Dynamics and Optimal Control of the Growth of the Gut Microbiome with Varying Nutrient

by

Brittni Hall

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
August 5, 2023

Keywords: chemostat, gut, microbiome, nonautonomous system, optimal control, shooting methods

Approved by

Maggie Han, Co-Chair, Professor of Mathematics and Statistics
Hans-Werner Van Wyk, Co-Chair, Associate Professor of Mathematics and Statistics
Peter E. Kloeden, Professor of Mathematics and Statistics
Georg Hetzer, Professor Emeritus of Mathematics and Statistics
Thi-Thao-Phuong Hoang, Associate Professor of Mathematics and Statistics
Raj Amin, Associate Professor of Drug Discovery and Development, University Reader

Abstract

The human gut microbiome consists of all the microbes that make up the human intestinal tract, and many diseases are associated with certain microbial compositions in the gut. First, a mathematical model describing the growth of gut microbiome inside and on the wall of the gut is developed based on the chemostat model with wall growth. Both the concentration and flow rate of the nutrient input are time-dependent, which results in a system of non-autonomous differential equations. First the stability of each meaningful equilibrium is studied for the autonomous counterpart. Then the existence of pullback attractors and its detailed structures for the nonautonomous system are investigated using theory and techniques of nonautonomous dynamical systems. In particular, sufficient conditions under which the microbiome vanishes or persists are constructed. Numerical simulations are provided to illustrate the theoretical results. Then a second model is developed describing the growth of one beneficial bacterial population with time-varying controlled rate of the input flow of the nutrient. First the stability of each meaningful equilibrium is studied for a constant input case. Then schemes are developed using optimal control theory to find an optimal time-varying rate of input flow. Numerical comparison simulations are provided.

Acknowledgments

This will already be much too long, and it won't include everyone, but to each person who has been along on this ride with me, I never could have done it without you.

- Maggie Han: Thank you for pushing me to be better than I think I am. Thank you for the endless hours you've put in on my behalf. Thank you for countless dinners, for celebrating the big and the small accomplishments, and for helping me enter the world of academia. I am forever indebted to you for the connections you've made for me, and all the ways you have uniquely personalized advising me. I'm so grateful that you have helped me achieve my personal goals, and have developed me into the mathematician and the academic that I am and am becoming.

- Hans-Werner van Wyk: Thank you for debugging my code even when Maggie said you were being too nice. Thank you for working late nights and early mornings, for always telling me both, "You CAN do this," and "You HAVE to do this." Thank you for one of the best graduate classes I've ever taken and propagating my love for applied math. Thank you for all the times you had to re-explain something to me, for being patient, and letting me brainstorm and ramble away in your office.

- Ricardo Cordero: I never *would* have done this without you. More than anyone else, you have supported my math career for the past 10+ years and I am so very grateful. Thank you for showing me what it's like to make much of Jesus in the way that I do math and the way that I teach. Thank you for not forgetting about me. Thank you for always being my biggest fan.

- Lora Merchant: Thank you for being a friend and a teacher. For all the ways you taught me to be a better educator. Thank you for always having a seat for me in your office to debrief or complain. Thank you for loving me and checking in on me through the years.

And thank you for forgiving me for that first week I didn't show up as your TA. I'm still sorry about that.

- Melinda Lanius: You were the missing piece to my mentors I didn't even know I needed. Thank you for leading by example, for all the snack breaks, and for helping me immensely in the job search. You have truly helped me to flourish, and I look forward to working with you for years to come.

- My committee members: A very special thank you to Drs. Kloeden, Hetzer, and Hoang for agreeing to serve on my committee. Thank you to Dr. Kloeden for the countless emails with links to interesting articles about the gut microbiome, even when they made me grimace.

- Dr. Raj Amin: Thank you for stepping up at the last minute and agreeing to be my University Reader. You truly saved the day and made this possible! Thank you for doing so with such enthusiasm.

- Dolores Williamson: Thank you for being the one to instill a love of math in me. Thank you for the foundation you laid and your passionate teaching which have been crucial to my success.

- Mom and Dad: Thank you for trusting me as I've embarked on this seemingly endless journey. Thank you for always encouraging me to do what I want to do, even when it's not the coolest or going to make me the most money. Thank you for teaching me to do what the Lord has called me to. Thank you for always being on my side, and never letting me doubt how proud you are of me and how much you love me.

- Zac and Allison Boman: Thank you for staying up too late with me to let me talk through everything. Thank you for being a balm to my anxiety and never letting me go without. And thank you to Rebekah and Ruthie for letting me borrow your mama even when we talked too long.

- Cara: I quite literally couldn't have done this without you. Thank you for being my best friend, for being someone I could look up to as a researcher and a follower of Christ. Thank you for loving STEM and Jesus alongside me. Thank you for all the times you've been the one I call crying. For understanding the unique joys and pressures of grad school. For sitting across from me right now as we both work on finishing our dissertations. For letting me into your family. You're my person and I am so grateful for you.

- Sidetrack, David, & Austin: Thank you for endless amounts of joy and cups of coffee. Thank you for all the times you started a pour-over for me before I even ordered. Thank you for an environment in which I could get a lot of this work done.

- Mary Carol: Thank you for celebrating every stepping stone, small and big. For countless prayers and post-it notes of encouragement. For dinners to celebrate passed prelims and to laugh and cry. Thank you for making our house a home I want to come back to at the end of a long day.

- Brigitte and Heather: To my aunts who have supported me emotionally, financially, and mentally throughout this entire process. Thank you for believing in me, for showing me girls could love math and excel at it, and for being there at 3AM when I was so riddled with anxiety I couldn't sleep. Thank you for regularly checking in on me from across the country.

- Jared: Thanks for always saying, "*If* she finishes," so that I would be motivated to prove you wrong. I love you.

- Grace Heritage Church: To my church family, there aren't enough words to thank you. You have been my rock and my grounding place, my home-away-from-home. Thank you for the prayers, the generous gifts, and the soul nourishment I so desperately longed for. Thank you for being the most bittersweet part of closing this chapter, and being such a hard place to leave.

To my faithful God, by Whose will and grace I've finally finished this dang thing. To You be all the glory.

Table of Contents

Abstract ii

Acknowledgments iii

1 Introduction 1

2 A nonautonomous chemostat model for the growth of gut microbiome with varying nutrient 4

2.1 Introduction of the Model and Basic Properties of Solutions 4

2.2 The Autonomous Chemostat Microbiome Model 11

2.3 The Nonautonomous Chemostat Microbiome Model 14

2.3.1 Preliminaries on Nonautonomous Dynamical Systems 15

2.3.2 Existence of a Pullback Attractor 17

2.3.3 Detailed Long Term Dynamics 19

2.4 Numerical Simulations 23

3 An optimal control analysis of a model of a chemostat model for the growth of gut microbiome with varying nutrient 29

3.1 Introduction of the Model and Basic Properties of Solutions 30

3.1.1 Existence and Uniqueness of Solutions 32

3.1.2 Model without the Probiotic 34

3.2 The Optimal Control Problem 37

3.2.1 Preliminaries on Optimal Control Theory 37

3.2.2 Optimal Control for the Gut Microbiome System 40

3.2.3 The Hamiltonian Boundary Value Problem and the Associated Linearized System 43

3.3 Shooting Methods 47

3.3.1 Sensitivity Equations 49

3.3.2 Single Shooting Method 53

3.3.3 Multiple Shooting Method 56

3.4 Numerical Simulations 63

3.4.1 Example 1: Single Shooting Method 63

3.4.2 Example 2: Single Shooting Fails, Multiple Shooting Converges 65

3.4.3 Example 3: Multiple Shooting Method 66

4 Conclusion and Future Work 68

References 70

List of Figures

1.1 Design of a chemostat model with continuous agitation and aeration [2] 3

2.1 Plot of the fraction $R(t)$ of bacteria in the gut for different initial conditions. . 24

2.2 Plot showing the joint dynamics of the nutrient $x(t)$ and the total bacteria $Y(t)$ over the time interval $[0, 2]$ under assumptions (**A1**) and (**A2**). The arrows indicate the trajectory directions. 25

2.3 The time evolution of the fraction $R(t)$ of bacteria in the gut and the total bacteria $Y(t)$ under nonautonomous, periodic forcing with parameters satisfying Assumptions (**A1**), (**A2**), and (**A3**), but not (**A4**). 25

2.4 Plots showing the convergence of the total bacteria $Y(t)$ to 0 and that of the nutrition level $x(t)$ to $x^*(t)$ under nonautonomous, periodic forcing and Assumptions (**A1**)–(**A4**). 26

2.5 Joint trajectories of the nutrition x, the bacteria in the gut y_1, and the bacteria on the wall y_2 subject to nonautonomous, periodic forcing, with parameters safisfying (**A1**)–(**A4**). 26

2.6 Plot of the joint dynamics of the nutrition level x and the total bacteria population Y under nonautonomous, periodic forcing, with parameters satisfying Assumptions (**A1**)–(**A3**), (**A5**), and (**A6**). 27

2.7 Plot of the trajectories in Figure 2.6 on a logarithmic scale. 27

2.8 Trajectory plot $(x(t), Y(t))$ in the time interval $[0, 50]$ of the nutrition x and the total bacteria population Y under constant nutrient injection for various initial conditions. 28

2.9 Plot of the fraction of bacteria in the gut over time under autonomous forcing and Assumptions (**A1**)–(**A3**), (**A5**), and (**A6**). 28

3.1 An example of multiple shooting method's convergence with parameter values given by $I = 2$, $\mu = 0.8$, $\alpha = 2$, $\beta = 2$, $m = 0.1$, $\nu = 0.5$, $c = 1$, and $u_{max} = 5$. 62

3.2 Plot of the joint dynamics of the nutrition level x and the bacterial population y, along with adjoint trajectories λ_1 and λ_2 with parameters $I = 2$, $\mu = 0.8$, $\alpha = 2$, $\beta = 2$, $m = 0.1$, $\nu = 0.5$, $c = 1$, and $u_{max} = 5$. 64

3.3 Plots from the Single Shooting Method of the optimal control u and the norm of the residual $\boldsymbol{r}$ after 40 iterations with parameters $I = 2$, $\mu = 0.8$, $\alpha = 2$, $\beta = 2$, $m = 0.1$, $\nu = 0.5$, $c = 1$, and $u_{max} = 5$. 64

3.4 A contrast between Single Shooting failure and Multiple Shooting convergence, both with time interval $t \in [0, .5]$ and parameters $I = 2$, $\mu = 0.8$, $\alpha = 2$, $\beta = 2$, $m = 0.1$, $\nu = 0.5$, $c = 1$, and $u_{max} = 5$. 65

3.5 An example of Multiple Shooting Method with convergence with parameters $I = 2$, $\mu = 0.8$, $\alpha = 2$, $\beta = 2$, $m = 0.1$, $\nu = 0.5$, $c = 1$, and $u_{max} = 5$. 66

3.6 The control functions are plotted over $[0, 1]$ for $u^*(t)$ and $\bar{u} = \frac{1}{T}\int_0^T u^*(t)dt$ with parameters $I = 2$, $\mu = 0.8$, $\alpha = 2$, $\beta = 2$, $m = 0.1$, $\nu = 0.5$, $c = 1$, and $u_{max} = 5$. Below that, the Lagrangian cost functions $L(t, \boldsymbol{x}(t), u^*(t))$ and $L(t, \boldsymbol{x}(t), \bar{u})$ are plotted, showing the area under the curve of the optimal control is more negative than the area under the curve of the constant control. 67

List of Tables

2.1 Wall growth model parameters 6

3.1 Optimal control model parameters 30

Chapter 1

Introduction

The human microbiome is the collection of microscopic organisms that live in the body, and it contains representatives from all domains of life [28]. The list of genes in the human microbiome discovered thus far is so extensive that biologists refer to the human microbiome as the "second genome" [54]. There is strong evidence that the microbiome interacts with the rest of a human body, immune system, and brain, and plays an important role in a variety of human diseases, such as infections, arthritis, food allergy, cancer, inflammatory bowel disease, neurological diseases, obesity, and diabetes (see, e.g., [23, 24, 26, 29, 31, 33, 41, 42, 43, 45] and references therein). Not only does the human microbiome play a role in physical well-being, but some studies in mice have shown that differing microbiome compositions are related to levels of social engagement and anxiety [54].

The human gut microbiome consists of all the microbes that make up the human intestinal tract, and many diseases are associated with certain microbial compositions in the gut (see, e.g., [26, 31, 32, 41, 43, 45]). For example, in a recent report by Curry [25], it was mentioned that correlations have been found between lower diversity of bacteria in the gut and higher rates of certain medical conditions. Over time, the makeup and diversity of the human gut microbiome change. In one recent archaeological finding, 1000-year old piles of feces indicate significant evolution in the diversity of the bacterial populations in the human gut [25]. With the advent of modern medicine, and specifically the discovery of penicillin in the twentieth century, come new health issues such as the deadly disease C diff [55].

Studies in Spain and the United States have shown there may be certain gut microbes that affect cardiovascular health and risk of strokes. According to Cyprien Rivier, a researcher at Yale University, "Bacteria can release toxins into the blood, they can also produce certain

proteins that interfere with physiological processes. There is also what we call the microbiota-gut-brain axis – a bidirectional pathway between the brain and the microbiome, whereby the brain is influencing the gut through the nerves, and the microbiome is in turn influencing the organs, including the brain, mainly through altering the blood pressure" [53]. In Israel, they are using faecal microbial transfers (FMT) to help cancer patients who have become resistant to treatment [52]. There is still much we do not know about FMT, but research has shown it helps to restore the metabolism of bile, and contributes to "resetting" the gut's immunology [55].

The critical role of gut microbes in health has stimulated research by scientists and clinicians. While the human gut microbiome is too diverse to analyze, a mathematical approach was developed in [30] to model interactions between gut bacteria in fruit flies. A model of networks was presented in [34] to study the role of microbiota in evolution. A method to infer microbial community ecology from time resolved metagenomics was introduced in [47]. Another important step toward designing bacteriotherapies was made by Gibson by mapping out microbial interactions and predicting population dynamics of the ecosystem [27].

It was indicated in a number of references mentioned above that understanding the population dynamics of these systems provides crucial insights into the complex mechanisms of the gut microbiome. The aim of this work is to study in the simplistic setting how the gut microbiome grows by feeding on intakes of nutrients, by using dynamical analysis and optimal control to explore how to promote persistence and maximization of good bacteria.

In both models considered in this dissertation, Dynamics and Optimal Control of the Growth of the Gut Microbiome with Varying Nutrient, we use a chemostat model. The chemostat model consists of microorganisms feeding on a single growth-limiting nutrient. It can be regarded as a laboratory idealization of nature to study competitions for the same resource, e.g. a common food supply of a growth-limiting nutrient, between two or more populations [16] and thus has been widely used in theoretical ecology [5, 6, 7, 8, 9, 10, 12, 13], waste water treatment [1, 4], recombinant problems in genetically altered organisms [14, 15], etc. A chemostat consists of three interconnected vessels, used to grow microorganisms in a cultured environment. See the figure for an illustration of a simple chemostat.

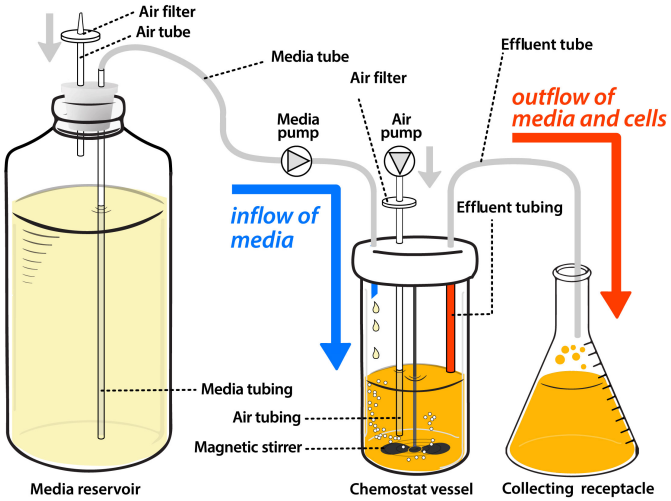


Figure 1.1: Design of a chemostat model with continuous agitation and aeration [2]

The first vessel is a feed bottle which contains nutrients to grow the microorganism(s), one of which is a limiting nutrient. The second is the culture vessel, where microorganisms grow feeding on the nutrients supplied from the feed bottle. We assume the culture is continuously stirred such that all the organisms have equal access to the nutrient. The last component of the chemostat is the collection vessel which contains nutrients, microorganisms and the products produced by the microorganisms. [17].

Chapter 2

A nonautonomous chemostat model for the growth of gut microbiome with varying nutrient

Noticing that the chemostat consists of microorganisms feeding on growth-limiting nutrients (see, e.g., [44, 46] and references therein), in Chapter 2, we develop a mathematical model based on a chemostat with a wall population for the evolution of one certain type of gut microbe that grows not only in the reservoir but also on the wall of gut. The input of the nutrient is considered to vary with respect to time, which gives rise to a nonautonomous chemostat system (see, e.g., [18, 19, 20]). We are interested in studying dynamical behavior of the resulting chemostat-microbiome model, and in particular, sufficient conditions under which the microbe vanishes or persists over time. In [11], the authors studied a model with antibiotics injection controlled by measured metagenomic data. In [3], they studied a chemostat model with an arbitrary number of competing species, looking for designed control laws to promote persistence. What makes our problem novel is the consideration of the sub-population that attaches to the wall of the intestine.

2.1 Introduction of the Model and Basic Properties of Solutions

For $t \in \mathbb{R}$, let $y_1(t)$ and $y_2(t)$ make up a total population of the same bacterial microorganism at time t, with $y_1(t)$ representing the amount of bacteria in the gut and $y_2(t)$ representing the amount of bacteria on the wall. These bacteria may switch their categories at any time, i.e., the microorganisms on the wall may join those in the gut, or the biomass in the medium may attach to the wall. Let r_1 and r_2 represent the rates at which the microorganisms stick onto and shear off from the wall, respectively.

The nutrients are assumed to be supplied to the gut through a flow at a time-dependent rate and concentration. Let $x(t)$ be the amount of nutrients in the gut at time t, and denote by $D(t)$ and $I(t)$ the input flow rate and input concentration of the aggregate nutrition, respectively. Both bacterial microorganisms inside the gut and on the gut wall consume the nutrient at a maximum consumption rate of a. The consumption function is assumed to take the Michaelis-Menten (or Holling Type II) form $\frac{x}{m+x}$ where m is the half-saturation constant. Let b be the factor that describes how efficient the consumption contributes to the growth. Due to the limitation of nutrients, intra-specific competitions are assumed to exist. In particular, let α and γ be the intra-specific competition rates of microbe population inside and on the wall of the gut.

The classical chemostat has the same in and out flow rate because the vessel is large enough that the flow rate remains almost unchanged by the microbes. But the gut has a much smaller volume and in the gut, nutrient or energy is extracted, and thus the out flow rate may not be the same as input flow rate. Considering that the aggregate output may consist of unused nutrition, produced energy used elsewhere in the organism and excrement, etc., here we model the outflow rate as proportional to the input flow rate $D(t)$ with a suitable scaling factor of μ.

Another difference from chemostat in laboratories is that the rate of flows that carry a nutrient through human bodies are low. Therefore the microorganisms may die naturally before being washed out, and moreover, the dead microorganisms in the reservoir are subject to bacterial decomposition which may lead to regeneration of the nutrient. Denote by ν the collective death rate of the microorganisms and δ the fraction of dead biomass that is recycled as nutrient. Expecting not a 100% of recycling, $\delta \in (0, 1)$. For the reader's convenience, interpretations of all model parameters are summarized in Table 2.1 below.

Summarizing the above, we obtain the following chemostat type model for the gut microbiome described by a system of ordinary differential equations

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = D(t)\big(I(t) - \mu x\big) - \frac{ax}{m+x}\big(y_1 + y_2\big) + \delta\nu y_1, \tag{2.1}$$

$$\frac{\mathrm{d}y_1(t)}{\mathrm{d}t} = -\big(\nu + \mu D(t)\big)y_1 + y_1\left(\frac{bx}{m+x} - \alpha y_1 - \gamma y_2\right) + r_2 y_2 - r_1 y_1, \tag{2.2}$$

$$\frac{\mathrm{d}y_2(t)}{\mathrm{d}t} = -\nu y_2 + y_2\left(\frac{bx}{m+x} - \alpha y_1 - \gamma y_2\right) - r_2 y_2 + r_1 y_1, \tag{2.3}$$

Table 2.1: Wall growth model parameters

Parameter	Meaning
$D(t) > 0$	input and output flow rate
$I(t) > 0$	input concentration of the nutrient
$\mu > 0$	ratio between output and input flow rates
$a > 0$	maximum consumption rate of the nutrient by the microorganisms
$b \geq 0$	growth rate of the microorganisms due to consumption
$m > 0$	half-saturation rate of the consumption function
$\delta \in (0, 1)$	recycle rate from dead microorganisms to new microbe biomass
$\nu \geq 0$	collective death rate of microorganisms
$r_1 \geq 0$	rate at which microbe attaches to the wall
$r_2 \geq 0$	rate at which microbe detaches from wall
$\alpha \geq 0$	intra-specific competition rate of microbe population in reservoir
$\gamma \geq 0$	intra-specific competition rate of wall population of microbe

where x, y_1, y_2 on the right hand sides of equations stand for $x(t)$, $y_1(t)$, and $y_2(t)$ in short, respectively. In this work, we will investigate detailed dynamics of the above system subject to the initial conditions

$$x(t_0) = x_0 \geq 0, \quad y_1(t_0) = y_{1,0} > 0, \quad y_2(t_0) = y_{2,0} > 0. \tag{2.4}$$

Throughout this work it is assumed that the input nutrient flow rate and concentration vary continuously in time, e.g., periodically or randomly, in bounded non-negative intervals. More precisely, assume that

(**A1**) the functions $D : \mathbb{R} \to [D_m, D_M]$ and $I : \mathbb{R} \to [I_m, I_M]$ are continuous, with $0 < D_m \leq D_M < \infty$ and $0 < I_m \leq I_M < \infty$.

For simplicity of notations, set $\boldsymbol{u}(t) = (x(t), y_1(t), y_2(t))^T \in \mathbb{R}^3$. Then the system (2.1)–(2.4) can be written in the matrix form as

$$\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}t} = \Gamma \boldsymbol{u} + \boldsymbol{f}(\boldsymbol{u}), \quad \boldsymbol{u}(t_0) = \boldsymbol{u}_0 := (x_0, y_{1,0}, y_{2,0})^T \tag{2.5}$$

where

$$\Gamma = \begin{bmatrix} -\mu D(t) & \delta\nu & 0 \\ 0 & -\nu - \mu D(t) - r_1 & r_2 \\ 0 & r_1 & -\nu - r_2 \end{bmatrix}, \ \boldsymbol{f}(\boldsymbol{u}) = \begin{bmatrix} D(t)I(t) - \frac{ax}{m+x}\Big(y_1 + x\Big) \\ y_1\Big(\frac{bx}{m+x} - \alpha y_1 - \gamma y_2\Big) \\ y_2\Big(\frac{bx}{m+x} - \alpha y_1 - \gamma y_2\Big) \end{bmatrix}.$$

Define the non-negative quadrant

$$\mathbb{R}^3_+ := \{(x, y_1, y_2) \in \mathbb{R}^3 : x \geq 0, y_1 \geq 0, y_2 \geq 0\}.$$

Consider the following definitions and theorems from classical theory for ordinary differential equations.

Let $f : (a, +\infty) \times \mathbb{R}^d \subseteq \mathbb{R}^{d+1} \to \mathbb{R}^d$ be a continuous mapping, and let (t_0, x_0) be a point in $(a, +\infty) \times \mathbb{R}^d$. The we can formulate the following IVP

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = f(t, x), \quad x(t_0) = x_0. \tag{2.6}$$

Definition 1. [20] Let $I \subseteq (a, +\infty)$ be a time interval. A *solution* to (2.6) on I is a mapping $\varphi : I \to \mathbb{R}^d$ which is continuously differentiable on I, i.e., $\varphi \in C^1(I; \mathbb{R}^d)$, and satisfies:

- $\frac{\mathrm{d}}{\mathrm{d}t}\varphi(t) = f(t, \varphi(t))$ for all $t \in I$;
- $\varphi(t_0) = x_0$

Theorem 1. *[20] Assume that $f : (a, +\infty) \times \mathbb{R}^d \to \mathbb{R}^d$ is continuously differentiable, i.e., its partial derivatives of first order are continuous functions, and there exist non-negative continuous mappings $h, k : (a, +\infty) \to \mathbb{R}$ such that*

$$|f(t, x)| \leq h(t)|x| + k(t), \text{ for all } (t, x) \in (a, +\infty) \times \mathbb{R}^d$$

Then, there exists a unique solution to (2.6) *which is defined globally in time.*

Remark 1. *[20] The existence and uniqueness of a local solution to* (2.6)*, that is defined on a finite time interval, can be proved by using a fixed point theorem. This local solution may only*

be defined on a small time interval but can always be extended to a larger time interval as long as it remains bounded. As a consequence, if the local solution does not blow up within finite time, then it can be defined globally in time.

We now apply these definitions and theorems from classical theory to our system and prove existence and uniqueness of solutions.

Lemma 1. *Let Assumption (***A1***) hold. In addition assume that*

(**A2**) $a(\nu + \mu D_m) > b\delta\nu$.

Then given any $(t_0, \boldsymbol{u}_0) \in \mathbb{R} \times \mathbb{R}^3_+$*, the system* (2.5) *admits a unique global solution. Moreover, the solution is non-negative and bounded for all* $t \geq t_0$.

Proof. First, since the functions $D(t)$ and $I(t)$ are bounded, the operator Γ generates an evolution system on $\mathbb{R}^3$. In addition, since $D(t)$ and $I(t)$ are continuous in t, the function $\boldsymbol{f}$ is continuously differentiable with respect to $\boldsymbol{u}$ for $x \neq -m$.

$$\frac{\partial \mathbf{f}}{\partial \mathbf{u}} = \left[\frac{\partial f_i}{\partial u_j}\right] = \begin{bmatrix} a\big(y_1 + x\big)\big(\frac{x}{(m+x)^2} - \frac{1}{m+x}\big) & -\frac{ax}{m+x} & -\frac{ax}{m+x} \\ y_1 b\big(\frac{1}{m+x} - \frac{x}{(m+x)^2}\big) & \frac{bx}{m+x} - 2\alpha y_1 & -\gamma y_1 \\ y_2 b\big(\frac{1}{m+x} - \frac{x}{(m+x)^2}\big) & -\alpha y_2 & \frac{bx}{m+x} - 2\gamma y_2 \end{bmatrix}$$

Then $\boldsymbol{f}$ is locally Lipschitz with respect to $\boldsymbol{u} \in \mathbb{R}^3_+$. Therefore by classical theory for ordinary differential equations (see Theorem 1), the system (2.5) has a unique local solution, $\boldsymbol{u}(t; t_0, \boldsymbol{u}_0) \in \mathcal{C}([t_0, T_{max}), \mathbb{R}^3)$.

We next show that the solution $\boldsymbol{u}(t; t_0, \boldsymbol{u}_0)$ is always non-negative on its existence interval. Notice that given $\boldsymbol{u}_0 := (x_0, y_{1,0}, y_{2,0})^T \in \mathbb{R}^3_+$ by continuity of solutions, each of $x(t)$, $y_1(t)$ and $y_2(t)$ has to take value 0 before it becomes negative. Now consider the scenarios when x, y_1, or y_2 first reaches zero, respectively, we have

$$\begin{aligned} \frac{\mathrm{d}x(t)}{\mathrm{d}t}\bigg|_{x=0, y_1, y_2 \geq 0} &= D(t)I(t) + \delta\nu\big(y_1 + y_2\big) \geq 0, \\ \frac{\mathrm{d}y_1(t)}{\mathrm{d}t}\bigg|_{y_1=0, x, y_2 \geq 0} &= r_2 y_2 \geq 0, \\ \frac{\mathrm{d}y_2(t)}{\mathrm{d}t}\bigg|_{y_2=0, x, y_1 \geq 0} &= r_1 y_1 \geq 0. \end{aligned}$$

This implies that any solution to the system (2.5) with $\boldsymbol{u}_0 \in \mathbb{R}^3_+$ will stay within $\mathbb{R}^3_+$ for all t in its existence interval.

It remains to show that the solution $\boldsymbol{u}(t; t_0, \boldsymbol{u}_0)$ is bounded. To that end, consider the linear combination $z(t) := bx(t) + ay_1(t) + ay_2(t)$ which satisfies the ODE

$$\begin{aligned}\frac{\mathrm{d}z(t)}{\mathrm{d}t} =& b\frac{\mathrm{d}x}{\mathrm{d}t} + a\frac{\mathrm{d}y_1}{\mathrm{d}t} + a\frac{\mathrm{d}y_2}{\mathrm{d}t} \\ =& bD(t)(I - x) + b^2\nu y_1 - a\alpha y_1^2 - a\gamma y_2 y_1 - a\nu y_2 - a\alpha y_2 y_1 \\ & - a\gamma y_2^2 - a(\nu + \mu D(t))y_1 \\ \leq& bD(t)(I - x) + b^2\nu y_1 - a\nu y_2 - a\alpha y_2 y_1 - a(\nu + \mu D(t))y_1\end{aligned}$$

By Assumption (**A1**),

$$\frac{\mathrm{d}z(t)}{\mathrm{d}t} \leq bD_M I_M - \mu D_m \cdot bx - \left(\nu + \mu D_m - \frac{b}{a}\delta\nu\right) ay_1 - \nu a y_2$$

Then by using Assumption (**A2**) and that $x, y_1, y_2 \geq 0$.

$$\frac{\mathrm{d}z(t)}{\mathrm{d}t} \leq bD_M I_M - \kappa z(t) \tag{2.7}$$

where

$$\kappa = \min\left\{\mu D_m, \nu + \mu D_m - \frac{b}{a}\delta\nu, \nu\right\} > 0. \tag{2.8}$$

Integrating the differential inequality (2.7) from t_0 to t with $z(t_0) = bx_0 + a\left(y_{1,0} + y_{2,0}\right)$ gives

$$z(t) \leq bx_0 + a\left(y_{1,0} + y_{2,0}\right) + \frac{bD_M I_M}{\kappa}, \quad t \in [t_0, T_{max}).$$

Now using $x(t), y_1(t), y_2(t) \geq 0$ we have

$$x(t) + y_1(t) + y_2(t) \leq \frac{z(t)}{\min\{a, b\}} \leq \frac{1}{\min\{a, b\}}\left(bx_0 + a\left(y_{1,0} + y_{2,0}\right) + \frac{bD_M I_M}{\kappa}\right), \tag{2.9}$$

which implies that $\boldsymbol{u}(t)$ is uniformly bounded for all t. Therefore the unique local solution $\boldsymbol{u}(t; t_0, \boldsymbol{u}_0) \in \mathcal{C}([t_0, T_{max}), \mathbb{R}^3_+)$ can be extended to a unique global solution $\boldsymbol{u}(t; t_0, \boldsymbol{u}_0) \in$

$\mathcal{C}([t_0, \infty), \mathbb{R}^3_+)$. Moreover, the global solution still satisfies the inequality (2.9). The proof is complete. □

Remark 2. *When the growth rate b is less than the maximum consumption rate a, i.e., the amount of nutrient consumed does not result in 100% of growth of the microorganisms, Assumption (***A2***) is automatically fulfilled because $\delta < 1$.*

Remark 3. *The parameter κ defined in* (2.8) *plays an important role in long term dynamics of the system* (2.5)*.*

Lemma 1 ensures that the chemostat-microbiome model (2.1)–(2.4) is biologically well-posed. In the following sections we will study detailed dynamics of the system (2.1)–(2.4). To facilitate analysis in the sequel, we first transform the system by defining two new variables $Y(t)$ and $R(t)$ by

$$Y(t) = y_1(t) + y_2(t), \quad Y(t)R(t) = y_1(t). \tag{2.10}$$

Notice that $Y(t)$ represents the total amount of microorganisms in the gut and on the wall, and $R(t)$ represents the percentage of the microorganisms that are inside the gut. $R(t) = 0$ corresponds to the case when $y_1(t) = 0$ and $y_2(t) \neq 0$, i.e., all microorganisms stay on the wall. On the other side, $R(t) = 1$ corresponds to the case when $y_1(t) \neq 0$ and $y_2(t) = 0$, i.e., all microorganisms stay in the reservoir.

With the transformation (2.10), the equations (2.1)–(2.3) become

$$\begin{aligned}
\frac{\mathrm{d}x(t)}{\mathrm{d}t} &= D(t)(I(t) - \mu x) - \frac{axY}{m+x} + \delta\nu RY, && (2.11)\\
\frac{\mathrm{d}Y(t)}{\mathrm{d}t} &= \gamma(R-1)Y^2 - \alpha RY^2 - \nu Y + \frac{bxY}{m+x} - \mu D(t)RY, && (2.12)\\
\frac{\mathrm{d}R(t)}{\mathrm{d}t} &= \mu D(t)R^2 - (\mu D(t) + r_1 + r_2)R + r_2, && (2.13)
\end{aligned}$$

where x, Y, R on the right hand sides of equations stand for $x(t)$, $Y(t)$, and $R(t)$ in short, respectively. The equivalent initial condition to (2.4) is

$$x(t_0) = x_0 \geq 0, \quad Y(t_0) = y_{1,0} + y_{2,0} := Y_0 > 0, \quad R(t_0) = \frac{y_{1,0}}{Y_0} := R_0 \tag{2.14}$$

Notice that now the equation for $R(t)$ is a Riccati equation which is decoupled from the equations for $x(t)$ and $Y(t)$. Moreover, under Assumption (**A1**)

$$\left.\frac{\mathrm{d}R(t)}{\mathrm{d}t}\right|_{R=1} = -r_1 < 0, \quad \left.\frac{\mathrm{d}R(t)}{\mathrm{d}t}\right|_{R=0} = r_2 > 0.$$

Therefore the interval $(0,1)$ is positive invariant for R. Notice that $R_0 \in (0,1)$ because $y_{1,0} \leq Y_0$, and thus

$$R(t) \in (0,1) \quad \text{for all } t \in \mathbb{R}. \tag{2.15}$$

Remark 4. *$R(t)$ represents the percentage of the microorganisms living in the reservoir of the gut over the total population inside and on the wall of the gut. Therefore $R(t) \in (0,1)$ makes biological sense.*

In the following sections we will analyze the equation of $R(t)$ first, and then use it to study dynamics of $x(t)$ and $Y(t)$.

2.2 The Autonomous Chemostat Microbiome Model

In this section we consider the special case where $D(t) \equiv D > 0$ and $I(t) \equiv I > 0$ and study the stability of each steady state of the system (2.11)–(2.13). Since the coefficients in the Riccati equation (2.13) are constants, then it can be solved explicitly to obtain

$$R(t) = \frac{R_1(R_0 - R_2) + R_2(R_1 - R_0)e^{\sqrt{(\mu D + r_1 + R_2)^2 - 4r_2\mu D}t}}{(R_0 - R_2) + (R_1 - R_0)e^{\sqrt{(\mu D + r_1 + R_2)^2 - 4r_2\mu D}t}}, \quad \text{with} \tag{2.16}$$

$$\begin{aligned} R_1 &= \frac{1}{2\mu D}\left(\mu D + r_1 + r_2 + \sqrt{(\mu D + r_1 + r_2)^2 - 4r_2\mu D}\right) > 1, \\ R_2 &= \frac{1}{2\mu D}\left(\mu D + r_1 + r_2 - \sqrt{(\mu D + r_1 + r_2)^2 - 4r_2\mu D}\right) \in (0,1). \end{aligned} \tag{2.17}$$

It is straightforward to check from (2.16) that $\lim_{t\to\infty} R(t) = R_2$ for any $R_0 \in [0,1]$, i.e., R_2 is a unique asymptotically stable equilibrium for $R(t)$.

Consequently the (x, Y) components of steady states of the system (2.11)–(2.13) satisfy the algebraic equations

$$\begin{aligned} 0 &= D(I-\mu x) - \frac{axY}{m+x} + \delta\nu R_2 Y, \\ 0 &= \gamma(R_2-1)Y^2 - \alpha R_2 Y^2 - \nu Y + \frac{bxY}{m+x} - \mu D R_2 Y, \end{aligned}$$

which has an axial root $(I/\mu, 0)$ and non-trivial roots (x^*, Y^*) satisfying

$$Y^* = \frac{\nu - \frac{bx^*}{m+x^*} + \mu D R_2}{\gamma(R_2-1) - \alpha R_2}, \tag{2.18}$$

$$D(I-\mu x^*) = \frac{\left(\frac{ax^*}{m+x^*} - \delta\nu R_2\right)\left(\nu - \frac{bx^*}{m+x^*} + \mu D R_2\right)}{\gamma(R_2-1) - \alpha R_2}. \tag{2.19}$$

To examine the stability of steady states, first calculate the Jacobian of the vector field for the sub-ODE system (2.11)–(2.12) at $R(t) = R_2$ to be

$$J = \begin{bmatrix} -\mu D - \frac{aYm}{(m+x)^2} & \delta\nu R_2 - \frac{ax}{m+x} \\ \\ \frac{bmY}{(m+x)^2} & 2\big(\gamma(R_2-1) - \alpha R_2\big)Y + \frac{bx}{m+x} - \nu - \mu D R_2 \end{bmatrix},$$

which has the trace and determinant

$$\mathrm{Tr}(J) = -\mu D - \frac{aYm}{(m+x)^2} + 2\big(\gamma(R_2-1) - \alpha R_2\big)Y + \frac{bx}{m+x} - \nu - \mu D R_2, \tag{2.20}$$

$$\begin{aligned} \mathrm{Det}(J) &= \left(\mu D + \frac{aYm}{(m+x)^2}\right)\Big(2\big(\gamma(1-R_2) + \alpha R_2\big)Y + \nu + \mu D R_2\Big) \\ &\quad - \frac{b\mu D x}{m+x} - \frac{b\delta m\nu R_2 Y}{(m+x)^2}. \end{aligned} \tag{2.21}$$

Notice that for the autonomous system, the Assumption (**A2**) now becomes

(**A2'**) $a(\nu + \mu D) > b\delta\nu$.

Theorem 2. *Let Assumption (***A2'***) hold and let R_2 be defined as in* (2.17)*. Then the steady state $(I/\mu, 0, R_2)$ is stable provided*

$$\nu + \mu D R_2 > \frac{bI}{m\mu + I}. \tag{2.22}$$

Proof. It follows immediately from (2.20)–(2.21) and (2.22) that

$$\begin{aligned} \mathrm{Det}(J)|_{(I/\mu,0)} &= \mu D \left(\nu + \mu D R_2 - \frac{bI}{m\mu + I} \right) > 0, \\ \mathrm{Tr}(J)|_{(I/\mu,0)} &= -\mu D - \nu + \frac{bI}{m\mu + I} - \mu D R_2 < 0. \end{aligned}$$

The proof is complete. □

Remark 5. *The assumption* (2.22) *can be fulfilled if a simpler but stricter assumption $\nu \geq \frac{bI}{m+I}$ is satisfied. This means that if the collective death rate is not smaller than the growth rate due to consumption, then both microorganisms in and on the wall of gut will eventually die out. The assumption* (2.22) *implies that even if the collective death rate is indeed smaller than the growth rate due to consumption, the microorganisms can still die out, and the larger the quantity $\mu D R_2$ is, the easier it becomes for the microorganisms to go extinct.*

We next investigate stability of the strictly positive steady states (x^*, Y^*, R_2), where x^* and Y^* satisfy the equations (2.18)–(2.19), and R_2 is defined as in (2.17). Notice that since $R_2 < 1$, the denominator of Y^* is negative. Therefore for the steady state to be meaningful, i.e., $Y^* > 0$, we focus our attention to steady states satisfying

$$\nu + \mu D R_2 < \frac{bx^*}{m + x^*}. \tag{2.23}$$

Theorem 3. *Let Assumption (***A2'***) hold and let R_2 be defined as in* (2.17)*. Then a steady state (x^*, Y^*, R_2) satisfying* (2.18)–(2.19) *and* (2.23) *is stable provided*

$$a\nu + a\mu D R_2 - b\delta\nu R_2 > 0. \tag{2.24}$$

Proof. First evaluating $\mathrm{Tr}(J)$ at (x^*, Y^*) and using (2.18) gives

$$\mathrm{Tr}(J)|_{(x^*,Y^*)} = -\mu D - \frac{aY^*m}{(m+x^*)^2} + \big(\gamma(R_2-1) - \alpha R_2\big)Y^* < 0,$$

due to the fact that $R_2 < 1$. We next estimate $\mathrm{Det}(J)$ at (x^*, Y^*). Using (2.18) again and simplifying the resultant equality we obtain

$$\begin{aligned}
\mathrm{Det}(J)|_{(x^*,Y^*)} &= \Big(\mu D + \frac{amY^*}{(m+x^*)^2}\Big)\Big(\frac{2bx^*}{m+x^*} - \nu - \mu D R_2\Big) \\
&\quad - \frac{b\mu D x^*}{m+x^*} - \frac{b\delta m\nu R_2 Y^*}{(m+x^*)^2} \\
&= \mu D\Big(\frac{bx^*}{m+x^*} - \nu - \mu D R_2\Big) \\
&\quad + \frac{mY^*}{(m+x^*)^2}\Big(\frac{2abx^*}{m+x^*} - a\nu - a\mu D R_2 - b\delta\nu R_2\Big).
\end{aligned}$$

It then follows from (2.23) and (2.24) that

$$\mathrm{Det}(J)|_{(x^*,Y^*)} > \frac{mY^*}{(m+x^*)^2}\Big(a\nu + a\mu D R_2 - b\delta\nu R_2\Big) > 0,$$

which implies the stability of the steady state (x^*, Y^*, R_2). □

Remark 6. *For the case when $b \le a$, Assumption* (2.24) *is automatically fulfilled, because $a\nu \ge b\delta\nu R_2$. The key condition for a non-trivial steady state to be stable is in fact the Assumption* (2.23).

Remark 7. *The system of algebraic equations* (2.18)–(2.19) *can have up to 3 different non-trivial roots (x^*, Y^*). While no theoretical proof is available, numerical experiments show that there is at least one, and likely only one, positive root that satisfies the condition* (2.23).

2.3 The Nonautonomous Chemostat Microbiome Model

In this section we study the long term dynamics of the nonautonomous system (2.11)–(2.13), using theory of nonautonomous dynamical systems. For the reader's convenience, a brief introduction of the process formulation of nonautonomous dynamical systems is given in Section

2.3.1. Then in Section 2.3.2 we show that the nonautonomous system (2.11)–(2.13) has a unique pullback attractor. Detailed long term dynamics of $R(t)$, $x(t)$ and $Y(t)$ will be studied in terms of structures of the pullback attractor in Section 2.3.3.

2.3.1 Preliminaries on Nonautonomous Dynamical Systems

In this subsection we provide some background information from the theory of nonautonomous dynamical systems [17, 37, 38] that we require in the sequel. Our situation is, in fact, somewhat simpler, but to facilitate the reader's access to the literature we give more general definitions here. Define

$$\mathbb{R}^2_{\geq} := \{(t, t_0) \in \mathbb{R}^2 \ : \ t \geq t_0\}.$$

Definition 2. A *process* ψ on space $\mathbb{R}^d$ is a family of mappings

$$\psi(t, t_0, \cdot) : \mathbb{R}^d \to \mathbb{R}^d, \qquad (t, t_0) \in \mathbb{R}^2_{\geq},$$

which satisfies

(i) initial value property: $\psi(t_0, t_0, x) = x$ for all $x \in \mathbb{R}^d$ and any $t_0 \in \mathbb{R}$;

(ii) two-parameter semigroup property: $\psi(t_2, t_0, x) = \psi\left(t_2, t_1, \psi(t_1, t_0, x)\right)$ for all $x \in \mathbb{R}^d$ and (t_2, t_1), $(t_1, t_0) \in \mathbb{R}^2_{\geq}$;

(iii) continuity property: the mapping $(t, t_0, x) \mapsto \psi(t, t_0, x)$ is continuous on $\mathbb{R}^2_{\geq} \times \mathbb{R}^d$.

Definition 3. Let ψ be a process on $\mathbb{R}^d$. A family $\mathcal{B} = \{B(t) : t \in \mathbb{R}\}$ of nonempty subsets of $\mathbb{R}^d$ is said to be ψ-*invariant* if $\psi\left(t, t_0, B(t_0)\right) = B(t)$ for all $(t, t_0) \in \mathbb{R}^2_{\geq}$ and ψ- positively invariant if $\psi\left(t, t_0, B(t_0)\right) \subseteq B(t)$ for all $(t, t_0) \in \mathbb{R}^2_{\geq}$.

Definition 4. Let ψ be a process on $\mathbb{R}^d$. A ψ-*invariant family* $\mathcal{A} = \{A(t) : t \in \mathbb{R}\}$ of nonempty compact subsets of $\mathbb{R}^d$ is called a forward attractor of ψ if it forward attracts all families $\mathcal{D} = \{D(t) : t \in \mathbb{R}\}$ of nonempty bounded subsets of $\mathbb{R}^d$, i.e.,

$$\text{dist}\left(\psi(t, t_0, D(t_0)), A(t)\right) \to 0 \quad \text{as } t \to \infty \quad (t_0 \text{ fixed}),$$

and is called a pullback attractor of ψ if it pullback attracts all families $\mathcal{D} = \{D(t) : t \in \mathbb{R}\}$ of nonempty bounded subsets of $\mathbb{R}^d$, i.e.,

$$\text{dist}\left(\psi(t, t_0, D(t_0)), A(t)\right) \to 0 \quad \text{as } t_0 \to -\infty \quad (t \text{ fixed}).$$

The existence of a pullback attractor follows from that of a pullback absorbing family, which is usually more easily determined.

Definition 5. A family $\mathcal{B} = \{B(t) : t \in \mathbb{R}\}$ of nonempty compact subsets of $\mathbb{R}^d$ is called a *pullback absorbing family* for a process ψ if for each $t_1 \in \mathbb{R}$ and every family $\mathcal{D} = \{D(t) : t \in \mathbb{R}\}$ of nonempty bounded subsets of $\mathbb{R}^d$ there exists some $T = T(t_1, \mathcal{D}) \in \mathbb{R}^+$ such that

$$\psi\left(t_1, t_0, D(t_0)\right) \subseteq B(t_1) \quad \text{for all } t_0 \in \mathbb{R} \text{ with } t_0 \leq t_1 - T.$$

The proof of the following theorem is well known, see e.g., [17, 37].

Theorem 4. *Suppose that a process ψ on $\mathbb{R}^d$ has a ψ-positively invariant pullback absorbing family $\mathcal{B} = \{B(t) : t \in \mathbb{R}\}$ of nonempty compact subsets of $\mathbb{R}^d$.*

Then ψ has a unique global pullback attractor $\mathcal{A} = \{A(t) : t \in \mathbb{R}\}$ with its component sets determined by

$$A(t) = \bigcap_{t_0 \leq t} \psi\left(t, t_0, B(t_0)\right) \quad \textit{for each } t \in \mathbb{R}.$$

If $\mathcal{B}$ is not ψ-positively invariant, then

$$A(t) = \bigcap_{s \geq 0} \overline{\bigcup_{t_0 \leq t-s} \psi\left(t, t_0, B(t_0)\right)} \quad \textit{for each } t \in \mathbb{R}.$$

A pullback attractor consists of *entire solutions*, i.e., functions $\xi : \mathbb{R} \to \mathbb{R}$ such that $\xi(t) = \psi(t, t_0, \xi(t_0))$ for all $(t, t_0) \in \mathbb{R}^2_{\geq}$. In special cases it consists of a single entire solution.

Definition 6. A nonautonomous dynamical system ψ is said to satisfy a *uniform strictly contracting property* if for each $r > 0$, there exist positive constants K and α such that

$$\|\psi(t, t_0, x_0) - \psi(t, t_0, y_0)\|^2 \;\leq\; K e^{-\alpha(t-t_0)} \cdot \|x_0 - y_0\|^2 \tag{2.25}$$

for all $(t, t_0) \in \mathbb{R}^2_{\geq}$ and x_0, $y_0 \in \mathbb{B}_r$, where $\mathbb{B}_r$ is the closed ball in $\mathbb{R}^d$ centered at the origin with radius $r > 0$.

This property suffices in combination with a pullback absorbing set to ensure the existence of an attractor in both the forward and pullback sense that consists of singleton sets, i.e., a single entire solution. The proof of the following result involves the construction of an appropriate Cauchy sequence which converges to a unique limit, see [35, 36].

Theorem 5. *Suppose that a process ψ on $\mathbb{R}^d$ is uniform strictly contracting on a ψ-positively invariant pullback absorbing family $\mathcal{B} = \{B(t) : t \in \mathbb{R}\}$ of nonempty compact subsets of $\mathbb{R}^d$. Then the process ψ has a unique global forward and pullback attractor $\mathcal{A} = \{A(t) : t \in \mathbb{R}\}$ with component sets consisting of singleton sets, i.e., $A(t) = \{\xi^*(t)\}$ for each $t \in \mathbb{R}$, where ξ^* is an entire solution of the process.*

2.3.2 Existence of a Pullback Attractor

Denote by $\tilde{\boldsymbol{u}}(t) = (x(t), Y(t), R(t))$. Recall that $x(t), Y(t) \geq 0$ and $R(t) \in (0, 1)$ for all $(t, t_0) \in \mathbb{R}^2_{\geq}$; we focus on the subspace $\Omega := \mathbb{R}^2_+ \times (0, 1)$ of $\mathbb{R}^3$. Then due to Lemma 1, given any $t_0 \in \mathbb{R}$ and $\tilde{\boldsymbol{u}}_0 \in \Omega$ the system (2.11)–(2.13) has a global solution $\tilde{\boldsymbol{u}}(\cdot; t_0, \tilde{\boldsymbol{u}}_0) \in \mathcal{C}\big([t_0, \infty), \Omega\big)$. For $(t, t_0) \in \mathbb{R}^2_{\geq}$ define a mapping $\psi(t, t_0, \cdot) : \Omega \to \Omega$ by

$$\psi(t, t_0, \tilde{\boldsymbol{u}}_0) = \boldsymbol{u}(t; t_0, \tilde{\boldsymbol{u}}_0), \quad (t, t_0) \in \mathbb{R}^2_{\geq} \tag{2.26}$$

where $\boldsymbol{u}(t; t_0, \tilde{\boldsymbol{u}}_0)$ is the solution to the system (2.11)–(2.13). It is straightforward to check that $\{\psi(t)\}_{(t,t_0)\in\mathbb{R}^2_{\geq}}$ is a process. From now on, it is referred to as the process defined by the system (2.11)–(2.13). The main goal of this subsection is to show that $\{\psi(t)\}_{(t,t_0)\in\mathbb{R}^2_{\geq}}$ has a unique pullback attractor. To that end we first construct an absorbing set for $\{\psi(t)\}_{(t,t_0)\in\mathbb{R}^2_{\geq}}$ in the following Lemma.

Lemma 2. *Let Assumptions (**A1**)–(**A2**) hold. Then the process $\{\psi(t)\}_{(t,t_0)\in\mathbb{R}^2_{\geq}}$ defined by the system* (2.11)–(2.13) *has an absorbing set defined as in* (2.32).

Proof. Consider first the decoupled Ricatti equation (2.13). Since $R(t) \in (0,1)$, then $R^2 - R < 0$ and along with Assumption (**A1**) we obtain

$$-(\mu D_M + r_1 + r_2)R(t) + R_2 < \frac{\mathrm{d}R(t)}{\mathrm{d}t} < -(r_1 + r_2)R(t) + r_2.$$

Integrating the differential inequalities above from t_0 to t gives

$$R(t) < \frac{r_2}{r_1 + r_2} + e^{-(r_1+r_2)(t-t_0)}\left(R_0 - \frac{r_2}{r_1 + r_2}\right), \tag{2.27}$$

$$R(t) > \frac{r_2}{r_1 + r_2 + \mu D_M} + e^{-(r_1+r_2+\mu D_M)(t-t_0)}\left(R_0 - \frac{r_2}{r_1 + r_2 + \mu D_M}\right). \tag{2.28}$$

It then follows immediately that given any $\epsilon > 0$ there exists $T_\epsilon > 0$ such that

$$\frac{r_2}{r_1 + r_2 + \mu D_M} - \epsilon \leq R(t) \leq \frac{r_2}{r_1 + r_2} + \epsilon, \quad t - t_0 \geq T_\epsilon. \tag{2.29}$$

Next, consider $z(t) := bx(t) + aY(t)$, and integrating the differential inequality (2.7) from t_0 to t results in

$$z(t) \leq \frac{bD_M I_M}{\kappa} + e^{-\kappa(t-t_0)}\left(bx_0 + aY_0 - \frac{bD_M I_M}{\kappa}\right). \tag{2.30}$$

Thus given any $\epsilon > 0$ there exists $T_\epsilon > 0$ such that

$$bx(t) + aY(t) \leq \frac{bD_M I_M}{\kappa} + \epsilon, \quad t - t_0 \geq T_\epsilon. \tag{2.31}$$

Now define the bounded set $\mathcal{B}$ in Ω by

$$\mathcal{B} := \Big\{(x, Y, R) \in \Omega : \quad bx + aY \leq \frac{bD_M I_M}{\kappa} + \epsilon \\ \frac{r_2}{r_1 + r_2 + \mu D_M} - \epsilon \leq R \leq \frac{r_2}{r_1 + r_2} + \epsilon\Big\}. \tag{2.32}$$

Then it follows directly from (2.29) and (2.31) that $\mathcal{B}$ is an absorbing set (in both pullback and forward sense) for the process $\{\psi(t)\}_{(t,t_0)\in\mathbb{R}^2_{\geq}}$. The proof is complete. □

Theorem 6. *Let Assumptions (***A1***)–(***A2***) hold. Then the process* $\{\psi(t)\}_{(t,t_0)\in\mathbb{R}^2_\geq}$ *defined by the system* (2.11)–(2.13) *possesses a unique pullback attractor* $\mathcal{A} = \{A(t) : t \in \mathbb{R}\}$ *with component subsets given by* $A(t) = \bigcap_{t_0\leq t} \psi(t, t_0, \mathcal{B})$ *for each* $t \in \mathbb{R}$ *where* $\mathcal{B}$ *is defined in* (2.32).

Proof. With Lemma 2, it remains to show that $\mathcal{B}$ is positive invariant for the process $\{\psi(t)\}_{(t,t_0)\in\mathbb{R}^2_\geq}$. In fact, given any $\tilde{\boldsymbol{u}}_0 = (x_0, Y_0, R_0) \in \mathcal{B}$, it follows immediately from (2.27), (2.28), and (2.30) that

$$\begin{aligned} \frac{r_2}{r_1 + r_2 + \mu D_M} - \epsilon &\leq R(t; t_0, \tilde{\boldsymbol{u}}_0) \leq \frac{r_2}{r_1 + r_2} + \epsilon, \quad \forall\ (t, t_0) \in \mathbb{R}^2_\geq, \\ bx(t; t_0, \tilde{\boldsymbol{u}}_0) + aY(t; t_0, \tilde{\boldsymbol{u}}_0) &\leq \frac{bD_M I_M}{\kappa} + \epsilon, \qquad\qquad \forall\ (t, t_0) \in \mathbb{R}^2_\geq, \end{aligned}$$

i.e., the set $\mathcal{B}$ is positive invariant under the process $\{\psi(t)\}_{(t,t_0)\in\mathbb{R}^2_\geq}$. Therefore by Theorem 4 the process $\{\psi(t)\}_{(t,t_0)\in\mathbb{R}^2_\geq}$ has a unique global pullback attractor $\mathcal{A} = \{A(t) : t \in \mathbb{R}\}$ consisting of nonempty compact subsets of Ω that are contained in $\mathcal{B}$. More specifically, the component subsets are given by $A(t) = \bigcap_{t_0\leq t} \psi(t, t_0, \mathcal{B})$ for each $t \in \mathbb{R}$. The proof is complete. □

2.3.3 Detailed Long Term Dynamics

In this subsection we investigate dynamics of $x(t)$, $Y(t)$, and $R(t)$ in greater detail. These dynamics characterize the structure of the pullback attractor $\mathcal{A}$ for the process $\{\psi(t)\}_{(t,t_0)\in\mathbb{R}^2_\geq}$ defined by the system (2.11)–(2.13). In particular, we first study the dynamics of $R(t)$, followed by the dynamics of $x(t)$ and $Y(t)$.

Recall that when $D(t) \equiv D$, the Riccati equation (2.13) has a steady state $R_2 \in (0, 1)$ that attracts all solutions of (2.13) with $R_0 \in (0, 1)$ as $t \to \infty$. Now for general time-dependent $D(t)$, the steady state no longer exists. Instead, the asymptotic dynamics can be characterized by a time-dependent singleton trajectory, if exists, that attracts all other solutions as $t \to \infty$ or $t_0 \to -\infty$. In Lemma 3 below, we prove the existence of such a trajectory.

Lemma 3. *Let Assumptions (***A1***)–(***A2***) hold. Then there exists a singleton trajectory* $R^*(t)$ *that both forward and pullback attracts all solutions of* (2.13)*, provided*

(A3) $\mu(2D_M - D_m) < r_1 + r_2$.

Proof. Consider any two solutions of the equation (2.13) with different initial values $R(t_0) = R_{0,1}$ and $R(t_0) = R_{0,2}$, denoted by $R(t; t_0, R_{0,1})$ and $R(t; t_0, R_{0,2})$, respectively. Set $\Delta_R(t) = R(t; t_0, R_{0,1}) - R(t; t_0, R_{0,2})$. Then $\Delta_R(t)$ satisfies the differential equation

$$\frac{\mathrm{d}\Delta_R(t)}{\mathrm{d}t} = \mu D(t)\big(R(t; t_0, R_{0,1}) + R(t; t_0, R_{0,2})\big)\Delta_R(t) - (\mu D(t) + r_1 + r_2)\Delta_R(t). \quad (2.33)$$

Without loss of generality, assume that $R_{0,1} > R_{0,2}$. Then by uniqueness of solutions, $\Delta_R(t) > 0$ for all $t \geq t_0$. Using (2.27) and Assumption (**A1**) in (2.33) gives

$$\frac{\mathrm{d}\Delta_R(t)}{\mathrm{d}t} \leq \left(\mu D_M(R_{0,1} + R_{0,2}) - \mu D_m - r_1 - r_2\right)\Delta_R(t)$$

which can be integrated to obtain

$$\Delta_R(t) \leq \Delta_R(t_0) e^{\left(-(\mu D_m + r_1 + r_2) + \mu D_M(R_{0,1} + R_{0,2})\right)(t - t_0)}, \quad (t, t_0) \in \mathbb{R}^2_{\geq}.$$

It then follows immediately from $R_{0,1}, R_{0,2} \in (0, 1)$ and Assumption (**A3**) that

$$\Delta_R(t) \leq \Delta_R(t_0) e^{\left(-\mu D_m + r_1 + r_2 + 2\mu D_M\right)(t - t_0)}, \quad (t, t_0) \in \mathbb{R}^2_{\geq},$$

i.e., $R(t)$ is uniformly strictly contracting. The desired assertion then follows from Theorem 5. □

We next study the dynamics of $x(t)$ and $Y(t)$ as a pair. First, it can be clearly observed that $Y = 0$ is still a steady state for the nonautonomous equation (2.12). In Lemma 4 below, a sufficient condition for $Y = 0$ to be stable is constructed.

Lemma 4. *Let Assumptions (**A1**) – (**A2**) hold. Then $Y(t; t_0, \tilde{\boldsymbol{u}}_0) \to 0$ as $t \to \infty$ and t_0 fixed, or t fixed and $t_0 \to -\infty$ if*

(A4) $\nu + \frac{\mu r_2 D_m}{r_1 + r_2 + \mu D_M} > b$.

Proof. First using $R < 1$ in the equation (2.12) gives

$$\frac{\mathrm{d}Y(t)}{\mathrm{d}t} < -\left(\nu + \mu D(t)R - \frac{bx}{m+x}\right) Y(t).$$

Given arbitrarily small $\epsilon > 0$, by (2.29) there exists $T_\epsilon > 0$ such that

$$\begin{aligned} \frac{\mathrm{d}Y(t)}{\mathrm{d}t} &< -\left(\nu + \mu D_m \left(\frac{r_2}{r_1 + r_2 + \mu D_M} - \epsilon\right) - \frac{bx}{m+x}\right) Y(t) \\ &< -\left(\nu + \mu D_m \left(\frac{r_2}{r_1 + r_2 + \mu D_M}\right) - b\right) Y(t). \end{aligned}$$

The desired assertion then follows immediately from Assumption (**A4**). □

Collecting results in Lemmas 3 and 4, setting $Y = 0$ in the equation (2.11) and integrating the resultant equation, we have following theorem.

Theorem 7. *Under the Assumptions (***A1***)–(***A4***) the pullback attractor $\mathcal{A}$ for the process ψ defined by the system* (2.11)–(2.13) *consists of a singleton trajectory* $(x^*(t), 0, R^*(t))$ *where*

$$x^*(t) = \int_{-\infty}^{t} D(\tau) I(\tau) e^{-\mu \int_\tau^t D(s)\mathrm{d}s} \mathrm{d}\tau.$$

The theorem above provides a sufficient condition under which all microorganisms die out. The singleton solution $R^*(t)$ of the equation (2.13) provides the important information on the ratio of bacteria inside and on the wall, even though they both approach zero. On the other side, when all microorganisms die out and no more nutrient is consumed, the asymptotic amount of nutrient in the gut simply becomes the accumulated nutrient input over time given by $x^*(t)$.

In the following theorem we construct sufficient conditions for the persistence of microorganisms, i.e., $Y(t)$ does not approach zero.

Theorem 8. *Let Assumptions (***A1***)–(***A2***) hold. The pullback attractor $\mathcal{A}$ for the process ψ defined by the system* (2.11)–(2.13) *contains points strictly inside the subspace* $(0, \infty)^2 \times (0, 1)$ *of* Ω *if the following two additional assumptions are satisfied*

(A5) $\frac{b\kappa D_m I_m}{mD_M(\kappa\mu+bI_M)+b\kappa D_m I_m} > \nu + \mu D_M \frac{r_1}{r_1+r_2}$,

(A6) $\kappa^2 D_m I_m < bD_M I_M(\kappa\mu D_M + bD_M I_M)$.

Proof. First we derive a lower bound for $x(t)$. Set $\ell_x = \frac{\kappa D_m I_m}{D_M(\kappa\mu+bI_M)}$. Then by Lemma 2 for any $\epsilon > 0$ there exists $T_\epsilon > 0$ such that $aY(t) \le \frac{bD_M I_M}{\kappa} + \epsilon$ for $t - t_0 > T_\epsilon$. Thus

$$\begin{aligned}\left.\frac{\mathrm{d}x(t)}{\mathrm{d}t}\right|_{x\le\ell_x} &\ge D_m I_m - \mu D_M \ell_x - \frac{\ell_x}{m+\ell_x}\left(\frac{bD_M I_M}{\kappa} + \epsilon\right)\\ &\ge D_m I_m - \left(\mu D_M + \frac{bD_M I_M}{\kappa}\right)\ell_x = 0, \quad t - t_0 > T_\epsilon,\end{aligned}$$

which implies that $x(t) \ge \ell_x$ for all $t - t_0 > T_\epsilon$.

Since $\frac{x}{m+x}$ is an increasing function of x, it then follows from Assumption (**A1**) and equation (2.12) that

$$\frac{\mathrm{d}Y(t)}{\mathrm{d}t} \ge Y\left(\frac{b\ell_x}{m+\ell_x} - \nu - \mu D_M R - (\gamma(1-R)+\alpha R)Y\right), \quad t - t_0 > T_\epsilon. \tag{2.34}$$

Notice that the Assumption (**A5**) implies that

$$\frac{b\ell_x}{m+\ell_x} > \nu + \mu D_M\left(\frac{r_1}{r_1+r_2} + \epsilon\right) \ge \nu + \mu D_M R(t), \quad t - t_0 > T_\epsilon. \tag{2.35}$$

Now set $\ell_Y := \frac{\frac{b\ell_x}{m+\ell_x} - \nu - \mu D_M \frac{r_1}{r_1+r_2}}{\frac{\gamma(r_2+D_M)}{r_1+r_2+D_M} + \frac{\alpha r_1}{r_1+r_2}}$. Clearly $\ell_Y > 0$, and by (2.27)–(2.28) we have for some T large enough

$$\begin{aligned}(\gamma(1-R)+\alpha R)\ell_Y &< \left(\frac{\gamma(r_2+D_M)}{r_1+r_2+D_M} + \frac{\alpha r_1}{r_1+r_2}\right)\frac{\frac{b\ell_x}{m+\ell_x} - \nu - \mu D_M \frac{r_1}{r_1+r_2}}{\frac{\gamma(r_2+D_M)}{r_1+r_2+D_M} + \frac{\alpha r_1}{r_1+r_2}}\\ &= \frac{b\ell_x}{m+\ell_x} - \nu - \mu D_M \frac{r_1}{r_1+r_2}, \quad \text{for } t - t_0 > T.\end{aligned} \tag{2.36}$$

Inserting (2.35)–(2.36) into (2.34), we obtain

$$\left.\frac{\mathrm{d}Y(t)}{\mathrm{d}t}\right|_{Y\le\ell_Y} \ge 0, \quad \text{for } t - t_0 \text{ large enough.} \tag{2.37}$$

The magnitude of ℓ_Y can be scaled by any positive parameter for which (2.37) still holds. In addition, Assumption (**A6**) ensures that the set

$$\tilde{\mathcal{B}} := \{(x, Y, R) \in \mathcal{B} : x \geq \ell_x, Y \geq \ell_Y\}$$

is non-empty. Moreover, $\tilde{\mathcal{B}}$ is positive invariant and absorbing for the process ψ. Therefore the attractor $\mathcal{A}$ has component sets within $\tilde{\mathcal{B}}$, which implies that both $x(t)$ and $Y(t)$ are strictly positive as $t - t_0 \to \infty$. □

2.4 Numerical Simulations

The numerical examples in this section aim to verify some of the theoretical properties established in the previous sections. Assume the system starts at time 0 and ends at time T. In comparisons between the autonomous and nonautonomous systems, we consider 2 different scenarios with the same average nutrient input, i.e., with the same value of $\frac{1}{T}\int_0^T D(t)I(t)\mathrm{d}t$:

1. constant injection rate and constant injection concentration

$$D(t) \equiv \tilde{D}, \quad I(t) \equiv \tilde{I};$$

2. periodic injection rate and injection concentration

$$\begin{aligned} D(t) &= \frac{D_m + D_M}{2} + \frac{D_M - D_m}{2} \sin \frac{2k\pi}{T} t, \\ I(t) &= \frac{I_m + I_M}{2} + \frac{I_M - I_m}{2} \sin \frac{2k\pi}{T} t, \quad k \in \mathbb{Z}, \end{aligned}$$

where $\tilde{D}$ and $\tilde{I}$ are chosen to ensure

$$\tilde{D}\tilde{I} = \left(\frac{D_m + D_M}{2}\right)\left(\frac{I_m + I_M}{2}\right) + \frac{1}{2}\left(\frac{D_M - D_m}{2}\right)\left(\frac{I_M - I_m}{2}\right).$$

Example 1. The first example illustrates the existence of absorbing sets for the nutrient x, total bacteria population Y, and the fraction R of bacteria in the gut, as established in Lemma 2 and

Theorem 6. The parameter values, given by $D_m = 0.5$, $D_M = 1$, $I_m = 1$, $I_M = 3$, $\mu = 5$, $a = 10$, $b = 5$, $m = 2$, $\delta = 0.85$, $r_1 = 5$, $r_2 = 1$, $\alpha = 1$, and $\gamma = 1$, are chosen to satisfy Assumptions (**A1**) and (**A2**). Figure 2.1 shows the fraction $R(t)$ of bacteria in the gut over the time interval $[0, 2]$ for a variety of initial conditions in $(0, 1)$, whereas Figure 2.2 plots the trajectories $(x(t), Y(t))$ of the nutrient x and the total bacteria Y over the same time interval. Note that, while the paths $(x(t), Y(t))$ intersect, those of the triplet $(x(t), Y(t), R(t))$ do not, in line with uniqueness results. Evidently, the path $(x(t), Y(t), R(t))$ enters and remains in the absorbing set defined in (2.32) for all initial conditions considered.

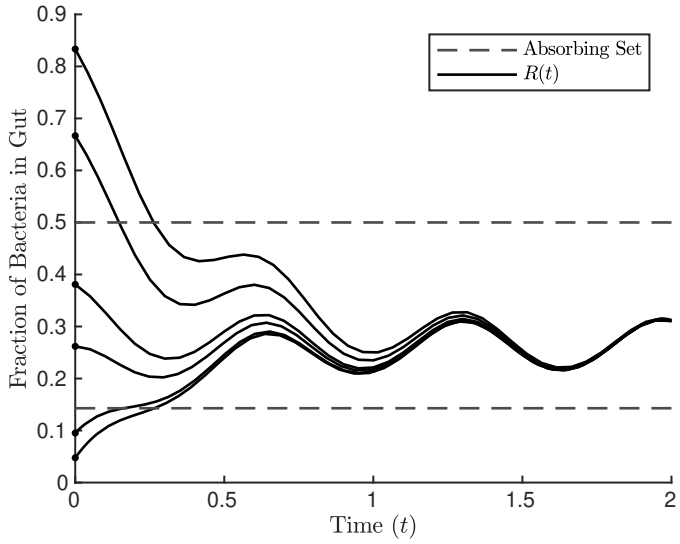


Figure 2.1: Plot of the fraction $R(t)$ of bacteria in the gut for different initial conditions.

Example 2. The second example explores the long term dynamics of the non-autonomous system established in Lemma 3, Lemma 4, and Theorem 7. The parameters $D_m = 0.5$, $D_M = 2$, $I_m = 1$, $I_M = 3$, $\mu = 3$, $a = 5$, $b = 1$, $m = 2$, $\delta = 0.85$, $\nu = 0.1$, $r_1 = 1$, $r_2 = 1$, $\alpha = 1$, and $\gamma = 1$ satisfy Assumptions (**A1**), (**A2**), and (**A3**) resulting, by Lemma 3, in the convergence of the function $R(t)$ to a singleton trajectory $R^*(t)$, as shown in Figure 2.3a. Since Assumption (**A4**) does not hold for these parameters, the total bacteria population need not die out, which is evident from Figure 2.3b. An increase in the collective death rate of the bacteria from $\nu = 0.1$ to $\nu = 0.9$, while leaving the remaining parameters unchanged, now guarantees that Assumptions (**A1**) – (**A4**) hold. Consequently, the total bacteria population $Y(t)$ dies out, according to Lemma 4, as shown in Figure 2.4a. Moreover, Figure 2.3b shows

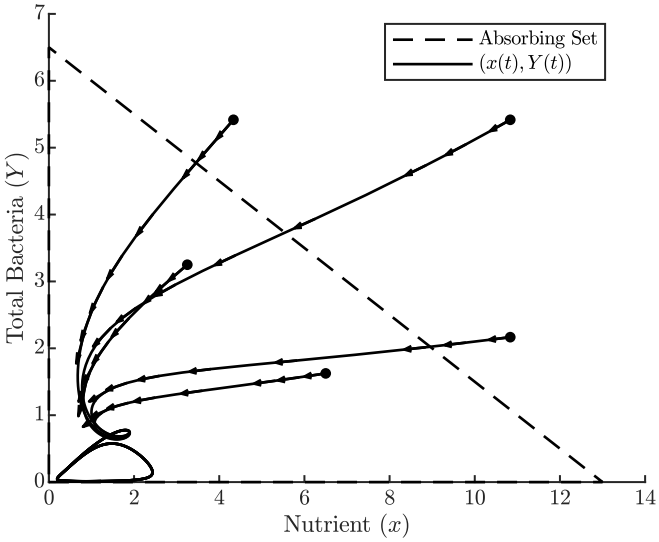
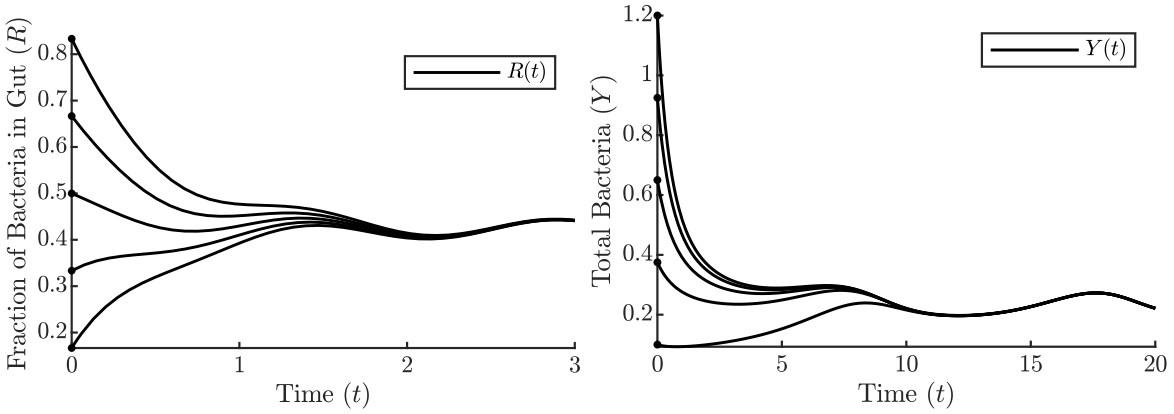


Figure 2.2: Plot showing the joint dynamics of the nutrient $x(t)$ and the total bacteria $Y(t)$ over the time interval $[0, 2]$ under assumptions (**A1**) and (**A2**). The arrows indicate the trajectory directions.



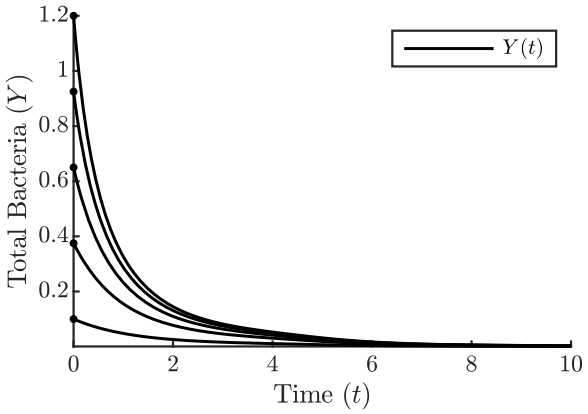
(a) $R(t)$ converges to $R^*(t)$

(b) $Y(t)$ persists.

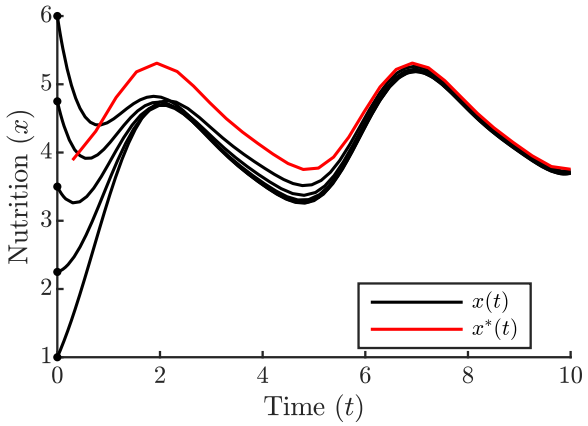
Figure 2.3: The time evolution of the fraction $R(t)$ of bacteria in the gut and the total bacteria $Y(t)$ under nonautonomous, periodic forcing with parameters satisfying Assumptions (**A1**), (**A2**), and (**A3**), but not (**A4**).

how the nutrition level $x(t)$ converges to the steady state function $x^*(t)$, defined in Theorem 7. Since the fraction $R(t)$ depends only on parameters $D(t)$, μ, r_1, and r_2, its dynamics remain unchanged. Lastly, Figure 2.5 shows the joint dynamics of the nutrient x, the bacteria in the gut y_1, and the bacteria on the wall y_2 in the time interval $[0, 10]$. Clearly, all trajectories converge towards the region the x-axis indicated in red.

Example 3. Finally, we verify numerically conditions under which the total population $Y(t)$ persists. In this simulation, the parameters were chosen as $D_m = 1.6$, $D_M = 1.7$, $I_M = 0.5$,



(a) $Y(t)$ dies out.



(b) $x(t)$ converges to $x^*(t)$.

Figure 2.4: Plots showing the convergence of the total bacteria $Y(t)$ to 0 and that of the nutrition level $x(t)$ to $x^*(t)$ under nonautonomous, periodic forcing and Assumptions (**A1**)–(**A4**).

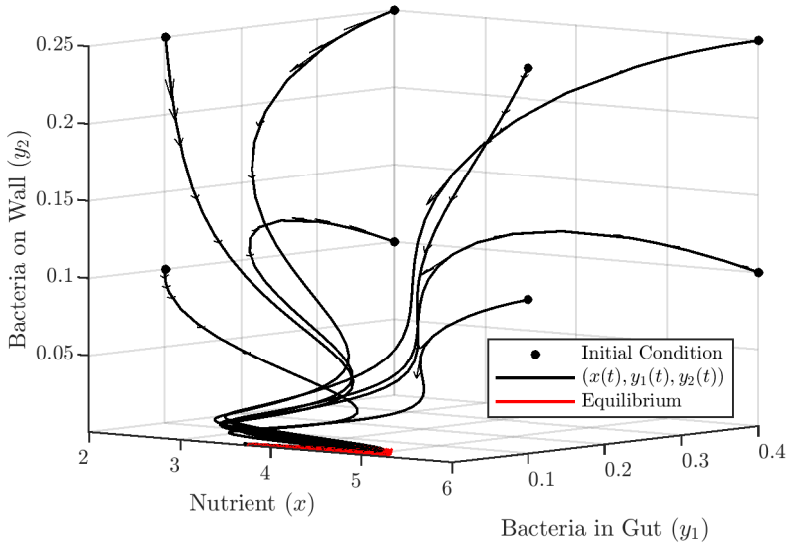


Figure 2.5: Joint trajectories of the nutrition x, the bacteria in the gut y_1, and the bacteria on the wall y_2 subject to nonautonomous, periodic forcing, with parameters safisfying (**A1**)–(**A4**).

$I_M = 1$, $\mu = 0.23$, $a = 2$, $b = 1$, $m = 0.18$, $\delta = 0.85$, $\nu = 0.2$, $r_1 = 0.1$, $r_2 = 5$, $\alpha = 1$, and $\gamma = 1$, ensuring that Assumptions (**A1**)–(**A3**), (**A5**), and (**A6**) hold. Figure 2.6 shows the joint trajectories of the nutrient x and the total bacterial population Y during the time interval $[0, 50]$ for a variety of initial conditions. Evidently, all trajectories enter and remain in the absorbing set defined in (2.32) and are bounded below by ℓ_x and ℓ_y defined in Theorem 8, respectively. Figure 2.7 plots these trajectories again on a logarithmic scale to show the total population's persistence behavior more clearly.

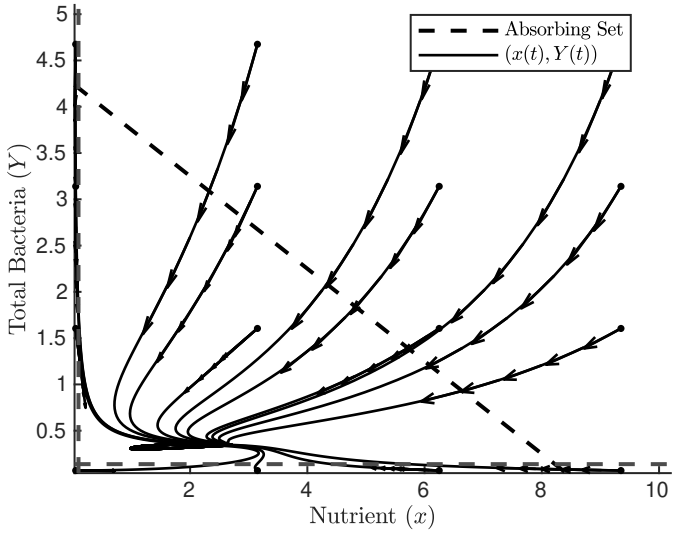


Figure 2.6: Plot of the joint dynamics of the nutrition level x and the total bacteria population Y under nonautonomous, periodic forcing, with parameters satisfying Assumptions (**A1**)–(**A3**), (**A5**), and (**A6**).

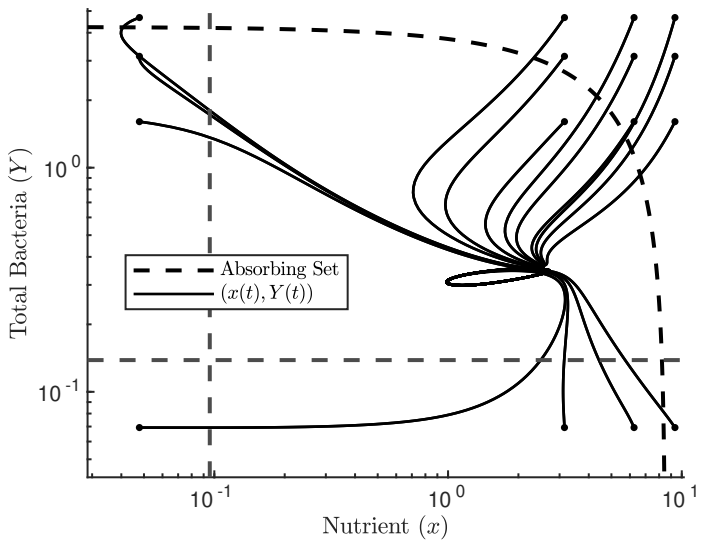


Figure 2.7: Plot of the trajectories in Figure 2.6 on a logarithmic scale.

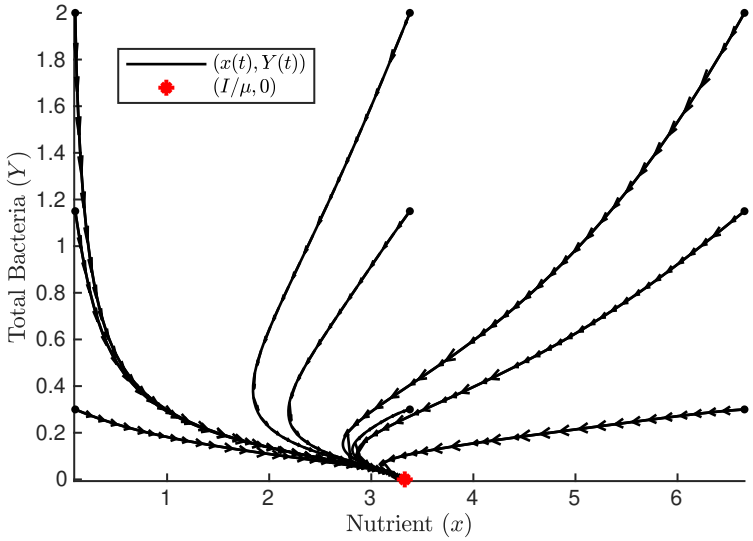


Figure 2.8: Trajectory plot $(x(t), Y(t))$ in the time interval $[0, 50]$ of the nutrition x and the total bacteria population Y under constant nutrient injection for various initial conditions.

We now demonstrate that the persistent long term behavior shown above for the nonautonomous system cannot be achieved by the equivalent autonomous system. Keeping all other parameters the same as above, we chose the constant injection rate $\tilde{I}$ and injection concentration $\tilde{D}$ to yield the same average nutrient input and to ensure that Assumptions (**A1**)–(**A3**), (**A5**), and (**A6**) hold. Figures 2.8 and 2.9 show how the solutions (x, Y, R) of system (2.11)–(2.13) corresponding to various initial conditions converge to the steady state $(\tilde{I}/\mu, 0, R_2)$ whose stability was established by Theorem 2.

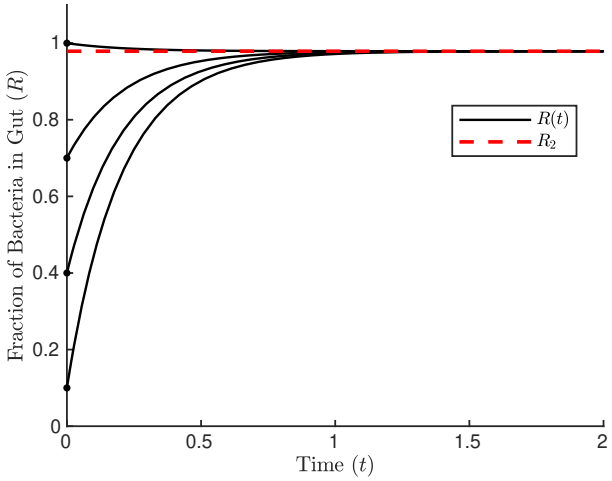


Figure 2.9: Plot of the fraction of bacteria in the gut over time under autonomous forcing and Assumptions (**A1**)–(**A3**), (**A5**), and (**A6**).

Chapter 3

An optimal control analysis of a model of a chemostat model for the growth of gut microbiome with varying nutrient

Motivated by the interesting discovery that time-dependent input of nutrient may result in asymptotic behavior different from constant input, we are interested in investigating how the growth of a particular type of bacteria can be boosted or inhibited to maintain a healthy gut. The second part of my dissertation project involves looking at a new aspect of analyzing the microbial population dynamics: optimal control. In particular, we consider a modified model for the dynamics of an intestinal microbiome with one beneficial bacterial microorganism that we are trying to promote by the infusion of some probiotic along with one limiting resource of nutrient population over a given finite time interval, $[0, T]$. In order to find an optimal controlled rate of the input flow, we turn to results from mathematicians during the Cold War. In the 1950s, a race began between the United States and the Soviet Union to answer a very important question: How can we move our aircrafts from their cruising position to a beneficial attack position optimally? [48] Pontryagin and his colleagues developed the Maximum Principle to apply to certain types of control problems. We use this theory for our system to find the optimal control.

The research to date for using optimal control chemostat models is primarily applied to waste-water treatment processes. In [39], they model a single microorganism species and substrate concentration and do time-optimal control analysis, seeking to steer the system to a target state in minimal time. In [40], they show using optimal control theory that the averaged conversion rate can be improved by a periodic dilution rate under certain conditions.

In our model, we consider a setting in which a microorganism's average concentration is maximized over a given time interval. This model could be used to promote maximal production of a given beneficial bacteria in the gut over a specific time interval. Similarly, the results could be used to minimize a harmful bacteria over a given time interval, for example, attempting to get rid of an infection.

3.1 Introduction of the Model and Basic Properties of Solutions

A probiotic is a substance which stimulates the growth of microorganisms, especially those with beneficial properties. At any time t, we let $x(t)$, $y(t)$, and $z(t)$ denote the concentration of the nutrient, the concentration of the microorganism, and the concentration of the probiotic inside the gut, respectively. The model that we consider reads

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = (I^x - \mu x(t))u(t) - y(t)\frac{\alpha x(t)}{m + x(t)}, \quad x(t_0) = x_0 \tag{3.1}$$

$$\frac{\mathrm{d}y(t)}{\mathrm{d}t} = y(t)\Big(\beta\frac{\alpha x(t)}{m + x(t)} - \mu u(t) - \nu\Big) + \gamma y(t)M(z(t)), \quad y(t_0) = y_0 \tag{3.2}$$

$$\frac{\mathrm{d}z(t)}{\mathrm{d}t} = (I^z - \mu z(t))u(t) - y(t)M(z(t)), \quad z(t_0) = z_0 \tag{3.3}$$

Table 3.1: Optimal control model parameters

Parameter	Meaning
$u(t) \in [0, u_{\max}]$	input flow rate
$I^x > 0$	input concentration of the nutrient
$I^z > 0$	input concentration of the probiotic
$\mu > 0$	ratio between output and input flow rates ("fudge" factor)
$\alpha > 0$	maximum consumption rate of the nutrient by the microorganism
$\beta \geq 0$	yielding factor
$m \geq 0$	half-saturation constant of the nutrient
$\gamma > 0$	gain factor of microorganism by probiotic consumption
$M(z(t))$	consumption function of probiotic by microorganism

where $u(t)$ is the controlled rate of the input flow, which brings the nutrient and probiotic into the gut. Note that the variable u has a different meaning in Chapter 2, where it denotes the state vector. Here I^x and I^z are constant input concentration of the nutrient and probiotic, respectively, α is the maximum consumption rate of the microorganism on the nutrient, m is the

half-saturation constant of the nutrient, β is the yielding factor, ν is the collective death rate of the microorganism, and μ is the fudge factor, same as in Chapter 2. The function $M(\cdot)$ describes the consumption of probiotic by the microorganism that results in a gain in the microorganism by a factor of γ in turn. We make the following assumptions on $M : \mathbb{R}^+ \to \mathbb{R}^+$ (see [19]):

- $M(0) = 0, \quad M(z) > 0$ for all $z > 0$;
- $\lim_{z\to\infty} M(z) = L$, where $L < \infty$;
- M is continuously differentiable;
- M is monotonically increasing.

Remark 8 (Notation on Derivatives)**.** *Since we will make extensive use of various derivatives of composite functions in the following sections, we would like to specify notation in this regard. Let $h : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ be a scalar multivariate function, mapping $(t, \boldsymbol{x}) \mapsto h(t, x)$, where $\boldsymbol{x} = (x_1, ..., x_n) \in \mathbb{R}^n$. The scalar function $\frac{\partial h}{\partial t} = h_t(t, \boldsymbol{x}) := \lim_{\delta t\to 0} \frac{h(t+\delta t, \boldsymbol{x}) - h(t, \boldsymbol{x})}{\delta t}$ denotes the partial derivative of h with respect to the t-variable, while the row vector $h_{\boldsymbol{x}}(t, \boldsymbol{x}) = [h_{x_1}(t, \boldsymbol{x}), ..., h_{x_n}(t, \boldsymbol{x})] \in (\mathbb{R}^n)^*$ denotes its Jacobian with respect to $\boldsymbol{x}$. For a vector-valued function $\mathbf{h} : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^m$, mapping $(t, \boldsymbol{x}) \mapsto \mathbf{h}(t, \boldsymbol{x}) = [h_1(t, \boldsymbol{x}), ..., h_m(t, \boldsymbol{x})]^T \in \mathbb{R}^m$ we use the column vector $\mathbf{h}_t(t, \boldsymbol{x}) = \left[\frac{\partial h_1}{\partial t}(t, \boldsymbol{x}), ..., \frac{\partial h_m}{\partial t}(t, \boldsymbol{x})\right]^T$ to denote the Jacobian of $\mathbf{h}$ with respect to the t-variable, while*

$$\frac{\partial \mathbf{h}}{\partial \boldsymbol{x}} = \mathbf{h}_{\boldsymbol{x}}(t, \boldsymbol{x}) := \begin{bmatrix} \frac{\partial h_1}{\partial x_1}(t, \boldsymbol{x}) & \dots & \frac{\partial h_1}{\partial x_n}(t, \boldsymbol{x}) \\ \vdots & \ddots & \vdots \\ \frac{\partial h_m}{\partial x_1}(t, \boldsymbol{x}) & \dots & \frac{\partial h_m}{\partial x_n}(t, \boldsymbol{x}) \end{bmatrix}$$

denotes the Jacobian of $\mathbf{h}(t, \boldsymbol{x})$ with respect to the variable $\boldsymbol{x} = (x_1, ..., x_n)$. Finally, suppose $t \mapsto \boldsymbol{x}(t)$ and $\mathbf{h}(t, \boldsymbol{x})$ is the vector-valued function defined above, then the derivative of the composite mapping $t \mapsto h(t, \boldsymbol{x}(t)$ with respect to t is determined by the multivariate chain rule, i.e.,

$$\frac{\mathrm{d}\mathbf{h}}{\mathrm{d}t}(t, \boldsymbol{x}(t)) = \mathbf{h}_t(t, \boldsymbol{x}(t))) + \mathbf{h}_{\boldsymbol{x}}(t, \boldsymbol{x}(t)) \frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}.$$

3.1.1 Existence and Uniqueness of Solutions

We let $\boldsymbol{x}(t) = (x(t), y(t), z(t))$ denote the state vector and summarize the autonomous initial value problem (3.1)–(3.3) as

$$\frac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t} = \mathbf{f}(\boldsymbol{x}(t), u(t)), \quad \boldsymbol{x}(0) = \boldsymbol{x}_0, \tag{3.4}$$

where $\boldsymbol{x}_0 = (x_0, y_0, z_0)$ denotes the initial condition.

In this subsection, we establish the existence and uniqueness of solutions to the IVP (3.4) for any continuous control function u within the set $\mathscr{U}$ of admissible controls given by

$$\mathscr{U} := \{u(t) | u(t) \in C([0, T]), 0 \leq u(t) \leq u_{\max}\}. \tag{3.5}$$

To this end, we separate the vector field $\mathbf{f}(\boldsymbol{x}, u)$ into a linear part $A(u)$ and a nonlinear part $\mathfrak{g}(\boldsymbol{x}, u)$. Specifically,

$$\mathbf{f}(\boldsymbol{x}, u) = A(u)\boldsymbol{x} + \mathfrak{g}(\boldsymbol{x}, u),$$

where

$$A = \begin{bmatrix} -u(t) & 0 & 0 \\ 0 & -(\mu u(t) + \nu) & 0 \\ 0 & 0 & \mu u(t) \end{bmatrix} \text{ and } \mathfrak{g}(\boldsymbol{x}) = \begin{bmatrix} I^x u(t) - \frac{\alpha x y}{m+x} \\ \frac{\alpha\beta x y}{m+x} + \gamma y M(z) \\ I^z u(t) - y M(z) \end{bmatrix}.$$

The initial value problem (3.4) can thus be rewritten as

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = A(u)\boldsymbol{x} + \mathfrak{g}(\boldsymbol{x}, u), \quad \boldsymbol{x}(0) = \boldsymbol{x}_0. \tag{3.6}$$

Define the non-negative octant

$$\mathbb{R}^3_+ := \{(x, y) \in \mathbb{R}^2 : x \geq 0, y \geq 0, z \geq 0\}.$$

Lemma 5. *Given any $\boldsymbol{x}_0 \in \mathbb{R}^3_+$ and $u \in \mathscr{U}$, the system* (3.6) *admits a unique global solution. Moreover, the solution is non-negative and bounded for all $t \in [0, T]$.*

Proof. First, since the function $u(t)$ is bounded, the operator A generates an evolution system on $\mathbb{R}^3$. In addition, since $u(t)$ is continuous in t, the function $\mathfrak{g}$ is continuously differentiable with respect to $\boldsymbol{x}$ for $x \neq -m$, with derivative

$$\frac{\partial \mathfrak{g}}{\partial \boldsymbol{x}} = \begin{bmatrix} \frac{\alpha m y}{(m+x)^2} & \frac{\alpha x}{m+x} & 0 \\ \frac{\alpha \beta m y}{(m+x)^2} & \frac{\alpha \beta x}{m+x} & \gamma y M'(z) \\ 0 & -M(z) & -y M'(z) \end{bmatrix}.$$

So $\mathfrak{g}$ is locally Lipschitz with respect to $\boldsymbol{x} \in \mathbb{R}^3_+$. Therefore by classical theory for ordinary differential equations (as stated in Theorem 1 in Section 2.1), the system (3.6) has a unique local solution, $\boldsymbol{x} \in \mathcal{C}([0, T), \mathbb{R}^3)$.

We next show that the solution $\boldsymbol{x}(t)$ is always non-negative on its existence interval $[0, T]$. Notice that given $\boldsymbol{x}_0 := (x_0, y_0, z_0)^T \in \mathbb{R}^3_+$ by continuity of solutions, each of $x(t)$, $y(t)$, and $z(t)$ has to take value 0 before it becomes negative. Now consider the scenarios when x, y, or z first reaches zero, respectively, we have

$$\begin{aligned} \left.\frac{\mathrm{d}x(t)}{\mathrm{d}t}\right|_{x=0, y\geq 0, z\geq 0} &= I^x u(t) \geq 0, \\ \left.\frac{\mathrm{d}y(t)}{\mathrm{d}t}\right|_{y=0, x\geq 0, z\geq 0} &= 0 \geq 0, \\ \left.\frac{\mathrm{d}z(t)}{\mathrm{d}t}\right|_{z=0, x\geq 0, y\geq 0} &= I^z u(t) \geq 0 \end{aligned}$$

This implies that any solution to the system (3.6) with $\boldsymbol{x}_0 \in \mathbb{R}^3_+$ will stay within $\mathbb{R}^3_+$ for all t in its existence interval.

It remains to show that the solution $\boldsymbol{x}(t)$ is bounded. To that end, consider the linear combination $\zeta(t) := \beta x(t) + y(t) + z(t)$ which satisfies the ODE

$$\frac{\mathrm{d}\zeta(t)}{\mathrm{d}t} = \beta I^x u(t) - \mu u(t) y(t) - \nu y(t) + (I^z - \mu z(t)) u(t) - y(t) M(z(t))$$

Then by using the assumption $0 \leq u(t) \leq u_{\max}$ and that $x, y, z \geq 0$ and the assumptions on $M(z(t))$ we obtain

$$\begin{aligned} \frac{\mathrm{d}z(t)}{\mathrm{d}t} &\leq \beta I^x u_{\max} - \mu u(t) y(t) - \nu y(t) + I^z u_{\max} - \mu z(t) u(t) - y(t) M(z(t)) \quad (3.7) \\ &\leq (\beta I^x + I^z) u_{\max} > 0, \quad (3.8) \end{aligned}$$

which implies that $\boldsymbol{x}(t)$ is uniformly bounded for all $t \in [0, T]$. Therefore the unique local solution $\boldsymbol{x} \in \mathcal{C}([0, T), \mathbb{R}^3_+)$ can be extended to a unique global solution $\boldsymbol{x} \in \mathcal{C}([0, \infty), \mathbb{R}^3_+)$. The proof is complete. □

3.1.2 Model without the Probiotic

As an initial step, we consider the special case without the probiotic, i.e. $z(t)$ is constant, so $\frac{\mathrm{d}z(t)}{\mathrm{d}t} = 0$. Then our system takes the following form:

$$\begin{aligned} \frac{\mathrm{d}x(t)}{\mathrm{d}t} &= (I - \mu x(t)) u(t) - y(t) \frac{\alpha x(t)}{m + x(t)}, \quad x(0) = x_0 \quad (3.9) \\ \frac{\mathrm{d}y(t)}{\mathrm{d}t} &= y(t) \Big(\beta \frac{\alpha x(t)}{m + x(t)} - \mu u(t) - \nu \Big), \quad y(0) = y_0, \quad (3.10) \end{aligned}$$

defined over a fixed time interval $[0, T]$, where I is the simplified notation for I^x. Additionally, we now will use $\boldsymbol{x}(t) = (x(t), y(t))$ to denote the state vector and we summarize the autonomous initial value problem (3.9)–(3.10) as

$$\frac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t} = \mathbf{f}(\boldsymbol{x}(t), u(t)), \quad \boldsymbol{x}(0) = \boldsymbol{x}_0, \quad (3.11)$$

where $\boldsymbol{x}_0 = (x_0, y_0)$ denotes the initial condition.

Analysis of Equilibria under Constant Control

In this subsection, we determine steady-states and analyze their stability in the special case where $u(t) \equiv u \in \mathbb{R}$, i.e., when the control function is constant. Then our system is autonomous and the steady-states can be found explicitly.

$$0 = (I - \mu x)u - \frac{y\alpha x}{m + x} \tag{3.12}$$

$$0 = y\big(\beta\frac{\alpha x}{m + x} - \mu u - \nu\big) \tag{3.13}$$

From Equation (3.13), we get our first steady-state, which is trivial, $y = 0$. Substituting $y = 0$ into Equation (3.12), we get $x = \frac{I}{\mu}$, so our axial steady-state is $\left(\frac{I}{\mu}, 0\right)$.
For the nontrivial steady-state, set $\beta\frac{\alpha x^*}{m+x^*} - \mu u - \nu = 0$ and solve explicitly for $x \equiv x^*$.
We get

$$x^* = \frac{m(\mu u + \nu)}{\alpha\beta - \mu u - \nu} \tag{3.14}$$

Then substituting x^* into Equation (3.12), we can explicitly solve for y^* and obtain

$$y^* = \beta u\left(\frac{I}{\mu u + \nu} - \frac{\mu m}{\alpha\beta - \mu u - \nu}\right) \tag{3.15}$$

This yields our nontrivial steady-state

$$(x^*, y^*) = \Big(\frac{m(\mu u + \nu)}{\alpha\beta - \mu u - \nu}, \beta u\Big(\frac{I}{\mu u + \nu} - \frac{\mu m}{\alpha\beta - \mu u - \nu}\Big)\Big) \tag{3.16}$$

In order for (x^*, y^*) to be positive, we need the following to be true:

$$\alpha\beta > \mu u + \nu \tag{3.17}$$

$$\frac{I}{\mu u + \nu} > \frac{\mu m}{\alpha\beta - \mu u - \nu} \tag{3.18}$$

Remark 9. *Note that assumptions* (3.17) *and* (3.18) *are easily satisified when* u *is small. However, in order for* y^* *to be maximized, we see in* (3.15) *that* u *needs to be large. Thus, the optimization of* y^* *depends on a trade-off of how large* u *is, while still needing* u *to be small in order for* (x^*, y^*) *to be stable.*

From our system (3.11), we have the following Jacobian matrix:

$$\mathbf{f}_{\boldsymbol{x}} = \begin{bmatrix} -\mu u - \frac{\alpha m y}{(m+x)^2} & \frac{-\alpha x}{m+x} \\ \frac{\alpha\beta m y}{(m+x)^2} & \frac{\alpha\beta x}{m+x} - \mu u - \nu \end{bmatrix}$$

And we have the following determinant and trace:

$$Det(\mathbf{f}_{\boldsymbol{x}}) = -\Big(\mu u + \frac{\alpha m y}{(m+x)^2}\Big)\Big(\frac{\alpha\beta x}{m+x} - \mu - \nu\Big) + \frac{\alpha^2\beta m x y}{(m+x)^3} \tag{3.19}$$

$$Tr(\mathbf{f}_{\boldsymbol{x}}) = -2\mu u - \frac{\alpha m y}{(m+x)^2} - \nu + \frac{\alpha\beta x}{m+x} \tag{3.20}$$

Lemma 6. *The axial steady-state* $(I/\mu, 0)$ *is stable provided that* $\mu u + \nu > \frac{\alpha\beta I}{\mu m + I}$.

Proof.

$$Det(\mathbf{f}_{\boldsymbol{x}})|_{(I/\mu,0)} = -\mu u\left(\frac{\alpha\beta I}{\mu m + I} - \mu u - \nu\right) > 0$$

$$Tr(\mathbf{f}_{\boldsymbol{x}})|_{(I/\mu,0)} = -2\mu u - \nu + \frac{\alpha\beta I}{\mu m + I}$$

□

Lemma 7. *The non-trivial steady-state* (x^*, y^*) *is stable provided* (3.17) *and* (3.18) *are true.*

Proof. Recall $\beta\frac{\alpha x^*}{m+x^*} - \mu u - \nu = 0$, so

$$Det(\mathbf{f}_{\boldsymbol{x}})|_{(x^*,y^*)} = \frac{\alpha^2\beta x^* y^*}{(m+x^*)^3} > 0$$

$$Tr(\mathbf{f}_{\boldsymbol{x}})|_{(x^*,y^*)} = -\mu u - \frac{\alpha m y^*}{(m+x^*)^2} < 0$$

□

3.2 The Optimal Control Problem

We are interested in finding a control $u(t)$ within an admissible set $\mathscr{U}$ such that the average concentration of the microorganism over time $[0, T]$ is maximized. We formulate this task as a minimization problem, according to the conventions of control theory literature, and add a regularization term to penalize large controls, i.e.,

$$\begin{aligned} &\min_{u \in \mathscr{U}} \mathscr{J}_T(u) = \int_0^T -y(t) + \frac{c}{2} u^2(t) dt, \\ &\text{subject to} \qquad \dot{\boldsymbol{x}} = \mathbf{f}(\boldsymbol{x}(t), u(t)), \qquad \boldsymbol{x}(0) = \boldsymbol{x}_0, \end{aligned} \tag{3.21}$$

where $c > 0$ is the regularization parameter and the admissible set $\mathscr{U}$ is given in Equation (3.5). In the following two sections, we will invoke the Pontryagin Maximum Principle (PMP) to establish a set of first order necessary conditions that are satisfied by the minimizer u^* of Problem (3.21). To this end, we will first recall some preliminary definitions and give a general statement of the PMP (Theorem 9) in Subsection 3.2.1. In Subsection 3.2.2, we will show that our control problem (3.21) satisfies the conditions of Theorem 9 and use the resulting necessary conditions to determine the optimal control u^*. Specifically, we show that u^* can be expressed explicitly in terms of the solution vector $\mathbf{p}$ of the Hamiltonian boundary value problem given by Equations (3.30)–(3.33).

3.2.1 Preliminaries on Optimal Control Theory

In this section, we will cover the preliminaries of optimal control theory. We first define various notions related to control systems and then state a general form of the Pontryagin Maximum Principle that is appropriate to Problem (3.21).

Definition 7 ([50])**.** A *control system* is a 4-tuple $\Sigma = (X, U, \mathbf{f}, \mathscr{U})$ consisting of a state space $X \subset \mathbb{R}^n$ representing the range of the state vector $\boldsymbol{x}(t)$, a control set $U \subset \mathbb{R}^m$ specifying the range of the control variable u, a vector field $\mathbf{f}$ specifying the dynamics of the state, and a class $\mathscr{U}$ of admissible control functions.

We make some assumptions about the the data defining the control system:

1. The *state space* X is an open and connected subset of $\mathbb{R}^n$.

2. The *control set* U is a subset of $\mathbb{R}^m$.

3. The *dynamics* $\dot{\boldsymbol{x}} = \mathbf{f}(t, \boldsymbol{x}, u)$ is defined by a family of time-varying vector fields $\mathbf{f}$ parametrized by the control values $u \in U$,

$$\mathbf{f} : \mathbb{R} \times X \times U \to \mathbb{R}^n, (t, \boldsymbol{x}, u) \mapsto \mathbf{f}(t, \boldsymbol{x}, u)$$

4. The class $\mathscr{U}$ of *admissible controls* is taken to be piecewise continuous functions u defined on a compact interval $I \subset \mathbb{R}$ with values in the control set U. Without loss of generality, we assume that controls are continuous from the left.

Definition 8. [50] Given a piecewise continuous control $u \in \mathscr{U}$ defined over some open interval $\mathcal{J}$, consider the following initial value problem (IVP):

$$\frac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t} = \mathbf{f}(t, \boldsymbol{x}(t), u(t)), \quad \boldsymbol{x}(t_0) = \boldsymbol{x}_0 \tag{3.22}$$

defined over some maximal interval $(\tau_-, \tau_+) \subset \mathcal{J}$ such that $t_0 \in (\tau_-, \tau_+)$.
Given an admissible $u \in \mathscr{U}$ defined over an interval $\mathcal{J}$, let $\boldsymbol{x}$ be the unique solution to the IVP (3.11) with a maximal interval of definition $I = (\tau_-, \tau_+)$. We call this solution $\boldsymbol{x}$ the trajectory corresponding to the control u and call the pair $(\boldsymbol{x}, u)$ an *admissible controlled trajectory* over the interval I.

An optimal control problem then consists of a control system Σ with an objective functional to be minimized:

$$\min_{u \in \mathscr{U}} \mathscr{J}(u) = \min_{u \in \mathscr{U}} \int_{t_0}^{T} \mathscr{L}(s, \boldsymbol{x}(s), u(s)) ds + \varphi(T, \boldsymbol{x}(T)), \tag{3.23}$$

$$\text{s.t.} \quad \frac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t} = \mathbf{f}(t, \boldsymbol{x}(t), u(t)), \quad \boldsymbol{x}(t_0) = \boldsymbol{x}_0 \tag{3.24}$$

$$\Psi(T, \boldsymbol{x}(T)) = 0, \tag{3.25}$$

where the Lagrangian functional $(t, u(t)) \mapsto \mathscr{L}(t, \boldsymbol{x}(t), u(t))$ represents the running cost over the time interval $[t_0, T]$ associated with a given control function, $\varphi(T, \boldsymbol{x}(T))$ represents a terminal cost, and the set $N = \{(t, \boldsymbol{x}) \in \mathbb{R} \times \mathbb{R}^2 : \Psi(t, \boldsymbol{x}) = 0\}$ represents a terminal constraint.

Definition 9. [50] The (control) *Hamiltonian* function H of the optimal control problem (3.23) is defined as

$$H : \mathbb{R} \times [0, \infty) \times (\mathbb{R}^n)^* \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$$

with

$$H(t, \lambda_0, \lambda, \boldsymbol{x}, u) = \lambda_0 \mathscr{L}(t, \boldsymbol{x}, u) + \boldsymbol{\lambda} \mathbf{f}(t, \boldsymbol{x}, u).$$

Theorem 9. *Pontryagin Maximum Principle [50] Consider a control system $(X, U, f, \mathscr{U})$. Let $(\boldsymbol{x}^*, u^*)$ be a controlled trajectory defined over the interval $[t_0, T]$ with the control u^* piecewise continuous. If $(\boldsymbol{x}^*, u^*)$ is optimal, then there exist a constant $\lambda_0 \geq 0$ and a covector $\boldsymbol{\lambda} : [t_0, T] \to (\mathbb{R}^n)^*$, the so-called* adjoint *variable, such that the following conditions are satisfied:*

1. Nontriviality of the multipliers*: $\big(\lambda_0, (\boldsymbol{\lambda}(t)\big) \neq 0$ for all $t \in [t_0, T]$.*

2. Adjoint equation*: the adjoint variable $\boldsymbol{\lambda}$ is a solution to the time-varying linear differential equation*

$$\frac{\mathrm{d}\boldsymbol{\lambda}}{\mathrm{d}t} = -\lambda_0 \mathscr{L}_{\boldsymbol{x}}(t, \boldsymbol{x}^*(t), u^*(t)) - \boldsymbol{\lambda}(t) f_{\boldsymbol{x}}(t, \boldsymbol{x}^*(t), u^*(t))$$

3. Minimum condition*: everywhere in $[t_0, T]$ we have that*

$$H(t, \lambda_0, \boldsymbol{\lambda}(t), \boldsymbol{x}^*(t), u^*(t)) = \min_{u \in U} H(t, \lambda_0, \boldsymbol{\lambda}(t), \boldsymbol{x}^*(t), u)$$

4. Transversality condition*: at the endpoint of the controlled trajectory, the covector $\Big($*$H + \lambda_0 \varphi_t, -\lambda + \lambda_0 \varphi_{\boldsymbol{x}}\Big)$ is orthogonal to the terminal constraint represented by the level set $\Psi(T, \boldsymbol{x}^*(T)) = 0$, i.e., the covector is parallel to $(\Psi_t(T, \boldsymbol{x}^*(T)), \Psi_{\boldsymbol{x}}(T, \boldsymbol{x}^*(T)))$.*

3.2.2 Optimal Control for the Gut Microbiome System

We now consider our optimal control problem (3.11) with objective functional (3.21). Under assumptions (3.17) and (3.18), we have that our optimal trajectory (x^*, y^*) is stable for $x_0 > 0, y_0 > 0$. Therefore, our $x(t), y(t)$ are strictly positive. So our state space is $X = \{(x, y) \in \mathbb{R}^2 : x > 0, y > 0\}$, which is an open and connected subset of $\mathbb{R}^2$. Our control set is $U = [0, u_{max}] \subset \mathbb{R}$. Our dynamics are given by the system of differential equations (3.11). From (3.21), we have $\mathscr{L}(t, \boldsymbol{x}(t), u(t)) = -y(t) + \frac{c}{2}u^2(t)$ and no terminal cost, i.e. $\varphi(t, \boldsymbol{x}) = 0$ for all $(t, \boldsymbol{x}) \in [0, T] \times X$. Moreover, we take the terminal constraint function to be $\Psi(t, \boldsymbol{x}) := t - T$. Since neither the Lagrangian functional $\mathscr{L}$ nor the vector field $\mathbf{f}$ depend explicitly on time t, the control Hamiltonian function for Problem (3.21) is a mapping $H : \mathbb{R}_+ \times (\mathbb{R}^2)^* \times \mathbb{R}^2_+ \times \mathbb{R}_+ \to \mathbb{R}$ is of the form,

$$H(\lambda_0, \boldsymbol{\lambda}, \boldsymbol{x}, u) = \lambda_0 \mathscr{L}(\boldsymbol{x}, u) + \boldsymbol{\lambda}\mathbf{f}(\boldsymbol{x}, u),$$

or more explicitly,

$$\begin{aligned} &H(\lambda_0, (\lambda_1, \lambda_2), (x, y)^T, u) \\ &\quad = \lambda_0 \left(-y + \frac{c}{2}u^2\right) + \lambda_1 \left((I - \mu x)u - \frac{y\alpha x}{m + x}\right) + \lambda_2 y \left(\beta \frac{\alpha x}{m + x} - \mu u - \nu\right). \end{aligned} \tag{3.26}$$

Let $((x^*, y^*)^T, u^*)$ be a controlled trajectory defined over the interval $[0, T]$ and associated with the optimal control $u^* \in \mathscr{U}$. Then, according to Theorem 9, there exist a constant $\lambda_0 > 0$ and a covector $(\lambda_1, \lambda_2) : [0, T] \to (\mathbb{R}^2)^*$ such that the following conditions are satisfied:

1. Nontriviality of the multipliers: $\big(\lambda_0, (\lambda_1(t), \lambda_2(t)\big) \neq 0$ for all $t \in [0, T]$.

2. Adjoint equation: the adjoint variable (λ_1, λ_2) is a solution to the differential equation

$$\dot{\boldsymbol{\lambda}} = -H_{\boldsymbol{x}}(\lambda_0, \boldsymbol{\lambda}, \boldsymbol{x}, u) = -\lambda_0 \mathscr{L}_{\boldsymbol{x}}(\boldsymbol{x}, u) - \boldsymbol{\lambda}\mathbf{f}_{\boldsymbol{x}}(\boldsymbol{x}, u),$$

or more explicitly

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}\lambda_1(t)\\ \lambda_2(t)\end{bmatrix} = -\lambda_0\begin{bmatrix}0\\ -1\end{bmatrix} - \begin{bmatrix}\lambda_1 & \lambda_2\end{bmatrix}\begin{bmatrix}-\mu u^* - \frac{\alpha m y^*}{(m+x^*)^2} & -\frac{\alpha x^*}{m+x^*}\\ \frac{\alpha\beta m y^*}{(m+x^*)^2} & \frac{\alpha\beta x^*}{m+x^*} - \mu u^* - \nu\end{bmatrix} \tag{3.27}$$

Note that all extremals are normal, since if $\lambda_0 = 0$, then $\lambda(t)$ is a solution to a homogeneous linear equation with zero boundary conditions (which result from the transversality conditions discussed below), hence identically zero, which contradicts the nontriviality statement. So without loss of generality, we normalize $\lambda_0 = 1$.

3. Minimum condition: at every $t \in [0, T]$ we have

$$H(\boldsymbol{\lambda}(t), \boldsymbol{x}^*(t), u^*(t)) = \min_{u\in[0,u_{\max}]} H(\boldsymbol{\lambda}(t), \boldsymbol{x}^*(t), u),$$

where $\lambda_0 = 1$ as established above. The mapping $u \mapsto H(\boldsymbol{\lambda}(t), \boldsymbol{x}^*(t), u)$ is quadratic in the variable u, specifically,

$$H(\boldsymbol{\lambda}(t), \boldsymbol{x}^*(t), u) = \frac{c}{2}u^2 + \big(\lambda_1(I-\mu x) - \lambda_2\mu y\big)u + \left(-y - \lambda_1\frac{y\alpha x}{m+x} + \lambda_2 y\big(\beta\frac{\alpha x}{m+x} - \nu\big)\right).$$

Moreover for any $t \in [0, T]$, its second partial derivative

$$\frac{\partial^2 H}{\partial u^2}(\boldsymbol{\lambda}(t), \boldsymbol{x}^*(t), u) = c > 0,$$

which implies that $H(\boldsymbol{\lambda}(t), \boldsymbol{x}^*(t), u)$ is strictly convex in u and its unconstrained minimizer is thus given by the stationary point $\tilde{u}$ computed as

$$\tilde{u} = \frac{1}{c}\big(\lambda_1(\mu x - I) + \lambda_2\mu y\big).$$

The minimizer $u^*(t)$ of $H(\boldsymbol{\lambda}(t), \boldsymbol{x}^*(t), u)$ over the set $[0, u_{\max}]$ can thus be written as

$$u^*(t) = \min(\max(0, \tilde{u}), u_{\max}). \tag{3.28}$$

This equation expresses the optimal control function explicitly in terms of the optimal state variables $\boldsymbol{x}^*$ and adjoint variables $\boldsymbol{\lambda}$, which in turn are determined dynamically by the state equations (3.11) and adjoint equations (3.27) respectively. It also shows that the optimal control $u^*(t)$ is in fact a continuous function.

4. Transversality condition: Finally, we use the transversality conditions to establish the terminal conditions satisfied by the adjoint variables $\boldsymbol{\lambda}$ at the endpoint T. Applying Theorem 9 to Problem (3.21) with $\Psi(t, \boldsymbol{x}) = t - T$ and $\varphi(t, \boldsymbol{x}) = 0$, we obtain that the co-vector $(H(\boldsymbol{\lambda}(T), \boldsymbol{x}^*(T), u^*(T)), -\boldsymbol{\lambda}(T))$ associated with the optimal controlled trajectory is parallel to $(\Psi_t(T, \boldsymbol{x}^*(T)), \Psi_{\boldsymbol{x}}(T, \boldsymbol{x}^*(T))) = (1, 0)$, implying that $\boldsymbol{\lambda}(T) = 0$.

Note that since $H_{\boldsymbol{\lambda}}(\boldsymbol{\lambda}, \boldsymbol{x}, u) = \mathbf{f}(\boldsymbol{x}, u)$, the first order necessary optimality conditions at the optimal controlled trajectory $(\boldsymbol{x}^*, u^*)$ arising from the Pontryagin Maximum Principle can be written abstractly as the Hamiltonian boundary value problem

$$\begin{aligned} \dot{\boldsymbol{x}} &= H_{\boldsymbol{\lambda}}(\boldsymbol{\lambda}, \boldsymbol{x}, u), \qquad \boldsymbol{x}(0) = \boldsymbol{x}_0, \\ \dot{\boldsymbol{\lambda}} &= -H_{\boldsymbol{x}}(\boldsymbol{\lambda}, \boldsymbol{x}, u), \qquad \boldsymbol{\lambda}(T) = 0, \end{aligned} \tag{3.29}$$

where u is given by Equation (3.28). More explicitly,

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = (I - \mu x(t))u^*(t) - y(t)\frac{\alpha x(t)}{m + x(t)}, \qquad x(0) = x_0 > 0, \tag{3.30}$$

$$\frac{\mathrm{d}y(t)}{\mathrm{d}t} = y(t)\Big(\beta\frac{\alpha x(t)}{m + x(t)} - \mu u^*(t) - \nu\Big), \qquad y(0) = y_0 \geq 0, \tag{3.31}$$

$$\frac{\mathrm{d}\lambda_1(t)}{\mathrm{d}t} = \lambda_1\left(\mu u^*(t) + \frac{\alpha m y}{(m + x)^2}\right) - \lambda_2\frac{\alpha\beta m y}{(m + x)^2}, \qquad \lambda_1(T) = 0, \tag{3.32}$$

$$\frac{\mathrm{d}\lambda_2(t)}{\mathrm{d}t} = 1 + \lambda_1\frac{\alpha x}{m + x} - \lambda_2\left(\frac{\alpha\beta x}{m + x} - \mu u^*(t) - \nu\right), \quad \lambda_2(T) = 0. \tag{3.33}$$

It now follows directly from Equation (3.29) that the Hamiltonian is constant in time along the optimal trajectory. Specifically, using the chain rule, we obtain

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}H(\boldsymbol{\lambda}(t), \boldsymbol{x}(t), u(t)) \\
&= H_{\boldsymbol{\lambda}}(\boldsymbol{\lambda}(t), \boldsymbol{x}^*(t), u^*(t))\dot{\boldsymbol{\lambda}}(t) + H_{\boldsymbol{x}}(\boldsymbol{\lambda}(t), \boldsymbol{x}^*(t), u^*(t))\dot{\boldsymbol{x}}(t) + H_u(\boldsymbol{\lambda}(t), \boldsymbol{x}^*(t), u^*(t))\dot{u^*}(t) \\
&= \dot{\boldsymbol{x}}(t)\dot{\boldsymbol{\lambda}}(t) - \dot{\boldsymbol{x}}(t)\dot{\boldsymbol{\lambda}}(t) + H_u(\boldsymbol{\lambda}(t), \boldsymbol{x}^*(t), u^*(t))\dot{u^*}(t).
\end{aligned}$$

It remains to show that the last term is zero for all $t \in [0, T]$. When the stationary point $\tilde{u}(t) \in [0, u_{\max}]$, then $H_u(\boldsymbol{\lambda}(t), \boldsymbol{x}^*(t), u^*(t)) = H_u(\boldsymbol{\lambda}(t), \boldsymbol{x}^*(t), \tilde{u}) = 0$ and hence the term vanishes. On the other hand, if $\tilde{u}(t) \notin [0, u_{\max}]$, then $u^*(t) \in \{0, u_{\max}\}$ is constant, which implies $\dot{u^*}(t) = 0$ and hence the term vanishes as well.

3.2.3 The Hamiltonian Boundary Value Problem and the Associated Linearized System

We re-write (3.30)–(3.33), letting $\boldsymbol{p}(t) = (x(t), y(t), \lambda_1(t), \lambda_2(t))$ and

$$\begin{aligned}
\Phi_1(\boldsymbol{p}; u) &= (I - \mu x(t))u(t) - y(t)\frac{\alpha x(t)}{m + x(t)}, & x(0) = x_0, \\
\Phi_2(\boldsymbol{p}; u) &= y(t)\Big(\beta\frac{\alpha x(t)}{m + x(t)} - \mu u(t) - \nu\Big), & y(0) = y_0, \\
\Phi_3(\boldsymbol{p}; u) &= \lambda_1\left(\mu u + \frac{\alpha m y}{(m + x)^2}\right) - \lambda_2\frac{\alpha\beta m y}{(m + x)^2}, & \lambda_1(T) = 0, \\
\Phi_4(\boldsymbol{p}; u) &= 1 + \lambda_1\frac{\alpha x}{m + x} - \lambda_2\left(\frac{\alpha\beta x}{m + x} - \mu u - \nu\right) & \lambda_2(T) = 0
\end{aligned}$$

using $\boldsymbol{\Phi}(\boldsymbol{p}) = [\Phi_1, \Phi_2, \Phi_3, \Phi_4]^T$ to denote the vector-valued function describing the dynamics of $\boldsymbol{p}$. Moreover, let $\mathbf{g} : \mathbb{R}^4 \times \mathbb{R}^4 \to \mathbb{R}^4$ be defined by $\mathbf{g}(\mathbf{a}, \mathbf{b}) = (a_1 - x_0, a_2 - y_0, b_3, b_4)^T$, so that the associated boundary conditions of (3.30)–(3.33) can be described by $\mathbf{g}(\boldsymbol{p}(0), \boldsymbol{p}(T)) = 0$. Note that since the optimal control u can be written explicitly as a function of the state and adjoint variables via Equation (3.28), we sometimes write $u(\boldsymbol{p})$ to emphasize this dependence. The Hamiltonian boundary value problem can thus be represented by

$$\frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} = \boldsymbol{\Phi}(\boldsymbol{p}; u(\boldsymbol{p})), \quad \mathbf{g}(\boldsymbol{p}(0), \boldsymbol{p}(T)) = 0. \tag{3.34}$$

In the ensuing discussion we will often require the linearization of the dynamical system $\mathbf{\Phi}(\boldsymbol{p}, u)$ along a given trajectory $\boldsymbol{p}$, which we will derive the linearized system matrix $\frac{d\mathbf{\Phi}}{d\boldsymbol{p}}$ in the following pages. First recall that by the chain rule,

$$\frac{d\mathbf{\Phi}}{d\boldsymbol{p}} = \frac{\partial \mathbf{\Phi}}{\partial \boldsymbol{p}} + \frac{\partial \mathbf{\Phi}}{\partial u}\frac{du}{d\boldsymbol{p}}.$$

Moreover,

$$\mathbf{\Phi}(\boldsymbol{p}, u(\boldsymbol{p})) = \begin{bmatrix} \mathbf{f}(\boldsymbol{x}, u(\boldsymbol{p})) \\ -H_{\boldsymbol{x}}(\boldsymbol{\lambda}, \boldsymbol{x}, u(\boldsymbol{p})) \end{bmatrix} = \begin{bmatrix} \mathbf{f}(\boldsymbol{x}, u(\boldsymbol{p})) \\ -\mathscr{L}_{\boldsymbol{x}}(\boldsymbol{x}, u(\boldsymbol{p})) - \mathbf{f}_{\boldsymbol{x}}(\boldsymbol{x}, u(\boldsymbol{x}))^T\boldsymbol{\lambda} \end{bmatrix}.$$

Now[1]

$$\frac{\partial \mathbf{\Phi}}{\partial \boldsymbol{p}} = \begin{bmatrix} \mathbf{f}_{\boldsymbol{x}}(\boldsymbol{x}, u(\boldsymbol{p})) & 0 \\ -H_{\boldsymbol{xx}}(\boldsymbol{\lambda}, \boldsymbol{x}, u(\boldsymbol{p})) & -\mathbf{f}_{\boldsymbol{x}}(\boldsymbol{x}, u(\boldsymbol{p}))^T \end{bmatrix}.$$

Differentiating the right hand side of Equations (3.9) and (3.10) gives

$$\mathbf{f}_{\boldsymbol{x}}(\boldsymbol{x}, u) = \begin{bmatrix} -\mu u - \frac{\alpha m y}{(m+x)^2} & -\frac{\alpha x}{(m+x)} \\ \frac{\alpha\beta m y}{(m+x)^2} & \frac{\alpha\beta x}{(m+x)} - \mu u - \nu \end{bmatrix}$$

while the derivative $\mathscr{L}_{\boldsymbol{x}}(\boldsymbol{x}, u)$ of the Lagrangian functional $\mathscr{L}(\boldsymbol{x}, u) = -y + \frac{c}{2}u^2$ is given by the row vector $\mathscr{L}_{\boldsymbol{x}}(\boldsymbol{x}, u) = [0, 1]$, which is constant, i.e.

$$H_{\boldsymbol{x}}(\boldsymbol{\lambda}, \boldsymbol{x}, u) = -\mathscr{L}_{\boldsymbol{x}}(\boldsymbol{x}, u(\boldsymbol{p})) - \mathbf{f}_{\boldsymbol{x}}(\boldsymbol{x}, u(\boldsymbol{x}))^T\boldsymbol{\lambda} = \left[\lambda_1\frac{\partial f_1}{\partial x} + \lambda_2\frac{\partial f_2}{\partial x}, 1 + \lambda_1\frac{\partial f_1}{\partial y} + \lambda_2\frac{\partial f_2}{\partial y}\right]^T.$$

[1]Note that in writing the system, we have converted the row vector $\boldsymbol{\lambda}$ to a column vector.

Using these formulas, we can compute $-H_{\boldsymbol{xx}}$ component-wise as

$$
\begin{aligned}
-\frac{\partial^2 H}{\partial x^2} &= \lambda_1 \frac{\partial^2 f_1}{\partial x^2} + \lambda_2 \frac{\partial^2 f_2}{\partial x^2} = -2\lambda_1 \frac{\alpha m y}{(m+x)^3} + 2\lambda_2 \frac{\alpha\beta m y}{(m+x)^3} = \frac{2\alpha m y(\beta\lambda_2 - \lambda_1)}{(m+x)^3} \\
-\frac{\partial^2 H}{\partial y^2} &= \lambda_1 \frac{\partial^2 f_1}{\partial y^2} + \lambda_2 \frac{\partial^2 f_2}{\partial y^2} = \frac{\partial}{\partial y}\left[1 + \lambda_1 \frac{\alpha m}{(m+x)} + \lambda_2(\mu u + \nu - \frac{\alpha\beta x}{(m+x)}\right] = 0 \\
-\frac{\partial^2 H}{\partial x \partial y} &= -\frac{\partial^2 H}{\partial y \partial x} = \lambda_1 \frac{\alpha m}{(m+x)^2} - \lambda_2 \frac{\alpha\beta m}{(m+x)^2} = \frac{\alpha m(\lambda_1 - \beta\lambda_2)}{(m+x)^2},
\end{aligned}
$$

so that

$$
H_{\boldsymbol{xx}} = \begin{bmatrix} \frac{2\alpha m y(\beta\lambda_2 - \lambda_1)}{(m+x)^3} & \frac{\alpha m(\lambda_1 - \beta\lambda_2)}{(m+x)^2} \\ \frac{\alpha m(\lambda_1 - \beta\lambda_2)}{(m+x)^2} & 0 \end{bmatrix}.
$$

Combining the expressions for $H_{\boldsymbol{xx}}$ and $\mathbf{f}_{\boldsymbol{x}}$ above, finally obtain

$$
\frac{\partial \boldsymbol{\Phi}}{\partial \boldsymbol{p}} = \begin{bmatrix} -\mu u - \frac{\alpha m y}{(m+x)^2} & -\frac{\alpha x}{(m+x)} & 0 & 0 \\ \frac{\alpha\beta m y}{(m+x)^2} & \frac{\alpha\beta x}{(m+x)} - \mu u - \nu & 0 & 0 \\ \frac{2\alpha m y(\beta\lambda_2 - \lambda_1)}{(m+x)^3} & \frac{\alpha m(\lambda_1 - \beta\lambda_2)}{(m+x)^2} & \mu u + \frac{\alpha m y}{(m+x)^2} & -\frac{\alpha\beta m y}{(m+x)^2} \\ \frac{\alpha m(\lambda_1 - \beta\lambda_2)}{(m+x)^2} & 0 & \frac{\alpha x}{(m+x)} & -\frac{\alpha\beta x}{(m+x)} + \mu u + \nu \end{bmatrix}. \quad (3.35)
$$

To compute the partial derivative $\frac{\partial \boldsymbol{\Phi}}{\partial u}$, we use the above expressions (or Equations (3.30)–(3.33) directly) to obtain

$$
\frac{\partial \boldsymbol{\Phi}}{\partial u} = \frac{\partial}{\partial u}\begin{bmatrix} \mathbf{f}(\boldsymbol{x}, u) \\ -\mathscr{L}_{\boldsymbol{x}}(\boldsymbol{x}, u) - \mathbf{f}_{\boldsymbol{x}}(\boldsymbol{x}, u)^T \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} (I - \mu x) \\ -\mu y \\ \mu\lambda_1 \\ \mu\lambda_2 \end{bmatrix}. \quad (3.36)
$$

Lastly, recall from Equation (3.28) that the optimal control $u^*(t)$ is defined in terms of the state $\boldsymbol{p}$ as

$$
u^*(\boldsymbol{p}) = \begin{cases} \tilde{u} := \frac{1}{c}(\lambda_1(\mu x - I) + \lambda_2 \mu y), & \text{if } \tilde{u} \in [0, u_{\max}] \\ 0, & \text{otherwise} \end{cases}
$$

Its derivative with respect to $\boldsymbol{p}$ is therefore

$$\frac{du}{d\boldsymbol{p}} = \begin{cases} \frac{d\tilde{u}}{d\boldsymbol{p}}, & \text{if } \tilde{u} \in [0, u_{\max}] \\ 0, & \text{otherwise} \end{cases}$$
$$= \begin{cases} \frac{1}{c}[\mu\lambda_1, \mu\lambda_2, \mu x - I, \mu y], & \text{if } \tilde{u} \in [0, u_{\max}] \\ 0, & \text{otherwise} \end{cases}. \tag{3.37}$$

The following lemma establishes well-posedness of the boundary value problem (3.34).

Lemma 8 (Isolated root, see [49])**.** *Problem* (3.34) *has a unique, isolated solution* $\boldsymbol{p}^* \in C([0,T], \mathbb{R}^4)$.

Proof. The linearized system is

$$\frac{\delta\boldsymbol{p}}{dt} = \frac{d\Phi}{d\boldsymbol{p}}\delta\boldsymbol{p}$$

where we define $\frac{d\Phi}{d\boldsymbol{p}}$ in Subsection 3.3.1, with boundary conditions

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \delta\boldsymbol{p}(0) + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \delta\boldsymbol{p}(T) = 0.$$

An equivalent way of stating these boundary conditions is

$$\delta p_1(0) = 1, \delta p_2(0) = 1, \delta p_3(T) = 1, \delta p_4(T) = 1.$$

This system is well-posed since $\frac{\mathrm{d}\Phi}{\mathrm{d}\boldsymbol{p}}$ is linear and therefore has a unique solution. Clearly, the trivial solution satisfies this BVP, so $\delta\boldsymbol{p} = \mathbf{0}$ is the only solution. Therefore, $\boldsymbol{p}^*$ is an isolated solution of (3.34). □

Since the Pontryagin Maximum Principle gives rise to a BVP, (3.30)–(3.33), we next look at some numerical methods for solving this BVP. In particular, we use will use shooting methods to solve the BVP.

3.3 Shooting Methods

The Pontryagin Maximum Principle applied to our system in the previous section provides a set of necessary optimality conditions that must be satisfied under the optimal control function u^*. Specifically, the minimum condition, Condition 3, allows us to express the optimal control explicitly in terms of the state $\boldsymbol{x} = (x, y)$ and adjoint vectors $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$, which jointly solve the two-point boundary value problem (BVP) (3.30)–(3.33). It is challenging to analytically find an explicit solution for thia BVP and must therefore be solved numerically. While there are numerical methods available to solve the BVP directly, it is often more convenient to formulate an equivalent initial value problem (IVP), i.e. one with the same solutions, which can then be solved using standard numerical integrators. Given a BVP, we can use shooting methods to transform the BVP into an IVP considering the boundary conditions as a function propogated by the different possible initial conditions.

Shooting methods aim to find a associated initial value problem (IVP) of the form

$$\frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} = \boldsymbol{\Phi}(\boldsymbol{p}; u(\boldsymbol{p})), \quad \boldsymbol{p}(0) = \mathrm{s}, \tag{3.38}$$

where $\boldsymbol{s} = (s_1, s_2, s_3, s_4)^T$ are the initial conditions, two of which are known, i.e., $s_1 = x_0$ and $s_2 = y_0$, and two of which are to be determined numerically, i.e., s_3 and s_4. Denote the solutions to (3.38) with initial conditions $\boldsymbol{s}$ as $\boldsymbol{p}(t; \boldsymbol{s})$.

Clearly, the solution $\boldsymbol{p}^*(t)$ of the boundary value problem (3.34) satisfies the initial value problem (3.38) for $\boldsymbol{s}^* = \boldsymbol{p}^*(0)$. We can evaluate the accuracy of any approximation $\boldsymbol{s} \approx \boldsymbol{s}^*$ by computing the residual, which is just the function g evaluated at the boundary conditions propagated by our guess $\boldsymbol{s}$, i.e.

$$\boldsymbol{r}(\boldsymbol{s}) := \mathrm{g}(\boldsymbol{s}, \boldsymbol{p}(T; \boldsymbol{s})).$$

Our goal is to find an $\boldsymbol{s}^* := [x_0, y_0, s_3^*, s_4^*]^T$ such that $\|\boldsymbol{r}(\boldsymbol{s}^*)\| = 0$. Then, by uniqueness we have that $\boldsymbol{p}(t; \boldsymbol{s}^*)$ is a solution to (3.34). We will approximate $\boldsymbol{s}^*$ iteratively by means of a Newton-Rhapson rootfinding method with initial guess $\boldsymbol{s}_0$. At the k-th iteration, the current

estimate $\boldsymbol{s}_k$ will thus be updated to

$$\boldsymbol{s}_{k+1} = \boldsymbol{s}_k - J(\boldsymbol{s}_k)^{-1}\boldsymbol{r}(s_k),$$

where $J(\boldsymbol{s}) = \frac{d\boldsymbol{r}}{d\boldsymbol{s}}$ is the 4×4 Jacobian matrix computed at the current iterate $\boldsymbol{s}_k$. Using the definition of residual

$$\mathbf{r}(\boldsymbol{s}, \mathbf{p}(T, \boldsymbol{s})) = \begin{bmatrix} s_1 - x_0 \\ s_2 - y_0 \\ \lambda_1(T, \boldsymbol{s}) \\ \lambda_2(T, \boldsymbol{s}) \end{bmatrix},$$

we can compute the Jacobian explicitly as

$$\frac{d\boldsymbol{r}}{d\boldsymbol{s}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ w_{3,1}(T) & w_{3,2}(T) & w_{3,3}(T) & w_{3,4}(T) \\ w_{4,1}(T) & w_{4,2}(T) & w_{4,3}(T) & w_{4,4}(T) \end{bmatrix},$$

where $w_{i,j} = \frac{dp_i}{ds_j}$ are the sensitivities of the solution $\boldsymbol{p}(t, \boldsymbol{s})$ with respect to the initial conditions, which are treated in detail in Subsection 3.3.1. Two questions that arise in this context are: 'For which values of $\boldsymbol{s}$ is the initial value problem (3.38) well-posed?' and 'Under what conditions are the sensitivies well-defined?' Lemma 9 shows that well-posedness is guaranteed for all $\boldsymbol{s}$ close enough to $\boldsymbol{s}^*$. To this end, let

$$S_\rho(\boldsymbol{s}^*) = \{\boldsymbol{s} \in \mathbb{R}^4 : \|\boldsymbol{s} - \boldsymbol{s}^*\| < \rho\} \tag{3.39}$$

denote a neighborhood of radius $\rho > 0$ around $\boldsymbol{s}^*$, and let

$$T_\delta\{\boldsymbol{p}^*\} := \{(t, \boldsymbol{p}) : 0 \leq t \leq T, ||\boldsymbol{p}^*(t) - \boldsymbol{p}|| \leq \delta\} \subset [0, T] \times \mathbb{R}^4 \tag{3.40}$$

be the δ-region around the optimal trajectory $\boldsymbol{p}^*(t)$.

Lemma 9 (Well-posedness of nearby IVPs and their linearizations)**.** *Let $\boldsymbol{p}^*(t)$ be a solution of* (3.34)*. Then there is a $\delta > 0$ so that the mapping $\boldsymbol{p} \mapsto \Phi(\boldsymbol{p}, u(\boldsymbol{p}))$ is continuously differentiable in $\boldsymbol{p}$ in $T_\delta\{\boldsymbol{p}^*\}$. Therefore*

(i) For each $\boldsymbol{s} \in S_\rho(\boldsymbol{s}^*)$, where $\rho = \delta e^{-KT}$, there exists a unique solution $\boldsymbol{p}(t, \boldsymbol{s})$ of (3.38) *on $0 \leq t \leq T$, where $K > 0$ is the Lipschitz constant so that*

$$||\boldsymbol{\Phi}(\boldsymbol{p}, u(\boldsymbol{p})) - \boldsymbol{\Phi}(\tilde{\boldsymbol{p}}, u(\tilde{\boldsymbol{p}}))|| \leq K||\boldsymbol{p} - \tilde{\boldsymbol{p}}||, \qquad \textit{for all } (t, \boldsymbol{p}), (t, \tilde{\boldsymbol{p}}) \in T_\delta\{\boldsymbol{p}^*\}.$$

(ii) The sensitivity $W(t) = \frac{d\boldsymbol{p}}{d\boldsymbol{s}}$ exists and is the solution of the linearized (sensitivity) equation

$$\dot{W} = \frac{d\boldsymbol{\Phi}}{d\boldsymbol{p}} W, \qquad W(0) = I.$$

Proof. It suffices to show that $p \mapsto \Phi(\boldsymbol{p}, u(\boldsymbol{p}))$ is continuously differentiable, in which case the result follows from Lemma 1.25 in [49]. Explicit expressions for $\frac{d\boldsymbol{\Phi}}{d\boldsymbol{p}}$, $\frac{\partial \boldsymbol{\Phi}}{\partial u}$, and $\frac{du}{d\boldsymbol{p}}$, given in Equations (3.35), (3.36), and (3.37) respectively, show that this can be achieved as long as $\delta > 0$ is chosen small enough to ensure that x remains bounded away from $-m$, e.g., by guaranteeing that $x \geq \frac{-m}{2}$. □

3.3.1 Sensitivity Equations

In this subsection, we formally derive the sensitivities $\frac{\partial \boldsymbol{p}}{\partial \boldsymbol{s}}$ of solutions $\boldsymbol{p}(t, \boldsymbol{s})$ to the the initial value problem (3.38) with respect to the initial conditions $\boldsymbol{s}$. These will take the form of solutions of the sensitivity equations (See Equation (3.41)). Let $W = \frac{\partial \boldsymbol{p}}{\partial \boldsymbol{s}}$ denote the matrix of sensitivities of the vector-valued function $\boldsymbol{p}(t, \boldsymbol{s})$ with respect to initial conditions $\boldsymbol{s}$. Specifically, let

$$W = \begin{bmatrix} \frac{\partial x}{\partial s_1} & \frac{\partial x}{\partial s_2} & \frac{\partial x}{\partial s_3} & \frac{\partial x}{\partial s_4} \\ \frac{\partial y}{\partial s_1} & \frac{\partial y}{\partial s_2} & \frac{\partial y}{\partial s_3} & \frac{\partial y}{\partial s_4} \\ \frac{\partial \lambda_1}{\partial s_1} & \frac{\partial \lambda_1}{\partial s_2} & \frac{\partial \lambda_1}{\partial s_3} & \frac{\partial \lambda_1}{\partial s_4} \\ \frac{\partial \lambda_2}{\partial s_1} & \frac{\partial \lambda_2}{\partial s_2} & \frac{\partial \lambda_2}{\partial s_3} & \frac{\partial \lambda_2}{\partial s_4} \end{bmatrix}.$$

As before, we will refer to the (i,j)-th entry of W as $w_{i,j} = \frac{\partial p_i}{\partial s_j}$. Moreover, we will use $\boldsymbol{w}_i = \frac{\partial p_i}{\partial \boldsymbol{s}}$ to refer to the i-th row of W. For each fixed value of $\boldsymbol{s}$, W represents a time-varying matrix-valued function. We will now derive a linear system of differential equations satisfied by W. To this end, we formally differentiate both sides of Equation (3.38) as well as its initial conditions with respect to $\boldsymbol{s}$. Interchanging the order of differentiation in the resulting system gives rise to the sensitivity equations. Indeed, recall that Equation (3.38) takes the form

$$\dot{\boldsymbol{p}} = \boldsymbol{\Phi}(\boldsymbol{p}, u(\boldsymbol{p})), \qquad \boldsymbol{p}(0) = \boldsymbol{s}.$$

We can differentiate the initial condition with respect to $\boldsymbol{s}$ to obtain

$$\frac{\partial}{\partial \boldsymbol{s}} \boldsymbol{p}(0, \boldsymbol{s}) = \frac{d\boldsymbol{s}}{d\boldsymbol{s}} = I,$$

where I represents the 4×4 identity matrix. Differentiating both sides of the differential equation and using the chain rule yields

$$\frac{\partial}{\partial \boldsymbol{s}} \left[\frac{\partial \boldsymbol{p}}{\partial t} \right] = \frac{\partial}{\partial \boldsymbol{s}} \left[\boldsymbol{\Phi}(\boldsymbol{p}, u(\boldsymbol{p})) \right] = \frac{d\Phi}{d\boldsymbol{p}} \frac{\partial \boldsymbol{p}}{\partial \boldsymbol{s}}.$$

Interchanging the order of differentiation then leads to the matrix system

$$\dot{W} = \frac{d\boldsymbol{\Phi}}{d\boldsymbol{p}} W, \qquad W(0) = I.$$

$W = [\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3, \boldsymbol{w}_4]^T = \left[\frac{\partial x}{\partial \boldsymbol{s}}, \frac{\partial y}{\partial \boldsymbol{s}}, \frac{\partial \lambda_1}{\partial \boldsymbol{s}}, \frac{\partial \lambda_2}{\partial \boldsymbol{s}} \right]^T$ and $= w_{i,j} = \frac{\partial \Phi_i}{\partial s_j}$ so then differentiating (3.38) with respect to $\boldsymbol{s}$ looks like:

Note that x and y are not dependent on s_3 or s_4, so $\begin{bmatrix} \frac{\partial x}{\partial s_3} & \frac{\partial x}{\partial s_4} \\ \frac{\partial y}{\partial s_3} & \frac{\partial y}{\partial s_4} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. Therefore, we have:

$$W = \begin{bmatrix} \frac{\partial x}{\partial s_1} & \frac{\partial x}{\partial s_2} & 0 & 0 \\ \frac{\partial y}{\partial s_1} & \frac{\partial x}{\partial s_2} & 0 & 0 \\ \frac{\partial \lambda_1}{\partial s_1} & \frac{\partial \lambda_2}{\partial s_4} & \frac{\partial \lambda_1}{\partial s_3} & \frac{\partial \lambda_1}{\partial s_4} \\ \frac{\partial \lambda_2}{\partial s_1} & \frac{\partial \lambda_2}{\partial s_2} & \frac{\partial \lambda_2}{\partial s_3} & \frac{\partial \lambda_2}{\partial s_4} \end{bmatrix}$$

To be able to evaluate the entries of $\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{s}}$, we want to be able to evaluate the entries of W, since

$$\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{s}} = \frac{\partial}{\partial \boldsymbol{s}} \begin{bmatrix} s_1 - x_0 \\ s_2 - y_0 \\ \lambda_1(T, \boldsymbol{s}) \\ \lambda_2(T, \boldsymbol{s}) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{\partial}{\partial s_3}\lambda_1(T, \boldsymbol{s}) & \frac{\partial}{\partial s_4}\lambda_1(T, \boldsymbol{s}) \\ 0 & 0 & \frac{\partial}{\partial s_3}\lambda_2(T, \boldsymbol{s}) & \frac{\partial}{\partial s_4}\lambda_2(T, \boldsymbol{s}) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & w_{3,3}(T) & w_{3,4}(T) \\ 0 & 0 & w_{4,3}(T) & w_{4,4}(T) \end{bmatrix}.$$

Therefore, we take the derivative our IVP, equation (3.38) with respect to $\boldsymbol{s}$.

$$\begin{aligned} &\frac{\partial}{\partial \boldsymbol{s}}\frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} = \frac{\partial}{\partial \boldsymbol{s}}\boldsymbol{\Phi}(\boldsymbol{p}, u), \quad \frac{\partial}{\partial \boldsymbol{s}}\boldsymbol{p}_0(0) = \frac{\partial}{\partial \boldsymbol{s}}[s_1, s_2, s_3, s_4]^T \\ &\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \boldsymbol{p}}{\partial \boldsymbol{s}} = \frac{\partial}{\partial \boldsymbol{s}}\boldsymbol{\Phi}(\boldsymbol{p}, u), \quad W(0) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{\partial s_3}{\partial s_3} & \frac{\partial s_3}{\partial s_4} \\ 0 & 0 & \frac{\partial s_4}{\partial s_3} & \frac{\partial s_4}{\partial s_4} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} := W^0 \\ &\frac{\mathrm{d}W}{\mathrm{d}t} = \mathbf{F}(\boldsymbol{p}, W, u), \quad W(0) = W_0 \end{aligned}$$

where $\mathbf{F}(\boldsymbol{p}, W, u) = \frac{\partial}{\partial \boldsymbol{s}}\boldsymbol{\Phi}(\boldsymbol{p}, u)$.

Thus we get our system of sensitivity equations

$$\frac{\mathrm{d}W}{\mathrm{d}t} = \mathbf{F}(\boldsymbol{p}, W, u), \quad W(0) = W_0. \tag{3.41}$$

To explicitly find our $\boldsymbol{F}$, we switch the order of differentiation to get the following:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial \Phi_1}{\partial \boldsymbol{s}}\Big) &= \frac{\partial}{\partial \boldsymbol{s}}\left((I-\mu x(t))u(t) - y(t)\frac{\alpha x(t)}{m+x(t)}\right) \\
\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial \Phi_2}{\partial \boldsymbol{s}}\Big) &= \frac{\partial}{\partial \boldsymbol{s}}\left(y(t)\Big(\beta\frac{\alpha x(t)}{m+x(t)} - \mu u(t) - \nu\Big)\right) \\
\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial \Phi_3}{\partial \boldsymbol{s}}\Big) &= \frac{\partial}{\partial \boldsymbol{s}}\left(\lambda_1\left(\mu u + \frac{\alpha m y}{(m+x)^2}\right) - \lambda_2\frac{\alpha\beta m y}{(m+x)^2}\right) \\
\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial \Phi_4}{\partial \boldsymbol{s}}\Big) &= \frac{\partial}{\partial \boldsymbol{s}}\left(1 + \lambda_1\frac{\alpha x}{m+x} - \lambda_2\left(\frac{\alpha\beta x}{m+x} - \mu u - \nu\right)\right)
\end{aligned}$$

Making the appropriate substitutions, $W = [\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3, \boldsymbol{w}_4]^T = \left[\frac{\partial x}{\partial \boldsymbol{s}}, \frac{\partial y}{\partial \boldsymbol{s}}, \frac{\partial \lambda_1}{\partial \boldsymbol{s}}, \frac{\partial \lambda_2}{\partial \boldsymbol{s}}\right]^T$, and differentiating, we get the following system of sensitivity equations:

$$\frac{\mathrm{d}\boldsymbol{w}_1}{\mathrm{d}t} = -\mu\boldsymbol{w}_1 u + (I-\mu x)\frac{\partial u}{\partial \boldsymbol{s}} - \frac{\alpha((m+x)\boldsymbol{w}_2 x + m\boldsymbol{w}_1 y)}{(m+x)^2} \tag{3.42}$$

$$\frac{\mathrm{d}\boldsymbol{w}_2}{\mathrm{d}t} = \boldsymbol{w}_2\left(\frac{\alpha\beta x}{(m+x)} - \mu u - \nu\right) + y\left(\frac{\alpha\beta\boldsymbol{w}_1 m}{(m+x)^2} - \mu\frac{\partial u}{\partial \boldsymbol{s}}\right) \tag{3.43}$$

$$\begin{aligned}
\frac{\mathrm{d}\boldsymbol{w}_3}{\mathrm{d}t} &= \mu u\boldsymbol{w}_3 + \frac{\alpha m}{(m+x)^2}(y\boldsymbol{w}_3 + \lambda_1\boldsymbol{w}_2 - \beta\boldsymbol{w}_4 y - \beta\lambda_2\boldsymbol{w}_2) \\
&\quad + \frac{2\alpha m y\boldsymbol{w}_1(\beta\lambda_2 - \lambda_1)}{(m+x)^3} + \mu\lambda_1\frac{\partial u}{\partial \boldsymbol{s}}
\end{aligned} \tag{3.44}$$

$$\begin{aligned}
\frac{\mathrm{d}\boldsymbol{w}_4}{\mathrm{d}t} &= \frac{\alpha x\boldsymbol{w}_3}{(m+x)} + \frac{\alpha m\lambda_1\boldsymbol{w}_1}{(m+x)^2} - \boldsymbol{w}_4\left(\frac{\alpha\beta x}{(m+x)} - \mu u - v\right) \\
&\quad - \lambda_2\left(\frac{\alpha\beta m\boldsymbol{w}_1}{(m+x)^2} - \mu\frac{\partial u}{\partial \boldsymbol{s}}\right)
\end{aligned} \tag{3.45}$$

where $u = u^*$ and $\frac{\partial u}{\partial \boldsymbol{s}} = -\frac{1}{c}(\boldsymbol{w}_3 I - \mu(\boldsymbol{w}_3 x + \boldsymbol{w}_1\lambda_1 + \boldsymbol{w}_4 y + \boldsymbol{w}_2\lambda_2))$ if $u^* = \tilde{u}$, but $\frac{\partial u}{\partial \boldsymbol{s}} = 0$ if $u^* \in \{0, u_{max}\}$.

This means $\frac{dW}{dt} = \boldsymbol{F}(\boldsymbol{p}, W, u)$ is a matrix-valued system of ordinary differential equations:

$$\frac{\mathrm{d}W}{\mathrm{d}t} = \begin{bmatrix} -\mu u + \frac{\alpha m y}{(m+x)^2} & \frac{\alpha x}{m+x} & 0 & 0 \\ \frac{\alpha\beta m y}{(m+x)^2} & \frac{\alpha\beta x}{(m+x)} - \mu u - \nu & 0 & 0 \\ \frac{2\alpha m y(\beta\lambda_2 - \lambda_1)}{(m+x)^3} & \frac{\alpha m(\lambda_1 - \beta\lambda_2)}{(m+x)^2} & \mu u + \frac{\alpha m y}{(m+x)^2} & \frac{-\alpha\beta m y}{(m+x)^2} \\ \frac{\alpha m(\lambda_1 - \beta\lambda_2)}{(m+x)^2} & 0 & \frac{\alpha x}{(m+x)} & -\frac{\alpha\beta x}{(m+x)} + \mu u + \nu \end{bmatrix} \begin{bmatrix} \boldsymbol{w}_1 \\ \boldsymbol{w}_2 \\ \boldsymbol{w}_3 \\ \boldsymbol{w}_4 \end{bmatrix} + \begin{bmatrix} I - \mu x \\ -\mu y \\ \mu\lambda_1 \\ \mu\lambda_2 \end{bmatrix} \frac{\partial u}{\partial \boldsymbol{s}},$$

with initial condition $W(0) = W_0$.

Let $W(t)$ denote the solution to equations (3.42)–(3.45) with initial condition W_0 at time t propogated by the initial condition to (3.38) equal to $\boldsymbol{s}$. These solutions evaluated at time $t = T$ give us the unknown entries of the Jacobian. We will denote the Jacobian associated with $\boldsymbol{s}$ as $J(\boldsymbol{s})$. Then

$$J(\boldsymbol{s}) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ w_{3,1}(T) & w_{3,2}(T) & w_{3,3}(T) & w_{3,4}(T) \\ w_{4,1}(T) & w_{4,2}(T) & w_{4,3}(T) & w_{4,4}(T) \end{bmatrix} \tag{3.46}$$

Note that J is dependent upon $\boldsymbol{s}$ since the solutions to (3.42)–(3.45) are dependent on finding $\boldsymbol{p}(t, \boldsymbol{s})$.

Therefore, given the kth guess for the initial conditions, $\boldsymbol{s}_k$, we can find the next guess

$$\boldsymbol{s}_{k+1} = \boldsymbol{s}_k - J(\boldsymbol{s}_k)^{-1}\boldsymbol{r}(\boldsymbol{s}_k) \text{ for } k \geq 1.$$

3.3.2 Single Shooting Method

Set desired tolerance $tol = \varepsilon$ and maximum desired number of iterations i_{max}.

Now we can define the Jacobian at the kth iteration to be:

$$J(\boldsymbol{s}_k) := \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ w_{3,1}(T) & w_{3,2}(T) & w_{3,3}(T) & w_{3,4}(T) \\ w_{4,1}(T) & w_{4,2}(T) & w_{4,3}(T) & w_{4,4}(T) \end{bmatrix} \tag{3.47}$$

Then we execute the following algorithm:

Algorithm 1 Single Shooting Method

Choose $\boldsymbol{s}_1 \in \mathbb{R}^4$, $\varepsilon > 0$, and $k_{\max} \in \mathbb{N}$.
for $k = 1, \ldots, k_{\max}$ **do**
    Solve (3.38).
    Compute $\boldsymbol{r}(\boldsymbol{s}_k)$.
    if $||\boldsymbol{r}(\boldsymbol{s}_k)|| < \varepsilon$ **then**
        Stop. Use $\boldsymbol{s}_k$ as adjoint initial conditions.
    else
        Compute J_k as defined in (3.47).
        $\boldsymbol{s}_{k+1} = \boldsymbol{s}_k - J(\boldsymbol{s}_k)^{-1}\boldsymbol{r}(\boldsymbol{s}_k)$.
    end if
end for

Remark 10. *Note that since the initial values $s_1 = x_0$ and $s_2 = y_0$ are known, the first two components of the remainder function $\mathbf{r}(\boldsymbol{s}_1)$ will be zero at the beginning of the first iteration. Writing the Jacobian in* (3.47) *in block matrix form*

$$J(\boldsymbol{s}) = \begin{bmatrix} I & 0 \\ W_{3:4,1:2} & W_{3:4,3:4} \end{bmatrix},$$

we can rewrite the linear system satisfied by $J(\boldsymbol{s})\delta s = \mathbf{r}(\boldsymbol{s})$ as

$$\begin{bmatrix} I & 0 \\ W_{3:4,1:2} & W_{3:4,3:4} \end{bmatrix} \begin{bmatrix} \delta s_1 \\ \delta s_2 \\ \delta s_3 \\ \delta s_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ r_3(\boldsymbol{s}_1) \\ r_4(\boldsymbol{s}_1) \end{bmatrix},$$

which implies

$$\begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \text{ and } \begin{bmatrix} s_3 \\ s_4 \end{bmatrix} = W^{(-1)}_{3:4,3:4} \begin{bmatrix} r_3(\boldsymbol{s}) \\ r_4(\boldsymbol{s}) \end{bmatrix},$$

so that

$$\boldsymbol{s}_2 = \boldsymbol{s}_1 - \begin{bmatrix} 0 \\ 0 \\ \delta s_3 \\ \delta s_4 \end{bmatrix}.$$

Thus the Newton method only updates the third and fourth components of $\boldsymbol{s}_k$.

The following theorem, taken from Theorem 1.29 in [49] guarantees that if we start close enough to the optimal initial conditions $\boldsymbol{s}^*$, we get *quadratic* convergence, provided that the Jacobian is Lipschitz continuous and boundedly invertible.

Theorem 10 (from [49])**.** *Let $\boldsymbol{s}^*$ be the isolated, optimal initial condition associated with the Hamiltonian boundary value problem* (3.34)*, and let $\rho_* > 0$ be such that*

(i) The inverse Jacobian $J^{-1}(\boldsymbol{s})$ is uniformly bounded on $S_{\rho_*}(\boldsymbol{s}^*)$, i.e.,

$$||J^{-1}(\boldsymbol{s}^*)|| \leq \beta, \qquad \boldsymbol{s} \in S_{\rho_*}(\boldsymbol{s}^*), \text{ and}$$

(ii) The Jacobian $J(\boldsymbol{s})$ is uniformly Lipshitz continuous on $S_{\rho^*}(\boldsymbol{s}^*)$, i.e., the is some $\gamma > 0$

$$||J(\boldsymbol{s}) - J(\boldsymbol{s}')|| \leq \gamma ||\boldsymbol{s} - \boldsymbol{s}'|| \text{ for all } \boldsymbol{s}, \boldsymbol{s}' \in S_{\rho_*}(\boldsymbol{s}^*), \text{ and}$$

(iii) The radius ρ^*, the bound β, and the Lipschitz constant γ satisfy the constraint $\rho_* \beta \gamma < \frac{2}{3}$.

Then for every initial guess $\boldsymbol{s}_1 \in S_{\rho_*}(\boldsymbol{s}^*)$ the iterates in the single shooting algorithm remain in $S_{\rho_*}(\boldsymbol{s}^*)$ and converge quadratically to $\boldsymbol{s}^*$; in fact, for all $k = 1, 2, ...$

$$||\boldsymbol{s}_{k+1} - \boldsymbol{s}^*|| \leqq a||\boldsymbol{s}_k - \boldsymbol{s}^*||^2, \qquad \text{where} \qquad a := \frac{\beta\gamma}{2(1 - \rho_*\beta\gamma)} < \frac{1}{\rho_*}.$$

In view of Remark 10, it suffices to consider the continuity and bounded invertibility of the submatrix $W_{3:4,3:4}$, consisting of sensitivies of the two adjoint variables with respect to both of their initial conditions. These in turn can be related, through classical results (e.g. see [51]), to the linear system defining the sensitivity equations. Indeed, as long as $x + m$ remains bounded away from 0 Theorem 7.4 and Theorem 7.5 in [51] imply that the sensitivities are continuously differentiable functions of the initial conditions $\boldsymbol{s}$. Moreover, using properties of the block matrix J, we have $\det(J) = \det(W_{3:4,3:4})$. We thus require the sensitivies

$$\begin{bmatrix} w_{3,3}(T) & w_{3,4}(T) \\ w_{4,3}(T) & w_{4,4}(T) \end{bmatrix}$$

to be linearly independent at the interval endpoint.

The major problem with the single shooting method is the radius of convergence. In particular, the radius $\rho^* > 0$ in Lemma 10 depends on the radius $\rho > 0$ in Lemma 9, which in turn is exponentially decaying in the variable T. Larger time intervals therefore guarantee a fairly small radius of convergence. The multple shooting method, discussed in the following section addresses this shortcoming by first subdividing the time interval into many shorter subintervals, thereby effectively increasing the radius of convergence.

3.3.3 Multiple Shooting Method

Part of the problem with the single shooting method is that when we transform our BVP into an IVP, the IVP may be unstable even though the related BVP was well-posed and stable. However, in some instance, the single shooting method works fine for a very small interval of time. The multiple shooting method essentially breaks up the entire time interval into sub-intervals and solves the system on each sub-interval. Now our residual does not only take into account the boundary conditions, but we also impose patching conditions for continuity amongst sub-intervals.

We start by partitioning our interval $[0, T]$ into N sub-intervals, calling the endpoints of the sub-intervals *nodes*.

$$0 = t_0 < t_1 < ... < t_N = T$$

Then our first guess $\boldsymbol{s}$ is actually a set of guesses for initial conditions on each sub-interval. So for the ith iteration, let $\boldsymbol{s}^i = [s_1, s_2, s_3, s_4, ..., s_{4N}]^T$. While we know the initial conditions on x and y at time $t_0 = 0$, we have to start with an initial guess for the initial conditions on each sub-interval. Let $\boldsymbol{s}_j^i = [s_j^i, s_{j+1}^i, s_{j+2}^i, s_{j+3}^i]^T$ denote the initial conditions for the jth interval on the ith iteration. Then we have $s_1^i = x_0, s_2^i = y_0$.

Let $x_j(t), y_j(t), \lambda_{1_j}(t), \lambda_{2_j}(t)$ denote the phase solutions on the jth sub-interval. Then for $1 \leq j \leq N$ the solutions for the jth sub-interval are found by solving the following IVP on $t \in [t_{j-1}, t_j]$:

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = (I - \mu x(t))u(t) - y(t)\frac{\alpha x(t)}{m + x(t)}, \qquad x(t_{j-1}) = s_{4j-3}, \tag{3.48}$$

$$\frac{\mathrm{d}y(t)}{\mathrm{d}t} = y(t)\Big(\beta\frac{\alpha x(t)}{m + x(t)} - \mu u(t) - \nu\Big), \qquad y(t_{j-1}) = s_{4j-2}, \tag{3.49}$$

$$\frac{\mathrm{d}\lambda_1(t)}{dt} = \lambda_1\left(\mu u + \frac{\alpha m y}{(m + x)^2}\right) - \lambda_2\frac{\alpha\beta m y}{(m + x)^2}, \qquad \lambda_1(t_{j-1}) = s_{4j-1}, \tag{3.50}$$

$$\frac{\mathrm{d}\lambda_2(t)}{dt} = 1 + \lambda_1\frac{\alpha x}{m + x} - \lambda_2\left(\frac{\alpha\beta x}{m + x} - \mu u - \nu\right), \quad \lambda_2(t_{j-1}) = s_{4j}, \tag{3.51}$$

Let $\boldsymbol{p}_j(t, \boldsymbol{s}_j) = [x_j(t), y_j(t), \lambda_{1_j}(t), \lambda_{2_j}(t)]^T$ be the phase solutions on the jth interval and let $\boldsymbol{p}_j^i(t, \boldsymbol{s}_j^i) = [x_j^i(t), y_j^i(t), \lambda_{1_i}^j(t), \lambda_{2_i}^j(t)]^T$ be the solutions on the jth interval at the ith iteration.

The following lemma establishes a radius around $\boldsymbol{s}^*$ around which the initial value problems (3.48)–(3.51) and their sensitivity equations are well-posed, showing explicitly how multiple shooting improves the radius of convergence.

Lemma 10 (Well-posedness of multiple shooting IVPs (see [49])**.** *Let the hypotheses of Lemma 9 hold, with $\delta > 0$ being the radius around $\boldsymbol{s}^*$ within which $\boldsymbol{p} \mapsto \Phi(\boldsymbol{p}, u(\boldsymbol{p}))$ is uniformly Lipschitz continuous with Lipschitz constant K. Then the initial value problems* (3.48)–(3.51)*, as well as the associated sensitivity equations are well-posed for any*

$$\boldsymbol{s} \in S_{\hat{\rho}}(\boldsymbol{s}^*), \qquad \hat{\rho} := \delta e^{-KT/N},$$

Specifically,

$$\hat{\rho} = e^{KT(1-\frac{1}{N})}\rho.$$

Now our residual function does not just measure whether the target boundary conditions are being met, but we also impose continuity conditions between sub-intervals. So our residual function looks like:

$$\boldsymbol{r}(\boldsymbol{s}) = \begin{bmatrix} \mathbf{g}(\boldsymbol{s}_1, \boldsymbol{p}_N(T, \boldsymbol{s}_N)) \\ \boldsymbol{p}_1(t_1, \boldsymbol{s}_1) - \boldsymbol{s}_2 \\ \vdots \\ \boldsymbol{p}_{N-1}(t_{N-1}, \boldsymbol{s}_{N-1}) - \boldsymbol{s}_N \end{bmatrix} \tag{3.52}$$

where $\mathbf{g}(\boldsymbol{s}_1, \boldsymbol{p}_N(T, \boldsymbol{s}_N)) = [s_1 - x_0, s_2 - y_0, \lambda_1^N(T, \boldsymbol{s}_N), \lambda_2^N(T, \boldsymbol{s}_N)]^T$.

And now when we evaluate the Jacobian, we differentiate with respect to $\boldsymbol{s}$, which yields the following:

$$J = \begin{bmatrix} \frac{\partial \mathbf{g}}{\partial \boldsymbol{s}} \\ \frac{\partial \boldsymbol{p}_1}{\partial \boldsymbol{s}} - \frac{\partial \boldsymbol{s}_2}{\partial \boldsymbol{s}} \\ \vdots \\ \frac{\partial \boldsymbol{p}_{N-1}}{\partial \boldsymbol{s}} - \frac{\partial \boldsymbol{s}_N}{\partial \boldsymbol{s}} \end{bmatrix}$$

For the first row of the Jacobian, the entries denote the partial derivative of $\mathbf{g}$ with respect to $\boldsymbol{s}$, which is derived by taking the derivative of g with respect to $\boldsymbol{s}_j$ for $1 \leq j \leq N$. So the entries of the first row are $\frac{\partial \mathbf{g}}{\partial \boldsymbol{s}_1}, \frac{\partial \mathbf{g}}{\partial \boldsymbol{s}_2}, \ldots, \frac{\partial \mathbf{g}}{\partial \boldsymbol{s}_N}$.

For each of the rest of the N rows, we have entries $\frac{\partial \boldsymbol{p}_{j-1}}{\partial \boldsymbol{s}_1} - \frac{\partial \boldsymbol{s}_j}{\partial \boldsymbol{s}_1}, \frac{\partial \boldsymbol{p}_{j-1}}{\partial \boldsymbol{s}_2} - \frac{\partial \boldsymbol{s}_j}{\partial \boldsymbol{s}_2}, \ldots, \frac{\partial \boldsymbol{p}_{j-1}}{\partial \boldsymbol{s}_j} - \frac{\partial \boldsymbol{s}_j}{\partial \boldsymbol{s}_j}$ for $2 \leq j \leq N$. Therefore our Jacobian takes the form:

$$J = \begin{bmatrix} \frac{\partial \mathbf{g}}{\partial \boldsymbol{s}_1} & \frac{\partial \mathbf{g}}{\partial \boldsymbol{s}_2} & \cdots & \frac{\partial \mathbf{g}}{\partial \boldsymbol{s}_N} \\ \frac{\partial \boldsymbol{p}_1}{\partial \boldsymbol{s}_1} - \frac{\partial \boldsymbol{s}_2}{\partial \boldsymbol{s}_1} & \frac{\partial \boldsymbol{p}_1}{\partial \boldsymbol{s}_2} - \frac{\partial \boldsymbol{s}_2}{\partial \boldsymbol{s}_2} & \cdots & \frac{\partial \boldsymbol{p}_1}{\partial \boldsymbol{s}_N} - \frac{\partial \boldsymbol{s}_2}{\partial \boldsymbol{s}_N} \\ \vdots & \ddots & \cdots & \vdots \\ \frac{\partial \boldsymbol{p}_{N-1}}{\partial \boldsymbol{s}_1} - \frac{\partial \boldsymbol{s}_N}{\partial \boldsymbol{s}_1} & \frac{\partial \boldsymbol{p}_{N-1}}{\partial \boldsymbol{s}_2} - \frac{\partial \boldsymbol{s}_N}{\partial \boldsymbol{s}_2} & \cdots & \frac{\partial \boldsymbol{p}_{N-1}}{\partial \boldsymbol{s}_N} - \frac{\partial \boldsymbol{s}_N}{\partial \boldsymbol{s}_N} \end{bmatrix} \tag{3.53}$$

But since $\mathbf{g} = [s_1 - x_0, s_2 - y_0, \lambda_1^N(T, \boldsymbol{s}_N), \lambda_2^N(T, \boldsymbol{s}_N)]^T$, we have $\frac{\partial \mathbf{g}}{\partial \boldsymbol{s}_j} = 0$ for all $j \notin \{1, N\}$, while

$$\frac{\partial \mathbf{g}}{\partial \boldsymbol{s}_N} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{\partial \lambda_{1_N}}{\partial s_{N-3}} & \frac{\partial \lambda_{1_N}}{\partial s_{N-2}} & \frac{\partial \lambda_{1_N}}{\partial s_{N-1}} & \frac{\partial \lambda_{1_N}}{\partial s_N} \\ \frac{\partial \lambda_{2_N}}{\partial s_{N-3}} & \frac{\partial \lambda_{2_N}}{\partial s_{N-2}} & \frac{\partial \lambda_{2_N}}{\partial s_{N-1}} & \frac{\partial \lambda_{2_N}}{\partial s_N} \end{bmatrix}$$

We also note that $\boldsymbol{s}_1 = [s_1, s_2, s_3, s_4]^T = [x_0, y_0, s_3, s_4]^T$, so that

$$\frac{d\mathbf{g}}{d\boldsymbol{s}_1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Let's denote $\mathbf{J_1} := \begin{bmatrix} \frac{\partial \lambda_{1_N}}{\partial s_{N-3}} & \frac{\partial \lambda_{1_N}}{\partial s_{N-2}} & \frac{\partial \lambda_{1_N}}{\partial s_{N-1}} & \frac{\partial \lambda_{1_N}}{\partial s_N} \\ \frac{\partial \lambda_{2_N}}{\partial s_{N-3}} & \frac{\partial \lambda_{2_N}}{\partial s_{N-2}} & \frac{\partial \lambda_{2_N}}{\partial s_{N-1}} & \frac{\partial \lambda_{2_N}}{\partial s_N} \end{bmatrix}$.

And then we can see now that our Jacobian so far looks like:

$$J = \begin{bmatrix} \begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix} & \begin{smallmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{smallmatrix} & \cdots & \begin{smallmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{smallmatrix} \\ \begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix} & \begin{smallmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{smallmatrix} & \cdots & \mathbf{J_1} \\ \frac{\partial \boldsymbol{p}_1}{\partial \boldsymbol{s}_1} - \frac{\partial \boldsymbol{s}_2}{\partial \boldsymbol{s}_1} & \frac{\partial \boldsymbol{p}_1}{\partial \boldsymbol{s}_2} - \frac{\partial \boldsymbol{s}_2}{\partial \boldsymbol{s}_2} & \cdots & \frac{\partial \boldsymbol{p}_1}{\partial \boldsymbol{s}_N} - \frac{\partial \boldsymbol{s}_2}{\partial \boldsymbol{s}_N} \\ \vdots & \ddots & \cdots & \vdots \\ \frac{\partial \boldsymbol{p}_{N-1}}{\partial \boldsymbol{s}_1} - \frac{\partial \boldsymbol{s}_N}{\partial \boldsymbol{s}_1} & \frac{\partial \boldsymbol{p}_{N-1}}{\partial \boldsymbol{s}_2} - \frac{\partial \boldsymbol{s}_N}{\partial \boldsymbol{s}_2} & \cdots & \frac{\partial \boldsymbol{p}_{N-1}}{\partial \boldsymbol{s}_N} - \frac{\partial \boldsymbol{s}_N}{\partial \boldsymbol{s}_N} \end{bmatrix}$$

Note that since the solution in the jth sub-interval is independent of the initial conditions in any other interval, and the initial conditions in the jth sub-interval are independent of the initial conditions in any other interval, we have: $\frac{\partial \boldsymbol{p}_j}{\partial \boldsymbol{s}_k} = 0$ and $\frac{\partial \boldsymbol{s}_j}{\partial \boldsymbol{s}_k} = 0$ for all $j \neq k, 1 \leq j, k \leq N$. Similarly, the initial conditions for $x, y, \lambda_1, \lambda_2$ in a given interval are independent of each other, so $\frac{\partial s_j}{\partial s_k} = 0$ for all $j \neq k, 5 \leq j, k \leq 4N$. Also, $\frac{\partial s_j}{\partial s_j} = 1$ for all $4 \leq j \leq 4N$. So along the diagonal beginning from the second row, we will have blocks of $-\mathbf{I}_4$,

where $\mathbf{I}_4$ is the identity matrix, $\mathbf{I}_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

So now we can see that so far our Jacobian looks like:

$$J = \begin{bmatrix} \begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix} & \begin{smallmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{smallmatrix} & \cdots & \cdots & \cdots & \begin{smallmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{smallmatrix} \\ \begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix} & \begin{smallmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{smallmatrix} & \begin{smallmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{smallmatrix} & \cdots & \begin{smallmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{smallmatrix} & \mathbf{J_1} \\ \frac{\partial \boldsymbol{p}_1}{\partial \boldsymbol{s}_1} & -\mathbf{I}_4 & \begin{smallmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{smallmatrix} & \cdots & \begin{smallmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{smallmatrix} & \begin{smallmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{smallmatrix} \\ \begin{smallmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{smallmatrix} & \frac{\partial \boldsymbol{p}_2}{\partial \boldsymbol{s}_2} & -\mathbf{I}_4 & \begin{smallmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{smallmatrix} & \cdots & \begin{smallmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{smallmatrix} \\ \vdots & \ddots & \cdots & \vdots & \ddots & \vdots \\ \begin{smallmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{smallmatrix} & \cdots & \cdots & \begin{smallmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{smallmatrix} & \frac{\partial \boldsymbol{p}_{N-1}}{\partial \boldsymbol{s}_{N-1}} & -\mathbf{I}_4 \end{bmatrix}$$

What's left is to define our unknown blocks, $\frac{\partial \boldsymbol{p}_1}{\partial \boldsymbol{s}_1}$ and $\frac{\partial \boldsymbol{p}_j}{\partial \boldsymbol{s}_j}$, for $2 \leq j \leq N-1$:

$$\frac{\partial \boldsymbol{p}_1}{\partial \boldsymbol{s}_1} = \begin{bmatrix} \frac{\partial x_1}{\partial s_3} & \frac{\partial x_1}{\partial s_4} \\ \frac{\partial y_1}{\partial s_3} & \frac{\partial y_1}{\partial s_4} \\ \frac{\partial \lambda_{1_1}}{\partial s_3} & \frac{\partial \lambda_{1_1}}{\partial s_4} \\ \frac{\partial \lambda_{2_1}}{\partial s_3} & \frac{\partial \lambda_{2_1}}{\partial s_4} \end{bmatrix} \text{ and } \frac{\partial \boldsymbol{p}_j}{\partial \boldsymbol{s}_j} = \begin{bmatrix} \frac{\partial x_j}{\partial s_{4j-7}} & \frac{\partial x_j}{\partial s_{4j-6}} & \frac{\partial x_j}{\partial s_{4j-5}} & \frac{\partial x_j}{\partial s_{4(j-1)}} \\ \frac{\partial y_j}{\partial s_{4j-7}} & \frac{\partial y_j}{\partial s_{4j-6}} & \frac{\partial y_j}{\partial s_{4j-5}} & \frac{\partial y_j}{\partial s_{4(j-1)}} \\ \frac{\partial \lambda_{1_j}}{\partial s_{4j-7}} & \frac{\partial \lambda_{1_j}}{\partial s_{4j-6}} & \frac{\partial \lambda_{1_j}}{\partial s_{4j-5}} & \frac{\partial \lambda_{1_j}}{\partial s_{4(j-1)}} \\ \frac{\partial \lambda_{2_j}}{\partial s_{4j-7}} & \frac{\partial \lambda_{2_j}}{\partial s_{4j-6}} & \frac{\partial \lambda_{2_j}}{\partial s_{4j-5}} & \frac{\partial \lambda_{2_j}}{\partial s_{4(j-1)}} \end{bmatrix}$$

Let's define $\mathbf{J}_2 := \frac{\partial \boldsymbol{p}_1}{\partial \boldsymbol{s}_1}$ and $\mathbf{J}_j := \frac{\partial \boldsymbol{p}_{j-1}}{\partial \boldsymbol{s}_{j-1}}$, for $3 \leq j \leq N$.

Recall $\left[\frac{\partial x}{\partial \boldsymbol{s}}, \frac{\partial y}{\partial \boldsymbol{s}}, \frac{\partial \lambda_1}{\partial \boldsymbol{s}}, \frac{\partial \lambda_2}{\partial \boldsymbol{s}}\right] = [w^1, w^2, w^3, w^4]^T = \boldsymbol{w}$, so for the jth sub-interval, we have $\left[\frac{\partial x}{\partial \boldsymbol{s}_j}, \frac{\partial y}{\partial \boldsymbol{s}_j}, \frac{\partial \lambda_1}{\partial \boldsymbol{s}_j}, \frac{\partial \lambda_2}{\partial \boldsymbol{s}_j}\right] = [w_j^1, w_j^2, w_j^3, w_j^4]^T = \boldsymbol{w}_j$. Therefore, the entries of $\mathbf{J}_j, 1 \leq j \leq N$ are found by solving the sensitivity equations on the interval $[t_{j-1}, t_j]$ under initial conditions $[1,0,0,0]^T, [0,1,0,0]^T, [0,0,1,0]^T, [0,0,0,1]^T$ when differentiating with respect to s_{4k-3}, s_{4k-2}, s_{4k-1}, and s_{4k} respectively, where $1 \leq k \leq N$, evaluated at the endpoint of the corresponding sub-interval.

Let's call the solutions to the sensitivity equations with these initial conditions as follows: $\boldsymbol{w}_1(t), \boldsymbol{w}_2(t), \boldsymbol{w}_3(t), \boldsymbol{w}_4(t)$.

Doing so, we see

$$\mathbf{J}_1 = \begin{bmatrix} w_1^3(T) & w_2^3(T) & w_3^3(T) & w_4^3(T) \\ w_1^4(T) & w_2^4(T) & w_3^4(T) & w_4^4(T) \end{bmatrix}, \tag{3.54}$$

$$\mathbf{J}_2 = \begin{bmatrix} w_3^1(t_1) & w_4^1(t_1) \\ w_3^2(t_1) & w_4^2(t_1) \\ w_3^3(t_1) & w_4^3(t_1) \\ w_3^4(t_1) & w_4^4(t_1) \end{bmatrix}, \tag{3.55}$$

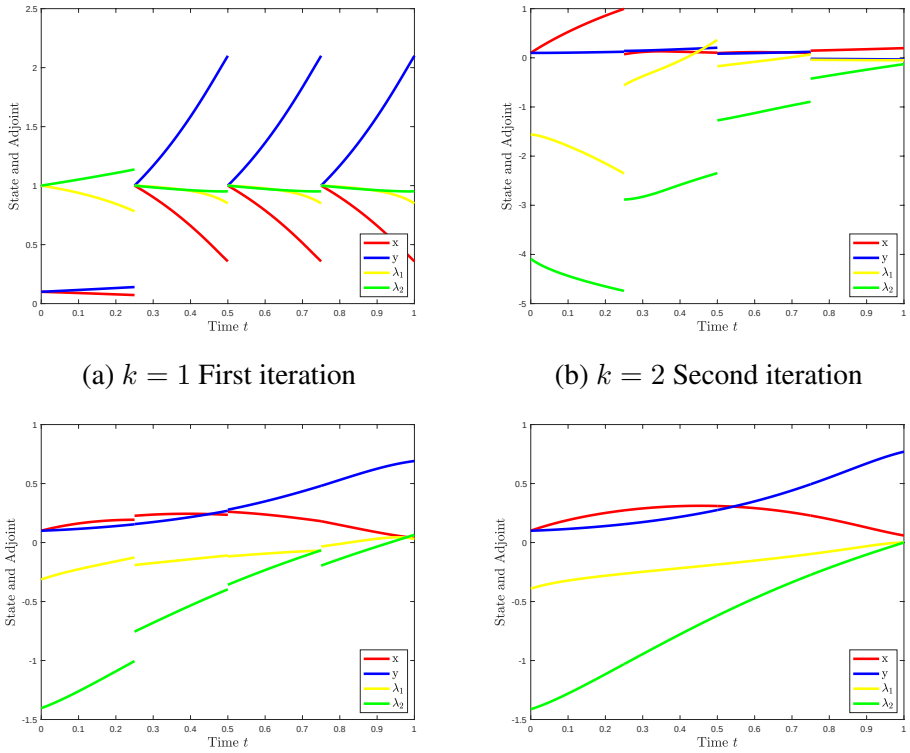
and for $3 \leq j \leq N$,

$$\mathbf{J}_j = \begin{bmatrix} w_1^1(t_{j-1}) & w_2^1(t_{j-1}) & w_3^1(t_{j-1}) & w_4^1(t_{j-1}) \\ w_1^2(t_{j-1}) & w_2^2(t_{j-1}) & w_3^2(t_{j-1}) & w_4^2(t_{j-1}) \\ w_1^3(t_{j-1}) & w_2^3(t_{j-1}) & w_3^3(t_{j-1}) & w_4^3(t_{j-1}) \\ w_1^4(t_{j-1}) & w_2^4(t_{j-1}) & w_3^4(t_{j-1}) & w_4^4(t_{j-1}) \end{bmatrix}. \tag{3.56}$$

Let $\mathbf{0}_{m\times n}$ be a matrix of dimension m rows by n columns with all entries $= 0$. Let $\boldsymbol{I}_n$ be the identity matrix of size $n \times n$. Then our Jacobian, J can be defined as follows:

$$J = \begin{bmatrix} \mathbf{I}_2 & \mathbf{0}_{2\times 4} & \dots & \dots & \dots & \dots & \mathbf{0}_{2\times 4} \\ \mathbf{0}_{2\times 2} & \mathbf{0}_{2\times 4} & \dots & \dots & \dots & \mathbf{0}_{2\times 4} & \mathbf{J}_1 \\ \mathbf{J}_2 & -\mathbf{I}_4 & \mathbf{0}_{4\times 4} & \dots & \dots & \dots & \mathbf{0}_{4\times 4} \\ \vdots & \dots & \ddots & \ddots & \ddots & \dots & \vdots \\ \mathbf{0}_{4\times 2} & \dots & \mathbf{J}_j & -\mathbf{I}_4 & \mathbf{0}_{4\times 4} & \dots & \mathbf{0}_{4\times 4} \\ \vdots & \dots & \ddots & \ddots & \ddots & \dots & \vdots \\ \mathbf{0}_{4\times 2} & \dots & \dots & \dots & \mathbf{0}_{4\times 4} & \mathbf{J}_N & -\mathbf{I}_4 \end{bmatrix} \tag{3.57}$$

Using this residual and Jacobian imposes patching conditions after each iteration and forces the solutions to converge to be continuous in addition to the previous need to satisfy the terminal conditions. For example, here is an example of the progression with $N = 4$ sub-intervals.



(a) $k = 1$ First iteration

(b) $k = 2$ Second iteration

(c) $k = 5$ Fifth iteration

(d) $k = 20$ Twentieth iteration

Figure 3.1: An example of multiple shooting method's convergence with parameter values given by $I = 2$, $\mu = 0.8$, $\alpha = 2$, $\beta = 2$, $m = 0.1$, $\nu = 0.5$, $c = 1$, and $u_{max} = 5$.

Let $\boldsymbol{s}_i^k$ be the initial conditions for the ith sub-interval on the kth iteration.We will denote by $J(\boldsymbol{s}^k)$ the Jacobian for the kth iteration. We then set desired tolerance $tol = \varepsilon$, number of sub-intervals $subs = N$, maximum desired number of iterations i_{max}. Then we execute the multiple shooting method by implementing the following algorithm.

Algorithm 2 Multiple Shooting Method

Set $N \in \mathbb{N}$, $\boldsymbol{s}^1 \in \mathbb{R}^{4N}$, $\varepsilon > 0$, and $i_{\max} \in \mathbb{N}$.
for $i = 1, \ldots, i_{\max}$ **do**
    Simultaneously solve (3.48) – (3.51) with initial conditions $\boldsymbol{s}^i = [\boldsymbol{s}_1^i, \boldsymbol{s}_2^i, \ldots, \boldsymbol{s}_N^i]$.
    Compute $\boldsymbol{r}(\boldsymbol{s}^i)$.
    if $||\boldsymbol{r}(\boldsymbol{s}^i)|| < \varepsilon$ **then**
        Stop. Use s_3^i, s_4^i as adjoint initial conditions.
    else
        Compute J_i as defined in (3.57).
        $\boldsymbol{s}^{i+1} = \boldsymbol{s}^i - J^{-1}(\boldsymbol{s}^i)\boldsymbol{r}(\boldsymbol{s}^i)$.
    end if
end for

3.4 Numerical Simulations

Here we illustrate some examples numerically justifying our results.

3.4.1 Example 1: Single Shooting Method

First, we show some simulations using the Single Shooting Method. Our initial conditions for the state variables are $x_0 = 0.1, y_0 = 0.1$. Our initial guess for the adjoint initial conditions is $\boldsymbol{s} = [-0.1, -3]^T$, and the method converges with a 8.944282×10^{-13} residual norm, ending with $\lambda_1(0) = -0.0245076454076917, \lambda_2(0) = -0.229827769589335$.

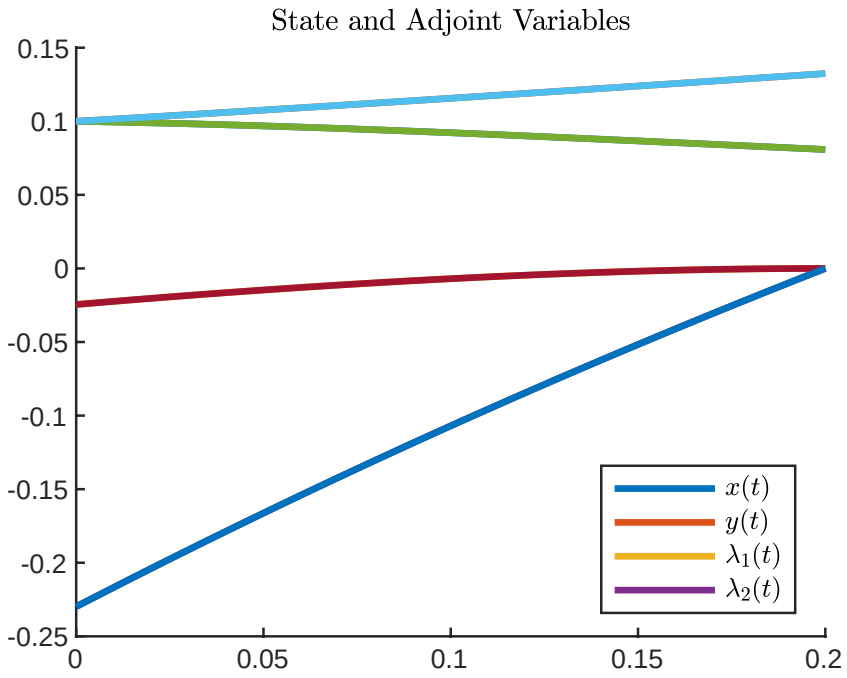
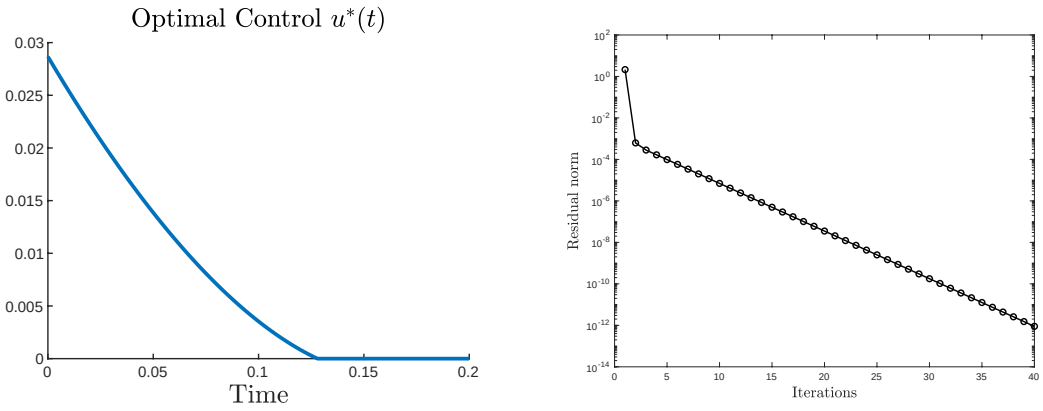


Figure 3.2: Plot of the joint dynamics of the nutrition level x and the bacterial population y, along with adjoint trajectories λ_1 and λ_2 with parameters $I = 2$, $\mu = 0.8$, $\alpha = 2$, $\beta = 2$, $m = 0.1$, $\nu = 0.5$, $c = 1$, and $u_{max} = 5$.



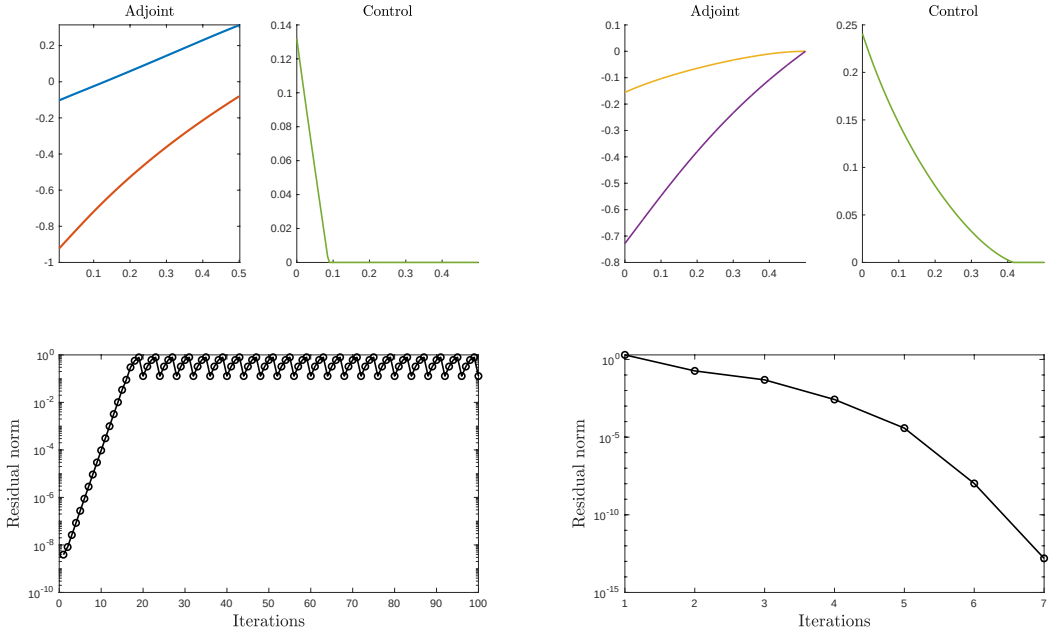
(a) Optimal Control $u^*(t)$

(b) Residual

Figure 3.3: Plots from the Single Shooting Method of the optimal control u and the norm of the residual $\boldsymbol{r}$ after 40 iterations with parameters $I = 2$, $\mu = 0.8$, $\alpha = 2$, $\beta = 2$, $m = 0.1$, $\nu = 0.5$, $c = 1$, and $u_{max} = 5$.

3.4.2 Example 2: Single Shooting Fails, Multiple Shooting Converges

Our next simulation illustrates an example where the Single Shooting Method fails to converge, but the Multiple Shooting Method does converge. For the Single Shooting method, we use initial guesses for the adjoint initial conditions $\boldsymbol{s} = [-0.155752318086413, -0.728016286737916]^T$ and the method fails to yield correct initial conditions for λ_1 and λ_2 to reach 0. Note that residual norm does not converge to 0, but actually begins to oscillate through the values $\{0.12837210010308, 0.323123408311658, 0.608671342895119, 0.793190350561888, 0.1283721001030640.323123408311807, 0.608671342895419\}$. In contrast, using the Multiple Shooting Method with $N = 10$ sub-intervals, the residual norm converges within only 7 iterations. We used $\boldsymbol{s}_j = [1, 1, 1, 1]^T$ for $1 \leq j \leq 10$ as the adjoint initial conditions on each sub-interval. The adjoint initial conditions produced by the algorithm are $\lambda_1(0) = -0.155752318086413, \lambda_2(0) = -0.728016286737916.$



(a) Single Shooting Method

(b) Multiple Shooting Method

Figure 3.4: A contrast between Single Shooting failure and Multiple Shooting convergence, both with time interval $t \in [0, .5]$ and parameters $I = 2$, $\mu = 0.8$, $\alpha = 2$, $\beta = 2$, $m = 0.1$, $\nu = 0.5$, $c = 1$, and $u_{max} = 5$.

3.4.3 Example 3: Multiple Shooting Method

Next we look at a simulation of the Multiple Shooting Method on a longer time interval, $[0, 1]$. Here we also plot the Hamiltonian to show it is conserved along the optimal u^* trajectory. We used $\boldsymbol{s}_j = [1, 1, 1, 1]^T$ for $1 \leq j \leq 10$ as the adjoint initial conditions on each sub-interval. The method converges after 8 iterations. The adjoint initial conditions produced by the algorithm are $\lambda_1(0) = -0.544639271150085, \lambda_2(0) = -2.77751948107959$.

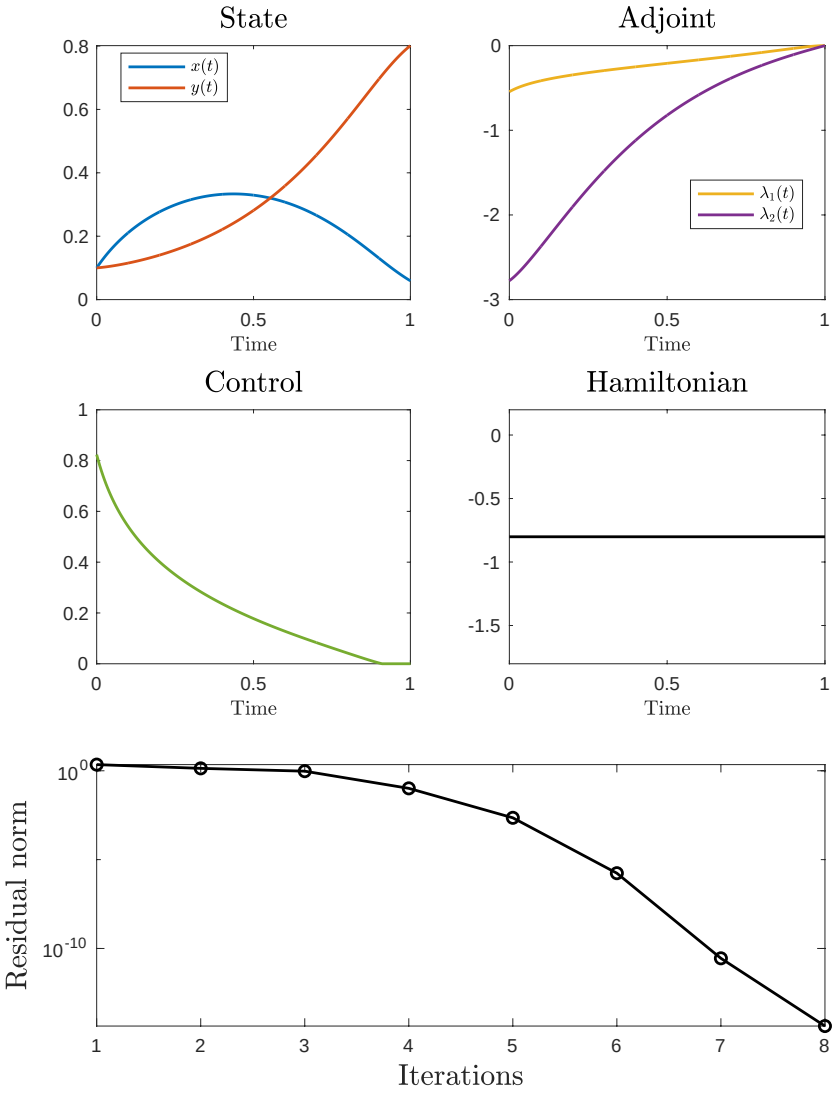


Figure 3.5: An example of Multiple Shooting Method with convergence with parameters $I = 2$, $\mu = 0.8$, $\alpha = 2$, $\beta = 2$, $m = 0.1$, $\nu = 0.5$, $c = 1$, and $u_{max} = 5$.

Next we contrast the cost yielded by the optimal $u^*(t)$ given above with a constant control $u(t) = \bar{u} \in \mathbb{R}$ where $\bar{u}$ is equal to the average value of $u^*(t)$ on $[0, T]$, i.e. $\bar{u} \equiv \frac{1}{T}\int_0^T u^*(t)dt$. We compute the cost functional, $J(u) = \int_0^T \big(- y(t) + \frac{c}{2}u^2(t)\big)dt$, for $u = \bar{u}$ and $u = u^*(t)$, and get $J(\bar{u}) = -0.261672237985602$, while $J(u^*(t)) = -0.752313227989741$. So we can see that the optimal control $u^*(t)$ minimizes the cost functional over a constant control with the same average value.

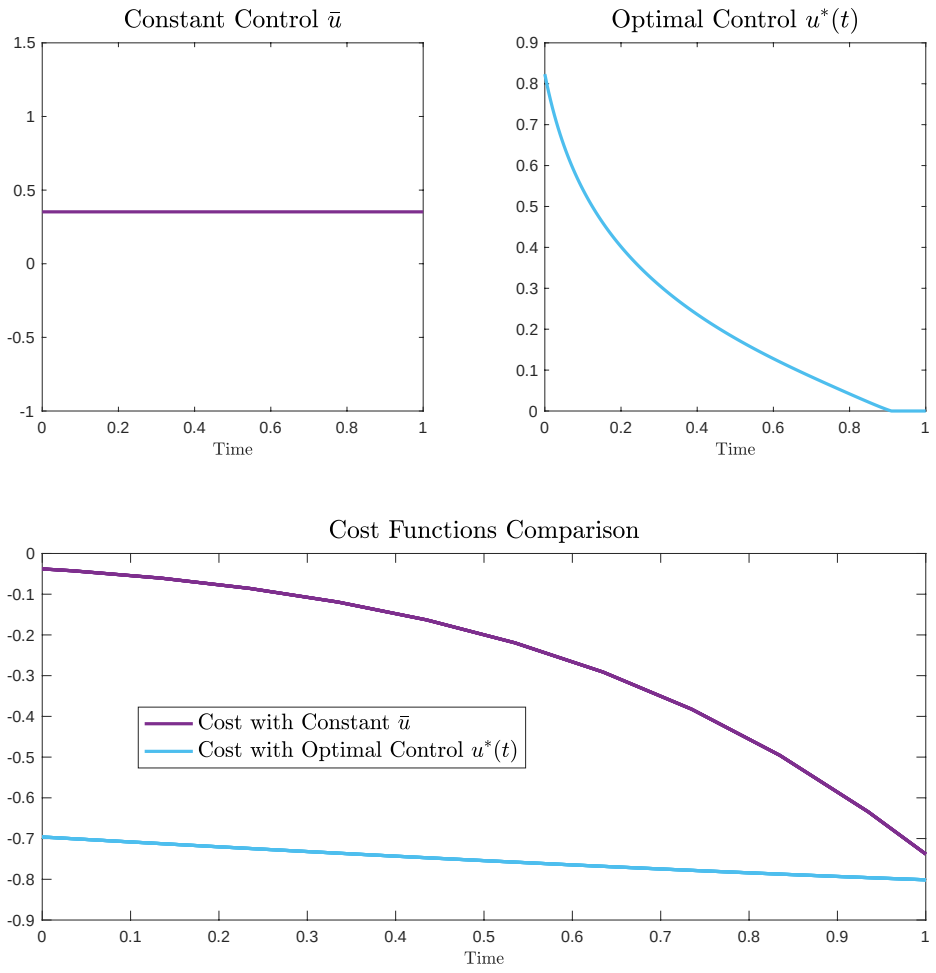


Figure 3.6: The control functions are plotted over $[0, 1]$ for $u^*(t)$ and $\bar{u} = \frac{1}{T}\int_0^T u^*(t)dt$ with parameters $I = 2$, $\mu = 0.8$, $\alpha = 2$, $\beta = 2$, $m = 0.1$, $\nu = 0.5$, $c = 1$, and $u_{max} = 5$. Below that, the Lagrangian cost functions $L(t, \boldsymbol{x}(t), u^*(t))$ and $L(t, \boldsymbol{x}(t), \bar{u})$ are plotted, showing the area under the curve of the optimal control is more negative than the area under the curve of the constant control.

Chapter 4

Conclusion and Future Work

In Chapter 2, a simple but novel model on the growth mechanisms of the gut microbiome was developed and studied. While in a simplistic setting, the model considers growth of the bacteria both on and inside the gut wall, intra-specific competition among bacterias in different groups, in addition to their growth due to consumption of nutrient and natural collective death. The highlight of this work is the consideration of variable nutrient inputs, that gives rise to a system of nonautonomous differential equations. Using theory and techniques of nonautonomous dynamical systems, sufficient conditions under which the bacteria dies out or persists are established and compared to those of the autonomous counterpart. In particular, it was both theoretically and numerically shown that time-dependent inputs may promote persistence of the bacteria. More precisely, with the same average nutrient input over a specific period of time, bacteria can persist when the inputs are time-dependent, but die out when the inputs are constant.

A natural extension of this problem would be to consider a larger model where we introduce the dynamics of another microorganism that can compete with our $Y(t)$. Another direction may be to model "leaky-gut syndrome," where the bacteria can not only attach to the wall but also break through the lining of the intestine.

Since Chapter 2 showed that time-dependent input of nutrient may result in asymptotic behavior different from constant input of nutrient, in Chapter 3 we investigated how the growth of a particular type of bacteria can be boosted or inhibited to maintain a healthy gut. In this chapter, a different model on the growth mechanisms of the gut microbiome was developed,

studied, and optimized. We used Pontryagin's Maximum Principle to find an optimal piecewise continuous control function, $u^*(t)$ for a given time interval $[0, T]$. We applied shooting methods to solve the boundary value problem that arises from PMP. Then we showed numerical simulations of the theoretical results that a time-varying input maximizes the growth of the beneficial microorganism over a constant input rate.

Further work on this problem will feature including the probiotic population. Another variation we are interested in is changing the optimality problem. In this dissertation, we considered one beneficial bacterial population and maximized the growth of this microorganism over a given time interval. Other interesting variations of the problem would be trying to eradicate a harmful bacteria efficiently while promoting persistence of the beneficial bacteria.

References

[1] G. D. ‘Ans, P. V. Kokotovic and D. Gottlieb, A nonlinear regulator problem for a model of biological waster treatment, *IEEE Transactions on Automatic Control* **AC-16** (1971), 341–347.

[2] D. Gresham and J. Hong, *The functional basis of adaptive evolution in chemostats.*,FEMS microbiology reviews, (2014), 39 1,2-16.

[3] F. Mazenc and Z.P. Jiang, *Persistence and time-varying nonlinear controllers for a chemostat with many species*, Dynamics of Continuous, Discrete and Impulse Systems, Series: A, 17, (2010), 737-763.

[4] J. W. M. La Riviere, Microbial ecology of liquid waste, *Advances in Microbial Ecology* **1** (1977), 215–259.

[5] H. R. Bungay and M. L. Bungay, Microbial interactions in continuous culture, *Advances in Applied Microbiology* **10** (1968), 269–290.

[6] A. Cunningham and R. M. Nisbet, Transients and oscillations in continuous cultures, *Mathematics in Microbiology*, Academic Press, New York (1983), 77–103.

[7] A. G. Fredrickson and G. Stephanopoulos, Microbial competition, *Science* **213** (1981), 972–979.

[8] H. W. Jannash and R. T. Mateles, Experimental bacterial ecology studies in continuous culture, *Advances in Microbial Physiology* **11** (1974), 165–212.

[9] P. A. Taylor and J. L. Williams, Theoretical studies on the coexistence of competing species under continuous flow conditions, *Canadian Journal of Microbiology* **21** (1975), 90–98.

[10] H. Veldcamp, Ecological studies with the chemostat, *Advances in Microbial Ecology* **1** (1977) 59–95.

[11] Y. Dong, Y. Takeuchi, and S. Nakaoka (2018). *A mathematical model of multiple delayed feedback control system of the gut microbiota—antibiotics injection controlled by measured metagenomic data*, Nonlinear Analysis. Real World Applications. An International Multidisciplinary Journal, 43, 1. (2018)

[12] P. Waltman *Competition Models in Population Biology*, Society for Industrial and Applied Mathematics, Philadelphia (1983).

[13] P. Waltman, S. P. Hubbel, and S. B. Hsu, Theoretical and experimental investigations of microbial competition in continuous culture, *Modeling and Differential Equations in Biology*, New York (1980) 107–152.

[14] R. Freter, Mechanisms that control the microflora in the large intestine, *in Human Intestinal Microflora in Health and Disease*, D. J. Hentges, ed., Academic Press, New York (1983), 33–54.

[15] R. Freter, An understanding of clonization of the large intestine requires mathematical analysis, *Microecology and Therapy* **16** (1986), 147–155.

[16] F. M. Williams, *Dynamics of microbial populations*. In Systems Analysis and Simulation in Ecology Vol. 1, pp. 197-265, B.O. Patten (Ed.), Academic Press, New York, 1971.

[17] T. Caraballo and X. Han, *Applied Nonautonomous and Random Dynamical Systems*, SpringerBriefs in Mathematics, Springer, (2016).

[18] T. Caraballo, X. Han, P. E. Kloeden, Chemostats with time-dependent input and wall growth, *Appl. Math. Inf. Sci.*, **9** (2015), 2283–2296.

[19] T. Caraballo, X. Han, P. E. Kloeden, Non-autonomous chemostats with variable delays, *SIAM Journal on Mathematical Analysis*, **47** (2015), 2178–2199.

[20] T. Caraballo, X. Han, P. E. Kloeden and A. Rapaport, Dynamics of non-autonomous chemostat models, *Continuous and Distributed Systems II* (eds. V. A. Sadovnichiy and M. Z. Zgurovsky), Spring-Verlag (2015), 103–120.

[21] F. Brauer and J.A. Nohel, *The Qualitative Theory of Ordinary Differential Equations, An Introduction*, Dover Publications (1989).

[22] J.H. Liu, *A First Course in the Qualitative Theory of Differential Equations*, Prentice Hall, Pearson Education, (2003).

[23] I. Cho and M. J. Blaser, The human microbiome: at the interface of health and disease, *Nat. Rev. Genet.*, **13** (2012), 260–270.

[24] A. V. Contreras, B. Cocom-Chan, G. Hernandez-Montes, T. Portillo-Bobadilla and O. Resendis-Antonio, Host-microbiome interaction and cancer: potential application in precision medicine, *Front. Physiol.*, **7** (2016), 606.

[25] A. Curry, Piles of ancient poop reavel 'extinction event' in human gut bacteria, Science AAAS (2021).

[26] J. J. Farrell *et al.*, Variations of oral microbiota are associated with pancreatic diseases including pancreatic cancer, *Gut*, **61** (2011), 582–588.

[27] T. Gibson and G. Gerber, Robust and Scalable Models of Microbiome Dynamics for Bacteriotherapy Design, (2018).

[28] J. Gilbert *et al.*, Microbiome-wide association studies link dynamic microbial consortia to disease, *Nature*, **535** (2016), 94–103.

[29] A. Gonzalez, J. Stombaugh, C. Lozupone, P. J. Turnbaugh, J. I. Gordon and R. Knight, The mind-body-microbial continuum, *Dialogues Clin. Neurosci.*, **13** (2011), 55–62.

[30] A. L. Gould *et al.*, Microbiome interactions shape host fitness, *PNAS*, **115** (2018) E11951–E11960.

[31] J. Halfvarson *et al.*, Dynamics of the human gut microbiome in inflammatory bowel disease, *Nat. Microbiol.*, **2** (2017), 1–7.

[32] L. V. Hooper, M. H. Wong, A. Thelin, L. Hansson, P. G. Falk and J. I. Gordon, Molecular analysis of commensal host-microbial relationships in the intestine, *Science*, **291** (2001), 881–884.

[33] L. V. Hooper, D. R. Littman and A. J. Macpherson, Interactions between the microbiota and the immune system, *Science*, **336** (2012), 1268–1273.

[34] S. Huitzil, S. Sandoval-Motta, A. Frank and M. Aldana, Modeling the role of the microbiome in evolution, *Front. Physiol* (2018).

[35] P.E. Kloeden and T. Lorenz, Mean-square random dynamical systems, *J.Differential Equations*, **253** (2012), 1422–1438.

[36] P.E. Kloeden and T. Lorenz, Pullback incremental attraction, *Nonautonomous & Random Dynamical Systems*, (2013), 53–60.

[37] P.E. Kloeden and M. Rasmussen, *Nonautonomous Dynamical Systems*, American Mathematical Society Providence, RI, 2011.

[38] P. E. Kloeden and M. Yang, *An Introduction to Nonautonomous Dynamical Systems and their Applications*, World Scientific, Singapore, 2020.

[39] T. Bayen, J. Harmand, M. Sebbah. Time-optimal control of concentrations changes in the chemostat with one single species. Applied Mathematical Modelling, Elsevier, 50, (2017), pp.257- 278.

[40] T. Bayen, A. Rapaport, F.-Z. Tani, Optimal periodic control of the chemostat with Contois growth function. IFAC PapersOnLine 51-2 (2018) 730–734.

[41] R. E. Ley, P. J. Turnbaugh, S. Klein and J. I. Gordon, Microbial ecology: human gut microbes associated with obesity, *Nature*, **444** (2006), 1022.

[42] R. E. Ley, Obesity and the human microbiome, *Curr. Opin. Gastroenterol*, **26** (2010), 5–11.

[43] X. C. Morgan *et al.*, Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment, *Genome Biol.*, **13** (2012) R79.

[44] V. S. H. Rao and P. R. S. Rao, *Dynamical models and control of biological Systems*, Springer-Verlag, Berlin (2009).

[45] G. Rogers, D. Keating, R. Young, M. Wong, J. Licinio and S. Wesselingh, From gut dysbiosis to altered brain function and mental illness: mechanisms and pathways, *Mole. Physchiatry*, **21** (2016) 738.

[46] H. L Smith and P. Waltman, *The Theory of the Chemostat: Dynamics of Microbial Competition*, Cambridge University Press, Cambridge, UK (1995).

[47] R. R. Stein, V. Bucci, N. C. Toussaint, C. G. Buffie, G. Rätsch, E. G. Pamer, C. Sander and J. B. Xavier, Ecological modeling from time-series inference: insight into dynamics and stability of intestinal microbiota, *PLOS Computational Biology*, **9** (2013) 31003388.

[48] H.J. Pesch, and M. Plail, The Cold War and the maximum principle of optimal control. *Documenta Mathematica* (2012).

[49] H.B. Keller, *Numerical Solution of Two Point Boundary Value Problems,* CSMS-NSF, Regional Conference Series in Applied Mathematics. (1976).

[50] H. Schattler and U. Ledzewicz, *Geometric Optimal Control: Theory, Methods and Examples*. Interdisciplinary Applied Mathematics **38**. Springer (2010).

[51] E. A. Coddington and N. Levinson, *Theory of ordinary differential equations* International series in pure and applied mathematics. McGraw-Hill (1955)

[52] J. Erdmann, How gut bacteria could boost cancer treatments *Nature* ISSN 1476-4687 (online). (2022).

[53] L. Geddes, Certain gut microbes may affect stroke risk and severity, scientists find, *The Guardian* (2022).

[54] R. Ley, The human microbiome: there is much left to do, *Nature*, **606** (2022), 435.

[55] J. Kinross, Ready for your crapsule? Faecal transplats could play a huge role in future medicine, *The Guardian*, Health, (2011).