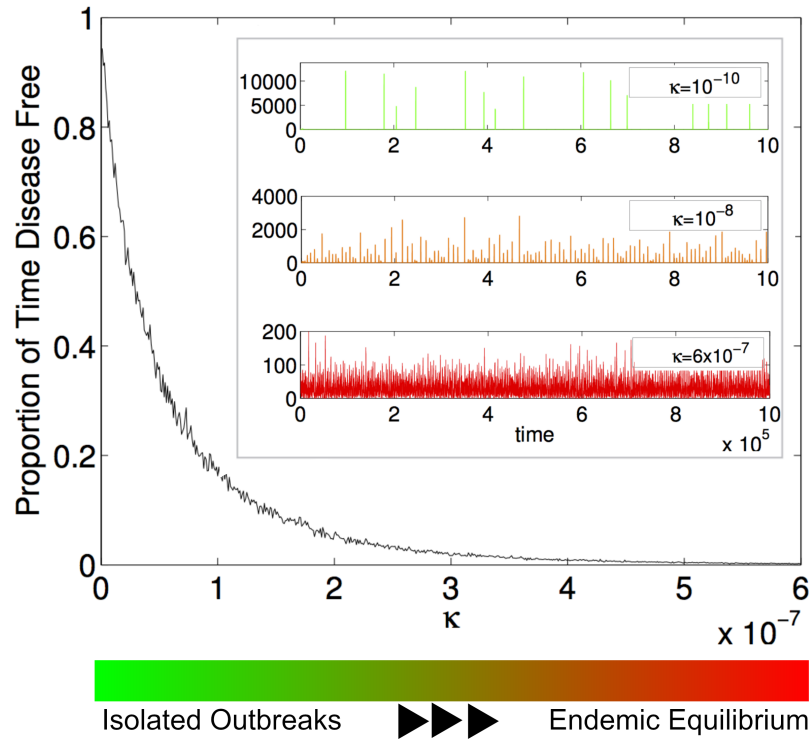
Figure 12: A measure of outbreak vulnerability as a function of the reservoir transmission rate κ. The figure, based on 10^3 stochastic simulations run for 10^6 days for a population $N = 500,000$, shows the proportion of disease-free time for a given κ. The inlay shows three representative time series of the number of cases of EVD over a sample of 10^5 days. All parameters are set to the values in Table 5 except κ, which is noted in each graph.

lation and the effectiveness of possible intervention strategies would help prepare health-

care workers and possibly prevent a breakdown of the system during an outbreak.

Stochastic models do not have a single solution like their deterministic counterparts.

In order to verify analytical results and gain both quantitative and qualitative insight into a

stochastic system, it is necessary to run in-silico numerical stochastic simulations. Large

collections of stochastic realizations are known as stochastic ensembles. Each individ-

ual stochastic realization represents a possible disease trajectory, and is produced using a

Monte Carlo method [Gillespie, 1976]. The algorithm uses a random number generator to

determine which event will occur as well as the time of occurrence. During each random time step exactly one event occurs. The probability of any particular event taking place is equal to its own transition rate divided by the sum of all transition rates. Representative stochastic realizations are shown in Fig. 12. These realizations were made using a sampling of reservoir transmission rates, with all other model parameters given in Table 5.

If there is very low transmission from the reservoir and most of the population is susceptible, the solution will exhibit large outbreaks, or spillover events, spread out in time so that there are long disease-free periods. This behavior is represented by the green time series in the inlay of Fig. 12. If the transmission rate from the reservoir is high, then the solution fluctuates about a non-zero steady state. In this case there are few (if any) disease-free periods. The color bar at the bottom of Fig. 12 shows the range of κ-values. The red, orange, and green colors are associated with different κ-values and with the representative time series in the upper right of Fig. 12. It can be seen that there are three different qualitative outbreak behaviors associated with the three different time series. In the green time series the disease exhibits relatively rare outbreak, in the orange time series the outbreaks become more frequent, and in the red time series the disease is practically always present.

The average proportion of disease-free time during a sample of 1,000 stochastic simulations was found to give a quantitative measure of outbreak vulnerability. The results for EVD parameters as κ is varied can be seen in Fig. 12. For populations with very weak connection to the disease reservoir (small κ), almost all of the time is spent disease-free.

As κ increases, the disease-free time decreases at a nearly exponential rate. The periods

of disease-free time can be described as disease fadeout and the decrease of fadeout is

associated with sustained outbreaks.

Both the contact rate with infected individuals (β_i) and the EVD burial rate (δ) are im-

portant factors for a population's outbreak vulnerability and can change over the course

of an outbreak, as they relate to human behavioral considerations that are likely to be af-

fected by the spread of information. One would assume that increasing δ and decreasing

β_i would be salient and practical strategies to reduce outbreak vulnerability. The stochas-

tic EVD model can provide an estimate of how the outbreak vulnerability changes with

these prevention measures, allowing for the assessment of how to achieve the maximum

impact. Figure 13 is a contour plot of the average proportion of disease-free time in sim-

ulations of the EVD model as β_i and δ are varied. These simulations follow the method-

ology described for Fig. 12. As expected, low values for the contact rate and high values

for the burial rate have the most disease free time and least outbreak vulnerability. Over-

laid as a dashed black curve is $R_0 = 1$ for the mean-field EVD model without reservoir

transmission, as described in Eq. (76). This curve is the boundary between the basins for

which the DFE and the endemic steady states are stable in the mean-field EVD model.

When $R_0 < 1$, the DFE is stable for low values for the contact rate and high values for the

burial rate. The $R_0 = 1$ curve is shown to approximate the boundary between low and high

outbreak vulnerability in the stochastic EVD model with reservoir transmission. There-

fore, it indicates intervention strategies that target β_i and δ parameter values for $R_0 < 1$
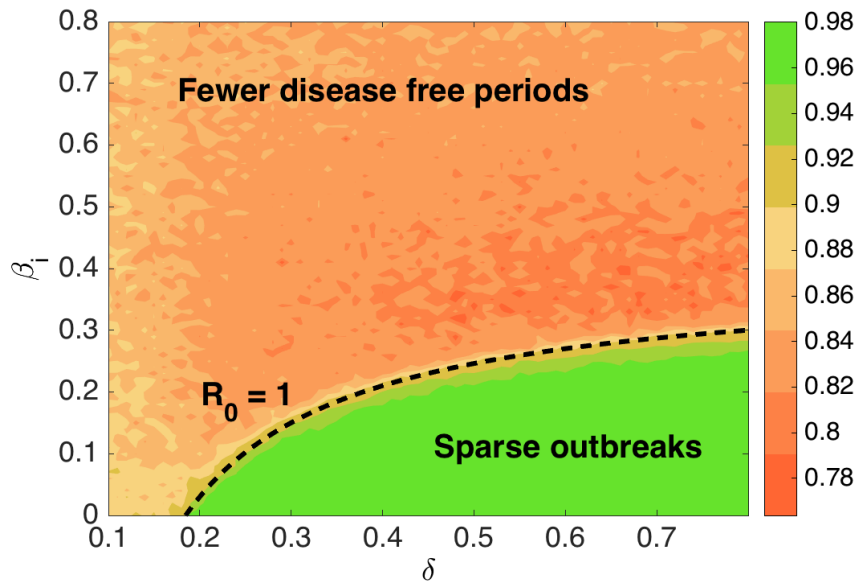
Figure 13: A measure of intervention effectiveness considering the impact of limiting the contact rate with infectious individuals β_i and increasing the burial rate for deceased EVD individuals δ. The figure shows a contour plot of the proportion of disease-free time for simulations as described in Fig. 12 with parameters given in Table 5 and a population $N = 500,000$. Higher values (green) represent infrequent outbreaks with long periods that are disease-free. Lower values (red) represent few disease-free periods and sustained outbreaks. Overlaid as a dashed black curve is $R_0 = 1$ for the mean-field EVD model without reservoir transmission, as described in Eq. (76).

(below the dashed line) which would have the greatest benefit for the population.

3.3.2   Extinction

The stochastic EVD model can exhibit periods of disease absence, or fadeout, despite

the fact that the basic reproduction rate is greater than one. As has been noted in Sec-

tion 3.3.1, outbreaks can be quick and infrequent or sustained and frequent, depending

on outbreak vulnerability. This distinction describes the two mechanisms that allow for a

solution to reach a disease-free state, which will be called an extinction event.

When the outbreak vulnerability is low, there are long disease-free periods in which
the susceptible population replenishes itself as the recovered individuals are naturally re-
moved. The small transmission rate from the reservoir has a low probability of causing an
outbreak, and finally when it happens, the susceptible compartment is the majority of the
population and the disease quickly spreads through the group. Here, the mean-field model
provides general insight for the underlying dynamics exhibited by the stochastic model.
The stable endemic state (assuming $R_0 > 1$) is a spiral sink and solutions spiral in towards
it in an anticlockwise fashion when projected on the Susceptible-Infectious plane. The
initial conditions for the invasion are far from the endemic state, which is stable but not
strongly attracting $(1 < R_0 < 2)$. Therefore, the solution will orbit the endemic state in a
large anticlockwise path (at a significant distance from the endemic state), quickly reach-
ing a disease-free state on the other side of the orbit. The solution will remain disease-free
as the susceptible population rebuilds and repeats the cycle. These outbreaks are random,
but as shown in Fig. 13, the average behavior follows a continuous measure of disease-
free time that depends on the system parameters. It is worth mentioning that the invasion
steady state is small and does not seem to play a significant role in these dynamics.

For large outbreak vulnerability, the invasion is a solution that escapes the disease-free
state and causes a large disease outbreak. In this case, the solution is similar to what one
would expect from the deterministic mean-field equations. What is unique about stochas-
tic systems is that the small random fluctuations allow escape from steady states, and in
simulations this will happen in finite time (although the time may be exponentially long).

To understand the mechanism of escape, one can use Hamilton's equations to identify the

metastable steady states in the stochastic system, as well as action-minimizing curves that

provide a path from one steady state to another. An action-minimizing curve identifies the

path that is probabilistically most likely to be taken during a switching event. This tra-

jectory is thus known as the optimal path [Dykman et al., 1994, Freidlin and Wentzell,

1984, Graham and Tél, 1984, Schwartz et al., 2011]. For the EVD model, the optimal path

to extinction is a curve in 12-dimensional space that connects the endemic state to the ex-

tinct state.

Noting that κ is much smaller than the other parameters in Table 5, solutions for the

stochastic EVD model near the endemic state can be approximated by assuming $\kappa = 0$.

The exception is during invasion events, when the reservoir brings the disease back from

extinction. Therefore, $\kappa = 0$ will result in an illustrative example for computing the opti-

mal path to extinction and comparing it to extinction dynamics in simulations.

3.3.3   Optimal Path

As has been shown previously in Chapter 2, the optimal path can be used to analyze the

dynamics of spontaneous escape from an endemic state. Recall that the optimal path is a

zero-energy curve for the Hamiltonian that connects two steady state saddle points. For

this high-dimensional model, an analytic solution is not tractable and so numerical meth-

ods are used to approximate the optimal path. The Iterative Action Minimizing Method

(IAMM) [Lindley and Schwartz, 2013] is used to find 2,400 points approximating the

12-dimensional optimal path of the stochastic EVD model, with a maximum error of

2.4487×10^{-10}.

The IAMM is useful in the situation where a path connecting steady states C_a and C_b

starts at C_a at $t = -\infty$ and ends at C_b at $t = +\infty$. A time parameter t exists such that $-\infty <$

$t < \infty$. A numerical approximation of the time needed to leave the region of C_a and arrive

in the region of C_b is needed for this method. Therefore, a time T_ε is defined such that

$-\infty < -T_\varepsilon < t < T_\varepsilon < \infty$. Additionally, $C(-T_\varepsilon) \approx C_a$ and $C(T_\varepsilon) \approx C_b$. In other words, the

solution stays very near the equilibrium C_a for $-\infty < t \leq -T_\varepsilon$, has a transition region from

$-T_\varepsilon < t < T_\varepsilon$, and then stays near C_b for $T_\varepsilon < t < +\infty$. The interval $[-T_\varepsilon, T_\varepsilon]$ is discretized

into n segments using a uniform step size $h = (2T_\varepsilon)/n$ or a suitable non-uniform step size

h_k. The corresponding time series is $t_{k+1} = t_k + h_k$.

The derivative of the function value $\mathbf{q}_k$ is approximated using central finite differences

by the operator δ_h given as

$$\frac{d}{dt}\mathbf{q}_k \approx \delta_h \mathbf{q}_k \equiv \frac{h_{k-1}^2 \mathbf{q}_{k+1} + (h_k^2 - h_{k-1}^2)\mathbf{q}_k - h_k^2 \mathbf{q}_{k-1}}{h_{k-1}h_k^2 + h_k h_{k-1}^2}, \quad k = 0,\dots,n. \tag{77}$$

Clearly, if a uniform step size is chosen then Eq. (77) simplifies to the familiar form given

as

$$\frac{d}{dt}\mathbf{q}_k \approx \delta_h \mathbf{q}_k \equiv \frac{\mathbf{q}_{k+1} - \mathbf{q}_{k-1}}{2h}, \quad k = 0,\dots,n. \tag{78}$$

Thus, one can develop the system of nonlinear algebraic equations

$$\delta_h \mathbf{x}_k - \frac{\partial H(\mathbf{x}_k, \mathbf{p}_k)}{\partial \mathbf{p}} = 0, \quad \delta_h \mathbf{p}_k + \frac{\partial H(\mathbf{x}_k, \mathbf{p}_k)}{\partial \mathbf{x}} = 0, \quad k = 0, \dots, n, \tag{79}$$

which is solved using a general Newton's method.

An extended vector of $2nN$ components is given by

$$\mathbf{q}_j(\mathbf{x}, \mathbf{p}) = \{\mathbf{x}_{1,j} \dots \mathbf{x}_{n,j}, \mathbf{p}_{1,j} \dots \mathbf{p}_{n,j}\}^T, \tag{80}$$

and contains the jth Newton iterate, where N is the number of populations. When $j = 0$,

$\mathbf{q}_0(\mathbf{x}, \mathbf{p})$ provides the initial "guess" as to the location of the path that connects C_a and C_b.

Given the jth Newton iterate $\mathbf{q}_j$, the new $\mathbf{q}_{j+1}$ iterate is found by solving the linear system

$$\mathbf{q}_{j+1} = \mathbf{q}_j - \frac{\mathbf{F}(\mathbf{q}_j)}{\mathbf{J}(\mathbf{q}_j)}, \tag{81}$$

where $\mathbf{F}$ is the function defined by Eq. (79) acting on $\mathbf{q}_j$, and $\mathbf{J}$ is the Jacobian. Equation (81) is solved using LU decomposition with partial pivoting.

Because the endemic state of the mean field for the parameters in Table 5 is a spiral

sink, one must use a continuation method to find the optimal path. This is done by picking

a large value of μ that results in an R_0 close to but larger than one, which decreases the

frequency of the oscillations about the endemic state. Therefore, the optimal path can be

found using an initial condition of a straight line connecting the endpoints. By slowly

decreasing the value of μ and using the previous optimal path as the initial condition, one

can repeat the process to find the optimal path at the desired parameter values. Although

the optimal path is in 12-dimensional space, and so it can not be easily visualized, the

progression of each individual compartment is shown in Fig. 14.

Figure 15 shows the resulting solution, starting at the endemic state and spiraling out

to the disease-free state. Our results are verified by comparing the optimal path with the

probability density of extinction prehistory in the Susceptible-Infectious plane. The prob-

ability density was numerically computed using $10,000$ simulations that ended in sponta-

neous extinction. Starting at the extinction point, the previous positions in state space for

each trajectory is binned and the frequency is plotted in the Susceptible-Infectious plane

in Fig. 15. The highest frequency regions are shown in red in the density plot, and the

optimal path is overlaid. Figure 15 shows that, among all the paths the stochastic system

can take to reach the extinct state, there is one path that has the highest probability of oc-

curring. This optimal path of extinction lies on the peak of the probability density of the

extinction prehistory.

Additional verification of the optimal path to extinction is achieved by projecting

the 12-dimensional optimal path onto the lower-dimensional stochastic center manifold

shown in Figure 16. Since the stochastic EVD model is an *SEIR*-type model, it is straight-

forward to compute the stochastic center manifold from the mean-field equations using

techniques developed in [Forgoston et al., 2009, Forgoston and Schwartz, 2013]. The
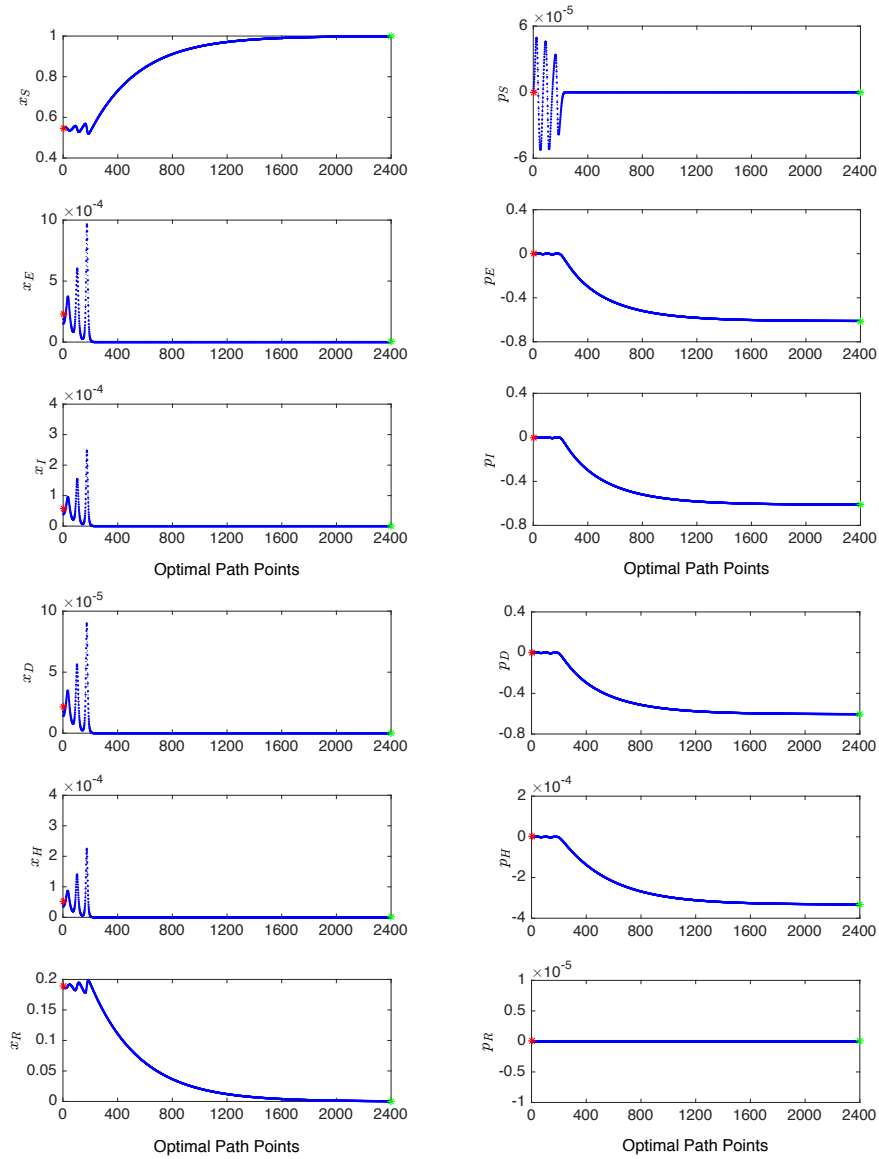
Figure 14: The parameters are given in Table 5, with the exception of $\kappa = 0$. Each set of 2,400 blue points represent the numerical approximation of the optimal path to extinction found by the IAMM method. The maximum error for this set is 2.4487×10^{-10}. The red star represents the location of the endemic state (left starting point) and the green star represents the location of the extinct state (right end point).
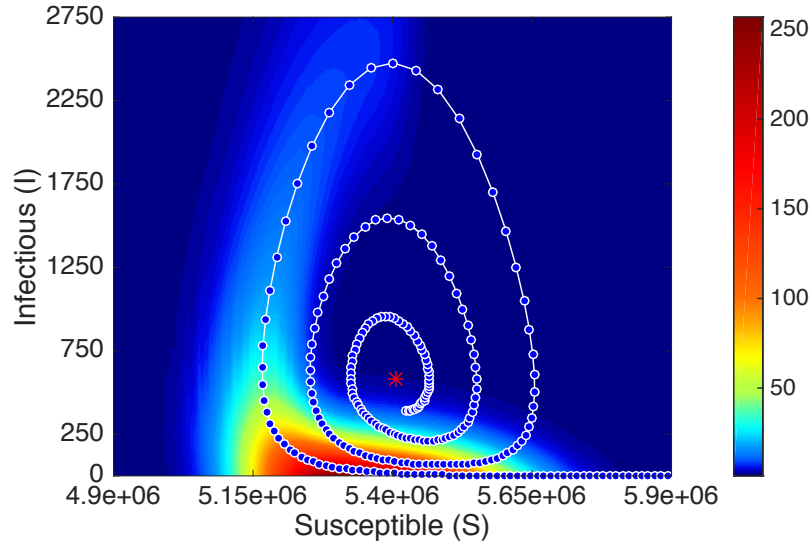
Figure 15: An optimal extinction path for a stochastic EVD system with $\kappa = 0$ and parameters given in Table 5. The blue points represent the numerical approximation of the optimal path to extinction. The path is overlaid on the probability density of extinction prehistories for 10^4 stochastic realizations for a population $N = 10^7$. Red indicates the highest frequencies.

stochastic center manifold, is characterized by the dimensionally reduced equation

$$x_I = \frac{\sigma(x_E - x_E{}^{(e)})}{(\gamma_{ir} + \tau + \mu_e + \mu)} + x_I{}^{(e)}. \tag{82}$$

The stochastic center manifold is the surface on which the long term dynamics of the

stochastic system will fall. Therefore it follows that the approximation of the optimal path

should lie on this hyperplane.

Knowledge of the optimal path is also useful because it allows one to find the average

time to travel from the endemic state to the extinct state. This is called the mean time to

disease extinction (MTE). Since the optimal path can be used to estimate the MTE it is
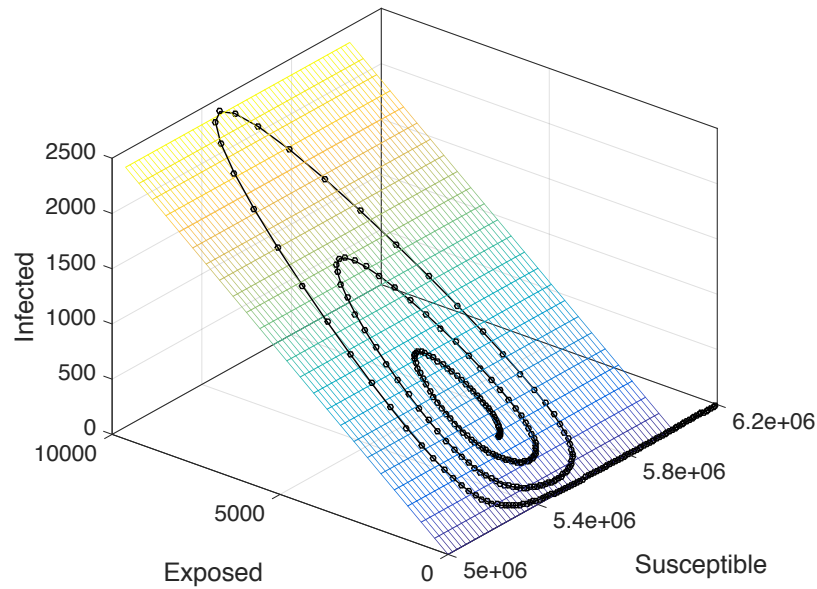
Figure 16: The projection of the numerically computed optimal path to extinction on the analytically determined stochastic center manifold. Parameter values are given in Table 5, with $\kappa = 0$.

useful for quantifying the efficacy of disease control such as vaccine and quarantine. In particular, an effective disease control program will shift the optimal path in a way that decreases the MTE.

For outbreak vulnerability parameters in between the two extremes, the system may exhibit a combination of quick and sustained outbreaks, sometimes utilizing the optimal path to escape from the endemic state. The MTE may provide insight to the frequency of disease-free periods and may also provide a measure of intervention effectiveness.

3.3.4   Dynamic Population Size

In all the previous sections, the birth rate was assumed to be equal to the death rate, which

results in a stationary mean population size. While population growth rates in many African

regions have declined steadily over time, the overall growth in Africa remains a net posi-

tive, with the population of many countries expected to double by the year 2050 [United

Nation Population Division Department of Economic and Social Affairs, 2015, Bongaarts,

2013]. When the birth rate is greater than the death rate, the total population will change

in time. If the population size is dynamic, then both the endemic state and the basic repro-

duction rate become dynamic. Higher birth rates increase the flow of susceptible individ-

uals into the population, which increases the probability of stochastic reintroduction and

the force of infection. Consequently, the likelihood of stochastic fadeout decreases and

the system is bound to transition to a dynamic endemic state. This implies that, even for

the subcritical initial dynamic condition $R_0 < 1$, a transition to endemic disease circulation

is likely. Moreover, it is dependent on various measures including both κ and birth rate.

This condition, when realized, may limit the impact of intervention. The linkage among

these effects is left to future work. By maintaining the natural death rate for all com-

partments at the value given in Table 5, and increasing the birth rate to $\mu = 9.5 \times 10^{-5}$,

one can compute stochastic realizations of this dynamic process. An example is shown

in Fig. 17, where one can see that there is a fundamental change in behavior. Initially,

the population is relatively small and invasion events are infrequent, as shown by the red

curve. Later in time, the inflow of susceptibles causes a shift to a persistent but increasing
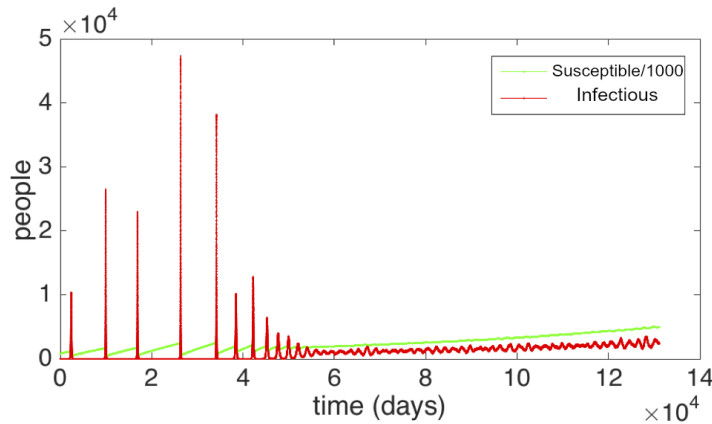
Figure 17: A growing population increases outbreak vulnerability and allows for sustained outbreaks. A sample time series with increasing population size (birth rate $\mu = 9.5 \times 10^{-5}$ and $\kappa = 1.8 \times 10^{-9}$), and a starting population of $N = 10^6$, is shown. The susceptible population has been scaled by a factor of 1000 so that both the infectious and susceptible time series can be clearly seen in the figure. Initially, the reservoir transmission triggers large outbreaks and result in fast extinction events. As the size of the population and the number of susceptible individuals increases, infectious individuals tend to a dynamic endemic state with sustained outbreaks. All other parameters are as in Table 5.

endemic state. Outbreak vulnerability has increased with population size, transitioning

from infrequent to sustained outbreaks.

3.4 Conclusions

The recent West African epidemic of the Ebola Virus Disease was a catastrophe that re-

sulted in a devastating loss of life. With its frighteningly low survival rate, along with our

apparent inability to predict the disease introduction and trajectory, EVD poses a con-

tinuing threat. For any such zoonosis, prevention prior to introduction is as important as

disease control during an outbreak, and so the disease must be taken as seriously dur-

ing its latent period as it is during an active period. In this chapter it is argued that EVD should be considered in a stochastic framework and not as an entirely predictable deterministic process. The study of zoonotic diseases benefits particularly from the application of stochastic modeling methods because of the unpredictable nature of the human-to-reservoir interactions.

The optimal path to disease extinction is a minimal action trajectory between an endemic and extinct state. This path is the most likely route to disease extinction in the presence of noise. The optimal path can therefore be used to estimate the mean time to disease extinction, and as a result may be used to quantitatively assess disease control measures.

When a disease intervention strategy is adopted, it will affect the optimal path and the expected time to disease extinction. In order to determine an optimal intervention strategy one must quantitatively compare different possible control methods. Control through an investigation of the optimal path has been successfully adopted in biochemical networks dynamics [Wells et al., 2015]. The mean time to extinction, as determined through the optimal path, is the quantitative metric that should be used to determine the efficacy of a control method. In particular, studying the relationship between the optimal path and disease intervention may allow for the optimal allocation of resources prior to or during an epidemic. For example, in [Khasin et al., 2010] it was identified that the resonant effect of vaccination pulses in an SIR model and derived an optimal vaccination protocol that can speed up extinction when the vaccine is in short supply. In the article [Billings et al.,

2013], the effects of treatment for the *SIS* model was considered and showed an exponential improvement in extinction times.

In the case of the EVD model considered in this chapter, the optimal path is a curve through 12-dimensional space, and is found by solving a 12-dimensional system of differential equations. Knowledge of the optimal path to extinction allows one to perturb our model to determine relative sensitivity to changes in human behavior and to disease control strategies. With these tools one could predict the relative effectiveness of disease control measures. Given appropriate cost functions for the implementation of the different control measures, one could minimize the mean time to extinction given a finite set of resources, thus developing an optimal strategy.

Once disease extinction has been achieved, then the Ebola virus will only be found in non-human populations and an understanding of the conditions for which invasion and outbreak are most likely can help us to identify at-risk populations. It can be seen using stochastic simulation techniques that there are populations where long lived endemic states are very unlikely. This means that the disease may invade consistently, but is always driven down to extinction in a relatively short amount of time. Human behavior as well as the attributes of the disease can affect the vulnerability of a population to outbreak. One way to quantify this risk is by using the basic reproduction number from the associated deterministic system. It has also been observed that in a dynamic population size model, where the birth rate is larger than the death rate, the ability of the disease to invade into a population is dependent on the size of the population.

Figure 13 shows that the basic reproduction number from the deterministic system can

be used as an indicator for outbreak vulnerability in a stochastic invasion model with a

weak connection to the disease reservoir. This fact allows one to extend important conclu-

sions from the investigation of deterministic systems to more realistic stochastic systems.

It has been shown that increasing the burial rate and decreasing the contact rate will, ul-

timately, result in controlling the disease. The frequencies that are reported in Fig. 13

show how the presence of the reservoir impacts the probability of controlling the outbreak

through standard intervention methods (self- or forced isolation and safe removal of in-

fected bodies) while also affording information about which parameter may be more ap-

propriate to modify in order to decrease the impact of the disease. It is worth noting that

the darker coloration in the region of medium contact rates and high burial rates is not a

consequence of the model, but is rather due to the population size and number of stochas-

tic realizations. An increase in either decreases the variance and results in a smoother,

more even coloration.

Since the size of the endemic state scales proportionally with the size of the total pop-

ulation, larger populations tend to be more susceptible to a long lived endemic state than

smaller populations. Consider the number of infectious individuals at the endemic state as

given in 3.2.2. As N increases, I increases proportionately. The deterministic system can

never escape from the stable endemic state, regardless of how large or small the endemic

state is. In the stochastic formulation, however, the endemic state becomes metastable.

This means that the natural variance of the stochastic system can force a transition be-

tween the endemic and extinct states. This favorable transition is less and less likely to

happen as the population size increases. In fact, the mean time to stochastic extinction as

a function of population size is known to be exponential. Therefore, larger populations

are more likely to have a persistent disease after an invasion event than are smaller pop-

ulations. Figure 17 shows that as the population size increases over time, the risk for an

endemic disease state increases as well. This suggests that the risk for disease invasion is

particularly dynamic in developing areas. Population growth is expected to continue for

decades in the developing world. The population of many African countries is predicted

to double by 2050 [Bongaarts, 2013, United Nation Population Division Department of

Economic and Social Affairs, 2015]. Fig. 17 shows that even a population with an en-

demic state small enough so that it will not be realized, but with a growing population,

may overcome a threshold and suddenly display an endemic disease state after a long pe-

riod of population growth. Although it is not explicitly investigated here, it would be rea-

sonable to hypothesize that a similar effect may be seen in an increasingly interconnected

and globalized region. In theory the phenomenon requires population growth, not nec-

essarily from an imbalance between the birth and death rates. In the future similar work

should be performed on a grid of interconnected populations, and in this way both spatial

spread, which is not considered here, and increased interconnectedness among popula-

tions can be investigated.

4   Outbreak Vulnerability and Connectedness in a Stochastic *SISκ* Endemic Disease Model

4.1   Introduction

The mathematical study of disease spread throughout human populations is largely built

upon the foundational assumption that groups of humans can be partitioned into distinct,

well-mixed, and spatially non-explicit populations [Malthus, 1888, Verhulst, 1838, An-

derson et al., 1992, May, 2001]. In this formulation of mathematical disease models a

population can be studied in isolation, which requires that populations are both disjoint

from one another and well-defined. Although these core assumptions are still commonly

used, the validity and usefulness of these assumptions depends on the particular scien-

tific inquiry. For instance, it is useful to study vaccination optimization strategies in both

isolated populations as well as in explicitly connected metapopulation models [Billings

et al., 2013, Burton et al., 2012]. For the study of global disease extinction or the spread

of disease between populations it is necessary to consider multiple separate populations or

sub-populations [Khasin et al., 2012, Arino and van den Driessche, 2006]. The connect-

edness and the disjointness of two such sub-populations, however, is not a well-defined

property.

   In the context of human populations, isolation is the exception, not the rule. For the

sake of infectious disease spread, people geographically near to one another are often a

part of the same population; people that are geographically near to one another and homo-

geneous for the sake of disease modeling will be described as sympatrically connected.

For example, the group of people that live in Paris and use the Metro transit system are

capable of spreading the influenza virus directly between one another. These people are

sympatrically connected since they reside in the same geographical location and they

spread disease directly amongst themselves. If some of those individuals live in Chatou

– a 10km and 15 minute train ride outside of Paris – and they transfer diseases between

Paris and Chatou, then one can say the people of Chatou are allopatrically connected to

the people of Paris. They are geographically separated, but are still capable of spreading

disease to one another.

Although sympatric groups of people are often connected, it need not be the case.

Given a particular disease and a particular research question, two groups of people can co-

exist in the same geographical area and not be a part of the same population. For instance,

if studying the spread of HIV throughout the people of a city, it would be necessary to

separate the intravenous drug users from the general population as the risk of infection to

an intravenous drug user is different from non-drug users. These two groups may be con-

sidered sympatrically disjoint – geographically coexisting, but not spreading the disease

from one another – for the sake of population level HIV modeling. Allopatrically disjoint

populations are the easiest to understand. They are both geographically distant, and un-

likely to spread a disease between them. The people of a small town in Arkansas are not

in much danger from a norovirus outbreak in North Korea; these places are geographi-

cally distant and they do not have appreciable amounts of human migration from one to

the other. These populations can be said to be entirely disjoint for the sake of mathematical modeling.

While the details are disease and population specific, the rarity of isolation means that most human populations are liable to have diseases invade from some external source. It is almost vanishingly rare for any human or animal population on Earth to be truly isolated. Therefore, the question is not "are two given populations connected?", but "with what strength are these two populations connected?".

This question has been tackled in spatially explicit models, where the connection is assumed to be inversely proportional to the square of the distance between the two populations. These models are called gravity models, after the inverse square law of the gravitational force [Viboud et al., 2006, Tuite et al., 2011]. In certain contexts, however, an inverse square law is not fully explanatory. It has even been shown that under vaccination, synchrony of measles outbreaks in sympatric metapopulations has become irregular [Rohani et al., 1999]. The contrapositive is also untrue; allopatric populations may be synchronous. This allopatric connectedness may be caused by population mixing, such as would happen with two cities that were major air travel hubs.

Additionally, we have a very poor understanding of the natural disease reservoir, and so the dynamic modeling of the disease reservoir is not practical [Karesh et al., 2012]. Similarly, it is impractical to consider all external human populations, so a subjective geographical boundary or threshold must be identified. Even if a rigorous mathematical or statistical method was used to identify useful physical boundaries, human interactions

are not entirely restricted by distance. Two places that are relatively close may be weakly

coupled to one another, and two places that are relatively distant may be strongly coupled.

Ease of travel, along with the complex nature of human interactions makes it difficult

to identify when human populations are genuinely distinct. Investigating disease invasion

relies on well-defined populations. A population's vulnerability to disease-outbreak (re-

ferred to here as outbreak vulnerability) is intrinsically tied to how strongly it is coupled

to peripheral disease-endemic populations; this is dependent on the rate at which the peo-

ple of the two populations mix. An example of this is the synchrony of disease outbreak

seen in New York City, New York; Philadelphia, Pennsylvania; Newark, New Jersey; and

Boston, Massachusetts during the early twentieth century. Based on the influenza data

shown in Fig. 18, disease synchrony and coupling strength are dependent on more than

just the distance between cities [van Panhuis et al., 2013].
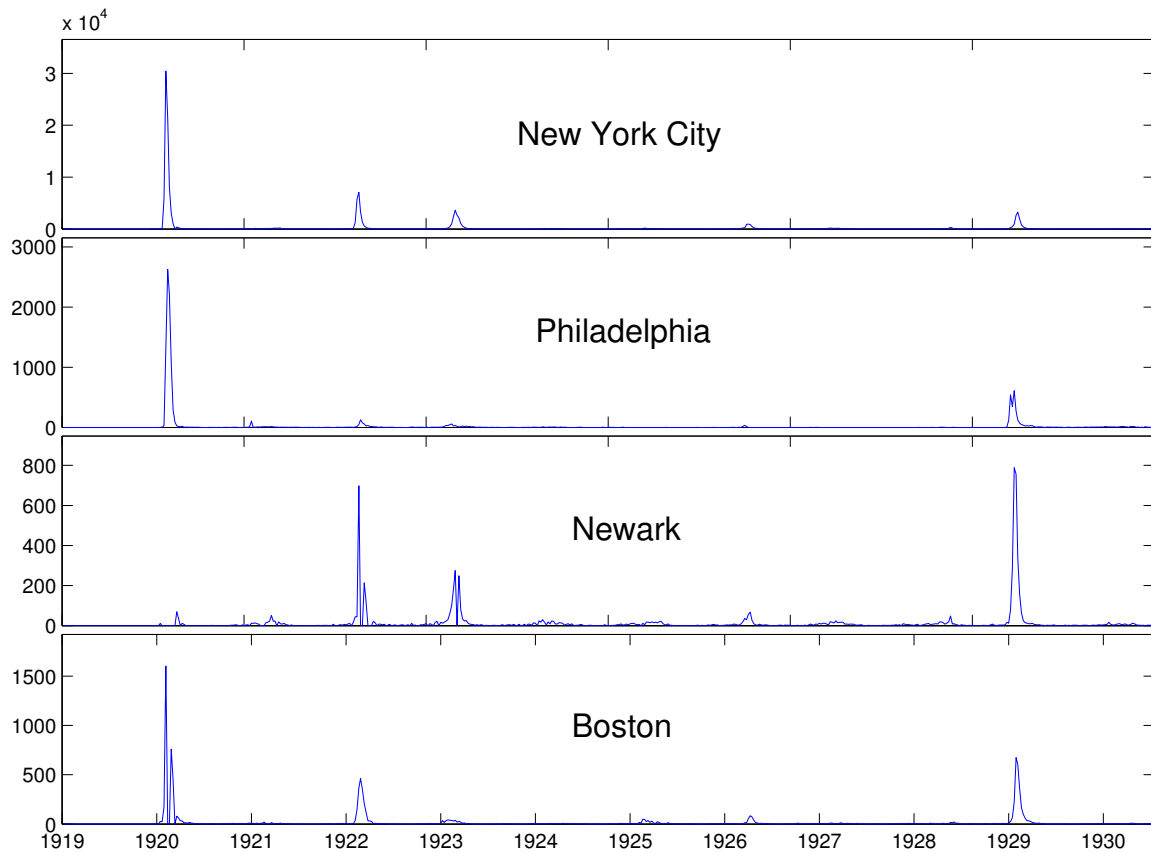
Figure 18: Influenza data for four cities: New York City, Philadelphia, Newark, and Boston. The number of infectious individuals is given by the vertical axis, while the year is provided on the horizontal axis. The correspondence of outbreaks synchronized in time suggests a strong coupling, regardless of the variable geographical distances between the cities. The data was downloaded from the Project Tycho database [van Panhuis et al., 2013].

Disease outbreak and extinction are dependent on the random interactions within and between populations. Disease spread throughout interconnected populations is usually considered from a deterministic perspective. Often the goal is to find an effective basic reproduction number, R_0, for the entire metapopulation [Arino and van den Driessche, 2006]. Other studies have investigated the response of metapopulations to injections of infectious individuals [Rho et al., 2008]. In order to make sense of outbreak and coloniza-

tion, thresholds can be used to denote the "uncolonized" or "disease-free" state [Potapov

and Rajakaruna, 2013, Be'er et al., 2015]. Some version of this idea is always necessary in

the deterministic case, where any degree of coupling to an endemic population will lead

to a guaranteed inflow of infectious individuals to all connected populations. In determin-

istic metapopulation models, extinction is either a global system state or is transient. In

both reality and stochastic models there can be long periods of disease die-out, even when

a population is coupled to disease-present populations.

This chapter presents an investigation of stochastic disease outbreak and connectivity

in a stochastic Susceptible-Infectious-Susceptible population model with connection to

a disease reservoir (*SISκ* model), where the model's basic reproduction number satisfies

$R_0 < 1$. Such a population will be called "non-endemic", in reference to the stability of

the endemic state in the corresponding deterministic non-invasion system; when $R_0 < 1$,

the endemic state of the deterministic system is unstable. The study will consider a sin-

gle population with a generic coupling to an external disease source. In section 4.2 the

stochastic model will be developed in terms of the possible demographic and infection-

related events, and a discrete master equation will be determined. Section 4.3 presents

a discussion of connectivity between an *SISκ* system and an external disease reservoir.

In Section 4.4 two metrics for outbreak vulnerability are discussed: (1) the basic repro-

duction number when a population is weakly coupled to the disease reservoir, and (2) the

normalization factor for the model's probability density function in the general case. In

Section 4.5 there is a discussion of the results in the context of practical disease manage-

ment strategy.

4.2   *SISκ* **Model**

For a given population and disease the external sources of disease exposure can be numerous. These sources, termed the disease reservoir, may be unknown or poorly understood. For instance, in the case of the Ebola virus disease it is unknown if the disease takes refuge in snakes or bats, or possibly some third natural carrier [Leirs et al., 1999, Pourrut et al., 2005]. Such a reservoir is impossible to effectively model dynamically. Instead, disease introduction from an external source can be accounted for using a generic infection event.

In this work the generic infection event will happen at a rate κS, where κ indicates how strongly the system is coupled with the external disease reservoir, and S is the size of the susceptible class, in the population being invaded. The population will be non-endemic with the disease, meaning that the population has an associated deterministic mean-field system with $R_0 < 1$. Figure 19 shows the flow of individuals between the compartments of the *SISκ* population, and Table 6 describes these transitions along with their associated rates and step sizes.

Description	Transition	Rate
Healthy Birth	$(S,I) \to (S+1,I)$	μN
Natural Death	$(S,I) \to (S-1,I)$	μS
Natural Death	$(S,I) \to (S,I-1)$	μI
Disease Transmission	$(S,I) \to (S-1,I+1)$	$\frac{\beta IS}{N} + \kappa S$
Recovery	$(S,I) \to (S+1,I-1)$	γI

Table 6: The model as described by random transitions and their associated rates. Here N is the average population size, I is the size of the infectious compartment, and S is the size of the susceptible compartment. The parameter μ is the birth and death rate, β is the contact rate within the population, γ is the recovery rate, and κ is the strength with which the population is coupled to the external disease reservoir.
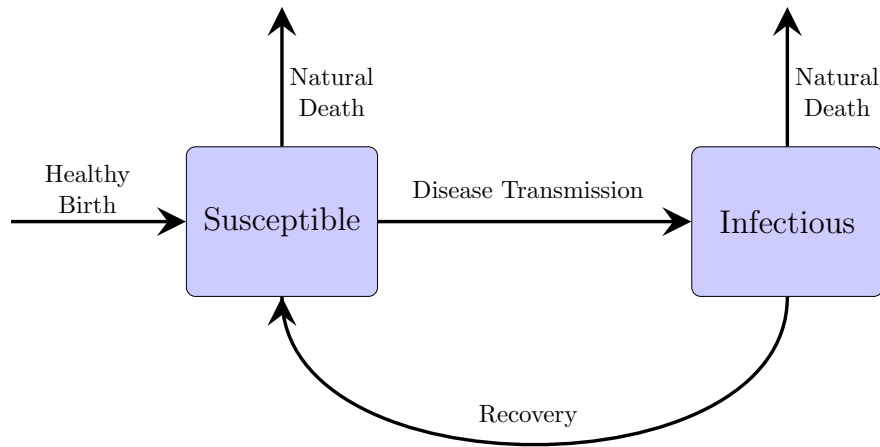
Figure 19: Flow chart showing the movement of individuals within an *SISκ* population.

An equal birth and death rate constrain the model to a stationary average population size N. If N is approximated as a constant with that average value, then the system is over-determined, and the equation $I + S = N$ is used to reduce the two-dimensional state vector of the *SISκ* system to a one-dimensional state variable. We will track the dynamics of the disease entirely through the evolution of the infectious class. Disease transfer along

with demographic events are assumed to be discrete continuous-time Markov processes,

with no autocorrelation in time. The population can thus be modeled using the discrete

stochastic master equations

$$\frac{d\mathbf{P(X,t)}}{dt} = \sum_r [W(\mathbf{X}-\mathbf{r};\mathbf{r})\mathbf{P}(\mathbf{X}-\mathbf{r},t) - W(\mathbf{X};\mathbf{r})\mathbf{P}(\mathbf{X},t)],\tag{83}$$

In Eq. 83, $W(\mathbf{X};\mathbf{r})$ is the transition rate from $\mathbf{X}$ to $\mathbf{X}+\mathbf{r}$, and $\mathbf{P}$ is the probability density

function for the state variable $\mathbf{X}$. In the *SISκ* system the state variable tracks the number

of infectious individuals I, and is real valued. Therefore $\mathbf{P} = [P_0, P_1, P_2, P_3, ..., P_N]$, and

each equation takes the form

$$\frac{dP_I}{dt} = [\beta(I-1)+N\kappa]P_{I-1} - [\beta I + N\kappa]P_I + \left[\frac{\beta(I+1)^2}{N} + (\kappa+\mu+\gamma)(I+1)\right]P_{I+1}.\tag{84}$$

Since there is zero probability of having fewer than zero individuals in the population, the

equation for the lower bound is

$$\frac{dP_0}{dt} = \left[\frac{\beta}{N} + \kappa + \mu + \gamma\right]P_1 - [N\kappa]P_0.\tag{85}$$

A classical *SIS* model possesses bistability in that there is a stable extinct state and a

metastable disease endemic state. However, the introduction of the κS term effectively

destroys the stable extinct state. The *SISκ* system has only a single physically relevant

equilibrium point, given by

$$I^* = \frac{N[\beta - (\mu + \gamma + \kappa)] + N\sqrt{[\beta - (\mu + \gamma + \kappa)]^2 + 4\beta\kappa}}{2\beta} \tag{86}$$

and has a stationary probability distribution. The mean value of the stationary distribution is given by Eq. (86). Setting $\frac{d\mathbf{P}}{dt} = 0$ and using Mathematica, one is able to find an analytical form for the PDF [Abramowitz and Stegun, 1967] as

$$P(I) = \Phi \frac{N^I \Gamma\left(\frac{N\kappa}{\beta} + I\right)}{\Gamma(I+1)\Gamma\left(\frac{N(\mu+\gamma+\kappa)}{\beta} + 1 + I\right)}, \tag{87}$$

where $\Phi = \frac{\pi_0 \Gamma\left(\frac{N(\mu+\gamma+\kappa)}{\beta} + 1\right)}{\Gamma\left(\frac{N\kappa}{\beta}\right)}$, and $(\Omega)_z = \frac{\Gamma(\Omega+z)}{\Gamma(\Omega)}$, with $\Gamma(n) = (n-1)!$ is the Gamma function. In Eq. 87 π_0 is the normalization factor, so that the total probability is exactly one.

4.3   Connectivity of the Reservoir

It is known that disease persistence is dependent on the size of the population. The minimum population size needed to have an endemic disease is called the critical community size (CCS) [Bartlett, 1960]. It has also been observed that the persistence behavior of a disease (as predicted by stochastic mathematical models) scales with a population's birth rate [Conlan and Grenfell, 2007]. The frequency and size of outbreaks vary continuously with changes of $N\kappa$, but the qualitative outbreak behavior can be divided into three classes: (1) random and rare; (2) random and frequent; and (3) the invading population

is perpetually endemic. These behaviors have been observed in both numerical simula-

tion (see Figs. 20 and 21) and in data [Louca et al., 2014, Conlan and Grenfell, 2007].

Although qualitatively distinct from one another, these three characteristic behaviors cor-

respond to indefinite regions in parameter space. The corresponding regions of parameter

space will be distinguished with the names: (1) rare outbreak zone (ROZ); (2) frequent

outbreak zone (FOZ); and (3) perpetually endemic zone (PEZ).

In 'Deterministic and Stochastic Models for Recurrent Epidemics' Bartlett aptly stated

that "This notion of a critical size is difficult to discuss quantitatively..." (pg. 97), before

giving what he described as a crude argument for a threshold [Bartlett, 1956]. In part this

difficulty is due to the indefinite nature of random processes. Invasion and persistence

behavior are closely related, and so it is not entirely surprising that a dependence on pop-

ulation size is seen in the invasion system given by the master equation (84). While CCS

has been shown to relate to birth rate, here it is shown that the outbreak behavior is analo-

gously dependent on the strength with which the system is coupled to a disease's external

reservoir.
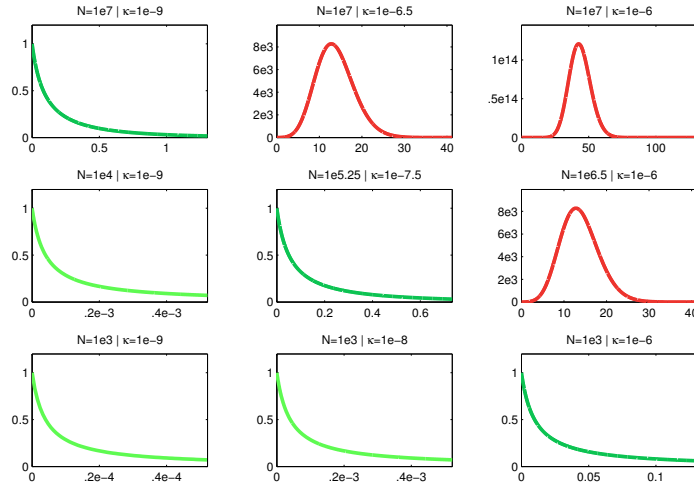
Figure 20: Nine unnormalized probability density functions, corresponding to the parameter sets $\{\mu, \gamma, \beta, \kappa, N\} = \{5 \times 10^{-5}, 0.33, 0.1, \kappa, N\}$. The values for N and κ are noted in the title to each individual PDF. These nine realizations are sampled over several orders of magnitude, from the same range of N and κ as from Fig. 22. The bright green realizations in the lower left come from the rare outbreak zone (ROZ). The three central realizations that are dark green in color show the transition into the frequent outbreak zone (FOZ). The red realizations in the upper right lie in the perpetually endemic zone (PEZ), beyond the transition zone shown in Figs. 22, 23, and 24.

Figures 20 and 21 show nine example PDFs, and characteristic stochastic realizations

for nine different parameter sets. In each of the subfigures in Fig. 20, Eq. 87 was used to

plot the stationary PDF for the *SISκ* system with the parameters as described in the figure

captions. The nine subfigures of Fig. 21 were produced with numerical stochastic Monte

Carlo simulation [Gillespie, 1977] and are representative of the actual disease dynamics

for the corresponding PDFs. The parameter sets were chosen to sample κ and N values

over several orders of magnitude. The smaller population sizes are found at the bottom of

Figs. 20 and 21, while smaller reservoir contact rates are found to the left side of Figs. 20

and 21.
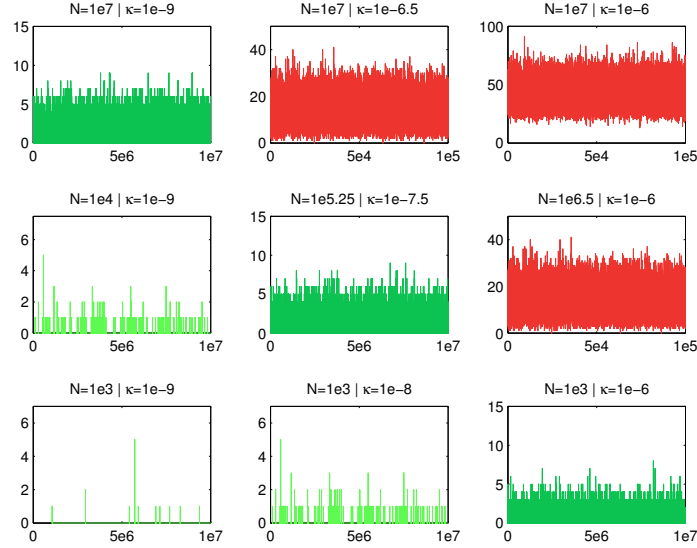
Figure 21: Stochastic realizations with $\{\mu, \gamma, \beta, \kappa, N\} = \{5 \times 10^{-5}, 0.33, 0.1, \kappa, N\}$. The values for N and κ are noted in the title to each individual realization. These nine realizations are sampled from the same range of N and κ as from Fig. 22. The bright green realizations in the lower left come from the rare outbreak zone (ROZ). The three central realizations that are dark green in color show the transition into the frequent outbreak zone (FOZ). The red realizations in the upper right lie in the perpetually endemic zone (PEZ), beyond the transition zone shown in Figs. 22, 23, and 24.

Given the exact PDF, the probability of the corresponding stochastic population being found within the interval of *I*-values (I_a, I_b) is given by

$$P_{ab} = \frac{\int_{I_a}^{I_b} \hat{P}(x)dx}{\int_0^\infty \hat{P}(x)dx},$$ (88)

where $\hat{P}$ is an un-normalized PDF. Notice that the same expression holds for a normalized PDF, but the denominator would reduce to unity.

Based on Eq. (88), it can be seen that for the three light green PDFs in the lower left of Fig. 20 that there is a very high probability – nearly a probability of one – that the

system is found at the disease extinction state at any given time. As should be expected,

the corresponding PDFs decay very quickly in I. These three parameter sets are sampled

from the ROZ. The three dark green PDFs decay, but at a much slower rate than the light

green PDFs. This results in a much lower probability of being found at the disease ex-

tinction state. The parameters for the dark green PDFs are sampled from the FOZ. The

three red PDFs in the upper right of Fig. 20 are no longer monotonically decreasing, and

therefore the probability of being found in the immediate vicinity of $I = 0$ is smaller than

for the decaying PDFs. The parameters used for the red PDFs are sampled from the PEZ.

Notice that the representative examples in Figs. 20 and 21 have corresponding parameter

sets, and display the expected behaviors of rare outbreak, frequent outbreak, and perpetu-

ally endemic.

The transition in the shape of the PDF is indicative of transition in the behavior of the

system. The subfigures in the lower left of Fig. 20 show quickly decaying PDFs and show

the expected rare outbreaks in Fig. 21. The subfigures up higher in the figure and more to-

wards the right correspond to larger N and κ values; in fact the important quantity seems

to be the product $N\kappa$, since outbreak behavior is found to be quite similar among param-

eter sets with consistent $N\kappa$-values. As $N\kappa$ transitions from the minimum value up to its

maximum value the PDF transitions from a decay rate that is much faster than exponen-

tial, to an exponential decay, then transitions to a PDF that is flat at $I = 0$, followed
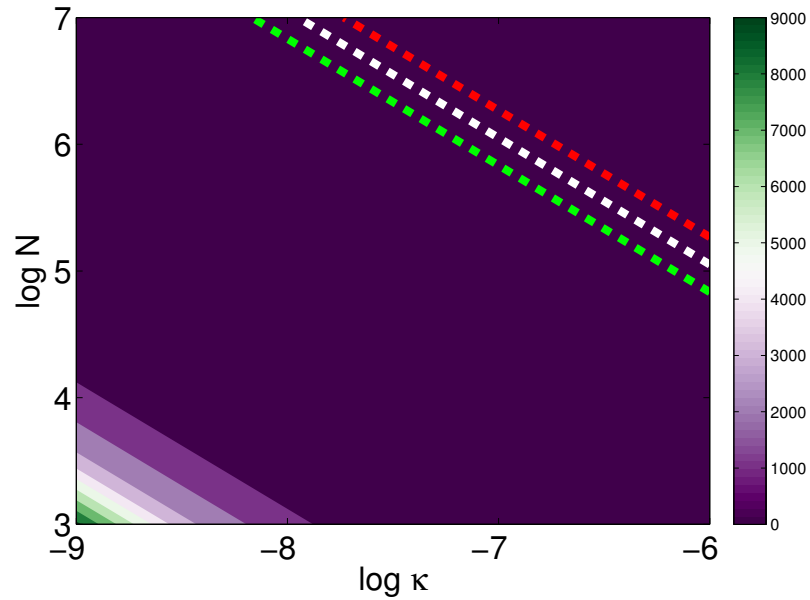
Figure 22: Contour plot showing the numerically calculated normalization factor π_0 for the parameter set $\{\mu, \gamma, \beta, \kappa, N\} = \{5 \times 10^{-5}, 0.33, 0.1, \kappa, N\}$. The red dashed line is given by Eq. (95), the green dashed line is given by Eq. (94), and the white dashed line is the arithmetic mean of the two.

by a PDF that becomes strongly peaked away from the $I = 0$ axis. The PDFs that are

peaked away from $I = 0$ are created by parameter sets in the PEZ. The PEZ is also the

region of non-validity for this kind of invasion model. A core assumption of the model

is that the disease is randomly but consistently being introduced into the system from an

inexhaustible external source. If the connection to that source is also very strong, then the

population becomes indistinguishable from the disease reservoir.
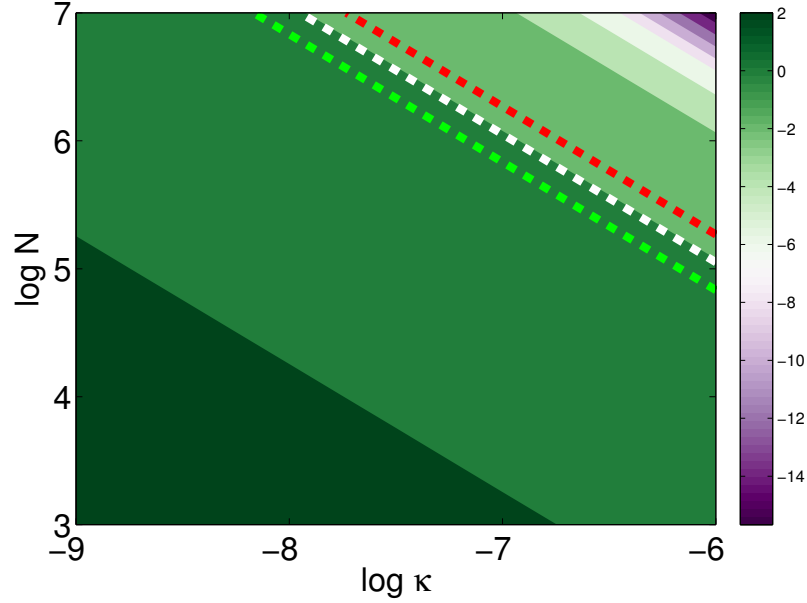
Figure 23: Contour plot showing the logarithm of the numerically calculated normalization factor π_0 for $\{\mu, \gamma, \beta, \kappa, N\} = \{5 \times 10^{-5}, 0.33, 0.1, \kappa, N\}$. The red dashed line is given by Eq. (95), the green dashed line is given by Eq. (94), and the white dashed line is the arithmetic mean of the two. Notice that the white line is a good approximation for the barrier between $\pi_0 > 1$ and $\pi_0 < 1$, over a change of several orders of magnitude in both κ and N.

By definition, an area of one will be enclosed by the normalized PDF

$$P_{\text{tot}} = \int_0^\infty P(z)dz = \pi_0 \int_0^\infty \hat{P}dz = 1, \tag{89}$$

where $P(I)$ is the normalized PDF, and $\hat{P}(I)$ is the un-normalized PDF.

Solving for π_0 leads to

$$\pi_0 = \left[\int_0^\infty \hat{P}(z)dz \right]^{-1}. \tag{90}$$

Figures 22 and 23 show the value of the normalization constant for 4800 different pa-
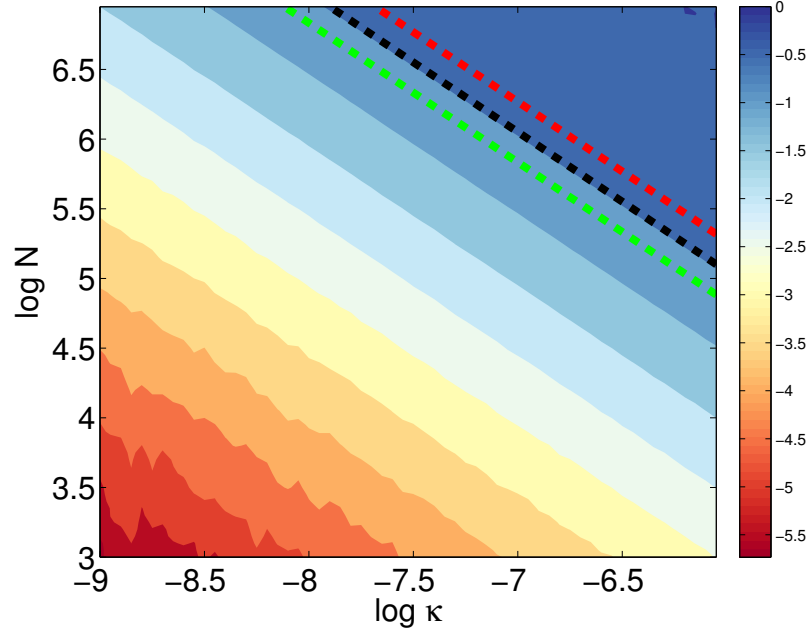
Figure 24: Contour plot showing the logarithm of the proportion of time spent with disease present using an ensemble of 4800 stochastic realizations. Each realization was allowed to progress until it had tracked the size of the infectious class for 10^7 days. The parameters used for these realizations are $\{\mu, \gamma, \beta, \kappa, N\} = \{5 \times 10^{-5}, 0.33, 0.1, \kappa, N\}$. The red dashed line is given by Eq. (95), the green dashed line is given by Eq. (94), and the black dashed line is the arithmetic mean of the two (color changed from white in Figs. 22 and 23 to improve visibility). Notice that the black line is a good approximation for the disease PEZ region border, and is shown over a change of several orders of magnitude in both κ and N.

rameter sets over several orders of magnitude in both N and κ. In each case the value of

π_0 was found numerically using the trapezoidal integration scheme to compute the integral given in Eq. (90) with the PDF defined in Eq. (87). Figures 22 and 23 show that the

normalization constant gets smaller as $N\kappa$ gets larger. This means that the area under the

un-normalized PDF gets larger as $N\kappa$ gets larger.

Figure 24 shows the logarithm of the proportion of time spent with the disease for

the same 4800 parameter sets used in Figs 22 and 23. For each parameter set a stochastic

Monte Carlo simulation was run for ten million simulation-days, and the time periods during which the population had a non-zero number of infectious individuals was measured and summed. This number was divided by the total time and the logarithm was taken. In the upper right of Fig. 24, where $N\kappa$ is largest, the population spends approximately all of its time with the disease, hence our use of the name perpetual endemic zone or PEZ. The barrier between the PEZ and the FOZ corresponds to the curve in Figs. 22 and 23 where π_0 is approximately one.

The boundary shown in Figs. 22, 23, and 24 gives one a good indication of what population size and coupling strength combinations result in a fully connected system. In the cases with $N\kappa$ combinations larger than the boundary the system of study has become indistinguishable from the disease reservoir, and so, the population of interest would then be considered entirely connected to the disease reservoir. To investigate interesting dynamics, one would need to re-define the system to include and dynamically model the disease reservoir. So, if $N\kappa$ is above the curve – white in Figs. 22 and 23 and black in Fig. 24 – then the system of interest is not a separate population suffering from disease invasion, but is only a fully connected subset of a larger population.

As discussed above, this boundary is coincident to where the PDF transitions away from a decaying form. In order to find a closed form equation that approximates this boundary, the first derivative will be used to describe the slope of the PDF at $I = 0$. The

first derivative of the PDF is given by

$$\frac{dP}{dI} = \frac{\Phi N^I \left[\ln(N) + \psi\left(\frac{N\kappa}{\beta} + I\right) - \psi(I+1) - \psi\left(\frac{N(\mu+\gamma+\kappa)}{\beta} + 1 + I\right) \right] \Gamma\left(\frac{N\kappa}{\beta} + I\right)}{\Gamma(I+1)\Gamma\left(\frac{N(\mu+\gamma+\kappa)}{\beta} + 1 + I\right)}, \quad (91)$$

where $\psi(z) = \Gamma'(z)/\Gamma(z)$ is the Digamma function. When I=0 this simplifies to

$$\frac{dP}{dI} = \pi_0 \left[\ln(N) + \psi\left(\frac{N\kappa}{\beta}\right) - \psi(1) - \psi\left(\frac{N(\mu+\gamma+\kappa)}{\beta} + 1\right) \right]. \quad (92)$$

In Eq. (92) $\psi(1) = -\hat{\gamma}$, where $\hat{\gamma} \approx 0.5772$ is the Euler-Mascheroni constant and the hat is

used to differentiate the Euler-Mascheroni constant from the disease recovery rate γ.

Using the first term of the asymptotic expansion of the Digamma function

[Abramowitz and Stegun, 1967]

$$\psi(z) = \ln(z) - \frac{1}{2z} - \frac{1}{12z^2} + \frac{1}{120z^4} + \cdots, \quad (93)$$

allows one to find a closed form approximation for the first derivative of the PDF. The

transition to the PEZ happens over a relatively small region of parameter space, but the

shape of the PDFs at $I = 0$ has been used to identify the parameter range over which the

transition occurs. The lower bound is the point at which the PDFs rate of change at $I = 0$

is approximately exponential; this is equivalent to setting Eq. (92) equal to π_0. Any point

beyond that curve in parameter space will have a PDF with weaker than exponential de-

cay at $I = 0$. The outer boundary is where the PDF is flat at $I = 0$; this is equivalent to

setting equation (92) equal to zero. These two equations can be arranged to implicitly de-

fine $N\kappa$ as

$$N\kappa = \frac{\mu + \gamma + \beta/N}{\exp(1 + \hat{\gamma}) - 1/N}, \qquad \text{and} \tag{94}$$

$$N\kappa = \frac{\mu + \gamma + \beta/N}{\exp(\hat{\gamma}) - 1/N}, \tag{95}$$

respectively. The arithmetic mean of these two equations is shown to be a good approxi-

mation for the precise boundary in Figs. 23 and 24.

4.4   Outbreak Vulnerability

In Section 4.3 we used a closed form approximation to the system in Eq. 84 to determine

a boundary in parameter space between outbreak dynamics and the PEZ. A region that

covers several orders of magnitude in both N and κ can now be identified as being prone

to disease invasion, although quantifying the outbreak vulnerability throughout this region

is non-trivial. It would be natural to look towards the deterministic basic reproduction

number R_0, because of its relevance to infectivity of a disease as described in determinis-

tic models.

   The formal definition of R_0 is the number of secondary infections that are caused by

a single infectious individual in an otherwise fully susceptible population. If R_0 is greater

than one, then the disease will spread, but if R_0 is less than one, then the disease will die

out. In a deterministic disease model without a source of external disease introduction, R_0

describes the stability of the disease-free equilibrium (DFE). For $R_0 > 1$ the DFE is unstable, while for $R_0 < 1$ the DFE is stable. A stable DFE corresponds to deterministic disease extinction and is used as a metric for disease infectivity; diseases with large R_0 values are considered very infective, while diseases with relatively small R_0 values are considered to be less infective. Since connection to a disease reservoir removes the DFE from both the deterministic and stochastic systems, there is no guarantee that R_0 has the same relevance. It has been shown in Chapter 3 that under the presence of weak connection to a disease reservoir, the R_0 of the related deterministic mean field model can be used as an indicator of outbreak vulnerability for a stochastic model. Figure 25 shows the base ten logarithm of the time an *SISκ* population spends with disease present over a one million simulation-day time span. The black dashed line is $R_0 = 1$, where R_0 is the basic reproduction number $R_0 = \frac{\beta}{\mu + \gamma}$ for the deterministic *SIS* model (*SISκ* with $\kappa = 0$). Parameters were chosen so that $\frac{1}{150} \leq R_0 \leq 15$. For any horizontal line on Fig. 25 the value of γ is fixed, and R_0 changes proportionately to β. For any vertical line, β remains fixed, and R_0 changes inversely with γ. It can be seen that the time spent disease-present has a general trend of increasing with increases of R_0, and has a drastic increase that happens over a very small region of parameter space, and that this drastic change roughly corresponds to the transition between $R_0 < 1$ and $R_0 > 1$.
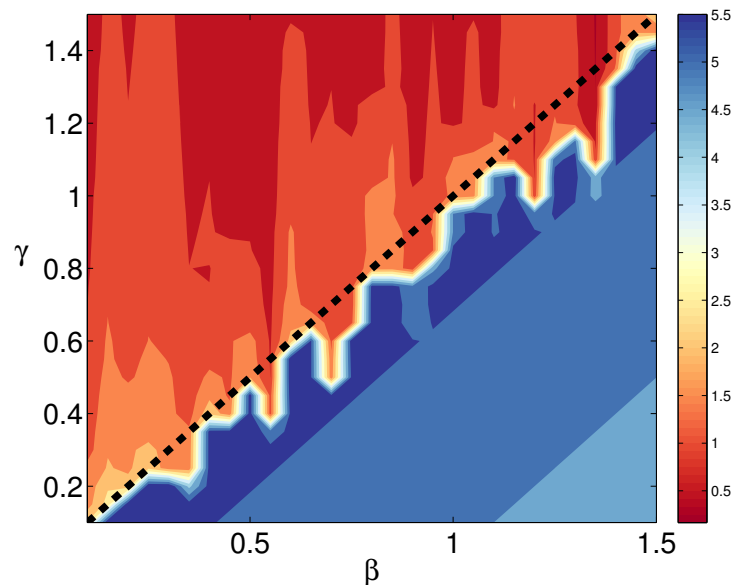
Figure 25: Contour plot showing the base ten logarithm of the time spent disease-present in an ensemble of 900 stochastic realizations. Each progressed until it had tracked the size of the infectious class for 10^6 days. The parameters used for these realizations are $\{\mu, \gamma, \beta, \kappa, N\} = \{5 \times 10^{-5}, \gamma, \beta, 1 \times 10^{-9}, 10000\}$, where γ and β are indicated by the axes. The black dashed line is given by $R_0 = 1$.

In models of dimension larger than one, with $R_0 > 1$ and $\kappa > 0$, it is possible that outbreaks occur without the endemic state being realized. This results in consistent outbreaks, in some cases devastatingly large. In a one dimensional disease model that includes connection to a disease reservoir, when $R_0 > 1$ and $\kappa > 0$, eventual disease presence and persistence is guaranteed. Notice that in Fig. 25 the total time spent with disease present is always less than 40%, which seems to contradict a claim that the disease will be perpetually present. The stochastic simulations that were used to make Fig. 25 had an initial state of $I = 0$. Based on the small $N\kappa$, the initial disease invasion takes quite a long time to occur. Since the perpetual state can not be realized before disease introduction, there will be a substantial amount of time at the beginning of the simulation for which

$I = 0$.

When R_0 is less than one, however, intermittent outbreaks will happen with frequency

dependent on both the population size (N) and reservoir-connection strength (κ). A use-

ful quantitative metric for outbreak vulnerability should be related to the probability of

disease presence. It was shown in Sec. 4.3 that there is a relation between the normaliza-

tion constant π_0 and the area under the PDF. It is also shown that there is a relationship

between the area under the PDF (about the point $I = 0$) and the frequency and severity of

disease outbreak.

Numerical realizations of the one-dimensional *SISκ* system have been computed for

parameter sets that span several orders of magnitude in both κ and N. Results from these

numerical solutions are shown in Fig. 22 and 23. Comparing Fig. 22 to Fig. 21 it can be

seen that π_0 appropriately predicts the transition throughout the ROZ and into the FOZ.

Throughout the ROZ (lower left of Fig. 22) π_0 shows a maximal sensitivity to changes

in N and κ. Although outbreak vulnerability seems to vary continuously as a function of

$N\kappa$, frequent outbreak behavior shown as dark green in Fig. 21 seems to be coincident

with the region between the area of maximal sensitivity and the FEZ. Therefore the value

of π_0 can be used as a quantitative metric for outbreak vulnerability, and is especially use-

ful for describing outbreak vulnerability in the region of maximal sensitivity, and can be

used to identify the barrier between the ROZ and the FOZ.

The normalization factor π_0 acts as a good indicator for probability of disease extinc-

tion because the proportion of the PDF-area located above an interval $[I = 0, I = b)$ to

the total PDF's area is exactly the probability of being found in the same interval about

$I = 0$ (see Eq. (88)). Given the form of the PDF in Eq. (87) it can be seen that $\pi_0 = P(0)$.

Therefore it is shown that the value of the unnormalized PDF at $I = 0$ is a quantitative

metric that can indicate the qualitative behavior of the system over a variety of parame-

ter sets. Large values of π_0 relate to a high probability of the disease being extinct, while

small values of π_0 relate to relatively low probability of extinction.

4.5   Conclusions

In this chapter a master equation approach was used to study the random Markovian pro-

cess of disease invasion into a non-endemic *SIS* population (termed an *SISκ* system af-

ter the introduction rate κ). In place of dynamically modeling the disease reservoir in

the common style of metapopulation studies, a generic invasion term was included. This

generic invasion term controls how strongly the population of study is coupled to an inex-

haustible external disease reservoir. In this chapter we use the stationary PDF of the *SISκ*

system to identify the boundary between intermittent outbreaks and the disease PEZ. The

intermittent zone includes the rare outbreak zone (ROZ) and the frequent outbreak zone

(FOZ). An analytical form for the normalization constant π_0 was found, and was numeri-

cally calculated for a sampling of parameter sets over several orders of magnitude in both

the population size N and the strength of the connection to the external disease reservoir

κ. The normalization constant π_0 is shown to be a good indicator for the outbreak behav-

ior of the system, particularly useful for identifying the borders between the ROZ, the

FOZ, and the PEZ.

Although disease outbreak and invasion are fundamentally random processes, most of the research being done on metapopulations is done from a deterministic perspective. In these studies there is a tendency towards making generalizations for metapopulations with an arbitrary number of subpopulations. Depending on the nature of the scientific inquiry, this approach can be useful, but it overlooks the issue of practical population segmentation that is necessary for the application of metapopulation results. Additionally, in practice it can be unnecessary or impossible to dynamically model all subpopulations of a metapopulation, making it useful to consider an approach that allows for isolation of a single subpopulation. Given that public policy is largely enacted at the local level, it may be superfluous to investigate the interplay between the decisions made by two or more cities or towns. The approach presented here takes the assumption that there will be a disease reservoir external to the system of interest, and that there is some rate of introduction from that external source. This might be the most practical view for a policy-maker, rather than trying to understand the intricacies of the larger population network.

It is reasonable to ask, how can one quantitatively determine the appropriate coupling strength? Although that question is not answered directly here, there is some literature cited here that suggests that geographical distance is an important factor, although is not entirely explanatory. There is no doubt that certain communities and groups have sympatric (geographically close) disease-connectivity. It is also clear that, given a particular population and disease, certain subpopulations can be sympatrically disjoint. For instance,

in the case of a sexually transmitted disease being spread among sex workers. While sex

workers would be counted as members of the general population in an influenza model,

it is not necessarily appropriate for the modeling of a sexually transmitted disease, which

will mostly be confined to the sex workers and their patrons. Primarily the individuals

that participate in that market will be sympatrically disjoint from those that do not; geo-

graphically close, but practically uncoupled for the sake of that disease.

It seems intuitive that public transportation between two geographic locations can

contribute to the strength of connectivity, and more than likely socioeconomic factors can

come into play as well. A gradient of economic opportunity will drive increased mixing

with external populations. For instance, some people will travel long distances to work in

New York City (NYC), because of a differential in wages between NYC and their home

town. Quantifying coupling strength is an interesting and complex topic. To address such

a difficult question I would expect that mathematical modeling techniques will be most

effective in coordination with data driven studies. The findings and framework here can

be used to help inform a study of that sort. Of particular interest to us would be a data

driven study of the factors that drive allopatric (non-geographical) disease-connectivity.

5   Conclusion

This dissertation is comprised of the analytical and numerical investigation of three non-linear stochastic population models. Chapter 2 quantifies the contribution of non-zero steady-state cycling to the longevity of a population with multiple metastable states, using stochastic theory and a series of optimal paths; the third chapter determines the optimal path to disease extinction in a stochastic Ebola virus disease (EVD) model, and additionally shows that the R_0 can be used to indicate outbreak vulnerability when the EVD-population is weakly coupled to the disease reservoir; the fourth chapter generalizes the concept of outbreak vulnerability beyond weakly coupled systems, and identifies a practical barrier in parameter space that divides the region into 'weak/strong coupling' and a 'fully connected zone' for which the two populations in question can not be said to be separate subpopulations.

The systems presented in this dissertation all have complex and nonlinear interactions that result in interesting and difficult to predict dynamics. Each system was constructed or modeled as simply as possible while preserving those interesting and important dynamics. This is done in order to understand the conditions for which catastrophic or important events (such as outbreak and extinction) are most likely to happen fortuitously, so that they can be predicted and effective control methods can be implemented to either divert or encourage those events. Each system was chosen because it displays hard to predict and potentially devastating behavior that happens in the presence of complex internal and external interactions (sometimes explicitly and sometimes implicitly). As dictated by the

nature of these hard to predict events, it has been necessary to use methods of analysis from both deterministic and stochastic dynamical systems. This is consistent with the long history of the intermingling of deterministic and stochastic methods being applied to problems in ecology and epidemiology. The catastrophic or important events are, however, often outside the purview of deterministic dynamical systems. Being catastrophic or otherwise important, they are of particularly high interest, and so stochastic modeling methods were applied.

5.1   Policy and Environmental Implications

In each chapter there is discussion of how to use the findings and methodology presented to help control the system of study. In Chapter 2 a stochastic population capable of two non-zero persistent states is presented. This behavior is seen in disease dynamics when a viral infection goes entirely into the lysogenic cycle and persists at much lower density than during an active infection. In an ecological setting it is sometimes a result of refugia, when a portion of a population will – through strategy or happenstance – live through a catastrophic die-off event. In that chapter a population control was applied, and extinction or eradication was sped up through the use of a culling mechanism. Thus a framework for measuring the efficacy of a population control method prior to implementation is presented. That population was considered to be a pest, and so quick extinction was desirable; hence the culling term. In practice the control method need not be used to speed up extinction, but some other strategy can be employed to delay extinction. Integration of a

cost function and a budget will allow for the comparison of different possible strategies in terms of expense and efficacy.

Chapters 3 and 4 are both concerned with the connection of populations for the sake of disease spread, and both indicate methods to help understand and combat disease outbreak into a vulnerable population. In each case a metric for outbreak vulnerability is described.

Zoonotic diseases such as EVD pose a continuing threat to human health and human life. In Chapter 3 a six-compartment stochastic EVD model was considered, and a twelve-dimensional stochastic optimal path to disease extinction was calculated. The optimal path can be used to determine mean time to disease extinction, and can therefore be used to assess the effectiveness of disease control strategies. Quantifying the effectiveness of a disease control strategy allows for an informed policy decision. Additionally, the deterministic R_0 was found analytically, and it was shown to be a good indicator for outbreak vulnerability in the EVD system while it was weakly coupled to the zoonotic disease reservoir. This reveals an intuitive result; that increasing safe burial and decreasing contact among population members will result in a lower likelihood of disease outbreak. These two factors, however, are not of equal importance for every population. Depending on current burial and contact rates, it may be more useful to affect one rather than the other.

In Chapter 4 a more basic and more general $SIS\kappa$ model is presented. While only weak coupling was considered in the EVD model, the $SIS\kappa$ model is analyzed for a range

of κ values that spans several orders of magnitude. The normalization constant from the associated probability density function was shown to be a good indicator of outbreak vulnerability. The disease reservoir is assumed to be inexhaustible and non-dynamic. It is shown that the larger $N\kappa$ is, the more infections occur from the external reservoir. For sufficiently large $N\kappa$ (and therefore sufficiently small π_0) the system can be overwhelmed by these infections. The PDF of the $SIS\kappa$ system was used to identify a boundary in parameter space that separates the true outbreak dynamics from the PEZ. If coupling is strong enough so that a population is found to be perpetually endemic with the disease because of an inexhaustible external source, then the model was poorly defined, and a new system that includes that disease reservoir must be considered. The PDF can be used to identify when a system will be overwhelmed. This finding can be used in coordination with data to design practical metapopulation models with appropriately defined subpopulations and connectivity.

In each case the findings presented here can be used to asses the need of preventative measures, the efficacy of control protocols, and can be used to devise effective control strategies.

6   Future Works

As has been discussed at several points in this dissertation, increased interconnectedness has been the trend for all of human history due to better communication, faster, cheaper, and safer transportation, and increased globalization with less tribalism. Although the topic of subpopulation connectivity is discussed here, there is no simple algorithmic way to define a disease network. Here the reservoir contact rate κ is discussed, but how to quantify the reservoir contact rate based on real world data is unclear. In order to build practical tools for disease outbreak prevention, it is necessary to determine what qualifies as a metapopulation or a subpopulation.

There is a large body of research that focuses on synchrony, and there is an opportunity to use that work to better quantify the frequency and strength with which a population interacts with disease reservoirs. Integration of disease incidence data, public transportation data, and census data into calculation of the external-contact term will help to develop population level disease treatment protocols. Additional information about social structure can be used to help define subpopulations and determine cross-contact rates. For instance, economic disparity can lead to the importation of workers, which can cause an allopatric connection between two geographically distant groups of people. On the other hand, there are subpopulations that are sympatrically disjoint, such as intravenous drug users that pass diseases to one another at a much higher rate than to those that do not use drugs. Although there are certain aspects that will always remain contextual, it is possible to increase our understanding through data-driven studies, and produce less vague results

that are still relatively general.

Finally, the studies should be formally generalized to ecological models, and should be supported with either collected field data or laboratory data. Disease outbreak is a fairly direct analogy to invasive species except that the factors favorable for disease outbreak are large population sizes and high contact rates, while the factors that permit an invasive species to thrive are more difficult to quantify. Once the factors that are important to invasion are well understood, control and eradication plans can be developed.

References

Abramowitz, M. and Stegun, I. A. (1967). Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables. Cambridge University Press.

Allee, W. C. (1931). Animal Aggregations, a Study in General Sociology. Univ. Chicago Press.

Althaus, C. L. (2014). Estimating the reproduction number of Ebola virus (EBOV) during the 2014 outbreak in West Africa. PLoS Currents, 6.

Anderson, R. M. and May, R. M. (1991). Infectious Diseases of Humans. Oxford University Press.

Anderson, R. M., May, R. M., and Anderson, B. (1992). Infectious Diseases of Humans: Dynamics and Control, volume 28. Wiley Online Library.

Arino, J. and van den Driessche, P. (2006). Disease spread in metapopulations. Nonlinear Dynamics and Evolution Equations, 48:1–13.

Assaf, M. and Meerson, B. (2010). Extinction of metastable stochastic populations. Physical Review E, 81:021116.

Baltensweiler, W. (1991). The folivore guild on larch (larix decidua) in the alps, y. baranchikov, w. mattson, f. hain, t. payne (eds.). forest insect guilds: patterns of interaction with host trees. USDA Forest Service. General Technical Report NE, 153:145–164.

Bartlett, M. (1949). Some evolutionary stochastic processes. Journal of the Royal Statistical Society. Series B (Methodological), 11(2):211–229.

Bartlett, M. (1956). Deterministic and stochastic models for recurrent epidemics. In Proceedings of the third Berkeley symposium on mathematical statistics and probability, volume 4, page 109.

Bartlett, M. (1960). The critical community size for measles in the United States. Journal of the Royal Statistical Society. Series A (General), pages 37–44.

Bauver, M., Forgoston, E., and Billings, L. (2016). Computing the optimal path in stochastic dynamical systems. Chaos, 26(8):083101.

Be'er, S., Assaf, M., and Meerson, B. (2015). Colonization of a territory by a stochastic population under a strong allee effect and a low immigration pressure. Physical Review E, 91(6):062126.

Berg, S., Christianou, M., Jonsson, T., and Ebenman, B. (2011). Using sensitivity analysis to identify keystone species and keystone links in size-based food webs. Oikos, 120(4):510–519.

Berndt, A., Yizhar, O., Gunaydin, L. A., Hegemann, P., and Deisseroth, K. (2009). Bistable neural state switches. Nature Neuroscience, 12(2):229–234.

Bernoulli, D. and Blower, S. (2004). An attempt at a new analysis of the mortality caused

by smallpox and of the advantages of inoculation to prevent it. Reviews in Medical Virology, 14(5):275–288.

Berryman, A. A. (1996). What causes population cycles of forest lepidoptera? Trends in Ecology & Evolution, 11(1):28–32.

Bialek, W. (2017). Perspectives on theory at the interface of physics and biology. Reports on Progress in Physics, 81(1):012601.

Billings, L., Bollt, E. M., and Schwartz, I. B. (2002). Phase-space transport of stochastic chaos in population dynamics of virus spread. Physical Review Letters, 88:234101.

Billings, L., Fiorillo, A., and Schwartz, I. B. (2008a). Vaccinations in disease models with antibody-dependent enhancement. Mathematical Biosciences, 211(2):265–281.

Billings, L., Mier-y Teran-Romero, L., Lindley, B., and Schwartz, I. B. (2013). Intervention-based stochastic disease eradication. PloS ONE, 8(8):e70211.

Billings, L., Schwartz, I. B., and Shaw, L. B. (2008b). The dynamics of antibody dependent enhancement in multi-strain diseases with vaccination. In Frontiers Of Applied And Computational Mathematics: Dedicated to Daljit Singh Ahluwalia on His 75th Birthday, pages 56–63. World Scientific.

Bongaarts, J. (2013). Human population growth and the demographic transition. Philosophical Transactions of the Royal Society of London B: Biological Sciences, 75:2985–2990.

Burton, J., Billings, L., Cummings, D. A., and Schwartz, I. B. (2012). Disease persistence in epidemiological models: the interplay between vaccination and migration. Mathematical Biosciences, 239(1):91–96.

Caswell, H. (2001). Matrix Population Models. Wiley Online Library.

Centers for Disease Control and Prevention (2016). Outbreaks chronology: Ebola Virus Disease.

Chowell, G., Hengartner, W., Castillo-Chavez, C., Fenimore, P. W., and Hyman, J. M. (2004). The basic reproductive number of Ebola and the effects of public health measures: the cases of Congo and Uganda. Journal of Theoretical Biology, 229:119–126.

Conlan, A. J. and Grenfell, B. T. (2007). Seasonality and the persistence and invasion of measles. Proceedings of the Royal Society of London B: Biological Sciences, 274(1614):1133–1141.

Coulson, T., Rohani, P., and Pascual, M. (2004). Skeletons, noise and population growth: the end of an old debate? Trends in Ecology and Evolution, 19(7):359–364.

Doering, C. R., Sargsyan, K. V., and Sander, L. M. (2005). Extinction times for birth-death processes: Exact results, continuum asymptotics, and the failure of the fokker–planck approximation. Multiscale Modeling & Simulation, 3(2):283–299.

D'Silva, J. P. and Eisenberg, M. C. (2015). Modeling spatial transmission of Ebola in West Africa. arXiv preprint arXiv:1507.08367.

Dykman, M. I., McClintock, P. V., Smelyanski, V., Stein, N., and Stocks, N. (1992). Optimal paths and the prehistory problem for large fluctuations in noise-driven systems. Physical Review Letters, 68(18):2718.

Dykman, M. I., Mori, E., Ross, J., and Hunt, P. M. (1994). Large fluctuations and optimal paths in chemical-kinetics. Journal of Chemical Physics, 100(8):5735–5750.

Dykman, M. I., Schwartz, I. B., and Landsman, A. S. (2008). Disease extinction in the presence of random vaccination. Physical Review Letters, 101(7):078101.

Elf, J. and Ehrenberg, M. (2004). Spontaneous separation of bi-stable biochemical systems into spatial domains of opposite phases. Systems Biology, 1(2):230–236.

Elgart, V. and Kamenev, A. (2004). Rare event statistics in reaction-diffusion systems. Physical Review E, 70:041106.

Feller, W. (1939). Die grundlagen der volterraschen theorie des kampfes ums dasein in wahrscheinlichkeitstheoretischer behandlung. Acta Biotheoretica, 5(1):11–40.

Forgoston, E., Bianco, S., Shaw, L. B., and Schwartz, I. B. (2011). Maximal sensitive dependence and the optimal path to epidemic extinction. Bulletin of Mathematical Biology, 73:495–514.

Forgoston, E., Billings, L., and Schwartz, I. B. (2009). Accurate noise projection for reduced stochastic epidemic models. Chaos, 19:043110.

Forgoston, E. and Schwartz, I. B. (2013). Predicting unobserved exposures from seasonal epidemic data. Bulletin of Mathematical Biology, 75(9):1450–1471.

Freidlin, M. I. and Wentzell, A. D. (1984). Random Perturbations of Dynamical Systems. Springer-Verlag.

Gang, H. (1987). Stationary solution of master equations in the large-system-size limit. Physical Review A, 36(12):5782–5790.

Gardiner, C. W. (2004). Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences. Springer-Verlag.

Gaveau, B. and Schulman, L. (1996). Master equation based formulation of nonequilibrium statistical mechanics. Journal of Mathematical Physics, 37(8):3897–3932.

Gillespie, D. (1976). A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. Journal of Computational Physics, 22(4):403–434.

Gillespie, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. The Journal of Physical Chemistry, 81(25):2340–2361.

Graham, R. and Tél, T. (1984). Existence of a potential for dissipative dynamical systems. Physical Review Letters, 52(1):9–12.

Graunt, J. (1663). Natural and Political Observations Made Upon the Bills of Mortality. Number 2. Royal Society.

Hu, K., Bianco, S., Edlund, S., and Kaufman, J. (2015). Social Computing, Behavioral-Cultural Modeling, and Prediction, chapter The Impact of Human Behavioral Changes in 2014 West Africa Ebola Outbreak, pages 75–84. Springer International Publishing.

Karesh, W. B., Dobson, A., Lloyd-Smith, J. O., Lubroth, J., Dixon, M. A., Bennett, M., Aldrich, S., Harrington, T., Formenty, P., Loh, E. H., et al. (2012). Ecology of zoonoses: natural and unnatural histories. The Lancet, 380(9857):1936–1945.

Kendall, D. G. (1948). On the generalize "birth-and-death" process. The Annals of Mathematical Statistics, pages 1–15.

Kessler, D. A. and Shnerb, N. M. (2007). Extinction rates for fluctuation-induced metastabilities: A real space WKB approach. Journal of Statistical Physics, 127(5):861–886.

Khasin, M. and Dykman, M. I. (2011). Control of rare events in reaction and population systems by deterministically imposed transitions. Physical Review E, 83(3):031917.

Khasin, M., Dykman, M. I., and Meerson, B. (2010). Speeding up disease extinction with a limited amount of vaccine. Physical Review E, 81(5):051925.

Khasin, M., Meerson, B., Khain, E., and Sander, L. M. (2012). Minimizing the population extinction risk by migration. Physical Review Letters, 109(13):138104.

Kubo, R., Matsuo, K., and Kitahara, K. (1973). Fluctuation and relaxation of macrovariables. Journal of Statistical Physics, 9(1):51–96.

Lande, R. (1993). Risks of population extinction from demographic and environmental stochasticity and random catastrophes. The American Naturalist, 142(6):911–927.

Landsman, A. S. and Schwartz, I. B. (2007). Stochastic extinction of epidemics in large populations and role of vaccinations. arXiv preprint arXiv:0710.4592.

Legrand, J., Grais, R. F., Boelle, P. Y., Valleron, A. J., and Flahault, A. (2007). Understanding the dynamics of Ebola epidemics. Epidemiology and Infection, 135:610–621.

Leigh, E. G. (1981). The average lifetime of a population in a varying environment. Journal of Theoretical Biology, 90(2):213 – 239.

Leirs, H., Mills, J. N., Krebs, J. W., Childs, J. E., Akaibe, D., Woollen, N., Ludwig, G., Peters, C. J., and Ksiazek, T. G. (1999). Search for the Ebola virus reservoir in kikwit, democratic republic of the congo: reflections on a vertebrate collection. The Journal of Infectious Diseases, 179(Supplement_1):S155–S163.

Leroy, E. M., Kumulungui, B., Pourrut, X., Rouquet, P., Hassanin, A., Yaba, P., Delicat, A., Paweska, J. T., Gonzalez, J.-P., and Swanepoel, R. (2005). Fruit bats as reservoirs of Ebola virus. Nature, 438(7068):575–576.

Lidicker Jr, W. Z. (2010). The Allee effect: Its history and future importance. Open Ecology Journal, 3:71–82.

Lindley, B. S. and Schwartz, I. B. (2013). An iterative action minimizing method for computing optimal paths in stochastic dynamical systems. Physica D, 255:22–30.

Louca, S., Lampo, M., and Doebeli, M. (2014). Assessing host extinction risk following exposure to batrachochytrium dendrobatidis. In Proceedings of the Royal Society B, volume 281, page 20132783.

Malthus, T. R. (1888). An Essay on the Principle of Population. Reeves & Turner.

May, R. M. (1971). Stability in multispecies community models. Mathematical Biosciences, 12(1-2):59–79.

May, R. M. (1972). Will a large complex system be stable? Nature, 238(5364):413.

May, R. M. (2001). Stability and Complexity in Model Ecosystems, volume 6. Princeton university press.

Meerson, B. and Sasorov, P. V. (2008). Noise-driven unlimited population growth. Physical Review E, 78:060103(R).

P. D. Walsh, R. Biek, L. A. R. (2005). Wave-like spread of Ebola Zaire. PLoS Biology, 3:e371.

Pigott, D. M., Golding, N., Mylne, A., Huang, Z., Henry, A. J., Weiss, D. J., Brady, O. J., Kraemer, M. U. G., Smith, D. L., Moyes, C. L., Bhatt, S., Gething, P. W., Horby,

P. W., Bogoch, I. I., Browstein, J. S., Mekaru, S. R., Tatem, A. J., Khan, K., and Hay, S. I. (2014). Mapping the zoonotic niche of Ebola virus disease in Africa. eLife, 3:e04395.

Pimm, S. L. (1984). The complexity and stability of ecosystems. Nature, 307(5949):321.

Potapov, A. and Rajakaruna, H. (2013). Allee threshold and stochasticity in biological invasions: colonization time at low propagule pressure. Journal of Theoretical Biology, 337:1–14.

Pourrut, X., Kumulungui, B., Wittmann, T., Moussavou, G., Delicat, A., Yaba, P., Nkoghe, D., Gonzalez, J., , and Leroy, E. M. (2005). The natural history of Ebola virus in Africa. Microbes and Infection, 7:1005–1014.

Rand, D. A. and Wilson, H. B. (1991). Chaotic stochasticity - A ubiquitous source of unpredictability in epidemics. Proceedings of the Royal Society B, 246(1316):179–184.

Reid, A. (1953). On stochastic processes in biology. Biometrics, 9(3):275–289.

Rho, Y.-A., Liebovitch, L. S., and Schwartz, I. B. (2008). Dynamical response of multipatch, flux-based models to the input of infected people: Epidemic response to initiated events. Physics Letters A, 372(30):5017–5025.

Rohani, P., Earn, D. J., and Grenfell, B. T. (1999). Opposite patterns of synchrony in sympatric disease metapopulations. Science, 286(5441):968–971.

Samoilov, M., Plyasunov, S., and Arkin, A. P. (2005). Stochastic amplification and signaling in enzymatic futile cycles through noise-induced bistability with oscillations. Proceedings of the National Academy of Sciences, 102(7):2310–2315.

Schwartz, I. B., Billings, L., and Bollt, E. M. (2004). Dynamical epidemic suppression using stochastic prediction and control. Physical Review E, 70(4):046220.

Schwartz, I. B., Billings, L., Dykman, M. I., and Landsman, A. (2009). Predicting extinction rates in stochastic epidemic models. Journal of Statistical Mechanics: Theory and Experiment, 2009(01):P01005.

Schwartz, I. B., Forgoston, E., Bianco, S., and Shaw, L. B. (2011). Converging towards the optimal path to extinction. Journal of the Royal Society Interface, 8(65):1699–1707.

Strogatz, S. H. (2014). Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering. Hachette UK.

Tsimring, L. S. (2014). Noise in biology. Reports on Progress in Physics, 77(2):026601.

Tuite, A. R., Tien, J., Eisenberg, M., Earn, D. J., Ma, J., and Fisman, D. N. (2011). Cholera epidemic in Haiti, 2010: using a transmission model to explain spatial spread of disease and identify optimal control interventions. Annals of Internal Medicine, 154(9):593–601.

United Nation Population Division Department of Economic and So-

cial Affairs (2015). World population prospects: 2015 revision, http://www.un.org/en/development/desa/population/.

van den Driessche, P. and Watmough, J. (2002). Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. Mathematical Biosciences, 180.1:29–48.

van Kampen, N. G. (1992). Stochastic Processes in Physics and Chemistry, volume 1. Elsevier.

van Panhuis, W. G., Grefenstette, J., Jung, S. Y., Chok, N. S., Cross, A., Eng, H., Lee, B. Y., Zadorozhny, V., Brown, S., Cummings, D., et al. (2013). Contagious diseases in the United States from 1888 to the present. The New England Journal of Medicine, 369(22):2152.

Verhulst, P.-F. (1838). Notice sur la loi que la population suit dans son accroissement. correspondance mathématique et physique publiée par a. Quetelet, 10:113–121.

Viboud, C., Bjørnstad, O. N., Smith, D. L., Simonsen, L., Miller, M. A., and Grenfell, B. T. (2006). Synchrony, waves, and spatial hierarchies in the spread of influenza. Science, 312(5772):447–451.

Wells, D. K., Kath, W. L., and Motter, A. E. (2015). Control of stochastic and induced switching in biophysical networks. Physical Review X, 5(3):031036.

Whittle, P. (1957). On the use of the normal approximation in the treatment of stochastic

processes. Journal of the Royal Statistical Society. Series B (Methodological), pages 268–281.

World Health Organization (2016). Ebola situation report - March 30, 2016.

Yang, W., Zhang, W., Kargbo, D., Yang, R., Chen, Y., Chen, Z., Kamara, A., Kargbo, B., Kandula, S., Karspeck, A., Liu, C., , and Shaman, J. (2015). Transmission network of the 2014-2015 Ebola epidemic in Sierra Leone. Journal of The Royal Society Interface, 12:20150536.

This page intentionally left blank.