Cancer Invasion in Time and Space

by

Lifeng Han

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2020 by the
Graduate Supervisory Committee:

Yang Kuang, Co-Chair
John Fricks, Co-Chair
Eric Kostelich
Steve Baer
Abba Gumel

ARIZONA STATE UNIVERSITY

May 2020

ABSTRACT

Cancer is a disease involving abnormal growth of cells. Its growth dynamics is perplexing. Mathematical modeling is a way to shed light on this progress and its medical treatments. This dissertation is to study cancer invasion in time and space using a mathematical approach. Chapter 1 presents a detailed review of literature on cancer modeling.

Chapter 2 focuses sorely on time where the escape of a generic cancer out of immune control is described by stochastic delayed differential equations (SDDEs). Without time delay and noise, this system demonstrates bistability. The effects of response time of the immune system and stochasticity in the tumor proliferation rate are studied by including delay and noise in the model. Stability, persistence and extinction of the tumor are analyzed. The result shows that both time delay and noise can induce the transition from low tumor burden equilibrium to high tumor equilibrium. The aforementioned work has been published (Han *et al.*, 2019b).

In Chapter 3, Glioblastoma multiforme (GBM) is studied using a partial differential equation (PDE) model. GBM is an aggressive brain cancer with a grim prognosis. A mathematical model of GBM growth with explicit motility, birth, and death processes is proposed. A novel method is developed to approximate key characteristics of the wave profile, which can be compared with MRI data. Several test cases of MRI data of GBM patients are used to yield personalized parameterizations of the model. The aforementioned work has been published (Han *et al.*, 2019a).

Chapter 4 presents an innovative way of forecasting spatial cancer invasion. Most mathematical models, including the ones described in previous chapters, are formulated based on strong assumptions, which are hard, if not impossible, to verify due to complexity of biological processes and lack of quality data. Instead, a nonparametric forecasting method using Gaussian processes is proposed. By exploiting the local

nature of the spatio-temporal process, sparse (in terms of time) data is sufficient for forecasting. Desirable properties of Gaussian processes facilitate selection of the size of the local neighborhood and computationally efficient propagation of uncertainty. The method is tested on synthetic data and demonstrates promising results.

to my wife, Julia

ACKNOWLEDGMENTS

First, I want to thank my advisors Dr. Yang Kuang and Dr. John Fricks. They have been exceptional mentors and taught me so much over the years. Their encouragement and kindness has been providing me energy since the first time I met them. Their enormous and continuous support came in various forms to which I am deeply grateful. They are very knowledgeable and every conversation with them was very inspiring. Particularly, I want to thank Dr. Kuang for teaching me big visions of science as well as practical approaches of research, and Dr. Fricks for many talks that spawned new ideas and thoughts. Their mentorship went beyond science and research to many aspects of lives. I truly appreciate everything they have done to me.

I want to thank my committee members, Dr. Eric Kostelich, Dr. Steve Baer and Dr. Abba Gumel, and external reviewer, Dr. Hao Wang, for their kind feedback and suggestions. Their involvement helped me improve the quality of my research and writing. My special thanks goes to Dr. Eric Kostelich who has also been an excellent collaborator and a group leader of ABRC meetings from which I benefited very much.

I am extremely lucky to be part of a team of many talented and friendly individuals (Dr. Steffen Eikenberry, Changhan, Tin, Elpi, Lauren, Penny, etc) and people from Barrow Surgical Institute and Mayo Clinics. I am also grateful for being part of the big ASU SoMSS community. Particularly, I want to thank Dr. Dieter Armbuster, Dr. Hal Smith and the school Director Dr. Al Boggess for working out my transfer to ASU in the middle of the graduate directorship transition, as well as everyone who met me in a short notice during my first visit to ASU. I also want to express my gratitude to the graduate coordinators who helped me with miscellaneous things.

My gratitude also goes to the Math Biology Group and the Math Department at University of Utah where I spent my first three years of PhD study. I want to thank Dr. Aaron Fogelson and Dr. Jim Keener for advising me on research. Dr. Fogelson

was always very patient with me and taught me to keep high standards. Dr. Keener taught me how to approach a problem from first principles and charge relentlessly for a solution. Moreover, I want to thank everyone in the Math Biology Group who made it a big supportive family. My special thanks goes to Huy who listened to my research and gave many constructive suggestions. I also want to thank Paula for being such a helpful graduate coordinator.

I am grateful as well to many people at UC Davis where I launched my graduate school journey. I have acknowledged them in my masters thesis. I just want to thank again Dr. John Goeschl and Dr. Bob Guy for introducing me to the wonderful research area of Math Biology. I am very grateful to have Dr. John Goesschl as a long-term collaborator and friend, whose passion and wide knowledge have been inspiring me for many years. I am also thankful to my advisors, Dr. Heiner Lieth and Dr. Matthew Gilbert, at UC Davis. Dr. Lieth influenced and taught me in many aspects. I have benefited so much from practicing his mantra of being proactive.

I am indebted to my parents who offer me endless support. My wife, Julia, who endured my years of hard work on pursuing PhD, has always been understanding, encouraging and patient. This cannot be done without her support and sacrifice. Thank you, Julia!

I also must thank Arizona Biomedical Research Commission (ABRC) for their funding support for the research done in Chapter 3.

I could keep going and fill many more pages with my gratitude. It is not possible to include everyone to whom I wish to extend my appreciation. I will stop here and say "thank you" again to everyone who helped me get to this point. I am deeply grateful.

TABLE OF CONTENTS

CHAPTER Page

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

Chapter 1

INTRODUCTION

1.1 Mathematical Oncology

Cancer generally refers to a group of diseases that involves uncontrolled and abnormally dividing cells which spread and destroy normal body tissues. Cancer has been around for all human history (Hajdu, 2011). The word "cancer" stems from the Greek word *karkinos* (crab or crayfish), which was used by Hippocrates to describe the appearance of the cut surface of a solid malignant tumor. This name itself reflects two hallmarks of cancer: its irregular structure resembles a crab; cancer is difficult to eradicate, just like a crab that grabs on and does not let go. In fact, even as the modern medicine advances, the success of treating or curing cancer is still limited. Moreover, as the average life span increases, cancer appears to more common and life-threatening.

To improve our chance of fighting cancer, there is no doubt about the importance of innovations in medicine and treatment. Towards this goal, deeper understanding of cancer biology and more quantitative approaches are imperative. Mathematical modeling is one of the most important tools that we can rely on in this endeavor. The evolution of cancer and its interaction with the host are often studied using dynamical systems, including discrete time maps and differential equations.

There are many reasons for the effectiveness of mathematical modeling in understanding cancer biology. I will give a brief discussion considering the definition of cancer. Cancer is usually identified with abnormal growth and loss of homeostasis. This growth is characterized by the the rate of increase in the tumor size and spread.

It should be noted that the term "tumor" refers the mass of cancerous cells. Even though it is seemly used interchangeably with cancer oftentimes, not all tumors are necessary a mass of cancerous cells. In contrast to malignant tumor, there are benign tumors which do not grow fast. The most important distinction is the rate of growth. So the time scale we are looking at the problem with matters. Moreover, cancer biology involves a lot of underlying processes, and hence poses very complicated problems. Dynamicists who study mathematical models are good at exploiting the separation of time scales of the underlying processes in order to simplify problems and gain insights. In some lucky cases, we can even derive a quantitative estimate of the rate of interests. For example, in Chapter 3, the traveling wave speed, which represents the growth rate of a tumor, is estimated using scaling techniques.

The homeostasis, from a dynamicist's point of view, can be thought of as a stable steady state of a dynamical system. First, a model is constructed with the necessary components that are believed to be relevant to the development of cancer based on a hypothesis. The disruption caused by the cancer can then be thought as loss of the stability of a steady state due to the change of some key parameters that represent the drive behind the cancer development. Along this thread, Chapter 2 is developed to illustrate how responsiveness of the immune system and stochastic growth play a role in cancer invasion.

Oncology is a branch of medicine that focuses on the treatment of cancer. Many mathematical models have been published that model treatments explicitly and address the questions regarding to the optimization of the treatments. Skipper *et al.* (1961); Skipper (1971) pioneered the modern theory of chemotherapy with the log-kill hypothesis. It is hypothesized that a fixed fraction of cells are killed by a drug of certain concentration. Later refinement of the theory recognize that the rapid dividing cancerous cells respond better to the chemotherapy (NORTON *et al.*, 1976;

Simon and Norton, 2006). An interesting line of mathematical modeling is focused on the sequencing of chemotherapeutic and surgical treatments. For example, Kohandel *et al.* (2006) tried to answer the question whether chemotherapy should be applied before surgery. In their model formulation with a assumption of Gompertz growth, it was shown that chemotherapy going before surgery leads to a tumor with smaller final size than the one resulted from the other way around. Another interesting question addressed by mathematical modeling fruitfully is about drug resistance in cancer treatment. A prime example is a surge of a series of modeling work on the intermittent treatment of prostate cancer (Ideta *et al.*, 2008; Hirata *et al.*, 2012; Portz *et al.*, 2012; Everett *et al.*, 2014; Baez and Kuang, 2016; Phan *et al.*, 2019). Studies along this line suggested that intermittent deprivation of androgen can reduce hormonal resistance of cancerous cells and improve quality of patients' life. Successful fit to clinical serum androgen data is a good validation of these models. Recent developments have focused on parameter estimation and uncertainty quantification (Wu *et al.*, 2019b).

However, in this thesis the cancer treatment is not modeled explicitly. This omission is intended due to my understanding that mathematical modeling included here has not matured enough to serve explicit guidance for medical treatment. Nevertheless, this work generates a lot of insights that are potentially useful in practice. For example, the vital parameters estimated in Chapter 3 can be useful indicators for personalizing treatment. The larger migration parameter may indicate a fast spreading tumor that requires an immediate surgical removal with a big margin. On the other hand, the birth and death rates are useful indicators for radiotherapy and chemotherapy since they target fast dividing cells. For another example, the forecasting method developed in Chapter 4 gives a prediction of tumor boundary with confidence interval. This would help target the tumor when administrating radiotherapy. That being

said, caution and further study are needed if readers are interested in applying the findings in this work in real clinical settings.

It should be noted that the dynamics and spatial patterns considered in this work are not unique in mathematical oncology. Especially the time delay and stochasticity are ubiquitous in many biological systems and can be interesting on itself mathematically. Hence a brief review is be given in the following sections. For a comprehensive account of mathematical oncology, Kuang *et al.* (2015) is a good resource.

1.2 Stochasticity, Time delay and Bistability in Biological Systems

1.2.1 Stochasticity

Randomness in a common phenomenon and mechanism in biological systems, which plays an important role in shaping their dynamics (Wilkinson, 2009). To make sense of randomness, biological systems are often modeled as a stochastic process (Wilkinson, 2011). Randomness can arise intrinsically. For example, in studying chemical reactions, it is noted that sometimes only a few copies of molecules are involved. They collide each other and react by chance. The number of reactions by a time can be modeled as a Poisson progress. In general, master equations that describe the change of the probability distribution over time can be written down. The master equations can be hard to solve even though numerical simulations of a sample path are widely available, e.g. Gillespie algorithm (Gillespie, 1976). Hence an approximation in certain limit is often sought. This gives rise to a type of stochastic differential equations (SDEs) known as chemical Langevin equations (Gillespie, 2000). The underlying justification is due to the functional central limit theorem (Donsker, 1951). This randomness at the discrete model is turned into a white noise term in the SDE. On the other hand, there is extrinsic noise that is due to extrinsic factors that is

not explicitly modeled, e.g. fluctuating of the environments. Sometimes the random noise is added in an ad hoc way to account for the imperfectness of a proposed model, as commonly seen in the field of data assimilation.

To have a reasonable interpretation of the noise term in SDEs, stochastic calculus is needed, which can be constructed by writing SDEs in a discrete version and then reaching a sum resembling the Riemann sum in the deterministic case. In contrast to Riemann sum for which the limit does not depend on the choice of the sample point in the discrete interval, there are two common interpretations of stochastic integrations depending on using the midpoint (Stratonovich) or left end point (Itô) (Øksendal, 2003). In this thesis, Itô interpretation is used and hence the Itô's formula is a workhorse which is repeatedly used throughout the analysis in Chapter 2.

1.2.2 Time Delay

Time delay is also common in in biological systems (Rihan *et al.*, 2018). In population dynamics, time delay often manifests as the time required for juvenile to mature in a population with non-overlapping generations or the time a predator takes to digest its prey. In a molecular scale, it takes time for substance to diffuse over distance. In a gene network, the synthesis of proteins also takes time and can cause time delay in a feedback loop. From modeling point of view, it oftentimes convenient to introduce a time delay instead of developing a full model for the underlying process that causes the time delay. In a continuous time model, time delays are usually introduced to appropriate terms in differential equations, giving arise to delayed differential equations (DDEs) or neutral differential equations. In some cases, time delay can cause loss of stability of a fixed point. For example, Parmar *et al.* (2015) reviewed DDE models of gene regulatory networks where standard local stability analysis was conducted.

1.2.3 Time Delay and Stochasticity in Bistable Systems

Of particular interests is how time delay and stochasticity play a role in a bistable system. Bistability is an important mechanism and phenomenon in biology, e.g. synthetic gene circuit (Gardner *et al.*, 2000), developmental biology (Tian *et al.*, 2013) and pattern formation (Umulis *et al.*, 2006). There are continuing interests in studying how time delay and stochasticity induce the transition from one state to another. Recent development on a technique based on stochastic sensitivity functions (SSFs) has shown a way to visualize confidence domains and chance of state transition (Bashkirtseva *et al.*, 2014; Bashkirtseva, 2017). It basically approximates an attractor as a quadratic potential well so that confidence domains can be easily drawn near it. This technique has been applied to predator-prey models (Wu *et al.*, 2019a). It would be interesting to apply the SSF techniques to a time-delayed SDE to illustrate how time delay enhance the state transition.

1.3 Traveling Wave in Reaction-diffusion Systems

Reaction-diffusion equations are a type of partial differential equations (PDEs) which are popular candidate models in spatial population dynamics. It was first used to study the disperse of advantageous genes in the seminal work of Fisher (1937). Kolmogorov *et al.* (1937) laid the foundation for the theory of traveling waves solutions of reaction-diffusion equations. Since then, there have been fruitful developments in its theory and applications (Volpert and Petrovskii (2009) and reference therein). In modeling population, the diffusion accounts for the random movement of individuals and the reaction terms account for the reproduction of or interaction between species. Traveling wave solutions are common in many reaction-diffusion models. The existence, stability and speed of the traveling wave are of primary interests in its analysis.

If there is only one species, to show the existence of the traveling wave, the scalar reaction diffusion equation can converted to a systems of two ordinary differential equations and hence is amendable to a phase plane analysis. For a reaction-diffusion system, if the reaction terms are monotone, maximum principle still holds and there are established theories to show the existence of a traveling wave (Volpert, 1994). However, if the reaction terms are not monotone, there does not exist general theories. Nevertheless, analysis sometimes can be done in the phase space similar to the one in the phase-plane analysis if the number of dimensions are not too large. In Chapter 3, approximated wave solution is sought. The full system analysis is challenging and remains an open problem. However, if using simple linear diffusion instead of the cross diffusion, a rigorous proof of existence of a traveling wave is possible and provided in the Appendix B.

It should be pointed out that time delay and stochasticity can complicate the analysis of traveling wave even more but offer interesting extension to the deterministic system. For example, Doering *et al.* (2003) considered a type of stochastic Fisher-KPP equation, for which they derived a duality relation to a birth-coagulation interacting particle system. A conjecture is made about the wave speed, which states that the noise intensity decrease the wave speed. Furthermore, there are some general theories that have been developed on reaction-diffusion systems with time-delayed reaction (Wu and Zou, 2001). In this thesis, however, it is decided to focus on deterministic PDE models, since the evidence and mechanism of the time delay and stochasticity are not clear in cancer invasion.

1.4 Filtering, Smoothing and Forecasting

In this section, state space modeling is reviewed, including methods of filtering, smoothing and forecasting. Many of the materials reviewed here have motivated

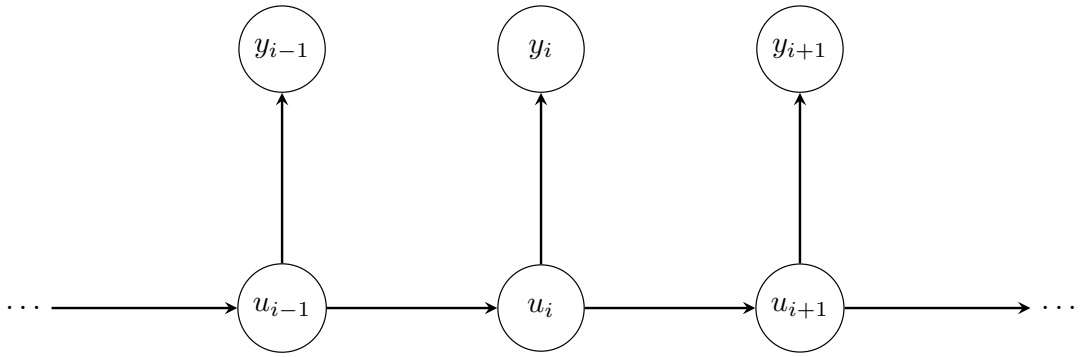

Figure 1.1: Graphical representation of a state space model where the arrows show the dependence and u's are hidden variable and y's are observations

the development of Chapter 4. By giving a broad background, it is hoped to make clear the gap in the literature our work attempted to fill as well as its limitations. Without doubt, further development of our method is needed to make it more general, i.e. capable of tasks of filtering, smoothing and forecasting. For a comprehensive introduction on state space modeling, see Durbin and Koopman (2012) and see Cappé *et al.* (2006) for a more theoretical exposition.

Figure 1.1 is a diagram illustrating the variables and the dependence structure in a general state space model, where u is the hidden process variable (a.k.a state variable) and y is the observation. The objective is to estimate certain property (e.g. mean and variance) of u based on observation y. There are three inference tasks: given observations $y_1, y_2 \cdots y_n$, the filtering concerns with $p(u_n|y_{1:n})$, the smoothing concerns with $p(u_t|y_{1:n})$ for $t \in \{1, 2, \cdots n-1\}$, the forecasting concerns with $p(u_t|y_{1:n})$ for $t \in \{n+1, n+2, \cdots\}$. The method developed in Chapter 4 focuses on the task of long term forecasting.

State space modeling is a very useful and general framework that has wide applications. For example, autoregressive–moving-average (ARMA) models (Box *et al.*,

2015) can be cast in the form of a state space model. Since the seimnal paper of Kalman (1960) , state space modeling has been heavily researched. Its use first became popular in engineering and now has applications across many fields, such as economics and biology.

Kalman (1960) showed that in the case of linear and Gaussian models, the analytical solution about u is available. The algorithm is later known as Kalman filter. Another important feature of Kalman filter is that the filtering is done in a recursive fashion. That is, as data come in sequentially, the analysis is updated based on previous analysis with a correction using the newly available data. Gaussian assumption is important for the Kalman filter to work. Gaussianity has many desirable features:

- Gaussianity is preserved under linear operations (linear combinations of normal random variables is also a normal random variables)
- for a multivariate normal random variable, the conditional on part of it is also Gaussian
- the integration of a Gaussian function is easy
- the product of two Gaussian functions is also Gaussian

The first two features are the foundations of Kalman filter. The last two features are taken advantage of in the development of our method in Chapter 4. To make Kalman filter work for nonlinear models has spur a lot of interests and been a major focus of modern development of Kalman filter. Refined Kalman filter that can deal with nonlinear models includes

- extended Kalman filters, which linearize the nonlinar model;
- unscented Kalman fitliers, which introduce sigma points and weights to get approximated first two moments under nonlinear maps (Wan and Van Der Merwe,

2000);

- ensemble Kalman filters, which propagate samples through a nonlinear model and then get approximated first two moments from the samples (Evensen, 2009).

However, almost all variants of Kalman filters still heavily rely on the Gaussian assumptions on state variables. This inevitably introduces approximation error if the underlying distribution is not normal. In particular, if the underlying distribution is multi-modal, it can cause significant errors and in that sense the mean does not convey useful information. For example, in cancer modeling, we may get into a situation in which the location of a tumor can be either in one or the other location, but it is unlikely that a tumor exists at both locations due to immune suppression. The inability to handle multi-modality may be mitigated by using a mixture of Gaussians. Nevertheless, a general method is to avoid the Gaussian assumption at all.

In the nonGaussian and nonlinear world, the particle filter is the "go to" method. Particle filters basically rely on a technique called sequential importance sampling. Assuming we can sample from a distribution $p(x)$, let $\{x_i\}$ denote samples (particles). Then we have the following approximation

$$\begin{aligned} p(x) &\approx \frac{1}{n}\sum_{i=1}^{n}\delta_{x_i}(x) \\ E[f(X)] &\approx \frac{1}{n}\sum_{i=1}^{n} f(x_i). \end{aligned}$$

In a filtering problem, it is often not possible to direct sample from $p(u_n|y_{1:n})$. Instead, we can use importance sampling. The importance distribution can be chosen to be the $p(u_t|u_{t-1})$, which results in the so-called bootstrap particle filter.

One of the drawbacks of particle filter is its suffering from the curse of dimensionality (Snyder *et al.*, 2008). In spatio-temporal modeling, the dimensionality of the state is often very large, which is particularly true in modeling brain tumors with

MRI data. Recently, progresses in local particle filters have shown some promise in extending particle filters to high dimensional problems (Rebeschini and van Handel, 2015). Nevertheless, the field of high-dimensional data assimilation is dominated by Kalman filters especially in weather models (Hunt *et al.*, 2007).

Both Kalman filters and particle filters require a formulation of the process model $p(u_t|u_{t-1})$. The formulation requires knowledge about the underlying process. Without specific knowledge, a random walker can be assumed. However, this approach makes high-quality long term forecasting impossible. Towards this goal, a mechanistic process model is needed. This in turn oftentimes introduces unknown parameters that need to be estimated. Even though parameter estimation is possible using maximum likelihood (e.g., EM algorithm) or Bayesian methods, it can be cumbersome and computational expensive. What is worse in biological applications is that the process model itself can be questionable. The necessary model validation can be another burden. This motivates a model-free (nonparametric) approach in handling $p(u_t|u_{t-1})$, for which we embarked in Chapter 4.

1.5 Outline of the Thesis

In Chapter 2, a model of tumor-immune interaction is proposed: one stable equilibrium representing a state of a tumor being contained; the other representing a state of a tumor that escaped the control of the immune system. The effects of time delay in immune response and stochastic noise in proliferation rate are studied. The conditions for number of equilibrium of the ODE system is include in Appendix A. In Chapter 3, a system of two reaction-diffusion equations is proposed to study a type of malignant brain tumor. The model demonstrate a traveling wave that represents tumor invasion. A novel method is developed to estimate vital parameters from MRI data for individual patient. In Chapter 4, Gaussian processes are used to forecast

a spatio-temporal process motivated by an application to predict growth of a brain tumor. The computer code used for simulation and generating figures is summarized in Appendix E.

Chapter 2

DYNAMICS OF A MODEL OF TUMOR-IMMUNE INTERACTION WITH TIME DELAY AND NOISE

2.1 Abstract

We propose a model of tumor-immune interaction with time delay in immune reaction and noise in tumor cell reproduction. Immune response is modeled as a non-monotonic function of tumor burden, for which the tumor is immunogenic at nascent stage but starts inhibiting immune system as it grows large. Without time delay and noise, this system demonstrates bistability. The effects of response time of the immune system and uncertainty in the tumor innate proliferation rate are studied by including delay and noise in the appropriate model terms. Stability, persistence and extinction of the tumor are analyzed. We find that delay and noise can both induce the transition from low tumor burden equilibrium to high tumor equilibrium. Moreover, our result suggests that the elimination of cancer depends on the basal level of the immune system rather than on its response speed to tumor growth.

2.2 Introduction

The immune system is a host defense system that distinguishes pathogens from one's own healthy cells and destroy them. Pathogens include bacteria and viruses as well as abnormal cells. There are evidence that immune systems can detect and eliminate cancer (Parish, 2003). Cancer immunology has seen renewed interests due to recent achievements in immunotherapy (Mahoney *et al.*, 2015). Unlike common pathogens, cancer cells are not as distinguishable. Moreover, cancer can evade the

control of immune systems by developing ways to achieve immune suppression (Whiteside, 2006). Recent progress in immunotherapy has focused on removing immune suppression so that cancer cells will be under attack of immune systems. Further development of immunotherapy hinges on a better understanding of tumor-immune system interaction.

The immune system consists of immune cells of several types and complex signaling network. For example, immune cells include cytotoxic T cells, helper cells, natural killer cells, and signaling networks involve cytokines, e.g., interlukin-2. The immune response falls into two categories: innate and adaptive. Though the adaptive response enables fast reaction to certain antigens, immune responses oftentimes involve recruiting immune cells from bone marrows and further training or activation. So it is inherently a delayed process. On the other hand, cancer is well-known for its heterogeneity and characterized by fast mutation, which can render it the ability to suppress immune response (Fisher *et al.*, 2013; Whiteside, 2006). In this process, stochasticity plays an important role.

Given the complexity of tumor-immune system interaction, mathematical modeling naturally offers some insights by capturing the key mechanisms. Literature on modeling non-spatial tumor growth under immune surveillance is abundant (see (Eftimie *et al.*, 2011) for a review). In (Kuznetsov *et al.*, 1994), the authors derived a system of five equations from a kinetic scheme and further reduced it to two equations tracking only effector cells and tumor cells. In their model, immune suppression is represented as annihilation by mass action of tumor cells and effector cells. Their paper inspired a number of later work. In (D'Onofrio, 2008), the author summarized

models of this type into a family as follows

$$x' = x(f(x) - \phi(x)y), \tag{2.1a}$$

$$y' = -\Psi(x)y + \sigma q(x) + \theta(t) \tag{2.1b}$$

where x is the size or density of tumor population, y is the size or density of immnue effector population and $\theta(t)$ represents treatment . With biological-relevant assumptions on $f(x), \phi(x)$ and $q(x)$, some general conclusions of were drew in (D'Onofrio, 2008).

In (Kirschner and Panetta, 1998), the authors took into account cytokines in their model in order to gain further insight into immunotherapy. In most cases, introducing more equations does not introduce new dynamics but is necessary to shed light on cancer treatments (Eftimie *et al.*, 2011). Signaling is arguably a fast process, so sometimes a quasi-steady state approximation can reduce more complicated models into the family of models studied by (D'Onofrio, 2008). For example, in a recent study on immune check point inhibitor (Nikolopoulou *et al.*, 2018), a variation of (2.1) is derived from a more complicated model consisting of a system of 14 partial differential equations (Lai and Friedman, 2017).

Building on aforementioned earlier work, the effect of time delay in tumor-immune interactions are also extensively studied (Banerjee and Sarkar, 2008; D'Onofrio *et al.*, 2010; Gałach, 2003; Rihan *et al.*, 2014). On the other hand, stochasticity in tumor-immune interaction is less often studied, and mostly focused on single-equation models of tumor population growth ignoring explicit interaction with immune system (Lefever and Horsthemke, 1979; Bose and Trimper, 2009; Li and Cheng, 2017). Models that take into account both time delay and stochasticity are even rarer. To the best of our knowledge, the work of (Guo and Mei, 2014) is the only exception. In (Guo and Mei, 2014), a single equation was considered and noise was introduced in an ad-hoc way.

To fill this gap, we will introduce time delay and noise one at a time to a variation of (2.1) with the novelty in its immune response term being non-monotonic.

The paper is organized as follows. In Section 2.3, we introduce our model formulation and state general properties of the model in absence of noise and time delay. In Section 2.4, we study the model with time delay where persistence, local stability and global stability results are presented. Section 2.5 is devoted to the stochastic version of the model where tumor extinction and persistence conditions are presented. In section 2.6, the paper comes to the culmination where both noise and time delay are present in the model, for which we study the effects of time delay on the stationary distribution. We conclude the paper with a discussion of future work in section 2.7.

2.3 ODE Model Formulation and Properties

We consider the following equations

$$\frac{dx}{d\hat{t}} = \hat{\rho}x(1-\frac{x}{\hat{K}}) - \hat{\gamma}xy, \tag{2.2}$$

$$\frac{dy}{d\hat{t}} = \hat{\beta} - \hat{\mu}y + \frac{\hat{\alpha}x}{\hat{\kappa}+x^2}, \tag{2.3}$$

where x denotes the tumor burden and y denotes the level of immune response. This is a special case of (2.1) for which we specify $\sigma q(x) = \hat{\beta} + \frac{\hat{\alpha}x}{\hat{\kappa}+x^2}$ to represent the immune response to tumor burden in a non-monotonic fashion. In spite of being phenomenological, it accounts for the fact that the cancer cells acquire mutations that can down-regulate immune response as the tumor grows bigger while the tumor is immuogenic at its nascent stage. Death term is assumed to be independent of tumor burden and thus we have $\Psi(x)y = -\hat{\mu}y$. Note that without tumor, immune system activity is maintained at the basal level β/μ. We assume logistic growth of tumor and that killing of tumor cells by the immune system is represented as a mass

action term so that we have $f(x) = 1 - \frac{x}{\hat{K}}$ and $\phi(x) = \hat{\gamma}x$. Both are common choices in modeling literature (Gałach, 2003; Nikolopoulou *et al.*, 2018). Even though Gompertz growth is also commonly used in modeling tumor growth and backed up by a lot of experimental data, it predicts unbound growth rate as tumor burden approaches zero (Eftimie *et al.*, 2011). Hence it is not suitable to our purpose.

Note that we represent the two populations generically as tumor burden and level of immune response, and intentionally avoid using specific units of cell density or volume. In some sense, it justifies our choice of $\Psi(x)y = -\hat{\mu}y$ since we are not tracking explicitly effector cells. The rationale behind our generic representation is that our objective is to study interesting dynamics that can arise from a simple model with time delay and noise. It is our intention to bring attention to the possible key mechanisms underlying cancer-immune system dynamics without making the wrong impression that it is capable of clinical prediction.

By nondimensionlization, we can reduce the number of parameters. Let $x = \sqrt{\kappa}u$, $y = \frac{\alpha}{\mu\sqrt{\kappa}}v$ and $t = \frac{1}{\mu}\hat{t}$. We get the dimensionless system

$$\frac{du}{dt} = \rho u(1 - \frac{u}{K}) - \gamma uv \equiv f(u, v), \tag{2.4a}$$

$$\frac{dv}{dt} = \beta - v + \frac{u}{1 + u^2} \equiv g(u, v). \tag{2.4b}$$

where $\rho = \frac{\hat{\rho}}{\hat{\mu}}$, $K = \frac{\hat{K}}{\sqrt{\hat{\kappa}}}$, $\gamma = \frac{\hat{\alpha}\hat{\gamma}}{\hat{\mu}^2\sqrt{\hat{\kappa}}}$ and $\beta = \frac{\hat{\beta}\sqrt{\hat{\kappa}}}{\hat{\alpha}}$ are dimensionless (positive) parameters. It can be shown that there is always a tumor free equilibrium $E_0 = \{0, \beta\}$ which is stable if and only if $\rho < \gamma\beta$ by linear stability analysis. If $\rho > \gamma\beta$, depending on the parameters, there is additional either one or three interior equilibria. In this paper, we focus on the case where there are three interior equilibria, namely low tumor equilibrium E_1, intermediate tumor equilibrium E_2 and high tumor equilibrium E_3 (see Figure 2.1). Since the intersections of nullclines are roots of a cubic polynomial,

any results involving analytical expressions of parameters for stability conditions, if possible, would be unwieldy. Nevertheless, qualitative argument similar to that of in ((Murray, 2002) Page 226-230) can be made regarding to the local stability . By doing so, we obtain the theorem below.

Theorem 2.3.1. *Suppose that $\rho > \gamma\beta$ and there are three interior equilibria E_1, E_2 and E_3 of (2.4), then E_1 and E_3 are stable, and E_2 is unstable.*

Proof. Consider the Jacobian matrix at the equilibrium

$$J = \begin{bmatrix} f_u & f_v \\ \\ g_u & g_v \end{bmatrix}$$

where f_u, f_v, g_u and g_v are the partial derivatives of $f(u, v)$ and $g(u, v)$ evaluated at the equilibrium. The signs of f_u, f_v, g_u and g_v can be determined by geometric arguments. For example, to see the sign of g_u at E_1, we first note that $g(u, v) < 0$ in the region above $g(u, v) = 0$ and $g(u, v) > 0$ in the region below. Thus as it moves across $g(u, v) = 0$ at E_1 along the direction of u-axis from left to right, $g(u, v)$ increases from being negative to being positive. Therefore, $g_u > 0$ at E_1. By the same argument, we find out the signs of the entries of Jacobian matrices J_1, J_2 and J_3 at E_1, E_2 and E_3 as follows

$$J_1 = \begin{bmatrix} - & - \\ \\ + & - \end{bmatrix} \quad J_2 = \begin{bmatrix} - & - \\ \\ - & - \end{bmatrix} \quad J_3 = \begin{bmatrix} - & - \\ \\ - & - \end{bmatrix} \tag{2.5}$$

Thus $Tr(J_1) < 0, Tr(J_2) < 0$ and $Tr(J_3) < 0$. To determine stability, we need to find signs of their determinants too. This can be done by noting the sign of the slopes of the nullclines. For example, at E_1, $\frac{dv}{du} > 0$ along $g(u, v) = 0$ and $\frac{dv}{du} < 0$ along

$f(u,v) = 0$. Implicit differentiating $g(u,v) = 0$ with respect to u gives $\frac{dv}{du} = -\frac{g_u}{g_v}$. Similarly, we have $\frac{dv}{du} = -\frac{f_u}{f_v}$ along $f(u,v) = 0$. Hence at E_1, we have $-\frac{g_u}{g_v} > 0 > -\frac{f_u}{f_v}$. Since $g_v < 0, f_v < 0$ at E_1, $\det(J_1) = f_u g_v - f_v g_u > 0$. Similarly, we have $\det(J_2) < 0$ and $\det_(\, J_3) > 0$. Therefore, we proved the claimed stability. □

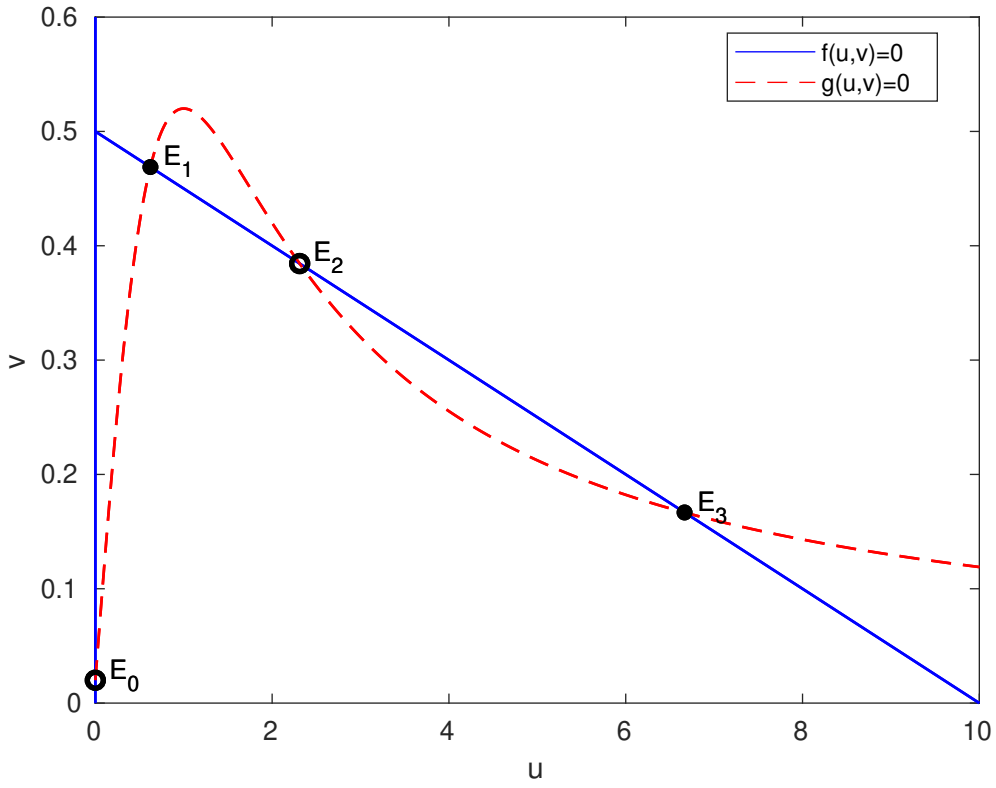

Figure 2.1: Nullclines of (2.4). Open circles denote unstable fixed points and filled circle denotes stable fixed points. Parameter values used:
$\rho = 2.5, \beta = 0.02, \gamma = 5, K = 10$

2.4 With Delay

It is known that it takes time for the immune system to respond and take effect. So we introduce time delay into the term representing immune response. The model now reads as follows

$$u' = \rho u(1 - \frac{u}{K}) - \gamma uv \tag{2.6a}$$

$$v' = \beta - v + \frac{u(t-\tau)}{1 + u(t-\tau)^2}. \tag{2.6b}$$

where $\tau \geq 0$ is the time delay of immune response. We are curious about any possible new dynamics due to the time delay in contrast to the ODE model, especially if the time delay in immune response can lead to oscillatory solutions that may represent cancer growth episodes.

2.4.1 Analysis

First we want to establish positivity and boundedness of the system (2.6) given appropriate initial values. For initial values, we assume $v(0) \geq 0$ and $u(\theta) = \phi(\theta)$ for $\theta \in [-\tau, 0]$ where $\phi(\theta) \in C([-\tau, 0]) : \phi(\theta) \geq 0$ and $\phi(0) > 0$.

Proposition 1. *The solution of (2.6) satisfies that* $u(t) > 0$ *and* $v(t) > 0$ *for all* $t > 0$. *Moreover,* $\limsup_{t\to\infty} u(t) \leq K$ *and* $\beta \leq \liminf_{t\to\infty} v(t) \leq \limsup_{t\to\infty} v(t) \leq \beta + 1/2$.

Proof. Suppose there exists $t_1 > 0$ such that $u(t_1) = 0$ and $u(t) > 0$ for $t \in (0, t_1)$. Note that (2.6a) is in the form $u' = uF(u, v)$. Then

$$u(t_1) = u(0) \exp(\int_0^{t_1} F(u(s), v(s)ds) > 0,$$

which is a contradiction. The positivity of v is obvious as if there is t_1 such that $v(t_1) = 0$, then $v'(t_1) > 0$. The boundedness is straightforward. □

The above proposition ensures that the solution of (2.6) is biological meaningful and will also be useful prerequisite for our later analysis. Next we present a theorem of the local stability of (2.6) as a counterpart to Theorem 2.3.1. In particular, we show that as τ increases there is a possible stability switching for low tumor equilibrium

E_1. Recalling the same definition for f_u, f_v, g_u and g_v as before, we have the following theorem.

Theorem 2.4.1. *Suppose that $\rho > \gamma\beta$. Then the tumor free equilibrium E_0 is unstable for for all $\tau \geq 0$. In addition, suppose that there are three interior equilibria E_1, E_2 and E_3 of (2.6). Then E_3 is stable, E_2 is unstable for all $\tau \geq 0$ and E_1 change its stability as τ increases if $f_u g_v + f_v g_u < 0$.*

Proof. Linearizing (2.6) at the equilibrium gives

$$\begin{aligned} u' &= f_u u + f_v v \\ v' &= g_u u(t-\tau) + g_v v \end{aligned}$$

Substituting $u = Ae^{\lambda t}, v = Be^{\lambda t}$ into the above equations yields a linear system of A and B for which the existence of a solution entails the characteristic equation

$$\lambda^2 - (f_u + g_v)\lambda + f_u g_v - f_v g_u e^{-\lambda\tau} = 0. \tag{2.7}$$

At E_0, (2.7) becomes $\lambda^2 - (\rho - \gamma\beta)\lambda - \rho + \gamma\beta = 0$, which always has a positive root and thus E_0 remains unstable for all $\tau \geq 0$.

Theorem 2.3.1 gives the stability results of E_1, E_2 and E_3 at $\tau = 0$. We want to study if their stability changes as time delay τ increases. If the equilibrium changes its stability as τ increases, there must exists $\omega > 0$ such that $\lambda = i\omega$ is a solution of (2.7) (Kuang, 1993). Substituting $\lambda = i\omega$ into (2.7) gives

$$-\omega^2 - (f_u + g_v)(i\omega) + f_u g_v - f_v g_u(\cos(\omega\tau) - i\sin(\omega\tau)) = 0.$$

Collecting real and imaginary parts gives

$$(f_u g_v - \omega^2 - f_v g_u \cos(\omega\tau)) - ((f_u + g_v)\omega + f_v g_u \sin(\omega\tau)i = 0.$$

Equating the real and imaginary parts to zeros respectively gives

$$\cos(\omega\tau) = \frac{f_u g_v - \omega^2}{f_v g_u}, \tag{2.8a}$$

$$\sin(\omega\tau) = -\frac{(f_u + g_v)\omega}{f_v g_u}. \tag{2.8b}$$

Squaring and adding the above equations gives

$$\omega^4 + (f_u^2 + g_v^2)\omega^2 + f_u^2 g_v^2 - f_v^2 g_u^2 = 0$$

which has a real solution $\omega = \omega_0$ if and only if

$$(f_u g_v - f_v g_u)(f_u g_v + f_v g_u) < 0. \tag{2.9}$$

From (2.5), we see that (2.9) is not satisfied for E_3 and hence it has no stability switching. Stability switching is possible for E_1 and E_2. For E_1, (2.9) entails an additional requirement

$$f_u g_v + f_v g_u < 0, \tag{2.10}$$

which is stated in the theorem. The critical value of τ can be found by substituting $\omega = \omega_0$ to (2.8), which leads to

$$\tau_0 = \frac{1}{\omega_0}[\arctan(\frac{f_u g_v - \omega_0^2}{(f_u + g_v)\omega_0})].$$

To ensure stability switching, We need further confirm that the imaginary root crosses the imaginary axis as τ increases beyond τ_0. To do so, we think λ as a function of τ and differentiating (2.7) gives

$$[2\lambda - (f_u + g_v) + f_v g_u e^{-\lambda\tau}\tau]\frac{d\lambda}{d\tau} = -f_v g_u e^{-\lambda\tau}\lambda.$$

It follows that

$$\begin{aligned} \operatorname{sgn}\left\{\left[\frac{d\Re(\lambda)}{d\tau}\right]_{\tau=\tau_0}\right\} &= \operatorname{sgn}\left\{\Re\left[(\frac{d\lambda}{d\tau})^{-1}\right]_{\tau=\tau_0}\right\} \\ &= \operatorname{sgn}\left\{\frac{(f_u - g_v)^2 + 2\omega_0^2}{f_v^2 g_u^2}\right\} = 1. \end{aligned}$$

So for E_1 which is stable at $\tau = 0$, the imaginary root indeed crosses the imaginary axis and stability switching is ensured. But for E_2 which is unstable at $\tau = 0$, the imaginary root does not cross the imaginary axis and there is no change in its stability. □

The above theorem indicates that the time delay in immune response can lead to tumor escaping the control of the immune system as a consequence of loss of stability of the low tumor equilibrium. The condition (2.10) for this stability switching is not explicit in terms of parameters. Nevertheless, it has a geometric interpretation that the slope of the v-nullcine is larger than the magnitude of the slope of u-nullcline at their interception E_1. In a biological sense, it means that there would be instability if the objective of immune response level is set to increase steeply with respect to tumor burden while immune response time is not fast enough to catch up.

Next we study the global stability of the tumor free equilibrium.

Theorem 2.4.2. (global stability of tumor free equilibrium)

If $\rho < \gamma\beta$, *then* $\lim_{t\to\infty}(u(t), v(t)) = (0, \beta)$.

Proof. We first shift v to make its equilibrium value to be zero by letting $w = v - \beta$. We consider the equivalent equation as below

$$\begin{aligned} u' &= u(\rho - \gamma\beta - \frac{\rho u}{K} - \gamma w), \\ w' &= -w + \frac{u(t-\tau)}{1 + u(t-\tau)^2}. \end{aligned}$$

Assuming $\rho < \gamma\beta$, then there exists a p such that $\rho - \gamma\beta \geq p > 0$. Define

$$V = \frac{1}{8p}u(t)^2 + \frac{1}{2}w(t)^2 + \frac{1}{4}\int_{t-\tau}^{t} u(s)^2 ds.$$

It follow that

$$\begin{aligned}\dot{V} &= \frac{1}{4p}u(t)^2(\rho-\gamma\beta-\frac{\rho}{K}u-\gamma w)+w(-w+\frac{u(t-\tau)}{1+u(t-\tau)^2})\\ &\quad +\frac{1}{4}(u(t)^2-u(t-\tau)^2)\\ &= (\frac{\rho-\gamma\beta}{4p})u(t)^2-\frac{\rho}{4pK}u(t)^3-\frac{\gamma}{4p}wu(t)^2\\ &\quad -w^2-\frac{1}{4}u(t-\tau)^2+\frac{wu(t-\tau)}{1+u(t-\tau)^2}.\end{aligned}$$

By construction, we have $\frac{\rho-\gamma\beta}{4p} < 0$ and by Proposition 1, we have $w = v - \beta \geq 0$ Hence,

$$\dot{V} \leq -(w-\frac{1}{2}u(t-\tau))^2-\frac{\rho}{4pK}u(t)^3 \leq 0$$

We also note that $\dot{V} = 0$ if and only if $u(t)$ is identically zero and $w = u(t-\tau)/2$ which implies that w must also be identically zero. Thus $\lim_{t\to\infty}(u(t), v(t)) = (0, \beta)$ by Liapunov-LaSalle theorem (see (Kuang, 1993) Theorem 2.5.3). □

If the stability condition of tumor-free equilibrium is not satisfied, we can prove that the tumor population is persistent.

Theorem 2.4.3. (tumor persistence) *If $\rho > \gamma\beta$, there is $m > 0$ such that* $\limsup_{t\to\infty} u(t) > m$.

Proof. Suppose not. Then $\lim_{t\to\infty)} u(t) = 0$. From this, it can be shown that $\limsup_{t\to\infty} v(t) \leq \beta$. There are two cases we need to consider. The first case is that u is eventually monotonically decreasing to zero. Then $\lim_{t\to\infty} v(t) = \beta$. It follows that there is a $t^* > 0$ such that $u(t^*)$ is sufficiently small and $v(t^*)$ is sufficiently close to β so that $u'(t^*) = u(\rho - \frac{\rho u}{K} - \gamma v) > 0$. So we have a contradiction. The second case is that u approaches zero in an oscillatory fashion. That is, there is a sequence of $\{t_i\}$ such that $u(t_i) \to 0$ and $u(t_i)$ is local minimum, i.e., $u'(t_i) = u(\rho - \frac{\rho u}{K} - \gamma v) = 0$. Then for any $\delta > 0$ there is an $N(\delta)$ such that $u(t_i) < \delta$ for $i > N(\delta)$. Since $\rho > \gamma\beta$,

there is a δ such that $\rho - \frac{\rho\delta}{K} - \gamma(\beta + \delta) > 0$. We see that for this δ, there is a T_1 such taht $t > T_1$ implies that $v(t) < \beta + \delta$ and an N such that for $i > N$, $t_i > T_1$. Hence for $i > N$, $u'(t_i) > 0$ which is a contradiction to $u'(t_i) = 0$. □

Putting the conditions of the above two theorems in original units, we find an important dimensionless value $R_0 = \frac{\hat{\rho}\hat{\mu}}{\hat{\gamma}\hat{\beta}}$. It is the ratio of the proliferation ability of the tumor verse the strength of the immune system, which can be viewed as the tumor reproduction number: if

$$R_0 = \frac{\hat{\rho}\hat{\mu}}{\hat{\gamma}\hat{\beta}} < 1, \tag{2.11}$$

the tumor can be totally cleaned out by immune system. The biological meaning of (2.11) is more clear by recognizing that it is equivalent to saying that at the onset of tumor growth, the killing rate of tumor by the immune system ($\hat{\gamma}\hat{\beta}/\hat{\mu}$) is larger than the tumor growth rate ($\hat{\rho}$). Theorem 2.4.3 tells us that otherwise ($R_0 > 1$) the tumor will always exist.

2.4.2 Numerical Simulation

We studied (2.6) numerically using dde23 in Matlab. We confirmed the criteria for tumor persistence and extinction in our numerical experiment. The existence of a stability switching for low tumor equilibrium as τ increases is also demonstrated (Figure 2.2). Moreover, we discovered that as τ is further increased, there is a transition to high tumor equilibrium (bottom right of Figure 2.2). Interestingly, the jerky transition to high tumor equilibrium is reminiscent of saltatory growth which is a commonly observed pattern in tumor growth (Kuang *et al.*, 2018). This suggests that a weakened immune system response may be indicated by its slow action which takes longer time delay, which may result in increased tumor growth.

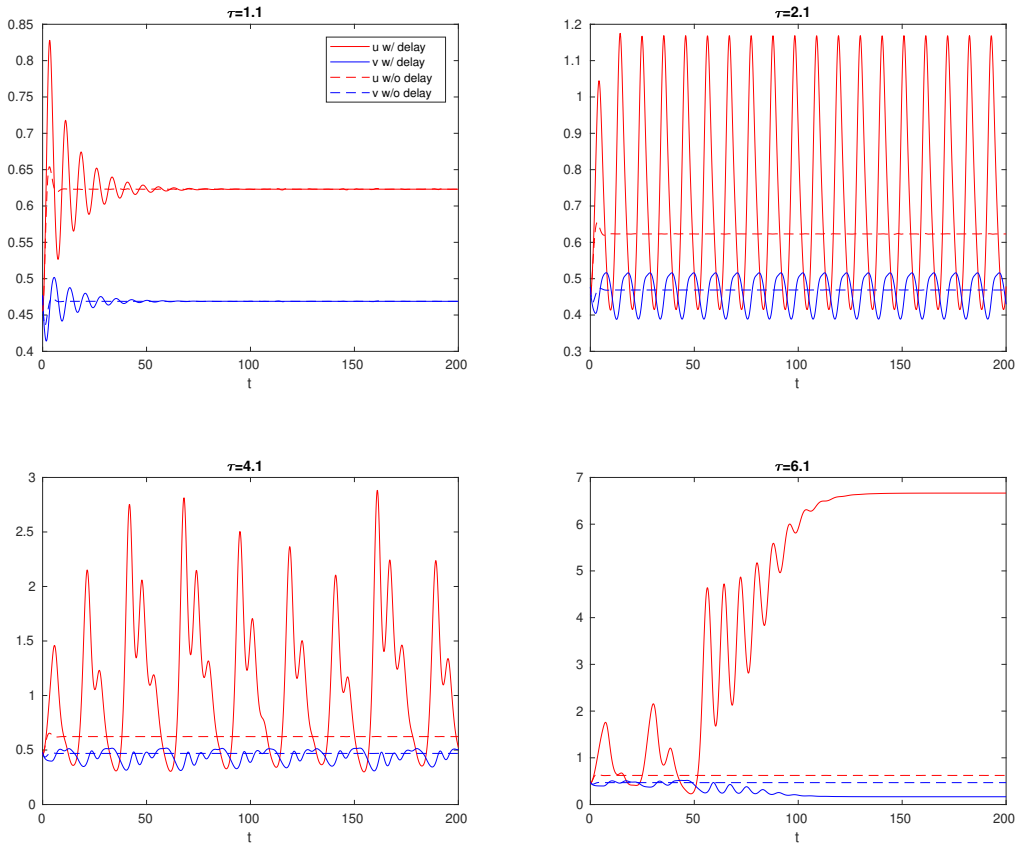

Figure 2.2: Time course of u and v with and without delay. There is a stability switch as τ increases and eventually the solution settles down to high tumor equilibrium. The observation indicates that responsiveness of the immune system is important to contain the tumor in its nascent size. If there is long time delay in immune response, the tumor can grow in oscillatory fashion and eventually escape the control of the immune system. Parameter values used to generate the plots: $\rho = 2.5, \beta = 0.02, \gamma = 5, K = 10$

2.5 With Noise

Tumor cell growth is known to lack of certain regulations compare to normal cells, we thus assume that the innate proliferation rate of tumor cells is subject to uncertainty. Adding white noise to it results in the following stochastic differential equations

$$du = (\rho u(1 - \frac{u}{K}) - \gamma uv)dt + \sigma u(1 - \frac{u}{K})dB(t), \tag{2.12a}$$

$$dv = (\beta - v + \frac{u}{1 + u^2})dt. \tag{2.12b}$$

where $B(t)$ is a scalar Brownian motion defined on the complete probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$ with filtration $\{\mathcal{F}_t\}_{t\geq 0}$. We use $a \wedge b$ to denote $\min(a, b)$, $a \vee b$ to denote $\max(a, b)$ and a.s. to denote almost surely.

2.5.1 Analysis

We first show existence and uniqueness of a global solution remains in $\mathcal{D} = (0, K)\times(\beta, \beta + u_m)$ where $u_m = \max_{u\in(0,K)}\{\frac{u}{1+u^2}\}$, whenever it starts in $\mathcal{D}$. Because the drift term and noise term in (2.12) do not satisfy linear growth condition, the general existence and uniqueness theorem (see (Mao, 2007) chapter 2) does not apply. The techniques employed here are standard and follow the similar lines as (Wang *et al.*, 2017; Gray *et al.*, 2011).

Theorem 2.5.1. *For any given initial value $(u(0), v(0)) \in \mathcal{D}$, there is a unique solution defined for all $t \geq 0$, and the solution remains in $\mathcal{D}$ a.s.*

Proof. Since the drift is locally Lipschitz continuous, there is a unique local solution on $t \in [0, \tau_e)$ where τ_e is the explosion time, i.e. when the solution exits $\mathcal{D}$. If we can show that $\tau_e = \infty$ a.s., then the solution stays in $\mathcal{D}$ a.s. for all $t > 0$.

Let $\mathcal{D}_n = (\frac{1}{n}, K - \frac{1}{n}) \times (\beta + \frac{1}{n}, \beta + u_m - \frac{1}{n})$ where $n \in \mathbb{N}^+$. Define the stopping time

$$\tau_n = \inf\{t \in (0, \tau_e) : (u(t), v(t)) \notin \mathcal{D}_n\}.$$

Clearly τ_n increases as n. Let $\tau_\infty = \lim_{n\to\infty} t_n$ and we have $\tau_e \geq \tau_\infty$. If we can show that $\tau_\infty = \infty$ a.s., then $\tau_e = \infty$ a.s. and we have the unique solution $(u(t), v(t)) \in \mathcal{D}$ for all t.

Suppose on contrary that there is a pair of constants $T > 0$ and $\epsilon \in (0, 1)$ such that $P(\tau_\infty \leq T) > \epsilon$. Then there is a n_1 such that $P(\tau_n \leq T) \geq \epsilon$ for $n > n_1$. Define $V : \mathcal{D} \to \mathbb{R}$ by

$$V(u, v) = \frac{1}{u} + \frac{1}{K - u} + \frac{1}{v - \beta} + \frac{1}{\beta + u_m - v}, \tag{2.13}$$

where we note it is nonnegative for $u, v \in \mathcal{D}$.

In general, given an Itô process $dX(t) = b(X(t)dt + \sigma(X_t)dB(t)$ and if $f \in C_0^2(\mathbb{R}^n)$, then its generator A acting on f gives

$$Af(x) = \sum_i b_i(x)\frac{\partial f}{\partial x_i} + \frac{1}{2}\sum_{i,j}(\sigma\sigma^T)_{i,j}(x)\frac{\partial^2 f}{\partial x_i \partial x_j}$$

(see Theorem 7.3.3 on Page 126 of (Øksendal, 2003)).

Consider the generator A of (2.12) and let A act on V, i.e., $AV : \mathcal{D} \to \mathbb{R}$. Some

calculation shows that

$$\begin{aligned} AV &= \begin{pmatrix} \rho u(1-u/K) - \gamma uv \\ \\ \beta - v + u/(1+u^2) \end{pmatrix} \cdot \begin{pmatrix} -1/u^2 + 1/(K-u)^2 \\ \\ -1/(v-\beta)^2 + 1/(\beta + u_m - v)^2 \end{pmatrix} \\ &\quad + \frac{\sigma^2}{2} u^2 (1 - \frac{u}{K})^2 (\frac{2}{u^3} + \frac{2}{(K-u)^3}) \\ &= (\gamma v - \rho(1 - \frac{u}{K})) \frac{1}{u} + \frac{\rho u}{K} \frac{1}{K-u} - \frac{\gamma uv}{(K-u)^2} \\ &\quad + (\beta - v + \frac{u}{1+u^2}) \frac{1}{(\beta + u_m - v)^2} + (v - \beta - \frac{u}{1+u^2}) \frac{1}{(v-\beta)^2} \\ &\quad + (1 - \frac{u}{K})^2 \frac{1}{u} + \frac{u^2}{K^2} \frac{1}{K-u} \\ &\leq CV, \end{aligned}$$

where $C = \frac{\rho}{k} \vee \gamma(\beta + u_m) \vee (\gamma(\beta + u_m) + 1)$ a positive constant. By Dynkin's formula (see, e.g., (Øksendal, 2003) Page 127), we have

$$\begin{aligned} E[V(u(\tau_n \wedge T), v(\tau_n \wedge T))] &= V(u(0), v(0)) + E[\int_0^{\tau_n \wedge T} AV(u(s), v(s) ds] \\ &\leq V(u(0), v(0)) + C \int_0^T E[V] ds. \end{aligned}$$

By Gronwall's inequality,

$$E[V(u(\tau_n \wedge T), v(\tau_n \wedge T))] \leq V(u(0), v(0)) e^{CT} < \infty.$$

On the other hand, consider the set $\Omega_n = \{\tau_n \leq T\}$ for $n > n_1$ for which we know $P(\Omega_n) \geq \epsilon$. For $\omega \in \Omega_n$, we know that $V \sim O(n)$. Thus

$$\begin{aligned} E[V(u(\tau_n \wedge T), v(\tau_n \wedge T))] &\geq E[\mathbf{1}_{\Omega_n} V(u(\tau_n \wedge T), v(\tau_n \wedge T))] \\ &= \epsilon V(u(\tau_n, \omega), v(\tau_n, \omega)) \to \infty, \end{aligned}$$

if we let $n \to \infty$. Hence we have a contradiction. □

Having established existence and uniqueness of a global solution, we present a theorem regarding to tumor extinction. Its proof involves an application of Itô's formula (see, e.g., (Øksendal, 2003) Page 49) to show $\limsup_{t\to\infty} \frac{1}{t}\log(u(t)) < 0$ a.s., namely, the tumor population tends to zero exponentially a.s..

Theorem 2.5.2. (tumor extinction) *If one of the conditions holds*

i) $\sigma^2 > \rho \vee \frac{\rho^2}{2\gamma\beta}$

ii) $\sigma^2 < \rho < \sigma^2/2 + \gamma\beta$

then for $(u(0), v(0)) \in \mathcal{D}$

$$\limsup_{t\to\infty} \frac{1}{t}\log(u(t)) < 0 \ \textit{a.s.}$$

Proof. By Itô's formula, we have

$$\begin{aligned}
\log(u(t)) &= \log(u_0) + \int_0^t \rho(1 - \frac{u(s)}{K}) - \gamma v(s) - \frac{\sigma^2}{2}(1 - \frac{u(s)}{K})^2 ds \\
&\quad + \int_0^t \sigma(1 - \frac{u(s)}{K}) dB(s) \\
&\leq \log(u_0) + \int_0^t f(1 - \frac{u(s)}{K}) ds + \int_0^t \sigma(1 - \frac{u(s)}{K}) dB(s)
\end{aligned}$$

where $f(y) = -\frac{1}{2}\sigma^2 y^2 + \rho y - \gamma\beta$. Note that $f(y)$ is a concave-down quadratic function with axis of symmetry $y = \frac{\rho}{\sigma^2}$ and $y(0) < 0$. There are two ways to ensure $\max_{y\in(0,1)} f(y) < 0$, i.e.,

$$\begin{cases} \frac{\rho}{\sigma^2} < 1 \\ f(\frac{\rho}{\sigma^2}) < 0, \end{cases} \quad \text{or} \quad \begin{cases} \frac{\rho}{\sigma^2} > 1 \\ f(1) < 0 \end{cases}$$

which corresponds to the stated conditions i) or ii). Thus,

$$\begin{aligned}
\limsup_{t\to\infty} \frac{1}{t}\log(u(t)) &\leq \max_{y\in(0,1)} f(y) + \limsup_{t\to\infty} \frac{1}{t}\int_0^t \sigma(1 - \frac{u(s)}{K}) dB(s), \\
&< 0 \text{ a.s.}
\end{aligned}$$

where the term with integral is zero a.s. by law of large numbers of martingale (see, e.g., (Mao, 2007)). □

To complement with the condition for tumor extinction, we present a theorem for tumor persistence in the following. The proof involves construction of a contradiction to show that $\limsup_{t\to\infty} u(t)$ is bounded from below, and again Itô's formula is the workhorse behind this type of proof.

Theorem 2.5.3. (tumor persistence) *If $\rho > \sigma^2/2 + \gamma(\beta + u_m)$, then for any given initial value $(u(0), v(0)) \in \mathcal{D}$, we have*

$$\limsup_{t\to\infty} u(t) > \xi$$

a.s. where ξ is the positive root of $g(u) = -\frac{1}{2}\sigma^2(1-\frac{u}{K})^2 + \rho(1-\frac{u}{K}) - \gamma(\beta+u_m)$.

Proof. Assume $\rho > \sigma^2/2 + \gamma(\beta + u_m)$. We note that $g(K) < 0$ and $g(0) = -\frac{1}{2}\sigma^2 + \rho - \gamma(\beta + u_m) > 0$. Thus ξ exists and $g'(u) < 0$ for u in a small neighborhood of ξ. Suppose on contrary that there is a small $\epsilon \in (0,1)$ such that $P(\Omega_1) > \epsilon$ where $\Omega_1 = \{\omega : \limsup_{t\to\infty} u(t,\omega) < \xi - 2\epsilon\}$. Then there is $T(\omega) > 0$ such that

$$u(t,\omega) < \xi - \epsilon \text{ for } t > T(\omega). \tag{2.14}$$

We choose ϵ small enough so that $g(u(t,\omega) > g(\xi-\epsilon)$ for $t > T(\omega)$. Moreover, by law of large numbers of martingale there is $\Omega_2 \subset \Omega$ with $P(\Omega_2) = 1$ such that for any $\omega \in \Omega_2$, $\lim_{t\to\infty} \frac{1}{t}\int_0^t \sigma(1-\frac{u(s)}{K})dB(s,\omega) = 0$.

Now considering $\omega \in \Omega_1 \cap \Omega_2$, we have

$$\begin{aligned}
\log(u(t,\omega)) &\geq \log(u_0) + \int_0^t g(u(s))ds + \int_0^t \sigma(1-\frac{u(s)}{K})dB(s,\omega) \\
&\geq \log(u_0) + \int_0^{T(\omega)} g(u(s)ds + (t-T(\omega))g(\xi-\epsilon) \\
&\quad + \int_0^t \sigma(1-\frac{u(s)}{K})dB(s,\omega).
\end{aligned}$$

Thus we have

$$\liminf_{t\to\infty} \frac{1}{t}\log(u(t,\omega)) \quad \geq \quad g(\xi-\epsilon) > 0,$$

which implies that $u(t,\omega) \to \infty$, which contradicts (2.14). □

We compare the conditions in the previous two theorems to their counterparts in section 3. In the original units, we find two noise magnitude regimes with $\frac{\hat{\rho}}{\hat{\mu}}$ as the threshold (we call $\sigma^2 < \frac{\hat{\rho}}{\hat{\mu}}$ small noise regime and $\sigma^2 > \frac{\hat{\rho}}{\hat{\mu}}$ big noise regime). In the small noise regime, the tumor extinction condition is weaker than the one for the deterministic system. The biological interpretation is that the small noise can possibly help eliminate the tumor. The condition for tumor persistence is stronger in the stochastic system. However, it is only a sufficient condition and there is a gap between between the persistence condition and the extinction condition.

2.5.2 Numerical Simulation

(2.12) is simulated using Milstein's method (Higham., 2001). As in Figure 2.3, the simulation confirmed our extinction criteria. Also shown is that the tumor persists in the monostability where the solution simply fluctuates about the equilibrium and bistablity where there is a noise-induced transition to high tumor equilibrium from low tumor equilibrium. The parameter values used in the simulations are summarized in Table 2.1.

2.6 With Delay and Noise

In this section, we study the system when both delay and noise are present as follows

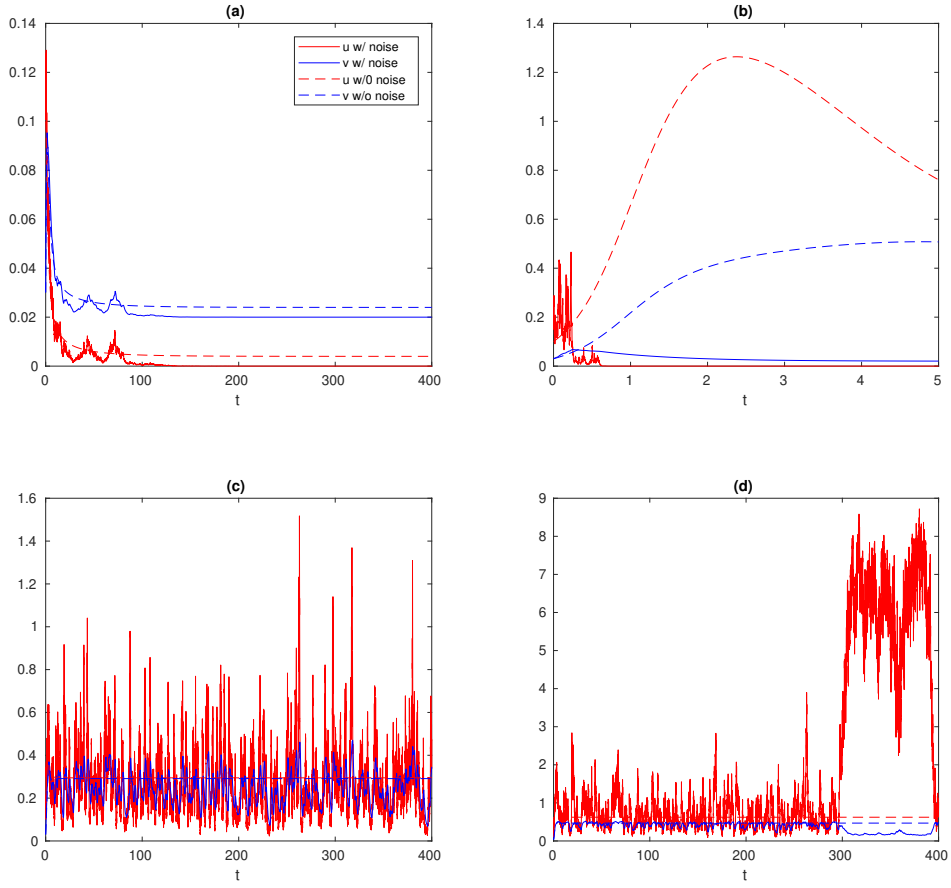

Figure 2.3: Computer simulated sample paths of the stochastic system (2.12) in comparison with its deterministic version (2.4). (a) tumor extinction in small noise regime; (b) tumor extinction in big noise regime; (c) monostable fluctuation; (d) bistable switching. Parameter values used here are summarized in Table 2.1

Table 2.1: Parameter values used in Figure 2.3

	small noise induced extinction (a)	big noise induced extinction (b)	monostable (c)	bistable (d)
ρ	0.12	2.5	1.5	2.5
β	0.02	0.02	0.02	0.02
γ	5	5	5	5
K	10	10	10	10
σ	0.3	5.66	0.65	0.65

$$du = (\rho u(1 - \frac{u}{K}) - \gamma uv)dt + \sigma u(1 - \frac{u}{K})dB(t), \tag{2.15a}$$

$$dv = (\beta - v + \frac{u(t-\tau)}{1 + u(t-\tau)^2})dt. \tag{2.15b}$$

2.6.1 Analysis

The analysis in section 4 can be extended to (2.15). Because of the time delay τ, here we need to supply suitable initial data on $[-\tau, 0]$. The extension of Theorem 2.5.1 is stated as follows

Theorem 2.6.1. *For any initial data* $\{(u(t), v(t)) : -\tau \leq t \leq 0\} \in C([-\tau, 0], \mathcal{D})$, *there is a unique solution defined for all* $t \geq -\tau$ *and remains in* $\mathcal{D}$ *a.s.*

Proof. The proof follows similar lines as the one for Theorem 2.5.1 with an application of method of steps. We keep the same definition of $\tau_n, \tau_\infty, \tau_e$ and $V(u,v)$. Same as before we want to show that $\tau_e = \infty$ by proving that $\tau_\infty = \infty$. First we want to show that $\tau_\infty > \tau$. For any $n \in \mathbb{N}^+$ and $t \in [0, \tau_\infty)$, it can be shown by Itô formula that

$$dV(u(t),v(t)) = LV(u(t),v(t),u(t-\tau),v(t-\tau))+\sigma(\frac{u}{K(K-u)}+\frac{1}{K}-\frac{1}{u})dB(t) \quad (2.16)$$

where

$$\begin{aligned}
&LV(u(t),v(t),u(t-\tau),v(t-\tau))\\
=&[\gamma v-\rho(1-\frac{u}{K})+(1-\frac{u}{K})^2]\frac{1}{u}+[\frac{\rho u}{K}+\frac{u^2}{K^2}]\frac{1}{K-u}+\frac{1}{\beta+u_m-v}+\frac{1}{v-\beta}\\
&-[u_m-\frac{u(t-\tau)}{1+u(t-\tau)^2}]\frac{1}{(\beta+u_m-v)^2}-[\frac{u(t-\tau)}{1+u(t-\tau)^2}]\frac{1}{(\beta-v)^2}\\
\leq& K_1 V
\end{aligned}$$

where K_1 is a positive constant. For $t_1 \in [0,\tau]$, integrating (2.16) from 0 to $t_1 \wedge \tau_n$ and taking expectation gives

$$E[V(u(t_1\wedge\tau_n),v(t_1\wedge\tau_n)] \;\leq\; V(u(0),v(0)) + \int_0^{t_1\wedge\tau_n} V(u(s),v(s))ds$$

By Gronwall's inequality, $E[V(u(t_1 \wedge \tau_n), v(t_1 \wedge \tau_n)] \leq V(u(0),v(0))e^\tau < \infty$. In particular,

$$E[V(u(\tau\wedge\tau_n),v(\tau\wedge\tau_n)] \leq V(u(0),v(0))e^\tau < \infty. \quad (2.17)$$

Suppose on contrary there is $\epsilon > 0$ such that $P(\tau_\infty < \tau) > \epsilon$. Then there is n_1 such that for $n > n_1$, $P(\Omega_1) \geq \epsilon$ where $\Omega_1 = \{\omega : \tau_n(\omega) < \tau\}$. Thus

$$\begin{aligned}
E[V(u(\tau\wedge\tau_n),v(\tau\wedge\tau_n)] &\geq E[\mathbf{1}_{\Omega_1}V(u(\tau_n),v(\tau_n)]\\
&\geq \epsilon n \to \infty
\end{aligned}$$

as $n \to \infty$, which is a contradiction to 2.17. Thus for $\tau_\infty \geq \tau$. By the same argument on $t \in [\tau, 2\tau]$,we have $\tau_\infty \geq 2\tau$. Repeating this procedure gives $\tau_\infty = \infty$. □

It is easy to see that $v(t) \in (\beta, \beta + \frac{1}{2})$ for all $t > 0$. This enables Theorem 2.5.2 and Theorem 2.5.3 be extended to the system (2.15) with similar arguments.

2.6.2 Numerical Simulation

From previous analysis, we notice that the system (2.15) bears a lot of similarity to the system (2.12). In this subsection, we instead focus on the effect of time delay on stationary distributions by numerical simulations. As seen in Figure 2.4, the stationary distribution is bimodal which is not a surprise since the underlying deterministic system is bistable. As the delay increases (larger τ), the more density shifts to the high tumor stable state and the bimodality becomes indistinguishable at $\tau = 2$. We also observe that the mean first passage time from the low tumor stable state to the high tumor stable state is reduced by increasing τ.

2.7 Discussion and Conclusion

In this paper, we presented a simple model of tumor-immune system interactions, in which the immune response to tumor is modeled as a non-monotonic function of tumor burden. We studied the effects of time delay in the immune response and the uncertainty in the innate proliferating rate of cancer cells on the tumor growth dynamics. The conditions of tumor extinction and persistence were proved in presence of noise or time delay. We also performed numerical experiments to confirm analytical results and to guide future analytical work.

We showed that the magnitude of the tumor reproduction number $R_0 = \frac{\hat{\mu}\hat{\rho}}{\hat{\gamma}\hat{\beta}}$ relative to 1 dictates the stability of tumor-free equilibrium. This condition does not depend on time delay. It suggests that the elimination of cancer depends on the basal level of the immune system rather than on its response speed to tumor growth. However, it maybe possible to have delay-induced stability switching for the low-tumor steady

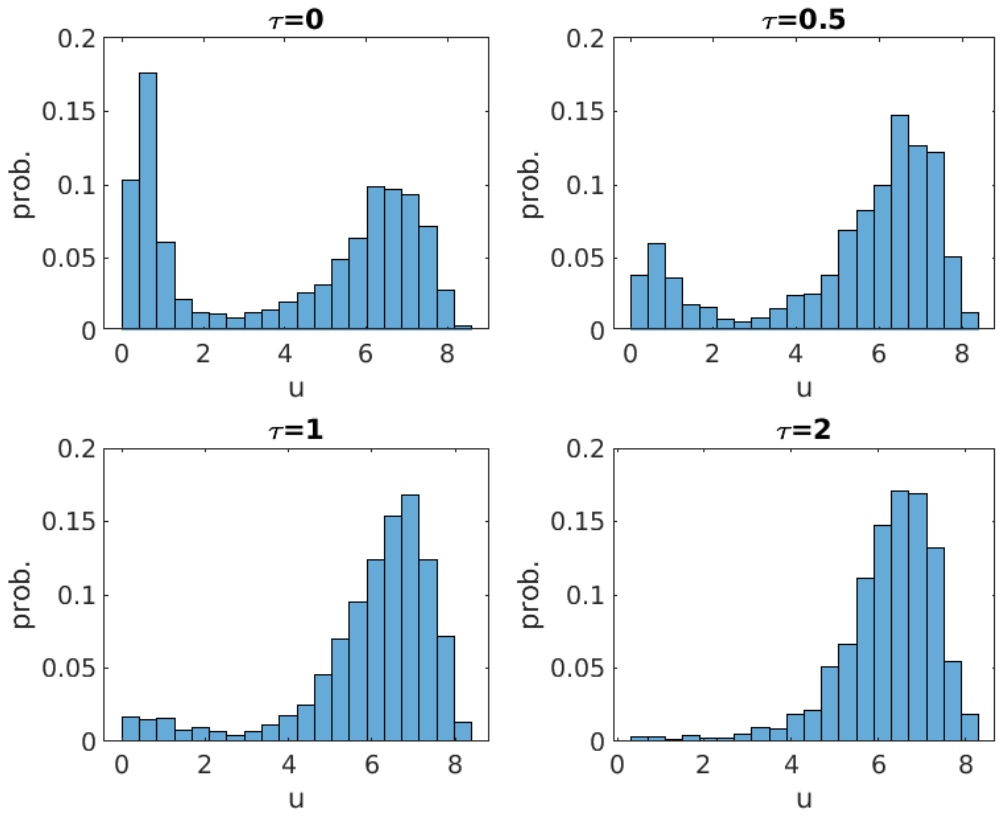

Figure 2.4: stationary distribution of (2.15) with $\tau = 0, 0.5, 1, 2$. The histogram is formed by 5000 samples of $u(1000)$.

state. We also established the global stability of tumor-free equilibrium, but that of interior equilibria is challenging and remains an open problem.

For stochastic version of the model, we showed that the noise can help eliminate cancer in either case of big or small noise. Similar results in an epidemic model were obtained by (Gray *et al.*, 2011). We note that the criteria for persistence in Theorem 2.5.3 is not a necessary condition and in fact far from being a sharp result. Indeed, parameter sets (c) (d) do not satisfy the persistence condition but appear to be persistent in the numerical simulation. The improvement of the current result will be deferred to future work. Also, as seen in the simulation, there is a stationary distribution of (2.12). However, it is challenging to show this analytically because

the diffusion matrix is degenerate, which makes the standard techniques employed in (Wang *et al.*, 2017; Gray *et al.*, 2011) not applicable. There have been recent studies on the stationary distribution resulting from a degenerate diffusion matrix (Lan *et al.*, 2018). It will also be the focus of our future work.

When including time delay with noise, we showed that the results we obtained for noise-only system (2.12) can be easily extended to (2.15). Moreover, we found numerically that the noise favors transition to high tumor stable state, which was also observed for an ecological model studied in (Zeng and Wang, 2012). The biological implication is that the less responsive the immune system is, the easier for the tumor to escape to high burden. Same as for the noise-only version of the model, the existence of a stationary distribution of (2.15) will be the focus of future work. There are some perturbation techniques applied to delayed Focker-Planck equation to study the effects of time delay on mean first passage time (Küchler and Mensch, 1992; Guillouzic *et al.*, 1999; Frank, 2005). Those techniques are limited to a scalar equation. Possible extension to study a system of equations will be carried out in the future.

Time delay and stochasticity are two hallmarks of cancer dynamics. Serious modeling work shall not shy away from them. The model we studied was kept minimal but nevertheless exhibited interesting dynamics. It is hoped that this paper will stimulate interests in stochastic delayed differential equations in modeling cancer dynamics.

Chapter 3

PATIENT-SPECIFIC PARAMETER ESTIMATES OF GLIOBLASTOMA MULTIFORME GROWTH DYNAMICS FROM A MODEL WITH EXPLICIT PROLIFERATION AND DEATH RATES

3.1 Abstract

Glioblastoma multiforme (GBM) is an aggressive primary brain cancer with a grim prognosis. Its morphology is heterogeneous, but prototypically consists of an inner, largely necrotic core surrounded by an outer, contrast-enhancing rim, and often extensive tumor-associated edema beyond. This structure is usually demonstrated by magnetic resonance imaging (MRI). To help relate the three highly idealized components of GBMs (i.e., necrotic core, enhancing rim, and maximum edema extent) to the underlying growth "laws," a mathematical model of GBM growth with explicit motility, proliferation, and death processes is proposed. This model generates a traveling-wave solution that mimics tumor progression. We develop several novel methods to approximate key characteristics of the wave profile, which can be compared with MRI data. Several simplified forms of growth and death terms and their parameter identifiability are studied. We use several test cases of MRI data of GBM patients to yield personalized parameterizations of the model, and the biological and clinical implications are discussed.

3.2 Introduction

Glioblastoma multiforme (GBM) is a highly aggressive primary brain cancer, with median survival time from diagnosis on the order of 15 months; long-term survival is

extremely rare (Norden and Wen, 2006). Such rapid progression is promoted by highly proliferative and diffusely invasive cancer cells, which makes complete surgical removal impossible. Magnetic resonance imaging (MRI) is conventionally used to identify the location and characteristics of the tumor pre-operatively, to guide surgery, and to monitor and track progression and treatment response. Perioperatively, MRI is used to guide the resection of the tumor mass, to assess post-operatively the volume of tumor resected, and to target other adjunct treatment such as radiation therapy.

GBMs morphologically typically appear (at least at initial diagnosis) as roughly spherical but highly heterogeneous masses that often exhibit a (crudely speaking) three-layer structure. Within the tumor there is usually extensive cell necrosis, often accompanied by tumor cells, and a cystic component as well. An outer region, which typically appears as contrast-enhancing on T1-weighted gadolinium contrast-enhanced MRI, is cytologically typified by proliferating cells that then infiltrate into surrounding brain tissue. The surrounding brain tissue is generally seen to be edematous on T2-weighted or T2-FLAIR MRI and at surgery, due to vasogenic edema. Prior statistical analyses have found that edema is a prognostic indicator of patient survival (Pope *et al.*, 2005), but the relationship is complex and appears to be mediated by the expression of vascular endothelial growth factor and the activity of related angiogenic genes (Carlson *et al.*, 2007) and various autocrine factors (Hoelzinger *et al.*, 2007).

The standard of care for GBM patients was largely established by the 2005 clinical trial by Stupp *et al.* (2005). It comprises maximal surgical resection of the primary tumor, followed by six weeks of radiation to the gross tumor volume, plus a 2–3 cm margin, with concomitant oral temozolamide (TMZ), and 6–12 months of maintenance TMZ chemotherapy (Gilbert *et al.*, 2013). Maximal surgical resection appears to offer some survival benefit. Nevertheless, the absolute survival benefit of even the

most effective therapy is typically on the order of months. The highly infiltrative nature of GBMs makes recurrence nearly inevitable, even with maximal resection and aggressive adjuvant therapy, although individual tumors vary in their degree of invasiveness.

Given the grim situation, mathematical modeling has been proposed as a method to better understand the biophysical rules underlying GBM growth, with the ultimate goal to provide more effective therapy. Mathematical models have been widely applied to a variety of cancers and to cancer treatment in general (Kuang *et al.*, 2015), and GBM is the focus of many such works (see (Martirosyan *et al.*, 2015) for a review). A popular class of cancer models takes the form of a system of reaction-diffusion equations. In many cases (Harley *et al.*, 2014; Gerlee and Nelander, 2016; Stepien *et al.*, 2018), such systems generate a traveling-wave solution, with the traveling-wave speed of great interest, as it is an indicator of how fast the cancer progresses.

Variants of the Fisher-Kolmogorov equation (Fisher, 1937), originally introduced in the 1930s, were first suggested (to our knowledge) as models for GBM growth by J. D. Murray and coworkers. The Fisher-Kolmogorov model is given by

$$\frac{\partial c}{\partial t} = \nabla \cdot (D \nabla c) + \rho c \left(1 - \frac{c}{K}\right), \tag{3.1}$$

where $c(x,t)$ is the cancer cell density at location x and time t, D is a diffusion coefficient, ρ is the intrinsic tumor cell growth rate, and K is the local carrying capacity. (Variants include a linear version that replaces the logistic growth term with a simple exponential growth rate, ρc.) Murray and coworkers have used them to explore the effect of chemotherapy (Tracqui *et al.*, 1995), to quantify patients' survival as a function of the extent of surgical resection (Woodward *et al.*, 1996), and to estimate the time of tumor initiation (Murray, 2012).

Other authors have suggested that the net growth and diffusion parameters of

model (3.1) may be estimated by image differencing when two sequential, pre-treatment patient MR series are available (Swanson *et al.*, 2008; Neal *et al.*, 2013; Jackson *et al.*, 2015). Such a procedure is problematic, however, because changes in the tumor in images taken a few days or weeks apart tend to be small and are convolved with image co-registration errors (van der Hoorn *et al.*, 2016). Some patients may be treated with steroids following initial diagnosis to reduce tumor-related edema and resulting neurological symptoms, which may alter the brain geometry and imaging appearance of the tumor at subsequent times (Watling *et al.*, 1994; Zaki *et al.*, 2004).

Model (3.1) can yield a dense tumor core with an advancing front, but it cannot capture the heterogeneity between live and necrotic tumor cells, as it assumes that all cells are equally viable. While several modeling efforts have taken into account various proliferating, migrating, and necrotic cell components (e.g., (Eikenberry *et al.*, 2009; Swanson *et al.*, 2011)), they are too complicated to be reliably parameterized by the limited number of patient MRI series in typical clinical cases. The motivation for this work is to extend model (3.1) to include necrotic cells in a simplified way, such that patient-specific model parameters can be estimated from suitable measurements of MR images acquired at a single time point.

T1-weighted MRI sequences of GBM often show a partially necrotic core surrounded by a bright enhancing rim that correlates with high blood vessel density and, presumably, with rapid cell proliferation. Neurosurgical and biopsy studies indicate that this core and rim are usually surrounded by a large expanse of edema, which is best visualized on T2-weighted MRI and has been found to correspond with a component of diffusely invasive GBM cells (Claes *et al.*, 2007). By approximating the tumor as a sphere, we may be able to identify three *idealized* digital marks from imaging: necrotic radius, enhancing radius, and what we shall call the "T2" or "maximum" radius. We hypothesize that a relatively simple mathematical model framework can

capture all these three digital marks and yield insights into the relative contributions of cellular proliferation, motility, and necrosis to the observed image features.

The next section describes our model and its assumptions. We demonstrate that the model has a traveling-wave solution and present the approximate wave profile. We describe a simple procedure to estimate patient-specific parameters by fitting the approximate wave profile to a tumor profile derived from patient MRIs. The identifiability of the model parameters is also discussed. We apply this parameter estimation procedure to obtain the key model parameters (consisting of the rate of cancer cell proliferation, death, and diffusion) for several patients.

3.3 Model and Method

3.3.1 Model Description

Our proposed model of the growth of GBM is a system of reaction-diffusion equations:

$$\frac{\partial p}{\partial t} = \nabla \cdot \left[\left(\frac{Dp}{p+q} \right) \nabla (p+q) \right] + \tilde{g}(w)p - \tilde{\delta}(w)p, \tag{3.2a}$$

$$\frac{\partial q}{\partial t} = \nabla \cdot \left[\left(\frac{Dq}{p+q} \right) \nabla (p+q) \right] + \tilde{\delta}(w)p, \tag{3.2b}$$

where

$$w = 1 - p - q, \tag{3.3}$$

and $p(x,t)$ and $q(x,t)$ represent the proliferating and quiescent cell densities at time t and location x, respectively; quiescent cells are functionally equivalent to necrotic cells in this framework. We assume that the flux of total population due to migration is $-D\nabla(p+q)$, where D is a constant diffusion coefficient. It is further assumed that the proportion of the total flux contributed by each cell type equals its proportion of the total population. This form of diffusion was used in (Sherratt and Chaplain, 2001)

to account for the key property of contact inhibition in cancer cell movement, with the underlying assumption that the two cell populations move together with equal motility, unaffected by necrotic cells. This type of model has successfully captured the structure of a growing tumor.

The per capita proliferation rate is $\tilde{g}(w)$; proliferating cells become quiescent at the per capita rate $\tilde{\delta}(w)$, where w (Eq. 3.3) represents the availability of space or some generic nutrient, which we will call *growth factor* henceforth. We have scaled the maximum cell density to be 1. In our model, necrosis is not explicitly included but can be regarded as being lumped into q. Insofar as quiescent cells cannot become proliferative, $\tilde{\delta}(w)$ can be viewed as a functional death rate. Our motivation in keeping the model framework relatively simple is to be able to estimate model parameters directly from clinical MRI imaging that is sparse in time.

To make the model biologically reasonable, we impose the following constraints on $g(w)$ and $\delta(w)$:

$$\tilde{g}'(w) \geq 0, \quad \tilde{\delta}'(w) \leq 0, \quad \tilde{g}(1) \geq \tilde{\delta}(1) = 0, \quad \tilde{\delta}(0) > \tilde{g}(0) = 0. \tag{3.4}$$

That is, proliferation (death) should increase (decrease) with the availability of the growth factor; there is more proliferation than death at maximum values of the growth factor; and there is only death with no growth in the absence of growth factor. It is also assumed that the death rate is negligible at maximum values of the growth factor. With these assumptions, we observe numerically that with suitable initial conditions, the solution of (3.2) stays positive and is bounded ($p + q \leq 1$) for all t.

We can estimate only up to three parameters based on the necrotic, enhancing, and maximum radii to be measured from MRI images. Therefore, we place a few more restrictions on $\tilde{g}(w)$ and $\tilde{\delta}(w)$ to simplify the estimation of model parameters and to ensure their identifiability. We assume that the proliferation rate at maximum growth

factor is ρ and that the death rate at zero growth factor is k and incorporate these parameters into $\tilde{g}$ and $\tilde{\delta}$, respectively; that is, $\tilde{g}(w=1;\rho)=\rho$ and $\tilde{\delta}(w=0;k)=k$. For reasons that will become clear later, we pick a functional form that can be written as $\tilde{g}(w;\rho)=\rho g(w)$ and $\tilde{\delta}(w;k)=k\delta(w)$. Some examples include the cumulative distribution function of the beta distribution family (cf. the left pane of Figure 3.4). These additional assumptions impose little impact on the generality of our model. The benefit of including them will become clear in Section 3.3.3.

3.3.2 Approximate Wave Profile

In most biological applications of reaction-diffusion models, solutions take the form of traveling waves. MRI images of GBM cancer growth suggest that we can approximate the evolution of the tumor by a traveling-wave solution of its growth model. To uniquely identify and accurately approximate GBM growth model parameters, it is highly desirable to obtain some analytic approximation of the traveling wave, to enable computational matching of the image wave profile and the approximate model wave profile. For this purpose, we consider one spatial dimension, which suffices insofar as the tumor is approximately spherical, and, at the time of diagnosis, its radius is large enough so that radial effects are negligible. With these assumptions, model (3.2) takes the form

$$\frac{\partial p}{\partial t} = \frac{\partial}{\partial x}\left[\left(\frac{Dp}{p+q}\right)\frac{\partial}{\partial x}(p+q)\right] + \rho g(w)p - k\delta(w)p \tag{3.5a}$$

$$\frac{\partial q}{\partial t} = \frac{\partial}{\partial x}\left[\left(\frac{Dq}{p+q}\right)\frac{\partial}{\partial x}(p+q)\right] + k\delta(w)p. \tag{3.5b}$$

We nondimensionlize the system using the characteristic length $\sqrt{D/k}$ and the characteristic time $1/k$ so that $x = \sqrt{D/k}\,\hat{x}$ and $t = \hat{t}/k$, which leads to

$$\frac{\partial p}{\partial \hat{t}} = \frac{\partial}{\partial \hat{x}}\left[\left(\frac{p}{p+q}\right)\frac{\partial}{\partial \hat{x}}(p+q)\right] + \hat{\rho}g(w)p - \delta(w)p \tag{3.6a}$$

$$\frac{\partial q}{\partial \hat{t}} = \frac{\partial}{\partial \hat{x}}\left[\left(\frac{q}{p+q}\right)\frac{\partial}{\partial \hat{x}}(p+q)\right] + \delta(w)p, \tag{3.6b}$$

where $\hat{\rho} = \rho/k$. We seek a traveling wave solution of the form $p(\xi) = p(\hat{x} - c\hat{t})$, $q(\xi) = q(\hat{x} - c\hat{t})$, where c is the wave speed. Substituting these into (3.6) gives

$$\frac{d}{d\xi}\left[\left(\frac{p}{p+q}\right)\frac{d}{d\xi}(p+q)\right] + c\frac{dp}{d\xi} + \hat{\rho}g(w)p - \delta(w)p = 0 \tag{3.7a}$$

$$\frac{d}{d\xi}\left[\left(\frac{q}{p+q}\right)\frac{d}{d\xi}(p+q)\right] + c\frac{dq}{d\xi} + \delta(w)p = 0. \tag{3.7b}$$

Linearizing at the wave head, i.e., substituting the ansatz $p = Ae^{-r\xi}$ and $q = Be^{-r\xi}$ into (3.7), gives $(r^2 - cr + \rho)A = 0$. For a biologically realistic wave front, we expect $A > 0$, $B > 0$, and $r > 0$. This requires that $c^2 > 4\rho$, which implies that the minimum speed of the wave is $c_{\min} = 2\sqrt{\hat{\rho}}$. It is numerically verified that the minimum speed is exactly the asymptotic speed, i.e., $c = c_{\min}$.

To obtain an approximate wave profile, we adopt a method first used by Canosa (Canosa, 1973). We rescale the wave coordinate as $z = -\xi/c$, which leads to

$$\frac{1}{c^2}\frac{d}{dz}\left[\left(\frac{p}{p+q}\right)\frac{d}{dz}(p+q)\right] - \frac{dp}{dz} + \hat{\rho}g(w)p - \delta(w)p = 0, \tag{3.8a}$$

$$\frac{1}{c^2}\frac{d}{dz}\left[\left(\frac{q}{p+q}\right)\frac{d}{dz}(p+q)\right] - \frac{dq}{dz} + \delta(w)p = 0. \tag{3.8b}$$

Assuming that $1/c^2$ is small, we neglect each first term of (3.8). Writing the resulting system in terms of p and w, we obtain the reduced system

$$\frac{dp}{dz} = p\left(\hat{\rho}g(w) - \delta(w)\right), \tag{3.9a}$$

$$\frac{dw}{dz} = -p\,\hat{\rho}g(w), \tag{3.9b}$$

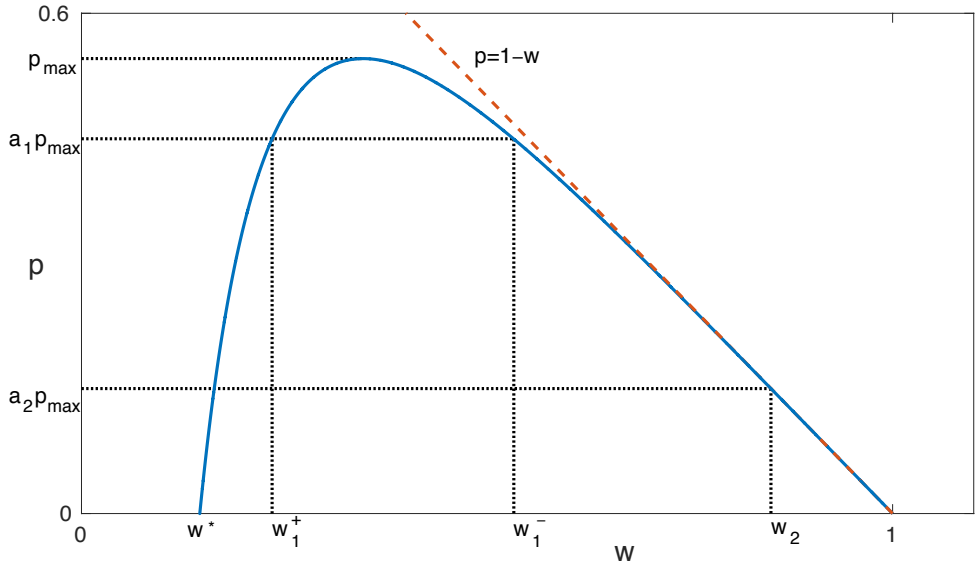

Figure 3.1: A typical trajectory that connects $(0, 1)$ and $(0, w^*)$ in the phase plane. Given $\delta(w)$ and $g(w)$, this trajectory can be found by integrating (3.10). It represents an approximate traveling-wave solution. See the appendix for a proof of its existence under general assumptions.

which is amenable to phase-plane analysis. The approximate wave solution corresponds to a trajectory that leaves $(0, 1)$ and ends at $(0, w^*)$, with $w^* \in [0, 1)$ (see Figure 3.1). (In the appendix, we show that such a trajectory exists, given the assumptions (3.4).) Dividing (3.9a) by (3.9b) yields

$$\frac{dp}{dw} = \frac{\delta(w)}{\hat{\rho}\, g(w)} - 1. \tag{3.10}$$

Upon integration, we obtain p as a function of w, which we will use in the next section.

3.3.3 Parameter Estimation

From clinical MRI data, we may derive three idealized radii: R_0, R_1, and R_2, representing respectively the radius of the inner necrotic core, the radius to the edge of the contrast-enhancing rim, and the radius to the outer edge of tumor-associated edema. Such data have been extracted from a series of anonymized patient MRI data

consisting of T1-contrast enhanced and T2-weighted MRIs at initial diagnosis. Using the publicly available MATLAB software package, Statistical Parametric Mapping 12 (SPM 12) (Penny *et al.*, 2007), MRIs are initially registered to a standard brain space, and then, using Slicer 3D (Fedorov *et al.*, 2012) software, the total necrotic core volumes, enhancing rim volumes, and tumor-associated edema volumes are determined from semi-manual tumor segmentation. Finally, these volumes are converted to radii assuming a spherical tumor geometry. The width of the proliferating rim, denoted as L_1, and the width of the edematous rim, denoted as L_2, can be calculated as $L_1 = R_1 - R_0$ and $L_2 = R_2 - R_1$, as demonstrated visually in Figures 3.2 and 3.3.

We assume that contrast-enhancing regions of T1-weighted images correspond to high densities of proliferating tumor cells and that edematous regions on T2-weighted imaging correspond to low densities. We denote the respective detection thresholds for T1 and T2 imaging as $a_1 p_{\max}$ and $a_2 p_{\max}$, where $0 < a_2 < a_1 < 1$ and $p_{\max} = \max_z p(z)$, i.e., the maximum density of proliferating cells given by the traveling-wave solution.

Often only a single MRI series is available before surgery, although in some cases, a diagnostic MRI followed some days or weeks later by a pre-surgery MRI may be available. In the latter case, the image-derived wave velocity V is the change in tumor radius divided by the length of the time interval.

From our approximate wave profile, we can compute the corresponding quantities to match with MR images (cf. Figure 3.3). The wave-solution based approximation of the width of the proliferating rim, denoted ℓ_1, and the width of the edematous rim, denoted ℓ_2, (in dimensional form) are computed as

$$\ell_1 = \frac{2\sqrt{D\rho}}{k} \int_{w_1^-}^{w_1^+} \frac{dz}{dw}\, dw, \tag{3.11a}$$

$$\ell_2 = \frac{2\sqrt{D\rho}}{k} \int_{w_2}^{w_1^-} \frac{dz}{dw}\, dw, \tag{3.11b}$$

respectively, where $w_1^{\pm}$ and w_2 satisfy, respectively, $p(w_1^{\pm}) = a_1 p_{\max}$ and $p(w_2) = a_2 p_{\max}$. Here $p(w)$ is obtained by integrating Eq. (3.10) (see appendix for details). Additionally, the model-derived wave speed $c = 2\sqrt{\rho D}$ can be matched with the image-derived speed V. Thus we have three nonlinear equations

$$\ell_1 = L_1, \quad \ell_2 = L_2, \quad c = V, \tag{3.12}$$

from which we hope to find the parameters D, ρ, and k. Given our assumptions, we can simply take the ratio of (3.11a) and (3.11b), which gives

$$f(\hat{\rho}) \equiv \frac{\displaystyle\int_{w_1^-}^{w_1^+} \frac{dz}{dw}\, dw}{\displaystyle\int_{w_2}^{w_1^-} \frac{dz}{dw}\, dw} = \frac{L_1}{L_2}, \tag{3.13}$$

insofar as the integrals are functions of $\hat{\rho}$. Equation (3.13) can be solved for $\hat{\rho}$ analytically in special cases or numerically in general. The monotonicity of $f(\hat{\rho})$ is important for the identifiability of parameters. Once we find $\hat{\rho}$, i.e., the ratio ρ/k, all parameters can be found by back substitution.

The above method requires two MR scans taken at two consecutive times prior to surgery to obtain an image-derived estimate of wave speed. If no second MR series is available, then tumor age may be estimated by the tumor radius divided by the wave speed. However, the estimate depends on which radius (R_1 or R_2) is used, because the tumor grows exponentially at first and linearly later on (Kuang *et al.*, 2015). This initial exponential growth stage needs to be taken into account as a correction to the aforementioned tumor age estimation. Suppose that for $0 \le t \le t^*$, quiescence is negligible and the proliferating cancer cells grow exponentially from a point source of density p_0, and that for $t > t^*$, the tumor grows as a traveling wave with speed $2\sqrt{\rho D}$. By equating the two age estimates, we obtain

$$\frac{R_1 - R_1^*}{2\sqrt{\rho D}} = \frac{R_2 - R_2^*}{2\sqrt{\rho D}}, \tag{3.14}$$

where

$$R_i^* = t^* \sqrt{4D\rho - \frac{4D}{t^*} \ln\left(\frac{a_i (4\pi D t^*)^{3/2}}{p_0}\right)}, \qquad i = 1, 2, \tag{3.15}$$

where R_1 and R_2 are respectively the T1 and T2 radii at $t = t^*$ (see details of R_1^* and R_2^* in the appendix).

Replacing the last equation in (3.12) with (3.14), we again have three equations. To solve them for the unknown parameters, we first take (3.11a) over (3.11b) as before to obtain (3.13). It can then be solved for the ratio $\hat{\rho} = \rho/k$. Substituting this expression back to either $\ell_1 = L_1$ or $\ell_2 = L_2$ gives the raito ρ/D. Finally, expressing D and k in terms of ρ, (3.14) can be solved for ρ, and D and k follow.

3.4 Results

Monotonicity is crucial for parameter identifiability, so we first investigate the monotonicity of f for some specific choices of $g(w)$ and $\delta(w)$. Given the restrictions described in Section 3.3.1, the cumulative density function (CDF) of the Beta distribution family suits our purposes. Therefore, we let $g(w) = B(w; \alpha_g, \beta_g)$ and $\delta(w) = 1 - B(w; \alpha_\delta, \beta_\delta)$, where $B(w; \alpha, \beta)$ is the CDF of the beta distribution with shape parameters α and β. By varying α and β, we can get linear, sigmoidal, and concave up/down curves (see the left pane of Figure 3.4). Our framework is robust to those choices, that is, the monotonicity of $f(\hat{\rho})$, defined by Eq. (3.13), is preserved (right pane of Figure 3.4). Sigmoidal-shaped growth and death functions (g and δ, respectively) may provide biologically realistic response functions to limited growth factors (most enzymatic reaction rates have sigmoidal shapes with respect to reactant concentration). Given this family of functions, we now consider the question of estimating patient-specific tumor growth and death rates from MR imaging.

We parameterize our model with patient data in which there is only one MRI scan before surgery. In Table 3.1, we summarize the image-derived tumor radii and

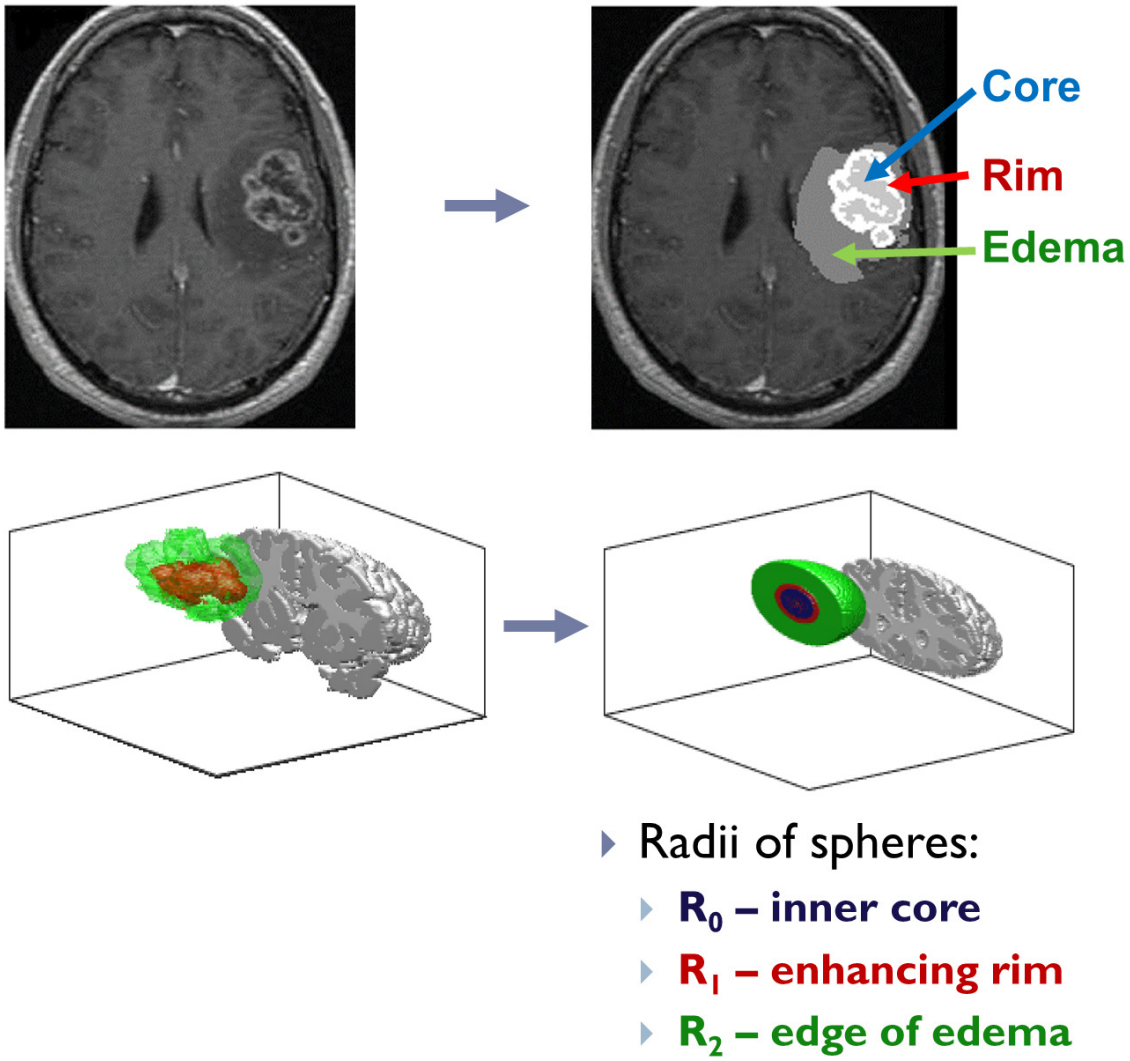

Figure 3.2: The top half of the figure shows an example patient MRI registered to the standard brain domain, with the three tumor segments, necrotic core, enhancing rim, and tumor-associated edema highlighted on a single 2-D slice. The full 3-D segmentation, and the equivalent tumor sphere with associated radii, R_0, R_1, and R_2, is shown in the lower half.

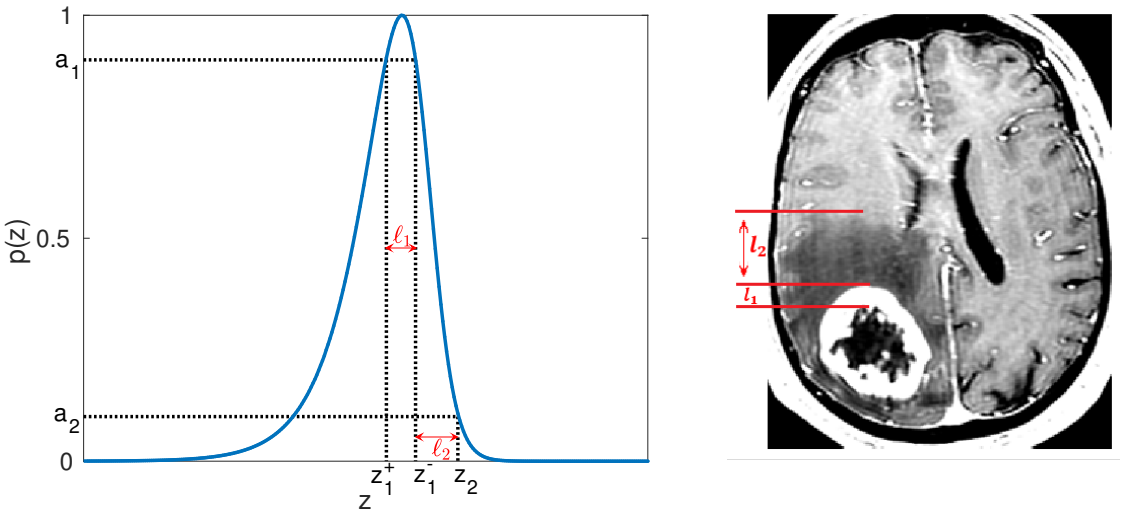

Figure 3.3: Left: normalized wave profile generated by the model in the z coordinate. Right: tumor profile seen in MR image. Parameter estimation is done by matching model-derived quantities, e.g., ℓ_1 and ℓ_2, to the corresponding image-derived ones.

the corresponding parameters estimated by the method introduced in the previous section. The parameters $a_1 = 0.9$ and $a_2 = 0.1$ are adapted from values found in the literature (Swanson *et al.*, 2008), while $p_0 = 0.02$ and $t^* = 60$ days are hypothetical values. The parameters vary considerably among individual patients.

We compare our approximate quantities to those obtained from the numerical solution of the model. As shown in Figure 3.5, the approximated results match well with the numerical results except for some discrepancy for L_2 when $\hat{\rho}$ is small. This result is not a surprise, because the approximation assumes that $c = 2\sqrt{\hat{\rho}}$ is large. Moreover, the numerical approximation of L_2 is prone to errors due to the fixed grid size and large rate of change around the threshold of L_2. Overall, we believe that our approximation is accurate for the parameter ranges estimated from the image data.

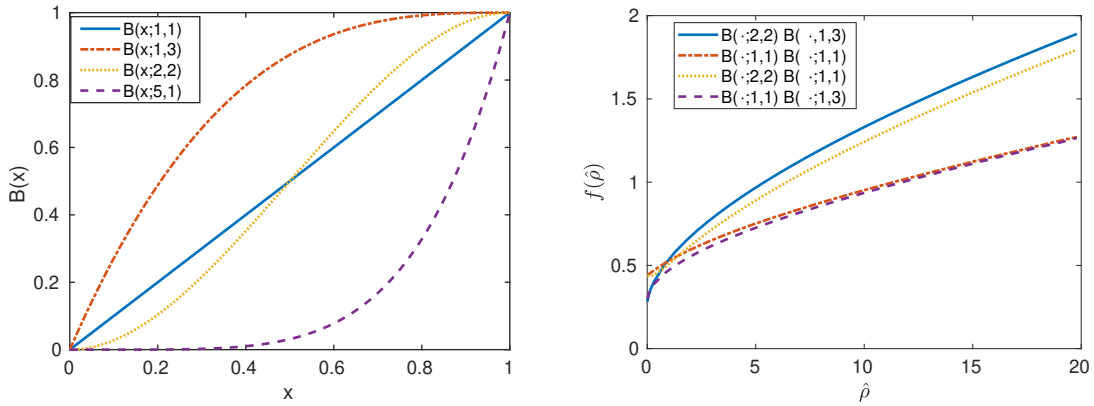

Figure 3.4: Left: Cumulative distribution functions of some beta distributions. These functions satisfy Eq. (3.4) and serve as candidates to represent biological response to limitation of growth factors. Right: monotonicity of $f(\hat{\rho})$ given different choices of $g(w)$ and $\delta(w)$ as indicated in the legend. All choices lead to a monotonic function $f(\hat{\rho})$ and hence identifiable parameters.

Table 3.1: Radii of equivalent tumor sphere derived from T1 and T2 images and the corresponding vital parameters estimated by our protocol. We have preset $a_1 = 0.9$, $a_2 = 0.1$, $p_0 = 0.02$ mm and $t^* = 60$ days.

Patient	R_0 (mm)	R_1 (mm)	R_2 (mm)	D (mm^{2}day^{-1})	ρ (day^{-1})	k (day^{-1})
1	14.87	20.73	27.77	0.2852	0.2102	0.0602
2	20.48	26.34	38.24	1.2791	0.2624	0.3537
3	6.61	10.91	15.24	0.0825	0.1736	0.0327
4	22.87	26.96	37.03	0.9825	0.2590	0.7819
5	8.17	14.20	25.10	0.9769	0.2520	0.2260
6	8.29	15.83	20.35	0.0687	0.1652	0.0106

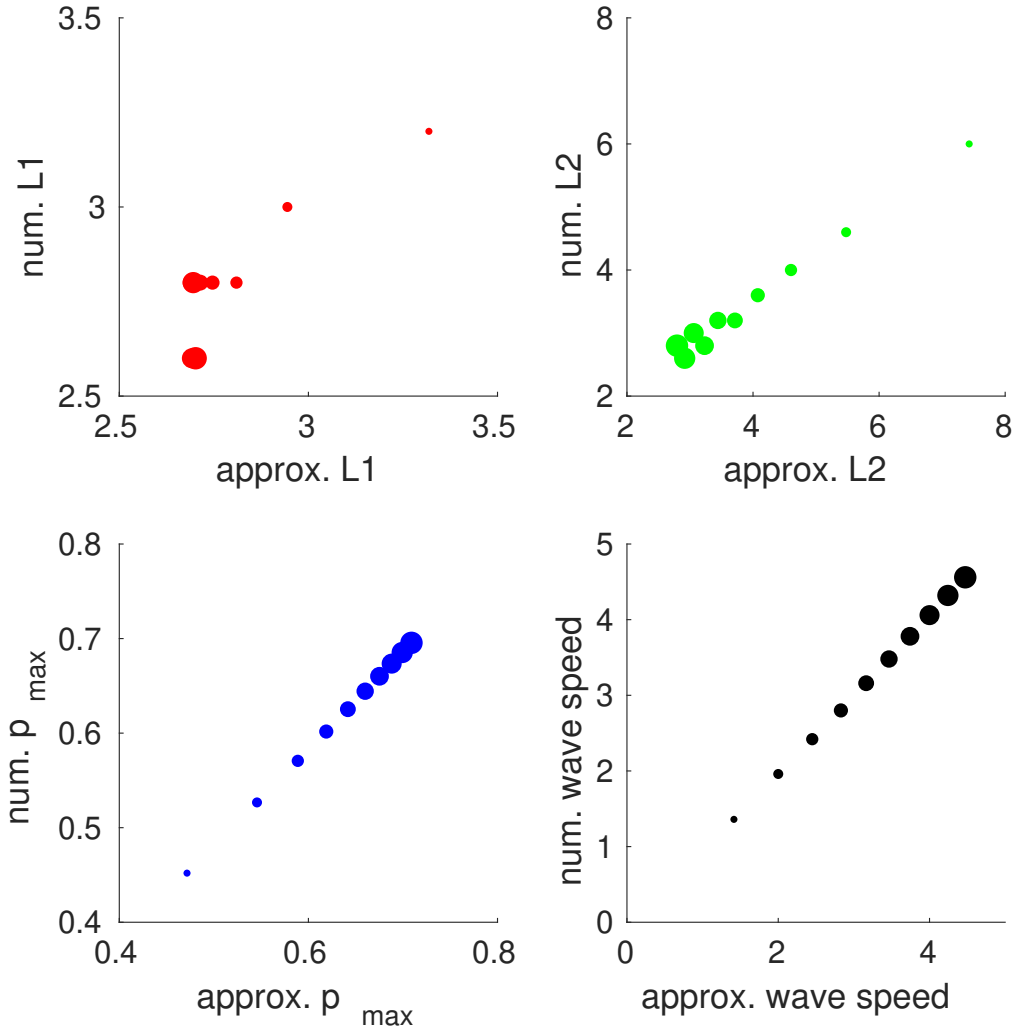

Figure 3.5: Scatter plots of approximate wave profile characteristics (on horizontal axis) versus the ones obtained by numerical simulation (on vertical axis) for a range of $\hat{\rho}$ from 0.5 to 5 by increments of 0.5. The size of the dot corresponds to the value of $\hat{\rho}$. The dots scatter closely to the diagonal line with slope 1, indicating agreement between the numerical solution and our approximation.

3.5 Discussion

In this work, we have extended the Fisher-Kolmogorov reaction-diffusion model of GBM growth, Eq. (3.1), to explicitly separate the cancer cell proliferation and death (or quiescence) processes, which are described in terms of generic functions that depend upon an implicit nutrient or growth factor. We specify the proliferation and death processes, $g(w)$ and $\delta(w)$, respectively, by the cumulative distribution function of a beta distribution, each uniquely specified by a single parameter, ρ and k. Thus, along with the diffusion coefficient, D, our model describes cancer growth via three parameters, D, ρ, and k, and yields a tumor morphology (in one dimension) consisting of a necrotic core, a high-density rim, and an outer low density rim, which we may correlate to three radii, R_0, R_1, and R_2, that can be estimated from a single patient MR image.

We have demonstrated that our reaction-diffusion system has a traveling-wave solution, which is common in such systems. Studies on this topic date back to the Fisher's work in the 1930s on the spread of advantageous genes (Fisher, 1937). Rigorous proof of the existence of a traveling-wave solution in a reaction-diffusion system often leads to phase-space analysis such as the one on the diffusive Lotka-Volterra equations (Dunbar, 1983). Although in general a rigorous proof of a traveling-wave solution is a daunting task, our reduced system is amenable to phase-plane analysis, and the orbit that represents the traveling-wave solution can be identified (see appendix).

Via traveling-wave analysis, we have developed a method to estimate D, ρ, and k from as few as a single magnetic resonance image, based on certain growth assumptions. We have estimated these parameters for six patient test cases, as shown in Table 3.1. Because of the sparsity of imaging data for a typical patient, parameter

identifiability in this case is provided by the monotonicity of the function $f(\hat{\rho})$ (as seen in left pane of Figure 3.4); our approach differs from more common statistical practices (Eisenberg and Jain, 2017) that are appropriate when more data are available.

Disaggregating the net cell proliferation into proliferation and death processes not only aids in relating (simplified) tumor appearance on MRI to the model parameters, but it may provide useful valuable information for personalized treatment design, insofar as chemotherapy and radiotherapy target proliferating cells. Moreover, the structural information is also potentially useful, as drug dosages might be selected to ensure penetration through the width of the proliferating rim. Research along this general line has been conducted using model (3.1) (Kim *et al.*, 2017).

Our analysis uses a more complex description of motility than simple diffusion. The diffusion term in Eq. (3.5) belongs to a more general category called *cross diffusion* (Madzvamuse *et al.*, 2017). It represents the phenomenon in which the gradient in the concentration of one species causes a flux of another species. The type of cross diffusion considered in this paper has been studied in a more general and theoretical context (Sherratt, 2000). In a modeling study of avascular tumour growth (Sherratt and Chaplain, 2001), the authors justified the adoption of a proportion-based cross diffusion in a tumor-growth model by recognizing that tumor cell migration is "contact inhibited": the presence of one type of cell halts the movement of the other. This type of cross diffusion can cause the solution to become negative (Madzvamuse *et al.*, 2017), but we have not encounted this difficulty so far in our numerical simulations. We conjecture that, for suitable initial conditions, solutions remain positive, which will be a topic for future study. Other types of density-dependent diffusion have been considered in modeling GBM migration (Stepien *et al.*, 2015).

Despite the existence of many possible diffusion terms, the exact form of diffusion

does not matter in the one-dimensional analysis, because the second derivatives are dropped in model (3.8). However, The diffusion coefficient does play an important role in the linearized wave head, where it affects the wave speed, and in the characteristic length where its square root scales the space. The scale-invariant part of the wave profile is mostly determined by the exact forms of the proliferation and death functions.

The major contribution of this paper is two novel methods to make patient-specific estimates of a three-parameter model of GBM grwoth from the limited MRI data that is typically available in clinical settings. It is possible to estimate the model parameters from a single pre-surgery image. Improved estimates may be possible when images are acquired at multiple time points.

As cancers progress, the underlying parameters describing their growth are unlikely to be static. Data assimilation refers to basic method of updating parameters as more data is acquired, and there has been research on applying full-fledged data assimilation to cancer modeling (Kostelich *et al.*, 2011; McDaniel *et al.*, 2013). The methods described in this paper can be incorporated as part of a future data assimilation system.

Chapter 4

SPATIO-TEMPORAL FORECASTING USING GAUSSIAN PROCESSES WITH APPLICATION TO PREDICT GLIOMA INVASION

4.1 Abstract

Mathematical models have been used to study spatio-temporal patterns of brain cancer invasion and make predictions. In order for any model predictions to be clinically useful, they need to be confronted with data, and their uncertainty has to be quantified. However, there is no "ground true" models in biology and all formulations tend to under-/over-fit data. We propose a nonparametric forecasting method to address these issues. A Gaussian process prior is assumed on the transition function. By exploiting the local nature of the spatio-temporal process, we can make reasonable predictions with very sparse time-series data. We test our method with synthetic data generated by a discretized version of bistable equation and show promising results.

4.2 Introduction

Glioma is a common primary tumor that occurs in the brain. It invades aggressively resulting in low patient survival rate and short life expectancy. The biological mechanism of glioma invasion is complex, including phenotypic plasticity, genetic variability, hypoxia-induced migration, immune system ligament, etc (Alfonso *et al.*, 2017). The spatio-tempoeral process of glioma invasion is partially revealed by MRI scans or biopsy in clinical settings. This usually guides treatment decisions, which oftentimes involves resection and radiation targeting the tumor. Improvement and innovation of clinical treatment can benefit from a better use of the information avail-

able and a model with prediction power.

Mathematical modeling has been applied to study glioma invasion (see Martirosyan *et al.* (2015); Alfonso *et al.* (2017) for a comprehensive review). Many of them are developed towards to clinical applications and can provide predictions of tumor growth. Despite effectiveness of mathematical modeling in physics, biological systems are usually very complex and there is no canonical models for biological systems that represent the ground truth. Most models in biology focus on one or a few aspects of the underlying biology. Oftentimes they are formulated based on heavy assumptions, which are hard to verify due to the complexity of biological processes and lack of data. There is no exception for mathematical models of glioma invasion. Moreover, for any models to be clinically useful, it has to be confronted with patients' data and the uncertainty of its predictions has to be quantified. There are two challenges in data-driven modeling of spatio-temporal processes: high dimensionality and over-/under-fitting. High dimensionality is inherent in the need to represent space. Compounding with this challenge is the scarcity of personal clinical data, e.g., only a few MRI scans of each patient are usually available. It is nontrivial how to learn the dynamics with such scarce data. As for data fitting, simple models may not capture all the variations in data and hence tend to under-fit while sophisticated models tend to treat noise as signals and hence over-fit. It is challenging to strike a balance to find the best model that maximize the information extracted from the data without over-fitting.

To tackle these challenges, we propose a state space model with a transition function as a Gaussian process (GP) acting locally. GP regression is a nonparmetric method that in theory avoids under-/over-fitting (reviewed in next paragraph). Representing transition function as a Gaussian process has gained a lot of interests in the literature of machine learning and system identification (Girard *et al.*, 2003; Deisen-

roth *et al.*, 2009), where it is sometimes called recurrent GP since GP is used recurrently to map one state to the next in time. It has been successfully used to predict pedestrian's movement for self-driving cars (Wang *et al.*, 2008). But applications to spatio-temporal process in high dimensionality settings are unprecedented to our knowledge, which we intend to embark in this work. By exploiting the local nature of the spatio-temperoal process, the transition function can be effectively learned using only two consecutive data in time. In reality, these two consecutive data corresponds to diagnostic and presurgical MRI scans. Based on these data, we can make forecasting onward by recurrently applying GP. By exploiting Gaussianity, uncertainty can be propagated with exact integration, which reduces the otherwise overwhelming computational costs.

We briefly review the GP regression. More details can be found in (Rasmussen and Williams, 2006). A GP $f(x)$, denoted as $f \sim \mathcal{GP}(m(x), k(x, x')$, is a stochastic process that is entirely determined by its mean function $m(x)$ and its covariance function $k(x, x')$. For any finite collection of x_i with $i = 1 \dots n$, the joint distribution of $f(x_1), f(x_2) \dots f(x_n)$ is a multivariate normal with mean $m(x_1) \dots m(x_n)$ and covariance matrix with entry being $k(x_i, x_j)$. The GP regression assumes a model $y = f(x) + \epsilon$ where iid $\epsilon \sim N(0, \sigma^2)$, and places a GP prior on the latent function f, denoted as $f \sim \mathcal{GP}(m(x), k(x, x')$. Given n training data $\mathcal{D} = \{x_r, y_r : r = 1 \dots n\}$ where $y_r = f(x_r) + \epsilon_r$ as input-output pairs, the GP regression makes inferences of $f(x_j^*)$ at the m test input x_j^* with $j = 1 \dots m$ by conditioning on $\mathcal{D}$. The conditional is easily worked out since the joint distribution is a multivariate normal, i.e.,

$$\begin{bmatrix} \mathbf{f} \\ \\ \mathbf{f}^* \end{bmatrix} \sim N(\begin{bmatrix} m(\mathbf{x}) \\ \\ m(\mathbf{x}^*) \end{bmatrix}, \begin{bmatrix} \Sigma & K_* \\ \\ K_*^T & K_{**} \end{bmatrix})$$

where $\mathbf{f}, \mathbf{f}^*, m(\mathbf{x})$ and $m(\mathbf{x}^*)$ are vectors, and $\Sigma = k(\mathbf{x}, \mathbf{x}) + \sigma^2 I$, $K_* = k(\mathbf{x}, \mathbf{x}^*)$, $K_{**} =$

$k(\mathbf{x}^*, \mathbf{x}^*)$ are matrices. Here for the ease of notation we have combined individual training data x_r, y_r and and test points x_j^* into bold notations so that the function acting on them are understood as vectors or matrices. For example, $m(\mathbf{x})$ is a vector with j-th component being $m(x_j)$ and $k(\mathbf{x}, \mathbf{x}^*)$ is a matrix with entries $[k(\mathbf{x}, \mathbf{x}^*)]_{ij} = k(x_i, x_j^*)$. As a result,

$$\mathbf{f}|\mathbf{f}^* \sim N(m(\mathbf{x}) + K_*^T \Sigma^{-1}(\mathbf{f} - m(\mathbf{x})), K_{**} - K_*^T \Sigma^{-1} \Sigma_*) \tag{4.1}$$

Note that in the standard GP regression, training input $\mathbf{x}$ and testing input $\mathbf{x}^*$ are noise-free. In the case of otherwise, the uncertainty in $\mathbf{x}$ and $\mathbf{x}^*$ can be taken into account by making appropriate assumptions and approximation.

The paper is organized as the following. In the Section 4.3, we motivate and describe the local GP method. In the section 4.4, we test the method with synthetic data. In section 4.5, assumptions made in this work are discussed and future directions are pointed out.

4.3 Method

We motivate our method by looking at a commonly used PDE model of brain cancer invasion: the Fisher-KPP equation

$$u_t = D u_{xx} + (1 - u)u$$

where $u(x, t)$ is the cancer cell density at position x and time t. Finite difference discretization gives

$$u_j^{n+1} = u_j^n + \frac{\Delta t}{\Delta x^2}(u_{j+1}^n - 2u_j^n + u_{j-1}^n) + \Delta t u_j^n (1 - u_j^n),$$

where superscripts indicate time and subscripts indicate space. This discretization reflects the local nature of PDEs.

Oftentimes we don't have exact measurement of u but a noisy version of it, denote as v at some discrete spatial points. To make use of these observation data, we let $\mathbf{u} \equiv (u_j)_{j\in\{1:n_x\}}$ and consider a state-space model

$$\mathbf{u}^n = f(\mathbf{u}^{n-1}) + \epsilon \tag{4.2}$$

$$\mathbf{v}^n = g(\mathbf{u}^n) \tag{4.3}$$

where $f(\cdot)$ is a deterministic transition function, $g(\cdot)$ is a probabilistic observation model, and we assume noise $\epsilon \sim N(0, \sigma_\epsilon^2 I)$ for now. We could specify $f(\cdot)$ to be the finite difference discretization of a PDE model such as Fisher-KPP or certain parametric formula (e.g. truncated basis expansion by Wikle and Hooten (2010)). However, there are many PDE models/basis expansion of various complexity and aspects. There are clearly no exact models in biology. These aforementioned modeling approaches oftentimes suffer from under-/over-fitting.

We instead take a nonparametric approach using Gaussian processes (GP). In a sense, it is a model-free approach, i.e., we do not formulate any models to fit data, but let the data dictate the prediction. To make use of data that is sparse in time, we exploit the local nature by making f a local map (Figure 4.1), i.e., $u_j^n = f(u_{\mathcal{N}(j)}^{n-1})$ where $\mathcal{N}(j)$ is an ordered set of local neighbors of j including j, and accordingly $u_{\mathcal{N}(j)}^{n-1}$ is a vector of u_i^{n-1} for $i \in \mathcal{N}(j)$ that preserve this order.

We digress here to expand an explanation of the ordering of $\mathcal{N}(j)$, which also exposes the essence of nonparametric methods and biological insight. First, the size of the local neighborhood should reflect the scale of biological iterations in a time step. From a modeling point of view, it is important to determine the size of the local neighborhood, i.e., how many neighbor sites should be included. In another word, we need perform variable selection. Luckily, Gaussian process regression with the squared exponential covariance function has built-in variable selection capability, also known

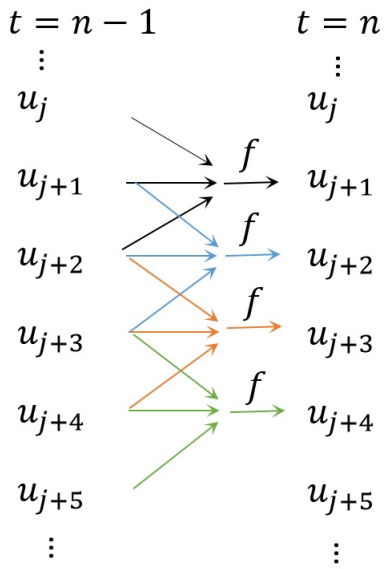

Figure 4.1: f as a local map. Here the neighborhood of a site includes itself and two adjacent sites, e.g. $\mathcal{N}(j+1) = \{j, j+1, j+2\}$

as automatic relevance determination (Neal, 2012; Williams and Rasmussen, 1996; Rasmussen and Williams, 2006). Second, the ordering of the neighborhood set should be chosen carefully. A straightforward approach would be taking the ordering same as the one of the spatial grids. However, it is beneficial to use an ordering of $\mathcal{N}(j)$ that makes biological sense or exploits certain symmetry of the problem (see next section for a comparison between the naive ordering and a biologically meaningful ordering). In the case of predicting cancer cell invasion, we imagine that the fate of a cell is solely determined by crowdedness of the site it occupies and its immediate neighbor sites. Our data tells us how the cell behaves in a variety of situations of local crowdedness. When we need to make a prediction in a new situation, it is natural to base the prediction on the situations we have seen before. This entails a measure of how similar the new situation is compared to the known situation offered by data. It is straightforward to compare the crowdedness of the occupied site directly by looking their difference. But how about the neighbor sites? We propose an ordering in which the neighbor sites are ranked by its cell density and then take the difference

component-wisely. This proposed ordering would enhance the data-use efficiency. The price is that it is incapable to capture a process that exhibits advection or anisotropic diffusion.

We also comment that the notation of f is slightly abused since f was used to denote the global map in (4.2) while it is also presents a local map (Figure 4.1); we will not distinguish them henceforth as it is easily understood in a context. We place a GP prior on f, i.e., $f(\mathbf{s}) \sim \mathcal{GP}(\mu(\mathbf{s}), k(\mathbf{s}, \mathbf{s}'))$. We could use a priori beliefs of the dynamics to specify the mean function $\mu(\cdot)$. For the time being, we simply assume $\mu(u_{\mathcal{N}(j)}) = u_j$, e.g, the tumor activity stays unchanged, to demonstrate the capability of the method without any specific knowledge of the dynamics. We use the squared exponential covariance function, i.e.,

$$k(\mathbf{s}, \mathbf{s}'; \mathbf{\Lambda}^{-1}) = \sigma_k^2 \exp(-\frac{1}{2}(\mathbf{s} - \mathbf{s}')^T \mathbf{\Lambda}^{-1}(\mathbf{s} - \mathbf{s}')) \tag{4.4}$$

where $\mathbf{\Lambda} = \text{diag}(\ell_i)$ is a diagonal matrix of the length scales. Here ℓ_i , σ_k, and possibly σ_ϵ, are hyperparameters and estimated by maximizing "marginal likelihood" (Rasmussen and Williams, 2006). It should be noted that the choice of covariance function (4.4) is made to facilitate the uncertainty propagation, which will be described later. Another benefit is that the squared exponential covariance function enables automatic relevance determination: the characteristic length $\ell_i's$ tells the relative importance the corresponding variable.

In this work, we focus on forecasting and defer filtering and smoothing to future study. Since there has not been any clearly established quantified relationship between cancer cell density and MRI enhancement, we choose to leave out the observation model (4.3). Instead, we assume availability of two consecutive data ($\mathcal{D}$) not far separated in time of exact observation of the latent variable $\mathbf{u}$, e.g., $\mathcal{D} = \{\mathbf{u}^0, \mathbf{u}^1\}$. The assumption on the small time interval is not needed in a strict sense: presumably

the larger time interval would lead to larger interaction neighborhood, which entails a more challenging task of learning a higher dimensional mapping. In reality, the two observations correspond to diagnostic and presurgical MRI scans which are oftentimes separated by a week or two. It should also be noted that (4.2) may lead to negative $\mathbf{u}$. So we cannot strictly interpret $\mathbf{u}$ as cell density. Instead we understand it as a measure of abundance of cancer cells. We could exponentiate $\mathbf{u}$ and use it as the parameter for Poisson observation model so that we can regain positivity. This approach has been adopted in (Hooten and Wikle, 2008).

Recall that $\mathbf{u}^2 = f(\mathbf{u}^1) + \epsilon$ so that $\mathbf{u}^2|\mathbf{u}^1, \mathbf{u}^0$ is a multivariate normal whose mean and variance are readily known by standard GP regression (see (4.1)). To move forward in time, we can apply the same inferences again but with test input being $\mathbf{u}^2$ this time. Since $\mathbf{u}^2$ is not deterministic and has uncertainty in itself, we need a way to do the same inference with uncertain test input. Fortunately, the mean and variance of $\mathbf{u}^2$ can be worked out analytically due to our choice of $k(\mathbf{s}, \mathbf{s}')$ and the fact that the product of Gaussian functions is also a Gaussian function. Even though $\mathbf{u}^2$ is not necessary Gaussian distributed any more, we approximate it as a multivariate Gaussian by matching the first two moments. In this way, we can repeat the procedure to propagate the uncertainty. The details on how to take into account the uncertainty of test input are given in Appendix D.1.

One more approximation is necessary when the number of states is large. Because the size of the covariance matrix grows quadratically with the number of the states, it quickly becomes impractical to keep track of a dense covariance matrix. For instance, a 2D 100 by 100 grids entails a covariance matrix of 10^8 entries. Since we believe that system dynamics is dominated by local interactions, we approximate the dense covariance matrix by only keep track of the covariance between states which are in a local neighborhood, effectively making the covariance matrix sparse. Similar

approaches have been taken in local Kalman Filters (Hunt *et al.*, 2007).

4.4 Results

4.4.1 Test on 1D Synthetic Data

We obtain the synthetic data motivated by the discretization of 1D bistable equation

$$u_j^{n+1} = u_j^n + \theta_1(u_{j+1}^n - 2u_j^n + u_{j-1}^n) + \theta_2 u_j^n(1 - u_j^n)(u_j^n - \alpha) + \epsilon_j^n \tag{4.5}$$

where $\epsilon_j^n \sim N(0, \sigma^2)$ is independent in space and time. The choice of (4.5) is due to its simplicity and regularity. The bistable equation is known to have a traveling wave solution given a suitable initial condition. The traveling wave resembles cancer progression. The zero steady state of the homogeneous system is stable. Since the additive white noise inevitably leads to negativity, this stability is favored in contrast to Fisher-KPP equation which may go to negative infinity. That being said, the slight negativity may pose some issues of interpretability. Here u can be interpreted as intensity of underlying tumor activity rather than cell density in a strict sense. We could also interpret u as a log of the cell density. Since the focus of this chapter is on forecasting, we forgo some of the rigor in model interpretation.

We use the numerical solution at two consecutive time points as training data. The hyperparameters found by maximizing marginal likelihood are reported in Table 4.1. It is obvious that only immediate neighbors are relevant and should be included in the neighborhood set. The neighbors that are two sites away have length scale several magnitude larger and hence does not contribute to prediction. It is as expected since the training data are generated from two consecutive time of (4.5).

Then we obtain our forecasting results and compare them with the realized sample paths. We pick two starting points for forecasting, namely fledged and burgeoning

data(see Figures 4.2 and 4.3). Our forecasting predictions captured the general trend of the true process and showed gradual deterioration of the prediction. The quality of the prediction depends on the starting point which controls what the training data is used. Obviously, we shall not expect the forecasting to work well on the region of unseen data. Even though the predicted mean was in disparity with the true mean, the uncertainty bounds were large. This showcases the strength of the method in uncertainty quantification.

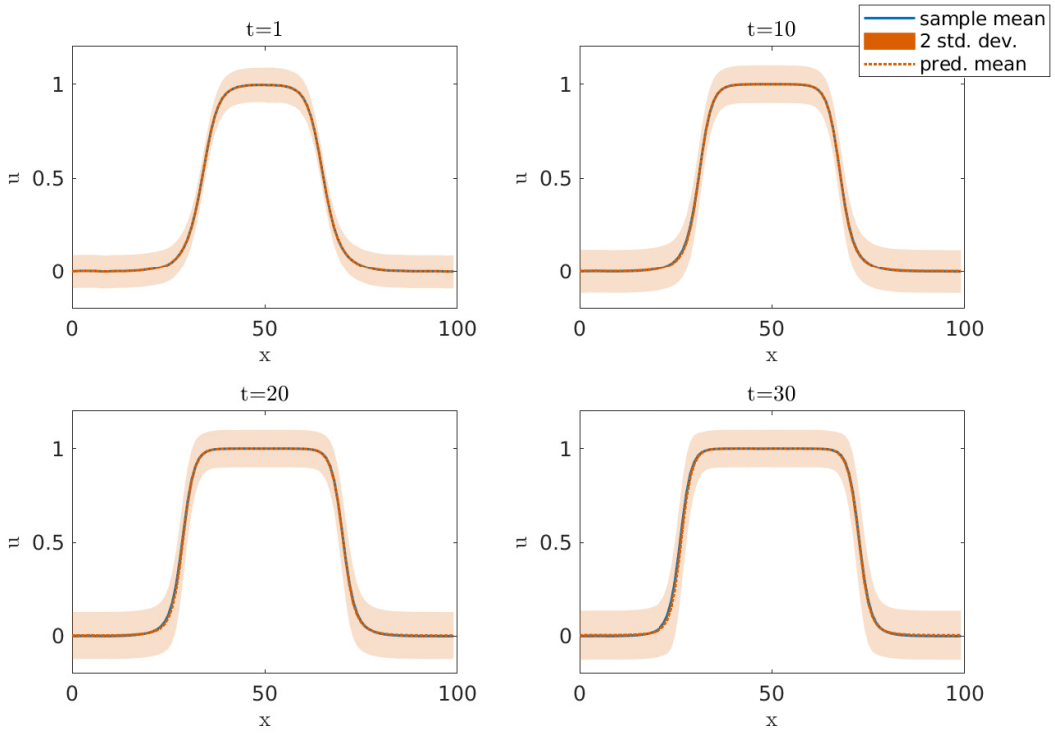

Figure 4.2: Forecasting to time $t = 2, 8, 14$ and 20 with GP trained with fledged synthetic data corrupted with noise $\sigma_\epsilon = 0.01$. The shaded bounds indicate the uncertainty (two standard deviation) of the prediction.

We compute the sample average of 1000 samples and defined it as the ground truth. Then we compute the mean squared error (MSE) of the predictive mean and the negative log likelihood (NLL). Loosely speaking, MSE indicates how good the predictive mean as a point estimate of the ground truth while NLL additionally takes

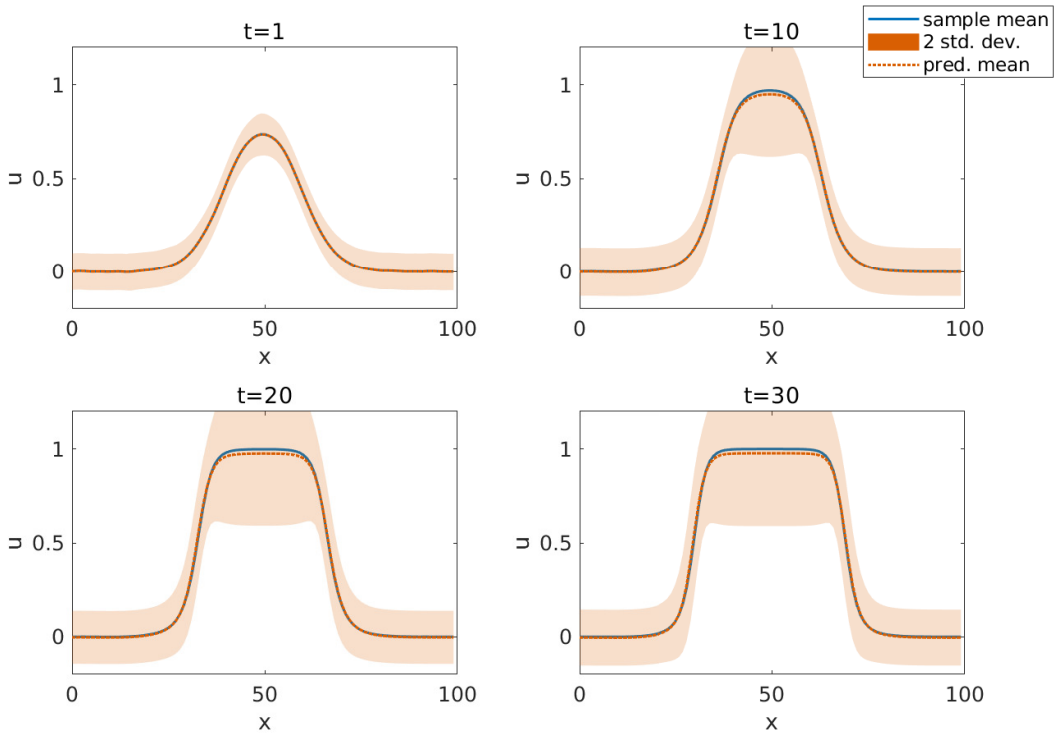

Figure 4.3: Forecasting to time $t = 2, 8, 14$ and 20 with GP trained with unfledged synthetic data corrupted with noise $\sigma_\epsilon = 0.01$. The shaded bounds indicate the uncertainty (two standard deviation) of the prediction.

into account the error bound that penalizes being too confident on an inaccurate point estimate or too conservative on an accurate point estimate. We see that biologically meaningful ordering result in smaller MSE and NLL than those obtained by naive ordering (Table 4.2), meaning that it is a better choice. We also notice that sparse approximation of covariance is prone to be over-confident about the point prediction but overall we are satisfied by its performance.

4.4.2 Test on 2D Synthetic Data

To get 2D synthetic data, we simply use the 2D version of (4.5)

$$\begin{aligned} u_{i,j}^{n+1} =& u_{i,j}^n + \theta_1(u_{i+1,j}^n + u_{i,j+1}^n - 4u_{i,j}^n + u_{i-,j}^n + u_{i,j-1}^n) \\ &+ \theta_2 u_j^n(1-u_j^n)(u_j^n - \alpha) + \epsilon_j^n \end{aligned} \tag{4.6}$$

Table 4.1: Hyperparamters found by maximizing marginal likelihood. The first row has neighborhood size being 3 and the second row has neighborhood size being 5.

σ_k^2	σ_ϵ^2	ℓ_{-2}	ℓ_{-1}	ℓ_0	ℓ_1	ℓ_2
0.1527	1.0277e-04	NA	6.0574	1.5658	6.9022	NA
0.1527	1.0277e-04	3.2606e+07	6.0574	1.5658	6.9022	3.6812e+07

Table 4.2: MSE and NLL of GP forecasting with or without local sparse covariance and naturally ordered neighborhood

	$\sigma_\epsilon = 0.01$		$\sigma_\epsilon = 0.02$	
	MSE	NLL	MSE	NLL
sparse ordered	1.6784e-04	-306.2233	6.3806e-04	-230.5547
non-sparse ordered	1.6783e-04	-340.6569	6.3796e-04	-272.3875
non-sparse non-ordered	1.8632e-04	-335.6242	6.5192e-04	-269.7519

with periodic boundary condition. Same as in section 4.4.1, we use the numerical solution at two consecutive time points as training data. It is obvious that only immediate neighbors are relevant and should be included in the neighborhood set. The neighbors that are two sites away have length scale several magnitude larger and hence does not contribute to prediction. Again, this meet our expectation since the data are generated from two consecutive time of (4.6).

In Figure 4.4, the prediction over time and its uncertainty is shown. As we can see, the prediction remains regular and its uncertainty gradually grows. The highest

uncertainty is at the wave front. The center at $t = 20$ shows considerable uncertainty, which is expected since it is the region unseen in training data.

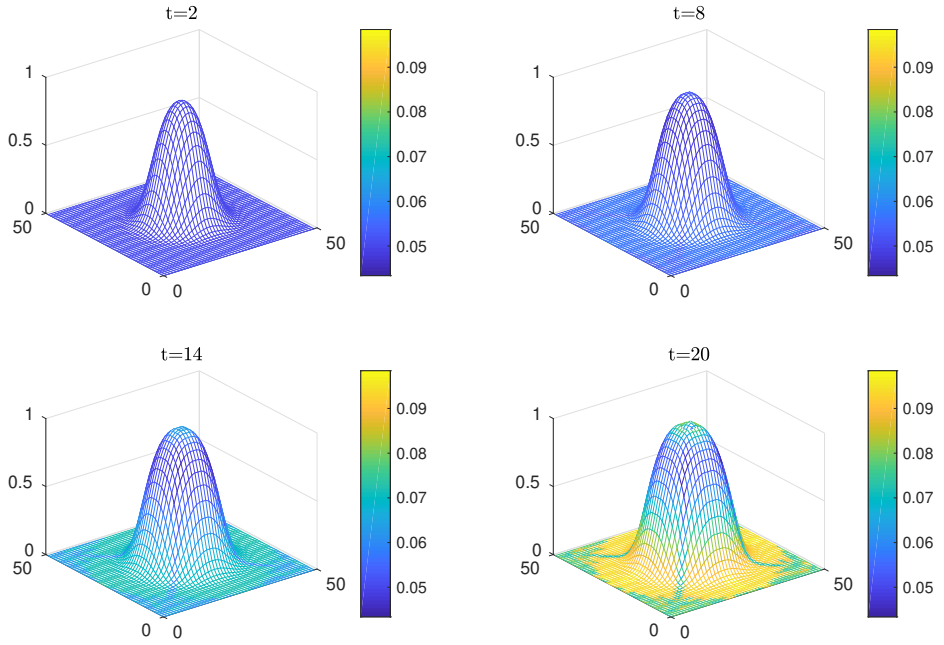

Figure 4.4: Forecasting to time $t = 2, 8, 14$ and 20 with GP trained with 2-D synthetic data corrupted with noise $\sigma_\epsilon = 0.001$. The colormap indicates the uncertainty (two standard deviation) of the prediction.

4.5 Discussion

Instead of formulating a model of cancer invasion, we take a nonparametric approach using Gaussian processes. The novelty of this work is exploiting the local nature of the spatio-temporal process so that we can use very sparse time-series data to make predictions. That being said, a local map with an input of a few dimensions is learned from the data instead of a global map from a high dimensional space. This improvement of data-efficient has the advantage in applications to glioma for which there are oftentimes only a few MRI scans of each patient. Data efficiently can be further improved if we order the neighborhood properly with reasonable assump-

tions. However, this is often comes at a price of generality, e.g., ignoring possibility of anisotropic diffusion or advection.

We also achieved computational efficiency in propagating uncertainty by choosing the squared exponential covariance function for which some integration can be done analytically. This choice of covariance function also makes variable selection easy due to automatic relevance determination. From the selected variables, we can determine the neighborhood size. In the tested cases, our method correctly indicated the relevant variables. Being able to identifying the neighborhood size also has biological significance since it tells the scales of spatial interaction in a time step. In the tested cases, the selected neighborhood is symmetric about its center due to the specific choice of the data-generating model. However, it should be noted that this is not necessary the case in practice. A asymmetric shape of a selected neighborhood can be an indicator of anisotropic diffusion or advection of the underlying process.

Even though automatic relevance determination correctly identified relevant variables in the case of synthetic data. But it only gives a relative importance between variables. It is foreseeable that in the case of real-world data, the cut-off of relevance will not be so clear. Ideally, we need a more informed variable selection criterion with the cut-off built-in. This will be one of the directions of future studies.

The squared exponential covariance function is a popular choice for Gaussian processes used in data modeling. It has many desirable features as discussed early. It also naturally arises as the covariance function as a limit of linear regression with Gaussian priors. It can be shown that a Gaussian process with the squared exponential covariance function has mean square derivatives of all orders (Papoulis and Pillai (2002); Cramer and Leadbetter (2004); Adler (2010), see details summarized in Appendix D.2). This strong smoothness seems unrealistic for modeling physical processes as argued by Stein (2012). In contrast, the Matern class of covariance functions

of the form

$$k(r) = \frac{2^{1-\nu}}{\Gamma(\nu)}(\frac{\sqrt{2\nu}}{\ell})^{\nu} K_{\nu}(\frac{\sqrt{2\nu r}}{\ell}),$$

where K_ν is a modified Bessel function, offers some flexibility of differentiability of the process: the process is k-time mean square differentiable if and only if $\nu > k$ (Rasmussen and Williams, 2006). Note that in the limit $\nu \to \infty$, The Matern covariance function becomes the squared exponential covariance function. In this work, the squared exponential covariance function is favored over the Matern covariance function in order to achieve the simplicity of integration of product of Gaussians.

We left out the observation model and tested our method on synthetic data. In the future, an observation model will be used to incorporate MRI data and brain geometry. We will compare our method to some existing studies of predicting cancer growth using state space framework (McDaniel *et al.*, 2013; Kostelich *et al.*, 2011; Lipková *et al.*, 2019). State space modeling is challenging for nonparametric methods because of loss of Markov property in (4.2) (note that $f(\cdot)$ itself depends on all historical $\mathbf{u}$.) There are some attempts aimed at resolving this issue (Frigola *et al.*, 2013; Ghosh *et al.*, 2014) but it is not clear how it can be extended to a high-dimensional spatio-temperoal process.

Moreover, the prediction deteriorates over time inevitably, intermediate data is needed to correct predictions. Thus we need a filtering algorithm which can assimilate data in an iterative fashion. This will be another focus of our future study. We will also devise a method for smoothing, which is less interesting in medical applications but is of interests in a retrospective study.

We approximated latent variables for all subsequent forecasting with normal distribution. This approximation is computationally efficient with price being complicated matrix algebra. We would like to know the error introduced and whether there are computational alternatives.

Another approximation we used to achieve the sparsity of the covariance matrix can also introduce error. The approximation is justified if the the covariance between remote sites decay rapidly. Since spatial independent noise is added in each time step, the assumption is expected to be met at least in the long term. A rigorous error analysis await for future studies. It should be pointed out that similar practice and rationale have been adopted in local Kalman filters and local particle filters. The localization of a Kalman filter makes parallelization of computation possible (Ott *et al.*, 2004; Hunt *et al.*, 2007; Szunyogh *et al.*, 2008). Moreover, it is believed that local analysis avoids spurious correlation that would otherwise allow obervation at one location to influence the analysis of a location an arbitrarily large distance away (Hunt *et al.*, 2007). It should be noted that in our case the correlation between remote sites are not necessarily spurious because of the nonparametric nature of the method. Rebeschini and van Handel (2015) gave a rigorous proof to justify a simple locality implementation called block particle filters. Since then, the local particle filter has been implemented by Morzfeld *et al.* (2018); Farchi and Bocquet (2018) and refined by Quinn (2019).

Estimating hyperparameter by maximizing marginal likelihood may not be ideal. It is known to lead to local optimum, and sometimes two local minima has equally good interpretations (see page 116 in Rasmussen and Williams (2006)). A full Bayesian approach would be more warranted to properly incorporate any prior beliefs on hyperparameters.

4.6 Conclusion

In this chapter, we proposed a method to make spatio-temporal forecasting using Gaussian processes. Instead of formulating a dynamic model for the process, we assumed it as a Gaussian process. By exploiting the local nature of the dynamics,

we improved the data efficiency and solved the problem of sparsity of temporal data. By exploiting Gaussianity, prediction and uncertainty can be propagated in a computationally efficient way. The method in its preliminary stage were tested on the synthetic data and showed promising results. Future directions were pointed out in order to make this method applicable in practice, such as forecasting glioma invasion with MRI data.

Chapter 5

CONCLUSION

This thesis was developed with an overarching goal to understand and predict cancer invasion in time and space. In Chapter 2, a model of stochastic delayed differential equations (SDDEs) was formulated with a focus on temporal dynamics of tumor-immune interaction. Immune response is modeled as a non-monotonic function of tumor burden, for which the tumor is immunogenic at nascent stage but starts inhibiting immune system as it grows large. Without time delay and noise, this system demonstrates bistability. We found that time delay and noise can both induce the transition from low tumor burden equilibrium to high tumor equilibrium. We also proved some conditions of stability, persistence and extinction of the tumor.

In Chapter 3, we studied Glioblastoma multiforme (GBM) using a partial differential equation (PDE) model. We proposed a mathematical model of GBM growth with explicit motility, birth, and death processes to help relate the three highly idealized components of GBMs (i.e. necrotic core, enhancing rim, and maximum edema extent) and to the underlying growth laws. The model demonstrated a traveling wave solution that mimics cancer progression. We developed a novel method to approximate key characteristics of the wave profile, which can be compared with MRI data. For several simplified forms of growth and death terms, we found evidence that support their parameter identifiability. We used several test cases of MRI data of GBM patients to yield personalized parameterizations of the model. We found large variations of parameters among patients. Future work on this front will focus on a rigorous study of positivity of the solution. This is non-trivial because of the cross-diffusion term in the model.

Chapter 4 presented an innovative way of forecasting spatial cancer invasion. It avoided hard-to-verify assumptions that have to be made by most mathematical models. Our approach employed a nonparametric forecasting method using Gaussian processes. High data efficiency was achieved by exploiting the local nature of the spatio-temporal process. We tested the method on synthetic data and the results were promising.

REFERENCES

Adler, R. J., *The Geometry of Random Fields* (Society for Industrial and Applied Mathematics, 2010).

Alfonso, J. C. L., K. Talkenberger, M. Seifert, B. Klink, A. Hawkins-Daarud, K. R. Swanson, H. Hatzikirou and A. Deutsch, "The biology and mathematical modelling of glioma invasion: a review", Journal of The Royal Society Interface **14**, 136, 20170490, URL `http://rsif.royalsocietypublishing.org/lookup/doi/10.1098/rsif.2017.0490` (2017).

Baez, J. and Y. Kuang, "Mathematical models of androgen resistance in prostate cancer patients under intermittent androgen suppression therapy", Applied Sciences **6**, 11, 352 (2016).

Banerjee, S. and R. R. Sarkar, "Delay-induced model for tumor-immune interaction and control of malignant tumor growth", Biosystems **91**, 1, 268–288, URL `https://www.sciencedirect.com/science/article/pii/S0303264707001499` (2008).

Bashkirtseva, I., "Stochastic sensitivity analysis: theory and numerical algorithms", in "IOP Conference Series: Materials Science and Engineering", vol. 192, p. 12024 (IOP Publishing, 2017).

Bashkirtseva, I., T. Ryazanova and L. Ryashko, "Confidence domains in the analysis of noise-induced transition to chaos for Goodwin model of business cycles", International Journal of Bifurcation and Chaos **24**, 08, 1440020 (2014).

Bose, T. and S. Trimper, "Stochastic model for tumor growth with immunization", Physical Review E **79**, 5, 051903, URL `https://link.aps.org/doi/10.1103/PhysRevE.79.051903` (2009).

Box, G. E. P., G. M. Jenkins, G. C. Reinsel and G. M. Ljung, *Time series analysis: forecasting and control* (John Wiley & Sons, 2015).

Britton, N. F., *Essential mathematical biology* (Springer, 2003).

Canosa, J., "On a Nonlinear Diffusion Equation Describing Population Growth", IBM Journal of Research and Development **17**, 4, 307–313, URL `http://ieeexplore.ieee.org/document/5391351/` (1973).

Cappé, O., E. Moulines and T. Rydén, *Inference in hidden Markov models* (Springer Science & Business Media, 2006).

Carlson, M. R., W. B. Pope, S. Horvath, J. G. Braunstein, P. Nghiemphu, C.-L. Tso, I. Mellinghoff, A. Lai, L. M. Liau, P. S. Mischel, J. Dong, S. F. Nelson and T. F. Cloughey, "Relationship between survival and edema in malignant gliomas: role of vascular endothelial growth factor and neuronal pentraxin 2", Clinical Cancer Research **13**, 9, 2592–2598 (2007).

Claes, A., A. J. Idema and P. Wesseling, "Diffuse glioma growth: A guerilla war", Acta Neuropathologica **114**, 5, 443–458 (2007).

Cramer, H. and M. R. Leadbetter, *Stationary and related stochastic processes : sample function properties and their applications* (Dover Publications, 2004).

Deisenroth, M. P., M. F. Huber and U. D. Hanebeck, "Analytic moment-based Gaussian process filtering", in "Proceedings of the 26th annual international conference on machine learning", pp. 225–232 (ACM, 2009).

Doering, C. R., C. Mueller and P. Smereka, "Interacting particles, the stochastic Fisher–Kolmogorov–Petrovsky–Piscounov equation, and duality", Physica A: Statistical Mechanics and its Applications **325**, 1-2, 243–259 (2003).

D'Onofrio, A., "Metamodeling tumor-immune system interaction, tumor evasion and immunotherapy", Mathematical and Computer Modelling **47**, 5-6, 614–637, URL `https://www.sciencedirect.com/science/article/pii/S0895717707001951` (2008).

D'Onofrio, A., F. Gatti, P. Cerrai and L. Freschi, "Delay-induced oscillatory dynamics of tumour-immune system interaction", Mathematical and Computer Modelling **51**, 5-6, 572–591, URL `https://www.sciencedirect.com/science/article/pii/S089571770900404X` (2010).

Donsker, M. D., *An invariance principle for certain probability limit theorems* (1951).

Dunbar, S., "Traveling wave solutions of diffusive Lotka-Volterra equations", Journal of Mathematical Biology **17**, 1, 11–32, URL `http://link.springer.com/10.1007/BF00276112` (1983).

Durbin, J. and S. J. Koopman, *Time series analysis by state space methods* (Oxford university press, 2012).

Eftimie, R., J. L. Bramson and D. J. D. Earn, "Interactions Between the Immune System and Cancer: A Brief Review of Non-spatial Mathematical Models", Bulletin of Mathematical Biology **73**, 1, 2–32, URL `http://link.springer.com/10.1007/s11538-010-9526-3` (2011).

Eikenberry, S. E., T. Sankar, M. C. Preul, E. J. Kostelich, C. J. Thalhauser and Y. Kuang, "Virtual glioblastoma: Growth, migration and treatment in a three-dimensional mathematical model", Cell Proliferation **42**, 4, 511–528 (2009).

Eisenberg, M. C. and H. V. Jain, "A confidence building exercise in data and identifiability: Modeling cancer chemotherapy as a case study", Journal of Theoretical Biology **431**, 63–78, URL `https://www.sciencedirect.com/science/article/pii/S0022519317303454?via{%}3Dihub` (2017).

Evensen, G., *Data assimilation: the ensemble Kalman filter* (Springer Science & Business Media, 2009).

Everett, R. A., A. M. Packer and Y. Kuang, "Can mathematical models predict the outcomes of prostate cancer patients undergoing intermittent androgen deprivation therapy?", Biophysical Reviews and Letters **9**, 02, 173–191 (2014).

Farchi, A. and M. Bocquet, "Comparison of local particle filters and new implementations.", Nonlinear Processes in Geophysics **25**, 4 (2018).

Fedorov, A., R. Beichel, J. Kalpathy-Cramer, J. Finet, J.-C. Fillion-Robin, S. Pujol, C. Bauer, D. Jennings, F. Fennessy, M. Sonka, J. Buatti, S. Aylward, J. V. Miller, S. Pieper and R. Kikinis, "3D Slicer as an image computing platform for the Quantitative Imaging Network", Magnetic Resonance Imaging **30**, 9, 1323–1341, URL `https://www.sciencedirect.com/science/article/pii/S0730725X12001816?via{%}3Dihub` (2012).

Fisher, R., L. Pusztai and C. Swanton, "Cancer heterogeneity: implications for targeted therapeutics", British journal of cancer **108**, 3, 479 (2013).

Fisher, R. A., "The wave of advance of advantageous genes", Annals of Eugenics **7**, 4, 355–369 (1937).

Frank, T. D., "Delay Fokker-Planck equations, perturbation theory, and data analysis for nonlinear stochastic systems with time delays", Physical Review E **71**, 3, 31106, URL `https://journals.aps.org/pre/abstract/10.1103/PhysRevE.71.031106` (2005).

Frigola, R., F. Lindsten, T. B. Schön and C. E. Rasmussen, "Bayesian inference and learning in Gaussian process state-space models with particle MCMC", in "Advances in Neural Information Processing Systems", pp. 3156–3164 (2013).

Gałach, M., "Dynamics of the Tumor—Immune System Competition—the Effect of Time Delay", International Journal of Applied Mathematics and Computer Science **13**, 3, 395–406, URL `https://pdfs.semanticscholar.org/88fb/d5af40f9ebdba3e4d262b0d5bf80963199b4.pdf` (2003).

Gardner, T. S., C. R. Cantor and J. J. Collins, "Construction of a genetic toggle switch in Escherichia coli", Nature **403**, 6767, 339 (2000).

Gerlee, P. and S. Nelander, "Traveling wave analysis of a mathematical model of glioblastoma growth", Mathematical Biosciences **276**, 75–81, URL `https://www.sciencedirect.com/science/article/abs/pii/S0025556416000602?via{%}3Dihub` (2016).

Ghosh, A., S. Mukhopadhyay, S. Roy and S. Bhattacharya, "Bayesian inference in nonparametric dynamic state-space models", Statistical Methodology **21**, 35–48, URL `https://www.sciencedirect.com/science/article/abs/pii/S1572312714000197` (2014).

Gilbert, M. R., M. Wang, K. D. Aldape, R. Stupp, M. E. Hegi, K. A. Jaeckle, T. S. Armstrong, J. S. Wefel, M. Won, D. T. Blumenthal, A. Mahajan, C. J. Schultz, S. Erridge, B. Baumert, K. I. Hopkins, T. Tzuk-Shina, P. D. Brown,

A. Chakravarti, W. J. Curran and M. P. Mehta, "Dose-dense temozolomide for newly diagnosed glioblastoma: a randomized phase III clinical trial.", Journal of clinical oncology : official journal of the American Society of Clinical Oncology **31**, 32, 4085–91, URL http://ascopubs.org/doi/10.1200/JCO.2013.49.6968http://www.ncbi.nlm.nih.gov/pubmed/24101040http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3816958 (2013).

Gillespie, D. T., "A general method for numerically simulating the stochastic time evolution of coupled chemical reactions", Journal of computational physics **22**, 4, 403–434 (1976).

Gillespie, D. T., "The chemical Langevin equation", The Journal of Chemical Physics **113**, 1, 297–306 (2000).

Girard, A., C. E. Rasmussen, J. Q. Candela and R. Murray-Smith, "Gaussian process priors with uncertain inputs application to multiple-step ahead time series forecasting", in "Advances in neural information processing systems", pp. 545–552 (2003).

Gray, A., D. Greenhalgh, L. Hu, X. Mao and J. Pan, "A Stochastic Differential Equation SIS Epidemic Model", SIAM Journal on Applied Mathematics **71**, 3, 876–902, URL http://epubs.siam.org/doi/10.1137/10081856X (2011).

Guillouzic, S., I. L'Heureux and A. Longtin, "Small delay approximation of stochastic delay differential equations", Physical Review E **59**, 4, 3970–3982, URL https://link.aps.org/doi/10.1103/PhysRevE.59.3970 (1999).

Guo, W. and D.-C. Mei, "Stochastic resonance in a tumor-immune system subject to bounded noises and time delay", Physica A: Statistical Mechanics and its Applications **416**, 90–98, URL https://www.sciencedirect.com/science/article/pii/S0378437114006748 (2014).

Hajdu, S. I., "A note from history: landmarks in history of cancer, part 1", Cancer **117**, 5, 1097–1102 (2011).

Han, L., S. Eikenberry, C. He, L. Johnson, M. C. Preul, E. J. Kostelich, Y. Kuang, M. C. Preul, E. J. Kostelich and Y. Kuang, "Patient-specific parameter estimates of glioblastoma multiforme growth dynamics from a model with explicit birth and death rates", Mathematical Biosciences and Engineering **16**, 5, 5307–5323, URL http://www.aimspress.com/article/10.3934/mbe.2019265 (2019a).

Han, L., C. He and Y. Kuang, "Dynamics of a model of tumor-immune interaction with time delay and noise", DCDS-S. (2019b).

Harley, K., P. van Heijster, R. Marangell, G. J. Pettet and M. Wechselberger, "Existence of Traveling Wave Solutions for a Model of Tumor Invasion", SIAM Journal on Applied Dynamical Systems **13**, 1, 366–396, URL http://epubs.siam.org/doi/10.1137/130923129 (2014).

Higham., D. J., "An Algorithmic Introduction to Numerical Simulation of Stochastic Differential Equations", SIAM Review **43**, 3, 525–546, URL http://epubs.siam.org/doi/10.1137/S0036144500378302 (2001).

Hirata, Y., K. Akakura, C. S. Higano, N. Bruchovsky and K. Aihara, "Quantitative mathematical modeling of PSA dynamics of prostate cancer patients treated with intermittent androgen suppression", Journal of molecular cell biology **4**, 3, 127–132 (2012).

Hoelzinger, D., T. Demuth and M. Berens, "Autocrine factors that sustain glioma invasion and paracrine biology in the brain microenvironment", Journal of the National Cancer Institute **99**, 21, 1583–1593 (2007).

Hooten, M. B. and C. K. Wikle, "A hierarchical Bayesian non-linear spatio-temporal model for the spread of invasive species with application to the Eurasian Collared-Dove", Environmental and Ecological Statistics **15**, 1, 59–70, URL http://link.springer.com/10.1007/s10651-007-0040-1 (2008).

Hunt, B. R., E. J. Kostelich and I. Szunyogh, "Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter", Physica D: Nonlinear Phenomena **230**, 1-2, 112–126, URL https://www.sciencedirect.com/science/article/pii/S0167278906004647 (2007).

Ideta, A. M., G. Tanaka, T. Takeuchi and K. Aihara, "A mathematical model of intermittent androgen suppression for prostate cancer", Journal of nonlinear science **18**, 6, 593 (2008).

Jackson, P. R., J. Juliano, A. Hawkins-Daarud, R. C. Rockne and K. R. Swanson, "Patient-Specific Mathematical Neuro-Oncology: Using a Simple Proliferation and Invasion Tumor Model to Inform Clinical Practice", Bulletin of Mathematical Biology **77**, 5, 846–856, URL http://link.springer.com/10.1007/s11538-015-0067-7 (2015).

Kalman, R. E., "A new approach to linear filtering and prediction problems", (1960).

Kim, M., J. Kotas, J. Rockhill and M. Phillips, "A Feasibility Study of Personalized Prescription Schemes for Glioblastoma Patients Using a Proliferation and Invasion Glioma Model", Cancers **9**, 12, 51, URL http://www.mdpi.com/2072-6694/9/5/51 (2017).

Kirschner, D. and J. C. Panetta, "Modeling immunotherapy of the tumor - immune interaction", Journal of Mathematical Biology **37**, 3, 235–252, URL http://link.springer.com/10.1007/s002850050127 (1998).

Kohandel, M., S. Sivaloganathan and A. Oza, "Mathematical modeling of ovarian cancer treatments: sequencing of surgery and chemotherapy", Journal of theoretical biology **242**, 1, 62–68 (2006).

Kolmogorov, A. N., I. G. Petrovsky and N. S. Piskunov, "Investigation of the equation of diffusion combined with increasing of the substance and its application to a biology problem", Bull. Moscow State Univ. Ser. A: Math. Mech **1**, 6, 1–25 (1937).

Kostelich, E. J., Y. Kuang, J. M. McDaniel, N. Z. Moore, N. L. Martirosyan and M. C. Preul, "Accurate state estimation from uncertain data and models: an application of data assimilation to mathematical models of human brain tumors", Biology Direct **6**, 1, 64, URL http://biologydirect.biomedcentral.com/articles/10.1186/1745-6150-6-64 (2011).

Kot, M., *Elements of mathematical ecology* (Cambridge University Press, 2001), URL https://www.cambridge.org/us/academic/subjects/life-sciences/ecology-and-conservation/elements-mathematical-ecology?format=PB{&}isbn=9780521001502.

Kuang, Y., *Delay differential equations: with applications in population dynamics*, vol. 191 (Academic press, 1993).

Kuang, Y., J. D. Nagy and S. E. Eikenberry, *Introduction to Mathematical Oncology* (Chapman and Hall/CRC, 2015), 1 edn.

Kuang, Y., J. D. Nagy and S. E. Eikenberry, *Introduction to mathematical oncology* (Chapman and Hall/CRC, 2018).

Küchler, U. and B. Mensch, "Langevins stochastic differential equation extended by a time-delayed term", Stochastics and Stochastic Reports **40**, 1-2, 23–42, URL https://www.tandfonline.com/doi/full/10.1080/17442509208833780 (1992).

Kuznetsov, V. A., I. A. Makalkin, M. A. Taylor and A. S. Perelson, "Nonlinear dynamics of immunogenic tumors: Parameter estimation and global bifurcation analysis", Bulletin of Mathematical Biology **56**, 2, 295–321, URL https://www.sciencedirect.com/science/article/pii/S0092824005802605 (1994).

Lai, X. and A. Friedman, "Combination therapy of cancer with cancer vaccine and immune checkpoint inhibitors: A mathematical model", PLOS ONE **12**, 5, e0178479, URL https://dx.plos.org/10.1371/journal.pone.0178479 (2017).

Lan, G., Z. Chen, C. Wei and S. Zhang, "Stationary distribution of a stochastic SIQR epidemic model with saturated incidence and degenerate diffusion", Physica A: Statistical Mechanics and its Applications **511**, 61–77, URL https://www.sciencedirect.com/science/article/pii/S0378437118309130 (2018).

Lefever, R. and W. Horsthemke, "Bistability in fluctuating environments. Implications in tumor immunology", Bulletin of Mathematical Biology **41**, 4, 469–490, URL https://www.sciencedirect.com/science/article/pii/S0092824079800038 (1979).

Li, D. and F. Cheng, "Threshold for extinction and survival in stochastic tumor immune system", Communications in Nonlinear Science and Numerical Simulation **51**, 1–12, URL https://www.sciencedirect.com/science/article/pii/S1007570417300850 (2017).

Lipková, J., P. Angelikopoulos, S. Wu, E. Alberts, B. Wiestler, C. Diehl, C. Preibisch, T. Pyka, S. E. Combs, P. Hadjidoukas, Others, J. Lipkova, P. Angelikopoulos, S. Wu, E. Alberts, B. Wiestler, C. Diehl, C. Preibisch, T. Pyka, S. E. Combs, P. Hadjidoukas, K. Van Leemput, P. Koumoutsakos, J. Lowengrub and B. Menze, "Personalized Radiotherapy Design for Glioblastoma: Integrating Mathematical Tumor Models, Multimodal Scans, and Bayesian Inference", IEEE transactions on medical imaging **38**, 8, 1875–1884, URL `https://ieeexplore.ieee.org/document/8654016/` (2019).

Madzvamuse, A., R. Barreira and A. Gerisch, "Cross-Diffusion in Reaction-Diffusion Models: Analysis, Numerics, and Applications", pp. 385–392 (Springer, Cham, 2017), URL `http://link.springer.com/10.1007/978-3-319-63082-3{_}61`.

Mahoney, K. M., G. J. Freeman and D. F. McDermott, "The next immune-checkpoint inhibitors: PD-1/PD-L1 blockade in melanoma", Clinical therapeutics **37**, 4, 764–782 (2015).

Mao, X., *Stochastic differential equations and applications* (Elsevier, 2007).

Martirosyan, N. L., E. M. Rutter, W. L. Ramey, E. J. Kostelich, Y. Kuang and M. C. Preul, "Mathematically modeling the biological properties of gliomas: A review", Mathematical Biosciences and Engineering **12**, 4, 879–905, URL `http://aimsciences.org/journals/displayArticlesnew.jsp?paperID=11020` (2015).

McDaniel, J., E. Kostelich, Y. Kuang, J. Nagy, M. C. Preul, N. Z. Moore and N. L. Matirosyan, "Data Assimilation in Brain Tumor Models", pp. 233–262 (Springer, New York, NY, 2013), URL `http://link.springer.com/10.1007/978-1-4614-4178-6{_}9`.

Morzfeld, M., D. Hodyss and J. Poterjoy, "Variational particle smoothers and their localization", Quarterly Journal of the Royal Meteorological Society **144**, 712, 806–825 (2018).

Murray, J., "Glioblastoma brain tumors: estimating the time from brain tumor initiation and resolution of a patient survival anomaly after similar treatment protocols", Journal of Biological Dynamics **6:sup2**, 118–127 (2012).

Murray, J. D., *Mathematical Biology: I. An Introduction* (Springer New York, 2002).

Neal, M. L., A. D. Trister, T. Cloke, R. Sodt, S. Ahn, A. L. Baldock, C. A. Bridge, A. Lai, T. F. Cloughesy, M. M. Mrugala, J. K. Rockhill, R. C. Rockne and K. R. Swanson, "Discriminating Survival Outcomes in Patients with Glioblastoma Using a Simulation-Based, Patient-Specific Response Metric", PLoS ONE **8**, 1, e51951, URL `https://dx.plos.org/10.1371/journal.pone.0051951` (2013).

Neal, R. M., *Bayesian learning for neural networks*, vol. 118 (Springer Science & Business Media, 2012).

Nikolopoulou, E., L. R. Johnson, D. Harris, J. D. Nagy, E. C. Stites and Y. Kuang, "Tumour-immune dynamics with an immune checkpoint inhibitor", Letters in Biomathematics **5**, sup1, S137–S159, URL `https://www.tandfonline.com/doi/full/10.1080/23737867.2018.1440978` (2018).

Norden, A. D. and P. Y. Wen, "Glioma therapy in adults", The neurologist **12**, 6, 279–92, URL `http://www.ncbi.nlm.nih.gov/pubmed/17122724` (2006).

NORTON, L., R. SIMON, H. D. BRERETON and A. E. BOGDEN, "Predicting the course of Gompertzian growth", Nature **264**, 5586, 542–545, URL `http://www.nature.com/doifinder/10.1038/264542a0` (1976).

Øksendal, B., "Stochastic differential equations", in "Stochastic differential equations", pp. 65–84 (Springer, 2003).

Ott, E., B. R. Hunt, I. Szunyogh, A. V. Zimin, E. J. Kostelich, M. Corazza, E. Kalnay, D. J. Patil and J. A. Yorke, "A local ensemble Kalman filter for atmospheric data assimilation", Tellus A: Dynamic Meteorology and Oceanography **56**, 5, 415–428 (2004).

Papoulis, A. and S. U. Pillai, *Probability, random variables, and stochastic processes* (Tata McGraw-Hill Education, 2002).

Parish, C. R., "Cancer immunotherapy: the past, the present and the future", Immunology and cell biology **81**, 2, 106–113 (2003).

Parmar, K., K. B. Blyuss, Y. N. Kyrychko and S. J. Hogan, "Time-Delayed Models of Gene Regulatory Networks", Computational and Mathematical Methods in Medicine **2015**, 1–16, URL `http://www.hindawi.com/journals/cmmm/2015/347273/` (2015).

Penny, W. D., K. J. K. J. Friston, J. Ashburner, S. Kiebel and T. Nichols, *Statistical parametric mapping : the analysis of funtional brain images* (Elsevier/Academic Press, 2007).

Phan, T., K. Nguyen, P. Sharma and Y. Kuang, "The impact of intermittent androgen suppression therapy in prostate cancer modeling", Applied Sciences **9**, 1, 36 (2019).

Pope, W., J. Sayre, A. Perlina, J. Villablanca, P. Mischel and T. Cloughesy, "{MR} imaging correlates of survival in patients with high-grade gliomas", American Journal of Neuroradiology **10**, 2474–2644 (2005).

Portz, T., Y. Kuang and J. D. Nagy, "A clinical data validated mathematical model of prostate cancer growth under intermittent androgen suppression therapy", AIP Advances **2**, 1, 011002, URL `http://aip.scitation.org/doi/10.1063/1.3697848` (2012).

Quinn, J., "A High-Dimensional Particle Filter Algorithm", URL `http://arxiv.org/abs/1901.10543` (2019).

Rasmussen, C. E. and C. K. I. Williams, *Gaussian processes for machine learning* (MIT Press, 2006), URL https://mitpress.mit.edu/books/gaussian-processes-machine-learning.

Rebeschini, P. and R. van Handel, "Can local particle filters beat the curse of dimensionality?", The Annals of Applied Probability **25**, 2809–2866, URL https://www.jstor.org/stable/24521616 (2015).

Rihan, F. A., D. H. Abdel Rahman, S. Lakshmanan and A. S. Alkhajeh, "A time delay model of tumour-immune system interactions: Global dynamics, parameter estimation, sensitivity analysis", Applied Mathematics and Computation **232**, 606–623, URL https://www.sciencedirect.com/science/article/pii/S0096300314001568 (2014).

Rihan, F. A., C. Tunc, S. H. Saker, S. Lakshmanan and R. Rakkiyappan, "Applications of delay differential equations in biological systems", Complexity **2018** (2018).

Sherratt, J. A., "Wavefront propagation in a competition equation with a new motility term modeling contact inhibition between cell populations", Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences **456**, 2002, 2365–2386, URL http://www.royalsocietypublishing.org/doi/10.1098/rspa.2000.0616 (2000).

Sherratt, J. A. and M. A. J. Chaplain, "A new mathematical model for avascular tumor growth", Journal of Mathematical Biology **43**, 4, 291–312, URL http://link.springer.com/10.1007/s002850100088 (2001).

Simon, R. and L. Norton, "The Norton–Simon hypothesis: designing more effective and less toxic chemotherapeutic regimens", Nature Clinical Practice Oncology **3**, 8, 406–407 (2006).

Skipper, H. E., "Kinetics of mammary tumor cell growth and implications for therapy", Cancer **28**, 6, 1479–1499 (1971).

Skipper, H. E., F. M. Schabel, M. W. Trader and J. R. Thomson, "Experimental evaluation of potential anticancer agents VI. Anatomical distribution of leukemic cells and failure of chemotherapy", Cancer research **21**, 9, 1154–1164 (1961).

Snyder, C., T. Bengtsson, P. Bickel and J. Anderson, "Obstacles to high-dimensional particle filtering", Monthly Weather Review **136**, 12, 4629–4640 (2008).

Stein, M. L., *Interpolation of spatial data: some theory for kriging* (Springer Science & Business Media, 2012).

Stepien, T. L., E. M. Rutter and Y. Kuang, "A data-motivated density-dependent diffusion model of in vitro glioblastoma growth.", Mathematical biosciences and engineering : MBE **12**, 6, 1157–72, URL http://www.ncbi.nlm.nih.gov/pubmed/26775861 (2015).

Stepien, T. L., E. M. Rutter and Y. Kuang, "Traveling Waves of a Go-or-Grow Model of Glioma Growth", SIAM Journal on Applied Mathematics **78**, 3, 1778–1801, URL https://epubs.siam.org/doi/10.1137/17M1146257 (2018).

Stupp, R., W. P. Mason, M. J. van den Bent, M. Weller, B. Fisher, M. J. Taphoorn, K. Belanger, A. A. Brandes, C. Marosi, U. Bogdahn, J. Curschmann, R. C. Janzer, S. K. Ludwin, T. Gorlia, A. Allgeier, D. Lacombe, J. G. Cairncross, E. Eisenhauer and R. O. Mirimanoff, "Radiotherapy plus Concomitant and Adjuvant Temozolomide for Glioblastoma", New England Journal of Medicine **352**, 10, 987–996, URL http://www.nejm.org/doi/abs/10.1056/NEJMoa043330 (2005).

Swanson, K. R., R. C. Rockne, J. Claridge, M. A. Chaplain, E. C. Alvord and A. R. A. Anderson, "Quantifying the Role of Angiogenesis in Malignant Progression of Gliomas: In Silico Modeling Integrates Imaging and Histology", Cancer Research **71**, 24, 7366–7375, URL http://www.ncbi.nlm.nih.gov/pubmed/21900399http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3398690http://cancerres.aacrjournals.org/cgi/doi/10.1158/0008-5472.CAN-11-1399 (2011).

Swanson, K. R., R. C. Rostomily and E. C. Alvord, "A mathematical modellng tool for predicting survival of individual patients following resection of glioblastoma: a proof of principle", British Journal of Cancer **98**, 1, 113–119, URL http://www.nature.com/articles/6604125 (2008).

Szunyogh, I., E. J. Kostelich, G. Gyarmati, E. Kalnay, B. R. Hunt, E. Ott, E. Satterfield and J. A. Yorke, "A local ensemble transform Kalman filter data assimilation system for the NCEP global model", Tellus A: Dynamic Meteorology and Oceanography **60**, 1, 113–130 (2008).

Tian, X.-J., H. Zhang and J. Xing, "Coupled reversible and irreversible bistable switches underlying TGFbeta-induced epithelial to mesenchymal transition", Biophysical journal **105**, 4, 1079–1089 (2013).

Tracqui, P., G. C. Cruywagen, D. Woodward, G. Bartoo, J. Murray and E. Alvord, "A mathematical model of glioma growth: The effect of chemotherapy on spatial-temporal growth", Cell Proliferation **28**, 17–31 (1995).

Umulis, D. M., M. Serpe, M. B. O'Connor and H. G. Othmer, "Robust, bistable patterning of the dorsal surface of the Drosophila embryo", Proceedings of the National Academy of Sciences **103**, 31, 11613–11618 (2006).

van der Hoorn, A., J. Yan, T. Larkin, N. Boonzaier, T. Matys and S. Price, "Validation of a semi-automatic co-registration of MRI scans in patients with brain tumors during treatment follow-up", NMR in Biomedicine **2**, 7, 882–889 (2016).

Volpert, A. I., "Traveling wave solutions of parabolic systems: translations of mathematical monographs", (1994).

Volpert, V. and S. Petrovskii, "Reaction–diffusion waves in biology", Physics of life reviews **6**, 4, 267–310 (2009).

Wan, E. A. and R. Van Der Merwe, "The unscented Kalman filter for nonlinear estimation", in "Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No. 00EX373)", pp. 153–158 (Ieee, 2000).

Wang, J., D. Fleet and A. Hertzmann, "Gaussian Process Dynamical Models for Human Motion", IEEE Transactions on Pattern Analysis and Machine Intelligence **30**, 2, 283–298, URL `http://ieeexplore.ieee.org/document/4359316/` (2008).

Wang, L., D. Jiang, G. S. K. Wolkowicz and D. O'Regan, "Dynamics of the stochastic chemostat with Monod-Haldane response function", Scientific Reports **7**, 1, 13641, URL `http://www.nature.com/articles/s41598-017-13294-3` (2017).

Watling, C., D. Lee, D. Macdonald and J. Cairncross, "Corticosteroid-induced magnetic resonance imaging changes in patients with recurrent malignant glioma", Journal of Clinical Oncology **12**, 9, 1886–1889 (1994).

Whiteside, T. L., "Immune suppression in cancer: Effects on immune cells, mechanisms and future therapeutic intervention", Seminars in Cancer Biology **16**, 1, 3–15, URL `https://www.sciencedirect.com/science/article/pii/S1044579X0500060X?via{%}3Dihub` (2006).

Wikle, C. K. and M. B. Hooten, "A general science-based framework for dynamical spatio-temporal models", TEST **19**, 3, 417–451, URL `http://link.springer.com/10.1007/s11749-010-0209-z` (2010).

Wilkinson, D. J., "Stochastic modelling for quantitative description of heterogeneous biological systems", Nature Reviews Genetics **10**, 2, 122–133 (2009).

Wilkinson, D. J., *Stochastic modelling for systems biology* (CRC press, 2011).

Williams, C. K. I. and C. E. Rasmussen, "Gaussian processes for regression", in "Advances in neural information processing systems", pp. 514–520 (1996).

Woodward, D., J. Cook, P. Tracqui, G. Cruywagen, J. Murray and E. Alvord, "A mathematical model of glioma growth: The effect of extent of surgical resection", Cell Proliferation **29**, 269–288 (1996).

Wu, D., H. Wang and S. Yuan, "Stochastic sensitivity analysis of noise-induced transitions in a predator-prey model with environmental toxins", Mathematical biosciences and engineering: MBE **16**, 4, 2141–2153 (2019a).

Wu, J. and X. Zou, "Traveling wave fronts of reaction-diffusion systems with delay", Journal of Dynamics and Differential Equations **13**, 3, 651–687 (2001).

Wu, Z., T. Phan, J. Baez, Y. Kuang and E. J. Kostelich, "Predictability and identifiability assessment of models for prostate cancer under androgen suppression therapy", Mathematical biosciences and engineering: MBE **16**, 5, 3512–3536 (2019b).

Zaki, H., M. Jenkinson, D. Du Plessis, T. Smith and N. Rainov, "Vanishing contrast enhancement in malignant glioma after corticosteroid treatment", Acta Neurochirurgica **146**, 8, 841–845 (2004).

Zeng, C. and H. Wang, "Noise and large time delay: Accelerated catastrophic regime shifts in ecosystems", Ecological Modelling **233**, 52–58, URL `https://www.sciencedirect.com/science/article/pii/S030438001200141X` (2012).

APPENDIX A

SUPPLEMENTAL MATERIAL FOR CHAPTER 2

We write (2.4) as following

$$\frac{du}{dt} = \gamma u[\eta(u - \frac{u}{K}) - v], \tag{A.1a}$$

$$\frac{dv}{dt} = \beta - v + \frac{u}{1+u^2} \tag{A.1b}$$

where $\eta = \rho/\gamma$. First, I will address a question: would a slack immune system itself cause loss of stability of the low tumor equilibrium? To put it precise, we assume γ is large, i.e. growth and killing of tumor cells are fast compared to the change of the immune system level. Note that instead of a large multiplier γ in (A.1a), we can rescale time to put a small ϵ to (A.1b) which is more commonly seen in the literature. Since this change does not make any difference to the nullclines and the geometric argument used in Theorem 2.3.1, we conclude that slack immune system by itself does not cause loss of stability of the low tumor equilibrium.

Next I will derive a condition for the number of interior equilibrium. Consider the nullclines

$$v = \eta(1 - \frac{u}{K}), \tag{A.2a}$$

$$v = \beta + \frac{u}{1+u^2} \equiv h(u) \tag{A.2b}$$

from which we can find the intersection(s) by solving a cubic polynomial. Since the formula for cubic polynomial is unwieldy, we instead derive a condition to determine the number of intersections by geometric argument. Note that the boundary cases happen when the nullcline (A.2a) is tangent to the nullcline (A.2b). That is, $-\frac{\eta}{K} = h'(u)$ which gives

$$\phi u^4 + (2\phi - 1)u^2 + \phi + 1 = 0$$

where $\phi = \eta/K$. Real roots exist iff $\phi < 1/8$. This inequity also makes sense by noting that $\min h'(u) = -1/8$. Assuming $\phi < 1/8$, we have two roots

$$u_\pm = \left(\frac{1 - 2\phi \pm \sqrt{1-8\phi}}{2\phi}\right)^{1/2}$$

where we note that the quantity inside the bracket is positive since $\phi < 1/8$. Let $v_\pm = h(u_\pm)$ so that $(u_\pm, v_\pm)$ are tangent points of the boundary cases. Let $\eta_\pm = v_\pm + \phi u_\pm$,

which are the v-intercepts of the tangent lines. In order to have three intersections, the following inequality has to be satisfied

$$\eta_+ < \eta < \eta_-.$$

Note that $\eta_\pm$ is also a function of η. The above inequality is implicit. Nevertheless, we can easily plot it. Figure A shows the parameter regions when $\beta - 0.02$, in which the nullclines corresponding to regions A and C are shown in Figure A and the nullclines corresponding region B are shown in Figure 2.1.

We further note that as the parameter values cross the boundary, there is a transcritical bifurcation, which is not very interesting. In contrast, when the time delay is introduced (see (2.6)), there is a Hopf bifurcation with the length of time delay τ as the bifurcation parameter (see Theorem 2.4.1). Furthermore, there is a global bifurcation which is not indicated in the local analysis: as the τ further increases, the limit cycle runs into a saddle point and disappears (See the bottom panes of Figure 2.2; see also Page 150 of Kot (2001) on this type of bifurcation in a predator-prey model).

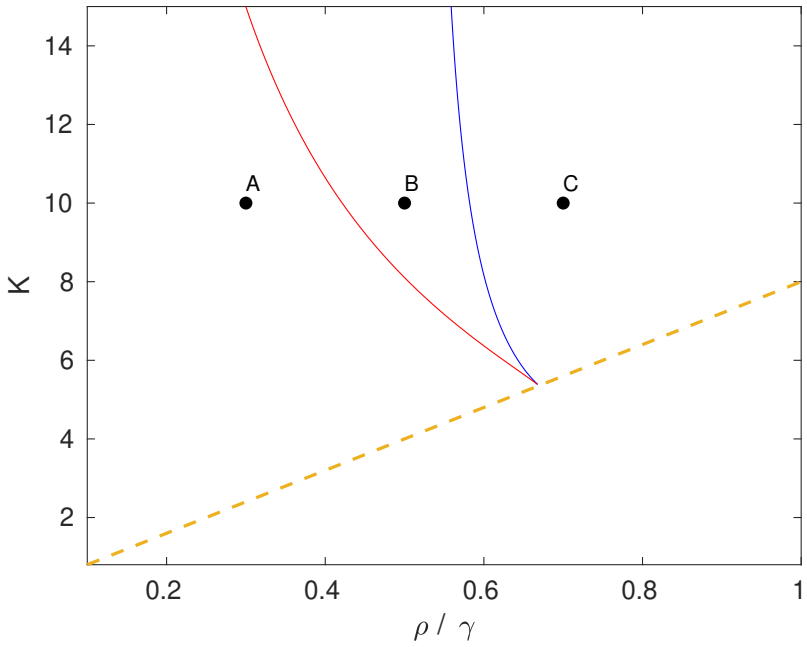

Figure A.1: Parameter regions for one and three interior equilibria when $\beta = 0.02$. In regions A and C, there is one interior equilibrium. In region B, there are three three interior equilibria. The red solid curve is from $\eta = \eta_+$; The blue solid curve is from $\eta = \eta_-$; The dashed yellow curve is from $\phi = 1/8$.

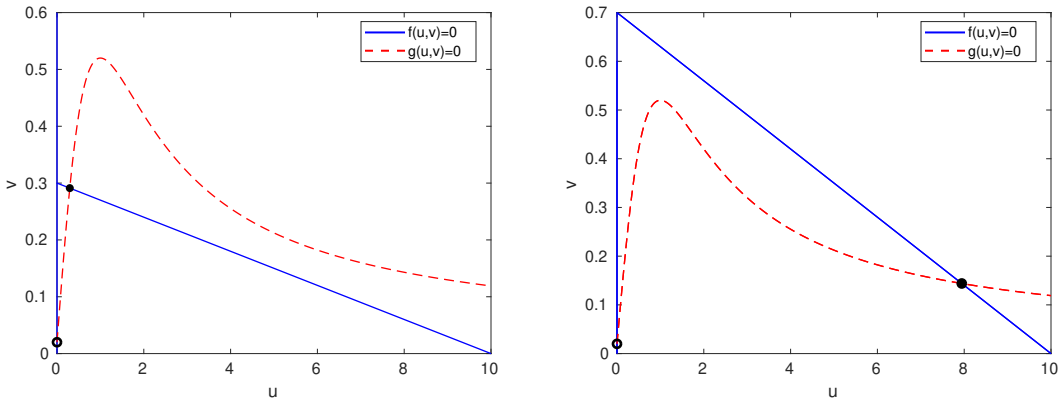

Figure A.2: Left: Nullclines corresponding to the dot in region A. Right: Nullclines corresponding the dot in region C.

APPENDIX B

EXISTENCE OF TRAVELING WAVE: A PHASE SPACE ANALYSIS

We want to show there is a traveling wave solution for the following

$$\begin{aligned} u_t &= u_{xx} + g(w)u - \delta(w)u \\ n_t &= \delta(w)u \end{aligned}$$

where

$$w = 1 - u - n. \tag{B.1}$$

For now we let $g(w) = w, \delta(w) = 1 - w$. Substituting $u(x,t) = u(x+ct)$ to the above gives

$$\begin{aligned} n' &= \frac{1}{c}(u+n)u \\ u' &= z \\ z' &= cz + (2u + 2n - 1)u \end{aligned} \tag{B.2}$$

We want to show the existence of a traveling wave equation by showing there is a trajectory from $(0,0,0)$ to $(n^*,0,0)$ where $n^* > 0$.

We do so in a similar line as Dunbar (1983) (and adopt similar notations). Define

$$\begin{aligned} Q &= \{n < 1, u < 0, z < 0\} \\ P &= \{n > 1, u > 0, z > 0\} \\ W &= \mathbb{R}^3 \setminus (P \cup Q) \end{aligned}$$

where we note that W is closed and will be shown to be a Wazeeski set. We define

$$\begin{aligned} J_1 &= \{(n,u,z) : n = 1, u \le -1, z \le 0\} \\ J_2 &= \{(n,u,z) : n \le 1, u \le 0, 2u + 2n - 1 \le 0, z = 0\} \end{aligned}$$

The immediate leaving set $W^- = \partial W \setminus (J_1 \cup J_2 \cup \{(n,u,z) : u = 0, z = 0\})$ which consists of two disjoint sets $\partial Q \setminus (J_1 \cup J_2 \cup \{u = 0, z = 0\})$ and $\partial P \setminus \{(1,0,0)\}$.

Lemma B.0.1. *Suppose $n(0) \ge 0$. Then $n(s) \ge 0$*

Proof. Suppose contrary that there is s s.t. $n(s) < 0$. Let $s_1 = \inf\{s : n(s) < 0\}$. Then $n(s_1) = 0, n'(s_1) = u^2 < 0$, which is impossible. □

Lemma B.0.2. *Let $c > 2$. Suppose $z(0) > \frac{c}{2}u(0), u(0) > 0, n(0) \ge 0$. Then $z(s) > \frac{c}{2}u(s), u(s) > 0 \ \forall s > 0$*

Proof. Suppose contrary that there is s s.t. $z(s) \leq \frac{c}{2}u(s)$. Let $s_1 = \inf\{s > 0 : z(s) \leq \frac{c}{2}u(s)\}$. For $0 \leq s \leq s_1$,$u' = z > \frac{c}{2}u$. Since $u(0) > 0$, $u(s) > 0$ for $0 \leq s \leq s_1$. Also, we have $z(s_1) = \frac{c}{2}u(s_1)$ and $z'(s_1) - \frac{c}{2}u'(s_1) \leq 0$. We will construct a contradiction to the last inequality. Henceforth, we suppress s_1:

$$\begin{aligned} z' - \frac{c}{2}u' &= cz + (2u + 2n - 1)u - \frac{c}{2}z \\ &= (\frac{c^2}{4} + 2u + 2n - 1)u > 0 \end{aligned}$$

since $u > 0, n \geq 0, c > 2$. So we have reached a contradiction. □

We draw a sufficient small circle around origin on the u-z plane. The circle intersects with the strong unstable manifold and positive u-axis. Define Σ be the arc connecting these two points in the first quadrant of u-z plane. Denote the solution of (B.2) starting at point $\boldsymbol{y}_1$ as $\boldsymbol{y}(s; \boldsymbol{y}_1)$. Then we have the following

Lemma B.0.3. *There is a $\boldsymbol{y}_1 \in \Sigma$ such that the solution $\boldsymbol{y}(s; \boldsymbol{y}_1)$ remains in W for all s*

Proof. Since we can make the circle as small as needed, the trajectory starting at the intersection between Σ and positive u-axis must enter Q (need to make more precise) and the other end point of Σ stratifies Lemma B.0.2. Then for the trajectory starting from the later point we have $u'(s), n'(s) > 0$ for all $s > 0$. Thus it must enter P.

To show existence of $\boldsymbol{y}(s; \boldsymbol{y}_1)$ remains in W, we suppose contrary that all trajectories starting from Σ exit W. We can show that W is a Wazeeski set (need to make more precise). Then by Proposition 1 in Dunbar (1983), the flow is a homeomorphism of the connected set Σ to a disjoint set W^-, which is impossible. Thus there must exist a $\boldsymbol{y}_1 \in \Sigma$ such that the solution $\boldsymbol{y}(s; \boldsymbol{y}_1)$ remains in W for all s □

Lemma B.0.4. *Suppose $\boldsymbol{y}(s)$ is a trajectory staying in W. Then it remains in the region $\mathcal{D} = \{0 \leq u \leq 1, 0 < n \leq 1, -\alpha u < z < du\}$ where $\alpha = 2, d$ are some positive numbers.*

Proof. We first show $u \geq 0$. Suppose contrary that $u(s) < 0$ for some s. Let $s_1 = \inf\{s : u(s) < 0\}$. Then $u(s_1) = 0$ and $z(s_1) = u'(s_1) < 0$. So it must enter Q which is a contradiction.

To show $u \leq 1$, suppose contrary that $u(s) > 1$ for some s. Let $s_1 = \inf\{s : u(s) > 1\}$. Then $u(s_1) = 1, z(s_1) = u'(s_1) > 0$. It follows that $n'(s) > 0, u'(s) > 0, z'(s) > 0$ for $s > s_1$. So it must enter P which is a contradiction. Similarly, $n \leq 1$.

To show $z < du$, we let $d > c$. Then $z(0) < du(0)$. Suppose contrary that $z(s) \geq du(s)$ for some s. Let $s_1 = \inf\{s : z(s) \geq du(s)\}$. Then $z(s_1) = du(s_1)$ and $z'(s_1) - du'(s_1) \geq 0$. We will construct a contradiction to the last inequality. Henceforth, we suppress s_1:

$$\begin{aligned} z' - du' &= cz + (2u + 2n - 1)u - dz \\ &= \Big((c-d)d + (2u + 2n - 1)\Big)u < 0 \end{aligned}$$

if we pick sufficient large d. So we have reached a contradiction.

To show $z > -\alpha u$, suppose contrary that there is s_1 s.t. $z(s_1) < -\alpha u(s_1)$. If there is $s_2 > s_1$ s.t. $z(s_2) = -\alpha u(s_2)$, then $z'(s_2) + \alpha u'(s_2) \geq 0$. However

$$\begin{aligned} z' + \alpha u &= cz + (2u + 2n - 1)u + \alpha z \\ &= \Big(-\alpha(c + \alpha) + (2u + 2n - 1)\Big)u < 0 \end{aligned}$$

if we pick sufficient large α, say $\alpha = 2$. Hence there is no such s_2, i.e., $z(s) < -\alpha u(s)$ for all $s > s_1$. It follows that $u'(s) < -\alpha u(s)$ and $z(s) < 0$ for $s > s_1$. Thus it must enter Q which is a contradiction. □

To make sure the availability of nutrients represented by (B.1) stays positive, we need the following

Lemma B.0.5. *The trajectory $\boldsymbol{y}(s)$ is bounded by $n + u \leq 1$*

Proof. Suppose contrary that $n(s) + u(s) > 1$ for some s. Let $s_1 = \inf\{s : n + u > 1\}$. Then $n(s_1) + u(s_1) = 1$, and

$$n'(s_1) + u'(s_1) = z(s_1) + \frac{1}{c}u(s_1) > 0. \tag{B.3}$$

There must exist some $s > s_1$ such that $u(s) + n(s) < 1$, otherwise $\boldsymbol{y}(s)$ cannot stay in $\mathcal{D}$. Let $s_2 = \inf\{s : n + u < 1, s > s_1\}$. Then $u(s_2) + n(s_2) = 1$ and

$$u'(s_2) + n'(s_2) = z(s_2) + \frac{1}{c}u(s_2) < 0.$$

Also, for $s \in (s_1, s_2)$, we have

$$(z + \frac{1}{c}u)' = cz + (2u + 2n - 1)u + \frac{z}{c} > c(z + \frac{1}{c}u).$$

Together with (B.3), we have $z(s_2) + \frac{1}{c}u(s_2) > 0$, which is a contradiction. □

Theorem B.0.6. *If $c > 2$, there is a trajectory from $(0,0,0)$ to $(n^*,0,0)$ where $n^* > 0$, i.e., there exists a traveling wave*

Proof. Define $V(n,u,z) = n - \log n + \frac{1}{2}u - \frac{1}{2c}z$. Note that V is continuous and bounded below in $\mathcal{D}$. Also,

$$\dot{V} = -\frac{1}{c}(\frac{u^2}{n} + \frac{u}{2}) \leq 0$$

The equality holds iff $u = 0$. To be invariant at $u = 0$, it further implies that $z = u' = 0$. Thus the trajectory must ends at $(0,0,n^*)$ □

APPENDIX C

SUPPLEMENTAL MATERIAL FOR CHAPTER 3

Traveling-wave solutions

In the following, we rigorously establish the existence of traveling-wave solutions in system (3.9). We first show that the solutions of system (3.9) with positive initial values are non-negative and bounded.

Lemma 1. Assume that $g(w) = wG(w)$, where $G(w)$ is a bounded function. Then the solutions of system (3.9) with positive initial values are positive and bounded.

Proof: If $x' = x\, f(t,x)$ and f is a bounded function, then $x(t) = x(t_0)\exp\left(\int_{t_0}^{t} f(s,x(s))\,ds\right)$, which is positive whenever $x(t_0) > 0$. Since

$$\frac{d(p+w)}{dz} = -\delta(w), \tag{C.1}$$

we see that $(p+w)$ is bounded, which implies that both p and w are also bounded. □

For any $w^* \in [0,1]$, the point $(0, w^*)$ is an equilibrium of (3.9).

Theorem 1. System (3.9) admits positive traveling-wave solutions that correspond to heteroclinic orbits connecting the steady state $(0,1)$ to another steady state, $(0, w^*)$.

Proof: We have performed a detailed phase-plane analysis of (3.9) to show the existence of a trajectory that starts from $(1,0)$ and ends at $(w^*, 0)$, where $w^* \in [0,1)$. (See Figure 3.1.) First we notice that

$$\frac{dp}{dw} = \frac{\delta(w)}{\hat{\rho} g(w)} - 1, \tag{C.2}$$

$$\frac{d^2p}{dw^2} = \frac{\delta'(w)g(w) - \delta(w)g'(w)}{\hat{\rho} g(w)^2}, \tag{C.3}$$

Since $g(w)$ and $\delta(w)$ are both positive functions and $g'(w) > 0$ and $\delta'(w) < 0$ on $w \in [0,1]$, it follows that $d^2p/dw^2 < 0$ for all $w \in [0,1]$.

We show first that the trajectory starting from $(1,0)$ will never cross the line $p = 1 - w$. Because $g(1) = 1$, we have $\delta(1) = 0$, and therefore, $dp/dw\big|_{w=1} = -1$. Since $d^2p/dw^2 < 0$, the slope of a trajectory with $w < 1$ will be greater than -1, which means it will not cross the line $p = 1 - w$. Also, because the solution components stay positive, the trajectory starting from $(1,0)$ will never cross the p-axis from right to left.

Because the functions g and δ are monotone, there is a unique value $w^\dagger \in (0,1)$ such that $dp/dw\big|_{w=w^\dagger} = 0$. Moreover, $dp/dw\big|_{1>w>w^\dagger} > 0$ while $dp/dw\big|_{0<w<w^\dagger} < 0$.

Insofar as w is strictly decreasing and bounded from below by 0, there exists some $w^* \in (0,1)$ such that $\lim_{z\to\infty} w(z) = w^*$. We claim that $w^* < w^\dagger$. Otherwise, p is

a non-decreasing function, which implies that the trajectory approaches a positive steady state $E^* = (p^*, w^*)$. However, the system (3.9) does not admit any positive steady states.

Let $a \in (w^*, w^\dagger)$ and $b = \hat{\rho} g(a) - \delta(a) < 0$. Since $\lim_{z\to\infty} w(z) = w^*$, there is a $z^* > 0$ such that for $z \geq z^*$ and $w < a$. Therefore, for $z \geq z^*$, $dp/dz < bp$, which implies that $\lim_{z\to\infty} p(z) = 0$. Hence, system (3.9) admits positive traveling-wave solutions that correspond to heteroclinic orbits connecting the steady state $(0, 1)$ to another steady state, $(0, w^*)$. □

Rim width

The trajectory in Figure 3.1 corresponds to a traveling-wave profile in the z coordinate as shown on the left pane of Figure 3.3. Because of our choice of nondimensionlization, $x = -\left(2\sqrt{D\rho}\,/k\right) z$, by definition we have the rim width in the dimensional form

$$\ell_1 = \frac{2\sqrt{D\rho}}{k}(z_1^+ - z_1^-) = \frac{2\sqrt{D\rho}}{k} \int_{z_1^-}^{z_1^+} dz \tag{C.4}$$

where $p(z_1^\pm) = a_1 p_{\max}$. We have taken p as a function of z (cf. the left pane of Figure 3.3). We can also take w as a function of z to make a change of variable to the above integral, which gives Eq. (3.11a). A similar argument applies to (3.11b).

Derivation of R_1^* and R_2^*

Tumor growth can be separated into two stages. First, the tumor cells grow exponentially until the cell density is high enough ($p_{\max}$) to form a stable wave profile. Afterward, the growth of tumor cells is described by (3.9). The radii R_1 and R_2 are observed from clinical MRI data and correspond respectively to the distance from the center of the tumor to the edge of the enhancing rim and to the edge of the edematous region on T2 imaging; they correspond to tumor dynamics before the stable wave profile has formed (cf. Figure C.1).

During the exponential growth phase, say from $0 < t < t^*$, quiescence is negligible and the governing equation of tumor cell density is

$$\frac{\partial p}{\partial t} = D\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial p}{\partial r}\right) + \rho p, \tag{C.5}$$

where spherical symmetry is used, as we assume that the tumor is spherical when its

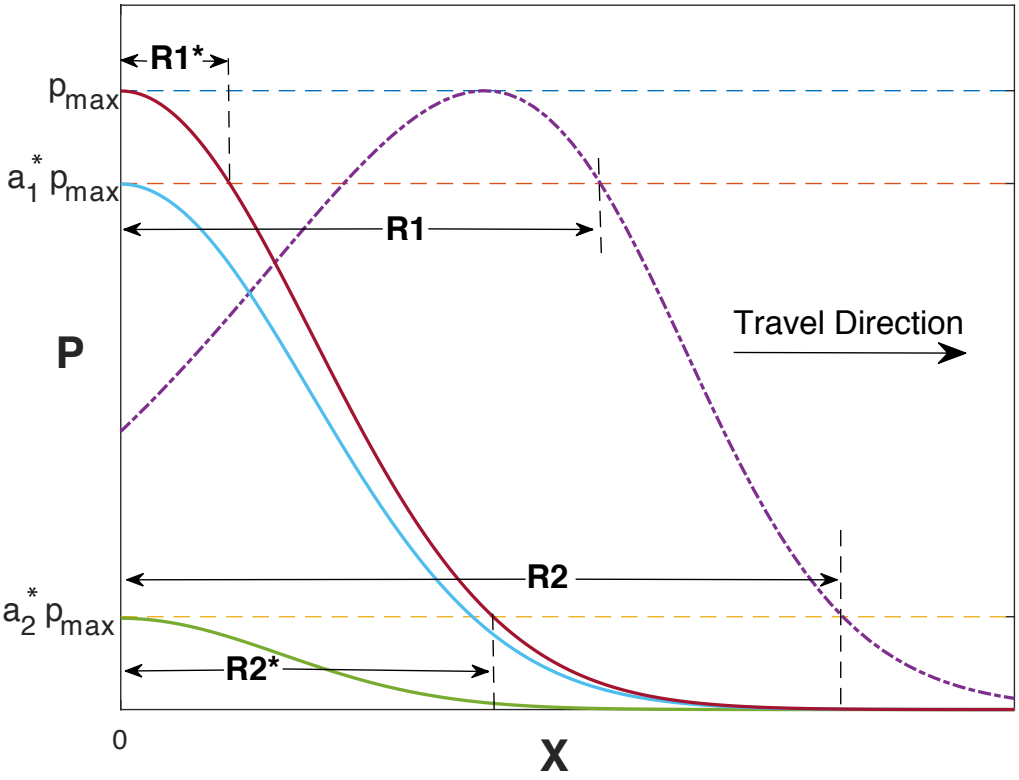

Figure C.1: The transition of wave profile based on the two-stage tumor growth model. The dotted (purple) curve represents a stable wave profile generated by system (3.9). The solid curves are generated by Eq. (C.5) and represent the tumor's exponential growth phase. The stable wave profile is formed after the exponential growth curve reaches $p_{\max}$ (red solid curve).

radius is small. This linear equation has the Green's function

$$p(r,t) = \frac{1}{(4\pi Dt)^{3/2}} \exp\left(\rho t - \frac{r^2}{4Dt}\right) \tag{C.6}$$

(see, e.g., Kot (2001) Page 314 and Britton (2003) Page 285). Suppose that the tumor starts at $t = 0$ as a point source with density p_0. It follows that at $t = t^*$, the position of the tumor front at density $p = a_i$ is given by Eq. (3.15).

APPENDIX D

SUPPLEMENTAL MATERIAL FOR CHAPTER 4

D.1 Derivation for Propagating Prediction and Uncertainty

Since product of Gaussian functions are again a Gaussian, we can do the following integration analytically. In general, for $\mathbf{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

$$\begin{aligned}&\int k(\hat{\mathbf{x}}, \mathbf{x}; \mathbf{V}) p(\mathbf{x}) d\mathbf{x} \\ =&\sigma_k^2 \left|\boldsymbol{\Sigma}\mathbf{V} + \mathbf{I}\right|^{-1/2} \exp\left(-\frac{1}{2}(\hat{\mathbf{x}} - \boldsymbol{\mu})^T (\boldsymbol{\Sigma} + \mathbf{V}^{-1})^{-1} (\hat{\mathbf{x}} - \boldsymbol{\mu}) \right)\end{aligned} \tag{D.1}$$

$$\begin{aligned}&\int k(\hat{\mathbf{x}}, \mathbf{x}; \mathbf{V}) k(\hat{\mathbf{y}}, \mathbf{x}; \mathbf{W}) p(\mathbf{x}) d\mathbf{x} \\ =&\frac{\sigma_k^4}{\left|\boldsymbol{\Sigma}(\mathbf{V} + \mathbf{W}) + \mathbf{I}\right|^{1/2}} \exp\left(-\frac{1}{2}\big(\hat{\mathbf{x}}^T \mathbf{V} \hat{\mathbf{x}} + \hat{\mathbf{y}} \mathbf{W} \hat{\mathbf{y}} - \mathbf{q}^T (\mathbf{V} + \mathbf{W}) \mathbf{q}\big) \right) \\ &\exp\left(-\frac{1}{2}(\mathbf{q} - \boldsymbol{\mu})^T (\boldsymbol{\Sigma} + (\mathbf{V} + \mathbf{W})^{-1})^{-1} (\mathbf{q} - \boldsymbol{\mu}) \right)\end{aligned} \tag{D.2}$$

where $\mathbf{q} = (\mathbf{V} + \mathbf{W})^{-1}(\mathbf{V}\hat{\mathbf{x}} + \mathbf{W}\hat{\mathbf{y}})$. In the following, we show how to take into account of the uncertainty of input in forecasting by using (D.1) and (D.2).

To simplify notations, we denote training data as $\mathcal{D} = (\hat{\mathbf{u}}, \tilde{\mathbf{v}})$ and test input $\mathbf{u}^* \sim N(\mathbf{m}, \mathbf{S})$. Note that here we deviate from our previous notations in main text in favor of presenting a general method. We want to find the mean and variance of the test output $\mathbf{u} \equiv (u_j)_{j \in \{1:n_x\}}$. The GP regression tells that

$$\mathbf{u} | \mathbf{u}^*, \mathcal{D} \sim N(\mathbf{u}^* + \mathbf{K}_*^T \boldsymbol{\Sigma}^{-1} (\tilde{\mathbf{v}} - \hat{\mathbf{u}}), \mathbf{K}_{**} - \mathbf{K}_*^T \boldsymbol{\Sigma}^{-1} \mathbf{K}_*)$$

where $(\mathbf{K}_*)_{ij} = k(\hat{\mathbf{u}}_{i-1:i+1}, \mathbf{u}^*_{j-1:j+1})$, $(\boldsymbol{\Sigma})_{ij} = k(\hat{\mathbf{u}}_{i-1:i+1}, \hat{\mathbf{u}}_{j-1:j+1}) + \sigma_\epsilon^2 \delta_{ij}$ and $(\mathbf{K}_{**})_{ij} = k(\mathbf{u}^*_{i-1:i+1}, \mathbf{u}^*_{j-1:j+1})$. To simplify notation we let $\boldsymbol{\beta} = \boldsymbol{\Sigma}^{-1}(\tilde{\mathbf{v}} - \hat{\mathbf{u}})$ and note that it is sorely determined by training data $\mathcal{D}$. The mean of $\mathbf{u}$ is easy, i.e.,

$$\begin{aligned} E[\mathbf{u}] &= E[E[\mathbf{u}|\mathbf{u}^*]] \\ &= E[\mathbf{u}^* + \mathbf{K}_*^T \boldsymbol{\beta}] \\ &= \mathbf{m} + \mathbf{L}^T \boldsymbol{\beta} \end{aligned}$$

where by simply applying (D.1), we have

$$\begin{aligned} (\mathbf{L})_{rj} &= \int k(\hat{\mathbf{u}}_{r-1:r+1}, \mathbf{u}^*_{j-1:j+1}; \boldsymbol{\Lambda}^{-1}) p(\mathbf{u}^*_{j-1:j+1}) d\mathbf{u}^*_{j-1:j+1} \\ &= \sigma_k^2 \left|\mathbf{S}_{\mathcal{N}(j)} \boldsymbol{\Lambda}^{-1} + \mathbf{I}\right|^{-1/2} \\ &\quad \exp\left(-\frac{1}{2}(\hat{\mathbf{u}}_{r-1:r+1} - \mathbf{m}_{j-1:j+1})^T (\mathbf{S}_{b_j} + \boldsymbol{\Lambda})^{-1} (\hat{\mathbf{u}}_{r-1:r+1} - \mathbf{m}_{j-1:j+1}) \right) \end{aligned}$$

where $\mathbf{S}_{\mathcal{N}(j)}$ is the sub-covariance matrix that concerns only $\mathcal{N}(j)$.

The variance-covariance matrix of $\mathbf{u}$ is slightly more complicated. Let's consider it component-wisely. By law of total variance,

$$cov(u_i, u_j) \quad = \quad E[cov(u_i, u_j|\mathbf{u}^*)] + cov(E[u_i|\mathbf{u}^*], E[u_j|\mathbf{u}^*])$$

We first compute

$$\begin{aligned}
&cov(E[u_i|\mathbf{u}^*], E[u_j|\mathbf{u}^*]) \\
=&E[E[u_i|\mathbf{u}^*]E[u_j|\mathbf{u}^*]] - E[E[u_i|\mathbf{u}^*]] \cdot E[E[u_j|\mathbf{u}^*]] \\
=&E[(u_i^* + \mathbf{k}_{*i}^T\boldsymbol{\beta})(u_j^* + \mathbf{k}_{*j}^T\boldsymbol{\beta})] - (m_i + \mathbf{l}_i^T\boldsymbol{\beta})(m_j + \mathbf{l}_j^T\boldsymbol{\beta}) \\
=&E[u_i^*u_j^*] + E[u_i^*\mathbf{k}_{*j}^T\boldsymbol{\beta}] + E[\mathbf{k}_{*i}^T\boldsymbol{\beta}u_j^*] + \boldsymbol{\beta}^T E[\mathbf{k}_{*i}\mathbf{k}_{*j}^T]\boldsymbol{\beta} - (m_i + \mathbf{l}_i^T\boldsymbol{\beta})(m_j + \mathbf{l}_j^T\boldsymbol{\beta}) \\
=&S_{ij} + E[u_i^*\mathbf{k}_{*j}^T\boldsymbol{\beta}] + E[\mathbf{k}_{*i}^T\boldsymbol{\beta}u_j^*] + \boldsymbol{\beta}^T(E[\mathbf{k}_{*i}\mathbf{k}_{*j}^T] - \mathbf{l}_i\mathbf{l}_j^T)\boldsymbol{\beta} - (m_i\mathbf{l}_j^T + m_j\mathbf{l}_i^T)\boldsymbol{\beta}
\end{aligned}$$

where $\mathbf{k}_{*i}$ is the i-th column of $\mathbf{K}_*$ and $\mathbf{l}_i$ is the i-th column of $\mathbf{L}$. We try to determine term by term. First,

$$E[u_i^*u_j^*] = S_{ij} + m_im_j.$$

Second and third, $E[u_i^*\mathbf{k}_{*j}^T]$ is a vector and its r-th entry is

$$\begin{aligned}
&\int u_i^* k(\hat{\mathbf{u}}_{\mathcal{N}(r)}, \mathbf{u}^*_{\mathcal{N}(j)}) p(\mathbf{u}^*_{\mathcal{N}(j)\cup i}) d\mathbf{u}^*_{\mathcal{N}(j)\cup i} \\
=&\int u_i^* k(\mathbf{P}^T\hat{\mathbf{u}}_{\mathcal{N}(r)}, \mathbf{u}^*_{\mathcal{N}(j)\cup i}; \mathbf{P}^T\boldsymbol{\Lambda}^{-1}\mathbf{P}) p(\mathbf{u}^*_{\mathcal{N}(j)\cup i}) d\mathbf{u}^*_{\mathcal{N}(j)\cup i} \\
=&\text{last entry of } (\mathbf{P}^T\boldsymbol{\Lambda}^{-1}\mathbf{P} + \mathbf{S}^{-1}_{\mathcal{N}(j)\cup i})^{-1}(\mathbf{P}^T\boldsymbol{\Lambda}^{-1}\hat{\mathbf{u}}_{r-1:r+1} + \mathbf{S}^{-1}_{\mathcal{N}(j)\cup i}\mathbf{m}_{\mathcal{N}(j)\cup i\cup i})
\end{aligned}$$

if $i \notin \mathcal{N}(j)$ where $\mathbf{P}\mathbf{u}^*_{\mathcal{N}(j)\cup i} = \mathbf{u}^*_{\mathcal{N}(j)}$. If $i \in \mathcal{N}(j)$, the r-th entry of $E[u_i^*\mathbf{k}_{*j}^T]$ is the one of the entries of

$$(\boldsymbol{\Lambda}^{-1} + \mathbf{S}^{-1}_{\mathcal{N}(j)})^{-1}(\boldsymbol{\Lambda}^{-1}\hat{\mathbf{u}}_{\mathcal{N}(r)} + \mathbf{S}^{-1}_{\mathcal{N}(j)}\mathbf{m}_{\mathcal{N}(j)})$$

that corresponds to the position of i in $\mathcal{N}(j)$.

Fourth, $E[\mathbf{k}_{*i}\mathbf{k}_{*j}^T] \equiv \tilde{\mathbf{L}}$ is a matrix with

$$\begin{aligned}
\tilde{L}_{rs} \quad &= \quad \int k(\hat{\mathbf{u}}_{\mathcal{N}(r)}, \mathbf{u}^*_{\mathcal{N}(i)}) k(\hat{\mathbf{u}}_{\mathcal{N}(s)}, \mathbf{u}^*_{\mathcal{N}(j)}) p(\mathbf{u}^*_{\mathcal{N}(i)\cup\mathcal{N}(j)}) d\mathbf{u}^*_{\mathcal{N}(i)\cup\mathcal{N}(j)} \\
&= \quad \int k(\mathbf{P}^T\hat{\mathbf{u}}_{\mathcal{N}(r)}, \mathbf{u}^*_{\mathcal{N}(i)\cup\mathcal{N}(j)}; \mathbf{P}^T\boldsymbol{\Lambda}^{-1}\mathbf{P}) k(\mathbf{Q}^T\hat{\mathbf{u}}_{\mathcal{N}(s)}, \mathbf{u}^*_{\mathcal{N}(i)\cup\mathcal{N}(j)}; \mathbf{Q}^T\boldsymbol{\Lambda}^{-1}\mathbf{Q}) \\
&\quad\quad p(\mathbf{u}^*_{\mathcal{N}(i)\cup\mathcal{N}(j)}) d\mathbf{u}^*_{\mathcal{N}(i)\cup\mathcal{N}(j)}
\end{aligned}$$

where $\mathbf{P}\mathbf{u}^*_{\mathcal{N}(i)\cup\mathcal{N}(j)} = \mathbf{u}^*_{\mathcal{N}(i)}$ and $\mathbf{Q}\mathbf{u}^*_{\mathcal{N}(i)\cup\mathcal{N}(j)} = \mathbf{u}^*_{\mathcal{N}(j)}$. We note that the above is in the form of (D.2) and hence can be readily evaluated.

Next we compute

$$E[cov(u_i, u_j|\mathbf{u}^*)] = E[k(\mathbf{u}^*_{i-1:i+1}, \mathbf{u}^*_{j-1:j+1})] - E[Tr(\mathbf{\Sigma}^{-1}\mathbf{k}_{*i}\mathbf{k}^T_{*j})]$$

The first term on right hand side

$$\begin{aligned} E[k(\mathbf{u}^*_{\mathcal{N}(i)}, \mathbf{u}^*_{\mathcal{N}(j)})] &= \int k(\mathbf{u}^*_{\mathcal{N}(i)}, \mathbf{u}^*_{\mathcal{N}(j)}) p(\mathbf{u}^*_{\mathcal{N}(i)\cup\mathcal{N}(j)}) d\mathbf{u}^*_{\mathcal{N}(i)\cup\mathcal{N}(j)} \\ &= \int k(0, \mathbf{u}^*_{\mathcal{N}(i)\cup\mathcal{N}(j)}; \mathbf{P}^T\mathbf{\Lambda}^{-1}\mathbf{P}) p(\mathbf{u}^*_{\mathcal{N}(i)\cup\mathcal{N}(j)}) d\mathbf{u}^*_{\mathcal{N}(i)\cup\mathcal{N}(j)} \end{aligned}$$

where $\mathbf{P}\mathbf{u}^*_{i-1:i+1\cup j-1:j+1} = \mathbf{u}^*_{i-1:i+1} - \mathbf{u}^*_{j-1:j+1}$. The second term

$$\begin{aligned} E[Tr(\mathbf{\Sigma}^{-1}\mathbf{k}_{*i}\mathbf{k}^T_{*j})] &= Tr(\mathbf{\Sigma}^{-1}E[\mathbf{k}_{*i}\mathbf{k}^T_{*j}]) \\ &= Tr(\mathbf{\Sigma}^{-1}\tilde{\mathbf{L}}) \end{aligned}$$

Putting together, we have

$$cov(u_i, u_j) = S_{ij} + E[u^*_i\mathbf{k}^T_{*j}\boldsymbol{\beta}] + E[\mathbf{k}^T_{*i}\boldsymbol{\beta}u^*_j] + \boldsymbol{\beta}^T(E[\mathbf{k}_{*i}\mathbf{k}^T_{*j}] - \mathbf{l}_i\mathbf{l}^T_j)\boldsymbol{\beta} - (m_i\mathbf{l}^T_j + m_j\mathbf{l}^T_i)\boldsymbol{\beta}$$

D.2 Mean Square Continuity and Differentiability of Gaussian Processes

In this section, we include some results of continuity and differentiability of Gaussian processes for completeness. These results can be found in Cramer and Leadbetter (2004); Adler (2010). What we present here is an extension to Cramer and Leadbetter (2004) and a complement to Adler (2010). Consider $f(s) \sim \mathcal{GP}(m(s), k(\cdot,\cdot))$ where $s \in \mathbb{R}^n$ and squared exponential covariance function $k(x,y) = \exp(-\frac{\|x-y\|^2}{2\ell^2})$. Without loss of generality, we assume $m(x) = 0$ (we could simply center the process by its mean function). We want to investigate the continuity and differentiability of $f(x)$. There are many types of convergence with which continuity and differentiability can be described. Here we focus on continuity and differentiability in the mean square which are defined below. The mean square limit is often denoted as $l.i.m.$. Next we define the mean square continuity.

Definition D.2.1. *(mean square convergence) Given a sequence of random variable X_n, we say that X_n converge to X in mean square if $E[(X_n - X)^2] \to 0$.*

Mean square convergence is preferred here because it is closely related to the covariance function.

Definition D.2.2. *(mean square continuity) Let f be a stochastic process indexed by $s \in \mathbb{R}^n$. If for any sequence $s_1, s_2, \cdots$ that converges to s, i.e., $\|s_n - s\| \to 0$, we have $f(s_n) \to f(s^*)$ in mean square, then we say that $f(s)$ is continuous in mean square at s^*.*

Theorem D.2.3. *A stochastic process $f(s)$ is continuous in mean square at s^* if and only if its covariance function $k(s,t)$ is continuous at s^*.*

Proof. The "if" part is straightforward. By applying definitions, we have

$$\begin{aligned} E[(f(s_n) - f(s))^2] &= E[f(s_n)^2 - 2f(s_n)f(s) + f(s)^2] \\ &= k(s_n, s_n) - 2k(s_n, s) + k(s, s) \to 0 \end{aligned}$$

The "only if" part is a little more complicated. Suppose $f(s)$ is mean square continuous at s^*. Let $\{s_n\}$ and $\{\hat{s}_n\}$ be two sequences that converge to s^*. Consider

$$\begin{aligned} & |k(s_n, \hat{s}_n) - k(s, s)| \\ = \; & |E[f(s_n)f(\hat{s}_n)] - E[f(s)^2]| \\ = \; & |E[\big(f(s_n) - f(s)\big)\big(f(\hat{s}_n) - f(s)\big) + \big(f(s_n) - f(s)\big)f(s) + \big(f(\hat{s}_n) - f(s)\big)f(s)] \\ \leq \; & \Big(E[\big(f(s_n) - f(s)\big)^2]E[\big(f(\hat{s}_n) - f(s)\big)^2] + E[\big(f(s_n) - f(s)\big)^2]E[f(s)^2] \\ & + E[\big(f(\hat{s}_n) - f(s)\big)^2]E[f(s)^2] \Big)^{1/2} \to 0 \end{aligned}$$

where we have applied Cauchy-Schwarz Inequality. □

Let δ_i be a vector in $\mathbb{R}^n$ of all zeros except the i-th element being 1. We define the mean square derivative below.

Definition D.2.4. *(mean square derivative) If the limit*

$$f_i(s) = l.i.m._{h \to 0} \frac{f(s + h\delta_i) - f(s)}{h}$$

exists, then $f_i(s)$ is called the mean square derivative of $f(s)$ at s.

Before we show conditions for differentiability in mean square, we state a lemma that is needed.

Lemma D.2.5. *Let X_n be a sequence of random variables. Then X_n converge in mean square if and only if $E[X_m X_n]$ coverage to a finite constant as $m \to \infty$ and $n \to \infty$.*

Theorem D.2.6. *If the derivative $\partial^2 k(s,t)/\partial s_i \partial t_i$ exists and is finite at (s,s), then the mean square derivative $f_i(s)$ exists.*

Proof. Assume that $\partial^2 k(s,t)/\partial s_i \partial t_i$ exists and is finite at (s,s). Consider

$$\begin{aligned} & E[\frac{f(s+h\delta_i)-f(s)}{h} \cdot \frac{f(s+h'\delta_i)-f(s)}{h'}] \\ =& \frac{k(s+h\delta_i, s+h'\delta_i) - k(s+h\delta_i, s) - k(s, s+h'\delta_i) + k(s,s)}{hh'} \\ & \to \partial^2 k(s,t)/\partial s_i \partial t_i \end{aligned}$$

as $h \to 0$ and $h' \to 0$. Then it follows by lemma D.2.5. □

Follow similar lines as Definition D.2.4 and Theorem D.2.6, higher order of mean square derivative can be defined and deduced from the differentiability of the covariance. For example, the second-order derivative

$$f_{ij}(s) = l.i.m. \frac{f(s+h\delta_i+h'\delta_j) - f(s+h\delta_i) - f(s+h'\delta_j) + f(s)}{hh'}$$

exists if $k(s,t)$ has derivative $\partial^2 k(s,t)/\partial s_i \partial t_i \partial s_j \partial t_j$. Since the squared exponential $k(x,y) = \exp(-\frac{\|x-y\|^2}{2\ell^2})$ is differentiable at all orders, it follows that $f \sim \mathcal{GP}(0, k(\cdot,\cdot))$ is mean square differentiable at all orders.

APPENDIX E

SUMMARY OF COMPUTER CODE

The computer code is summarized in the three tables below for each chapter. All code is available on the author's Github (https://github.com/juhang62).

Table E.1: computer code files and summary for Chapter 2

file name	usage
`nullclines.m`	plots Figure 2.1
`bistablefinal.m`	plots Figure 2.2; DDE is solved with the built-in `dde23`
`sdefinal.m`	plots Figure 2.3; SDE is solved using the Milstein method
`sdefinal_statdist_delayed.m`	plots Figure 2.4; SDDE is solved using the Milstein method combined with the method of steps.

Table E.2: computer code files and summary for Chapter 3

file name	usage
`plotphaseplane.m`	plots Figure 2.1
`uw-phaseplane.mat`	stores the trajectory data of Figure 2.1
`plotwaveprofile.m`	plots the left pane of Figure 3.3
`waveprofile.mat`	stores wave profile data of Figure 3.4 3.3
`l1l2.m`	the function that generates data for Figure and 3.5
`VtoR.m`	converts the volume data from MRI to radii listed in Table 3.1
`findparams.m`	the function that estimate parameters in Table 3.1.
`wavprof_num.m`	compute the numerical estimate for Figure 3.5
`wavprof.m`	computes analytical estimate for for Figure 3.5
`scatterplot.mat`	stores wave profile characteristics data for Figure 3.5

Table E.3: Computer code files and summary for Chapter 4

file name	usage
`main.m`	the main file
`predictorGau.m`	a function takes mean and variance of the state variable as input and compute the forecasted mean and variance of the state variable in the next time step
`plotstuff.m`	plots Figures 4.2 and 4.3
`plot2d.m`	plots Figure 4.4

APPENDIX F

CO-AUTHOR PERMISSIONS

I certify that my co-authors, Changhan He and Dr. Yang Kuang have given me permission, in writing, to include all material in my PhD thesis for Chapter 2.

I certify that my co-authors, Dr. Steffen Eikenberry, Changhan He, Lauren Johnson, Dr. Mark C. Preul, Dr. Eric J. Kostelich and Dr. Yang Kuang have given me permission, in writing, to include all material in my PhD thesis for Chapter 3.