ABSTRACT

ATTARIAN, ADAM R. Patient Specific Subset Selection, Estimation and Validation of an HIV-1 Model with Censored Observations under an Optimal Treatment Schedule. (Under the direction of Hien T. Tran.)

This work centers around several analyses of an in-vivo human immunodeficiency virus (HIV)-1 dynamic model. We first consider the model in an inverse problem context, performing model validation and verification using clinically collected data. After calibration, we then investigate the optimal control problem of designing patient specific treatment schedules for the administration of drug therapy.

Within the model calibration problem, we perform a sensitivity analysis to determine which parameters contribute most to model output. Because sensitivity may not constitute identifiability, we explore different methods of computing identifiable subsets of parameters, including QR and sensitivity based methods before settling on a subspace selection algorithm to find the identifiable parameters. The model validation problem is complicated by the presence of censored clinical data due to quantification methods. Statistical methods are employed to estimate the truth observations so that the parameter estimates can be further refined. The inverse problem is in turn solved on a patient specific basis, without the use of population averages.

Finally, we consider the construction of patient specific treatment strategies using optimal control methodologies, including the installation of a feedback control mechanism through receding horizon control. We discuss structured treatment interruptions in the context of the optimal control and demonstrate how customized treatment protocols can be designed through the variation of control parameters on a patient specific basis.

Patient Specific Subset Selection, Estimation and Validation of an HIV-1 Model with Censored Observations under an Optimal Treatment Schedule

by
Adam R. Attarian

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

Applied Mathematics

Raleigh, North Carolina

2012

APPROVED BY:

Stephen Campbell	John Harlim
Donald Bitzer	Hien T. Tran Chair of Advisory Committee

DEDICATION

To my parents, who taught me to never stop paddling.

BIOGRAPHY

Adam R. Attarian was born 19 March, 1984 in Durham, North Carolina. After spending a few years in Oregon as a child, he grew up in Raleigh, North Carolina, where he went to Enloe High School. Adam enrolled at North Carolina State University initially in political science, before coming to his senses and switching to electrical engineering before a single lecture.

After a seminal trip to San Francisco when he was 18, Adam returned wanting nothing more than to move to San Francisco and go to film and art school. Clearly lacking foresight and perspective, this became an overwhelming theme through his freshman and sophomore year. To compromise with himself and his parents, he became the first engineering student to dual enroll at the Center for Documentary Studies at Duke University in Durham. Here he earned the certificate in documentary studies, having created photography portfolios and documentary videos that screened at several film festivals.

In his junior year, Adam saw the guiding light of mathematics and realized how it can be used to model and explain the world around us. Wanting more, he added applied mathematics as a second major and continued on to graduate with Bachelors of Science in both Electrical Engineering and Applied Mathematics in Spring of 2007. He continued on at NC State for graduate school, ultimately receiving his M.S and Ph.D in applied mathematics, but not before completing an REU and two summer internships at MIT Lincoln Labs in 2008 and 2010.

Throughout college and grad school, Adam kept himself busy by working in the Raleigh food and beverage industry. This provided an education unto itself, learning how to talk to anyone and how to make and keep people happy even while under duress. In the final years of his graduate work, he found a love and passion for enology, ultimately being elevated to wine buyer and sommelier at his little culinary outpost on Hillsborough street.

As Adam pauses to reflect on the last ten years of academic engagement, he very much looks forward to leaving it behind. He'll be heading up north to work at MIT Lincoln Labs as a full time technical staff member. He anxiously awaits to see what is next.

ACKNOWLEDGEMENTS

What a ride this has been. If you were to have told me as a freshman that in just about eight years I would have been taking 800-level mathematics classes, I would have told you that you were crazy; I barely made it out of my 100 level classes. Yet here I sit, looking back and asking myself, "How did I get here?". Somehow, I have made it to the other side and lived to tell the tale. That somehow is, of course, the support and generosity of my parents, friends, and professors along the way. You have all allowed me to maintain my sanity (at varying degrees of successfulness) through the last few years, and for that I am grateful.

Writing these acknowledgments wouldn't be complete without talking about Karen, Jim, and Susan. The three of you have become my closest friends over the last few years, and I would not be the person I am today without you all in my life. From getting through Campbell's control exams, to surviving the qualifier summer (remember that Thursday at Raleigh Times?), to off campus lunch day, you all have had a profound impact on me. Without a doubt, you guys have been the brightest part about the last 5 years; you have been my retreat. I'm sure that when we meet again in the future it will be the same as it ever was.

Finally, I would like to thank my committee members for agreeing to give their time to facilitate the completion of this degree. This entire journey couldn't have been made without my advisor, Hien T. Tran, whose guidance has helped shape me into the researcher that I am today and will become tomorrow. Perhaps I'll run into him at an international conference sometime soon.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Introduction

There are few diseases in our society today that are more polarizing, politicized, and personal than human immunodeficiency virus/acquired immune deficiency syndrome, or HIV/AIDS. The story of HIV/AIDS in America is a story that touches on race, social class, sexuality, morality, and the political climate of the 1980's through today. In the nearly 30 years since being identified, tens of millions have died, and millions more are infected every year. It is a disease that seemingly came from nowhere; before we knew what it was, it was already with us. Everyone was taken off guard, and the scientific community has been playing catch up ever since.

According to UNAIDS, there were 34 million people living with HIV/AIDS in 2010, 22.9 million of which live in sub-Saharan Africa. Advances in treatment protocols have seen annual HIV infections decrease from nearly 3.1 million new infections in 2001 to 2.7 million new infections in 2010 (WHO, 2011). The first treatment, AZT (zidovudine) wasn't released until 1987, nearly six years after the effects of HIV were first being identified. Tolerance of this drug was mixed at best throughout the infected community, and the first combination therapy that would go on to constitute the new standard of care was not introduced until 1996. The effect

of these drugs were so dramatic for the chronically infected that it was soon referred to as the "Lazarus Syndrome" (Andriote, 1999).

Despite the prevalence and availability of effective treatment of HIV, even as late as 2010 the best practices for management of acute HIV infection are still unknown (Bell et al., 2010). The time to initiate treatment, and the manner of treatment are still the subject of many ongoing clinical trials that are investigating the treatment benefits versus risks.

Throughout the last few decades, mathematical models have been developed to further the understanding of the progression of the virus as it replicates in-host. These models focus on the interaction between the virus and the $CD4^+$ cells of the immune system, one of the virus' primary targets (Perelson et al., 1996; Perelson and Nelson, 1999; Callaway and Perelson, 2002). The early models of progression were critical in the understanding of the virus and the development of early treatments, as the science and understanding were lagging behind the impact of the spreading epidemic.

The models of in-vivo dynamics are most often formulated as a system of nonlinear ordinary differential equations, though recently several size structured partial differential equation models have been developed. These models allow for variations in viral production rates as a function of time, since it is thought that as infected cells age their overall fitness decreases (Nelson et al., 2004; Rong et al., 2007). The ordinary differential equation models that we utilize are based on a compartmental and mass-balance analysis of the biological system. These models are developed to incorporate as much relevant physiological principles as possible, though at the cost of increased model complexity.

A primary goal of using modern models of HIV is to influence treatment decisions and construct better treatment protocols for infected patients. The construction of optimal control methodologies has been investigated by many authors, (Brandt and Chen, 2001; Joshi, 2002; Adams et al., 2004; Ge et al., 2005; David et al., 2008), however in almost all of these cases the authors worked with simulated data and were not patient specific in nature. To use these dynamic models in a patient specific capacity the models must first be calibrated to each

individual patient. While much attention has been paid to the forward simulation problem over the years, less attention has been focused on the so-called inverse problem. Inverse problems are formulated when we wish to determine the parameters that characterize the system using measurements. One of the main obstacles in solving the inverse problem is that biological and physiological models are often nonlinear and contain many parameters. Additionally, the data for validation is often sparse, or possibly representative of only a portion of the model. Before parameter estimation methods can be applied to an ODE model, an investigation into the identifiability of the parameters, given available data, must be performed.

The investigation into parameter identifiability is not unique to the mathematical sciences, and many papers in engineering, statistics, and biomedical engineering have shed light into this area (Burth et al., 1999; Audoly et al., 2001). Identifiability in an HIV model context has been examined in (Xia and Moog, 2003; Xia, 2003) where structural and observer perspectives were used with a less complex model and in (Guedj et al., 2007), where the methodologies were heavily statistical in nature. In this work we study the parameter identifiability of a seven state, twenty two parameter HIV-1 model based on clinically obtained measurements of both $CD4^+$ count and viral load. We utilize both sensitivity analysis and identifiability analysis methods to compute patient specific best-estimatable subsets of parameters. In (Adams et al., 2007) an identifiability analysis was conducted on an identical model, however the authors used a different, ad-hoc subset selection procedure.

An additional component of this work is the handling of clinical data that are censored below. To obtain robust parameter estimates, the censored data must be estimated in a reliable manner, rather than ignored which is commonplace. The data used in validation consists of both $CD4^+$ count as well as viral load statistics collected through clinical trials at Massachusetts General Hospital. Specifically, the viral load measurements have a lower limit of quantification, and so these data have a lower threshold value. For robust estimates, the censored data can not be ignored. This data set was first analyzed in (Adams, 2005) where the methodology for accounting for the censored viral load measurements was developed.

Finally in this work we investigate the construction of optimal control and treatment therapies in a patient specific context. We develop treatments known as structured treatment interruptions, wherein the effective treatment is terminated at specified times and then restarted in an attempt to increase an in-host immune response and provide a drug holiday for the patient. These treatment methods have been studied in detail, though a lack of double-blind, randomized clinical trials has lead to a lack of consensus on the effectiveness of these types of treatments (Lisziewicz and Lori, 2002; Lisziewicz et al., 1999; Lori et al., 2000; Lori and Lisziewicz, 2001).

This manuscript is organized as follows. In chapter 2 we introduce the biology of HIV and the different avenues of treatment, and then describe a variety of HIV models before offering a detailed analysis of the model used throughout this work. We also describe the data that has been collected by our collaborators at Massachusetts General Hospital, noting the censored nature of the viral load observations and their impact on the inverse problem methodology. Chapter 3 describes the methodologies for model validation and calibration, wherein we discuss the process for solving the inverse problem. Included here is a discussion of sensitivity analysis and subset selection, the two procedures for determining which parameters can be estimated from the clinical data. In chapter 4 the results from the model validation are presented, and we show how the model can well-fit both the $CD4^+$ count and viral load. We also list the patient specific identifiable subsets and note the similarities in subsets between each patient. The control problem is considered in chapter 5, where background on both linear and nonlinear feedback control is provided before offering the control formulation used in this work to compute treatment strategies of an STI type. We provide computed optimal control simulations along side the clinical data and discuss the differences. Finally, we look forward in chapter 6 and discuss where this research can be taken next.

CHAPTER 2

Model and Data Discussion

Models consisting of systems of ordinary differential equations have been widely used to model physical and biological phenomena. Recently, ODE models have played a significant role in modeling the in-host dynamics and epidemics of infectious diseases, including HIV, hepatitis, and influenza. These models are used to investigate the puzzling features of infection as well as model different treatment protocols. Models of HIV have changed and shifted the way that HIV is treated and thought of. It was through modeling that it was discovered HIV replication occurs on the order of 10^{10} viral particles per day, leading to the realization that HIV evolution and mutation occurs so rapidly that a monothearpy was doomed to fail (Perelson and Nelson, 1999). Modeling has also illustrated that HIV occurs on multiple time scales, from hours to weeks; the half life of HIV in plasma is on the order of minutes to days (Brandt and Chen, 2001). This was in contrast to the original prevailing paradigm that because AIDS can take on average 10 years to develop that all of the underlying dynamics were slow. In this chapter we provide a brief history of HIV modeling before analyzing the model that we have chosen to work with.

2.1 HIV Biology and Drug Therapy

HIV is an RNA based virus, part of a larger class of viruses called retroviruses. The primary target of HIV are the $CD4^+$ cells in the immune system, though other classes of cells are also infected and contribute to viral progression. The $CD4^+$ cells are those that identify foreign bodies or antigens; they have no cytotoxic ability on their own. When an HIV particle infects a cell, the enzyme reverse transcriptase (RT), which HIV carries, makes a DNA copy of the HIV RNA genome. This DNA copy is then integrated into the cellular DNA through the enzyme integrase. This new DNA is then reproduced in every new generation of the cell, and so the cell remains infected for life. The viral DNA is termed provirus, and can remain latent for months, even years showing no evidence of its presence. While we do not model these latently infected cells, a model that incorporates the so called "latent reservoir" is given in (Rong and Perelson, 2009). When new virus is being formed, the viral DNA is read and viral RNA is created. The new viral RNA must be cut into different parts by way of an enzyme called protease.

Current drug therapies operate by inhibiting some combination of the actions of protease, integrase, and reverse transcriptase. If RT is inhibited, then HIV will be able to enter a cell but not infect the cell. The conversation of RNA to DNA will not occur, and the cell will not produce any new HIV particles. The introduced viral RNA is unstable, and will ultimately degrade and be cleared from the cell. If protease is inhibited, then the viral copies that are produced by the infected $CD4^+$ cell will not be infectious. Insights from research and mathematical modeling have led to the prevailing paradigm of HIV treatment consisting of highly active antiretroviral therapy (HAART). In HAART treatment two or more anti-retroviral agents are taken in combination by the patient upwards up three times per day. It wasn't until 1996 that these drugs were prescribed in combination, and truly marked a turning point in the AIDS epidemic. There are currently six available combination so called "fixed-dose combinations" where the individual drugs are combined into a single pill, alleviating some of the burdens of HAART treatment. These drugs are broadly classified by the phase of the HIV life-cycle that the drug inhibits. Some examples of these drugs include entry inhibitors, reverse transcriptase inhibitors (RTI), protease inhibitors

(PI), and integrase inhibitors, among others. In our work, we will model both RTIs and PIs. A more detailed description of treatment in the context of our model is given in section 2.3.

2.2 Survey of Prior and Existing Mathematical Models

In modeling HIV, or any other complex biological system a modeler must first decide which biologic processes are suitable and able to be modeled. A survey of the development of en-vivo mathematical models is best presented in (Perelson et al., 1993), which we follow here. When the HIV epidemic first started, there were no mathematical models in place, and both clinicians and mathematical biologists were constantly refining their modeling techniques. The data showed that under effective anti-retroviral therapy the viral load would decrease exponentially. This suggested a simple model of the form

$$\dot{V} = P - cV, \tag{2.2.1}$$

where the viral production $V(t)$ is governed by some function $P(t)$, and is cleared at rate c. If the therapy completely blocked viral production, then $P(t) = 0$ and V will fall exponentially, i.e. $V(t) = V_0 e^{-ct}$. To account for viral production and inter-cellular infection, an uninfected population T is introduced. Despite the population dynamics of T cells not being well understood, a reasonable model is given by

$$\dot{T} = s + pT\left(1 - \frac{T}{T_{max}}\right) - d_T T, \tag{2.2.2}$$

where s is the rate at which new T cells are produced, and have a natural death rate of d_T. The generation of T cells are governed by a logistic function in which p governs the production rate up to the saturation point T_{max}, where proliferation shuts off. The simplest and most common method to model infection is through mass-action dynamics, where the rate of infection is given as kVT, with k being the rate of infection. With the addition of the mass-action term, a simple

model of HIV infection is given by

$$\dot{T} = s + pT\left(1 - \frac{T}{T_{max}}\right) - d_T T - kVT \tag{2.2.3}$$

$$\dot{T}^* = kVT - \delta T^* \tag{2.2.4}$$

$$\dot{V} = N\delta T^* - cV. \tag{2.2.5}$$

A full stability analysis of this model is given in (Perelson et al., 1993), including the identification and discussion of the associated steady states. This is a simple model that does not include treatment terms (addressed later in (Perelson et al., 1993), and in the next section), nor immune response. The dynamics represented here are what most other ODE models are based upon, and eventually extend.

The simple nonlinear compartment model is extended further in (Bailey et al., 1992; Perelson et al., 1993). For an excellent review of the various types of models developed, see (Callaway and Perelson, 2002; Culshaw et al., 2004; Perelson and Nelson, 1999). More recently several size structured partial differential equation models have been developed (Rong et al., 2007; Nelson et al., 2004), based on the idea that viral fitness is dependent on age of the viral particles. There have also been delay differential equation models developed (Culshaw and Ruan, 2000; Nelson and Perelson, 2002) to account for the delay between initial infection and viral reproduction through the process of reverse transcription, integration, and production of new viral particles.

An over-arching focus in the development of recent models has been to investigate the lag between initial infection and the onset of AIDS. The reason for this is still unknown, but appears to be tied to the number of $CD4^+$ cells in the patient. To this end, models have been developed to describe the immune system, the affecting dynamics of HIV, and the corresponding decline in $CD4^+$ cells. These have included a wealth of deterministic models as well as stochastic models (Tan and Wu, 1998; Yuan and Allen, 2011); the latter able model very well early infection dynamics when there are few viral particles and infected cells, or when population statistics are required.

The first HIV models were aimed on the $CD4^+$ dynamics as the virus progressed, and focused on explaining the $CD4^+$ decline. Some models incorporated treatment strategies, but were limited to monotherapy RTI treatments as these were the first drugs developed. Additionally the early models were linear approximations to the nonlinear cell decay (Wei et al., 1995; Ho et al., 1995; Perelson et al., 1996) and couldn't capture the biological realities beyond short time initial time periods. To model longer term behavior, especially long term viral decay, a nonlinear model is needed (Bonhoeffer et al., 1997).

The models described so far have not included any in-host HIV activated immune response. From early on in clinical trials and observations from the medical community it was noticed that individuals that retain a high level of CD8 CTLs remained healthier longer (Carr et al., 1996), and this motivated the development of models that included a stimulated immune response. Early models (Menezes Campello de Souza, 1999) for CD8 populations were based on very simple predator prey relationships which were mathematically convenient but couldn't accurately produce observed dynamics. The immune response compartment is expanded in (Adams et al., 2005), which constructs the immune effector state based on comments and observations made in (Bonhoeffer et al., 2000), utilizing Michaelis-Menten nonlinearity saturations and a non-zero steady state. Other immune response models such as those presented in (Wodarz, 2001; Wodarz and Nowak, 1999) model a more complicated process, tracking not only lysing CTL cells, but also their immature precursor cells that are incapable of lysing.

2.3 Model Development and Analysis

For the purposes of this work, we settle on a model developed in (Adams et al., 2005), and then later modified in (Adams, 2005) and elsewhere. The dynamics of the HIV model are described by the system of nonlinear ordinary differential equations

$$\dot{T}_1 = \lambda_1 - d_1 T_1 - (1-\varepsilon_1) k_1 V_I T_1 \tag{2.3.1a}$$

$$\dot{T}_2 = \lambda_2 - d_2 T_2 - (1-f\varepsilon_1) k_2 V_I T_2 \tag{2.3.1b}$$

$$\dot{T}_1^* = (1-\varepsilon_1) k_1 V_I T_1 - \delta T_1^* - m_1 T_1^* E \tag{2.3.1c}$$

$$\dot{T}_2^* = (1-f\varepsilon_1) k_2 V_I T_2 - \delta T_2^* - m_2 T_2^* E \tag{2.3.1d}$$

$$\dot{V}_I = (1-\varepsilon_2) N_T \delta (T_1^* + T_2^*) - \big(c + (1-\varepsilon_1)\rho_1 k_1 T_1 + (1-f\varepsilon_1)\rho_2 k_2 T_2\big) V_I \tag{2.3.1e}$$

$$\dot{V}_{NI} = \varepsilon_2 N_T \delta (T_1^* + T_2^*) - c V_{NI} \tag{2.3.1f}$$

$$\dot{E} = \lambda_E + b_E \frac{T_1^* + T_2^*}{T_1^* + T_2^* + K_b} E - d_E \frac{T_1^* + T_2^*}{T_1^* + T_2^* + K_d} E - \delta_E E, \tag{2.3.1g}$$

with an initial condition vector

$$\big[T_1(0)\ T_2(0)\ T_1^*(0)\ T_2^*(0)\ V_I(0)\ V_{NI}(0)\ E(0)\big]^\top.$$

The model compartments are denoted by T_1 (type 1 target cells, e.g. $CD4^+$ T-cells, cells/μl blood), T_2 (type 2 target cells, e.g. macrophages, cells/μl blood), V_I (infectious free viron particles, RNA copies/μl blood), V_{NI} (non-infectious free viron particles, RNA copies/μl blood) and E (CD8 cytotoxic T-lymphocytes (CTL), cells/μl blood). Infection occurs readily in a well-mixed environment through interaction between free viral particles and uninfected target cells. We denote the infected cell populations with asterisks, e.g. T_1^* and T_2^*. A summary of the model states are given in Table 2.1.

The two target cell populations have possibly different source rates λ_1, λ_2 with natural death rates d_1, d_2. The populations are infected at different rates k_1, k_2, and is proportional to the amount of virus in the system. Having two different infection rates could account for the different activation rates between $CD4^+$ population and macrophage population. The presence of a second target cell type, T_2 satisfies a modeling requirement suggested in (Callaway and Perelson, 2002) and others: a reasonable HIV model predicts a non-zero steady-state viral load even in the

Table 2.1: Summary of state variables used in the HIV model.

Variable	Description	Units
T_1	Uninfected type 1 target cells (e.g $CD4^+$ T-cells)	$\frac{\text{cells}}{\mu\text{l}}$
T_2	Uninfected type 2 target cells (e.g macrophages)	$\frac{\text{cells}}{\mu\text{l}}$
T_1^*	Infected type 1 target cells	$\frac{\text{cells}}{\mu\text{l}}$
T_2^*	Infected type 2 target cells	$\frac{\text{cells}}{\mu\text{l}}$
V_I	Infectious free virus	$\frac{\text{RNA copies}}{\text{ml}}$
V_{NI}	Non-infectious free virus	$\frac{\text{RNA copies}}{\text{ml}}$
E	Cytotoxic T-lymphocytes (e.g CD8 cells)	$\frac{\text{cells}}{\mu\text{l}}$

presence of effective therapy. Because of this, even in the best case with ideal treatment one can not expect the viral load to be driven to zero, only to be reduced significantly. Infected cells are removed either by natural death or by a cell-mediated immune response E, described below.

The original Callaway-Perelson model describes several states that we do not include in our model. These states include a quiescent cell population, chronically infected cells, and a state representing follicular dendritic cells, which reside in secondary lymphoid organs and have been found to impede HIV clearance. We omit these model states, as the essential qualitative behaviors are preserved without these extra states. The important features of the model, such as a low steady state viral dominated equilibrium and high viral load sensitivities to treatment are retained even when the other states are removed.

During the primary infection of HIV, large numbers of cells are affected, leading to high concentrations of free viral particles in blood plasma as well as other cells. These concentrations can be as high as 10^7 copies/ml. Free virus particles are produced by both types of infected cells at a rate that we assume to be matched, though this could easily be generalized. In the original Callaway-Perelson model, virus only leaves the V compartment through natural death at rate c. In our model, we assume virus also is removed from V due to infection of the cell populations. The lack of this clearance mechanism is justified in (Perelson and Nelson, 1999): in a normal HIV patient the term $k_i T_i V$ is small compared to cV. They further state that if T_i is

near constant, then cV accounts for nearly all of the viral clearance from V.

A second important aspect of the model is the immune response state E. This HIV specific immune response consists of CTLs lysing antigens to HIV infected cells, killing them at rates m_1 and m_2 for each class of target cell. The immune response model used in this work does not directly clear viral particles, and so there is no interaction between the immune compartment and the virus compartments. In (Banks et al., 2008) the authors extended the model by adding a second immune effector state that represents CTL cells with memory. While that model incorporates a greater degree of HIV dynamics, it has been shown in (Adams et al., 2007) that the model used in this work is sufficient for clinical data fitting and model prediction. A systems level overview of the model is given in Figure 2.1.

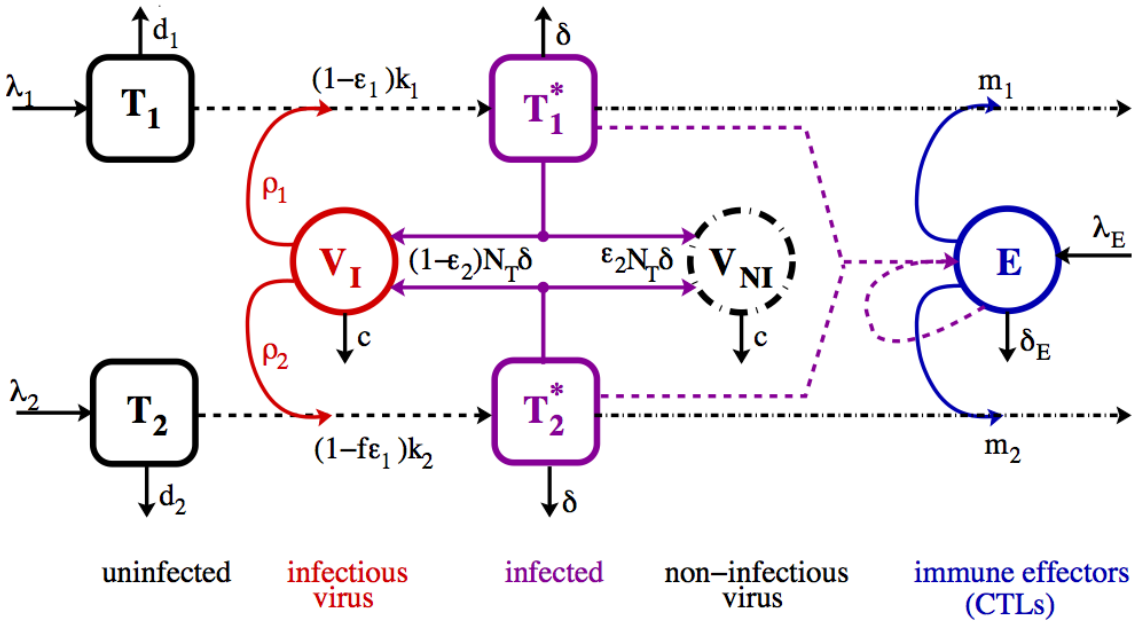

Figure 2.1: Overview schematic of the compartment based HIV model.

External mechanisms by which the virus is controlled (e.g. treatment) are incorporated into the model. We allow for treatment by reverse transcriptase inhibitors (RTIs) and protease

inhibitors (PIs). AZT, one of the very first drugs prescribed for the treatment of HIV is an RTI. PIs cause infected cells to produce non-infectious virions, hence the noninfectious viral state V_{NI}. Thus when considering treatment that includes PIs, we consider both infectious and noninfectious viral particles. We note however that even in the absence of a PI there can exist noninfectious viral particles.

Recently however these multiple drugs have been combined into a single drug, alleviating some of the stresses of HAART. These treatment factors are given as $\varepsilon_1(t) = \overline{\varepsilon}_1 u(t)$ and $\varepsilon_2(t) = \overline{\varepsilon}_2 u(t)$, consisting of efficacy $\overline{\varepsilon}_1 \in [0, 1]$ corresponding to a reverse transcriptase inhibitor and $\overline{\varepsilon}_2 \in [0, 1]$ modeling the effectiveness of a protease inhibitor with a time dependent function $u(t), 0 \leq u(t) \leq 1$. Because treatment has a reduced efficacy on the second target cell type, we model treatment for the population of T_2, as $f\varepsilon_1(t), f \in (0, 1)$. Treatment protocols are always considered as a combination of both PIs and RTIs, so monotherapy is not considered.

2.4 Nominal Parameters and Model Stability Analysis

The model (2.3.1) contains 22 biologically relevant parameters that must be specified before numerical simulations can be performed. We evaluate the model at a nominal parameter value that is justified through literature and past clinical studies; these values are given in Table 2.2. The parameter values are taken primarily from work done by (Callaway and Perelson, 2002; Bonhoeffer et al., 2000). Several of the parameters are not available from human or animal data. The nominal values for $\lambda_1, k_1, \lambda_2$, and k_2 are chosen so that several conditions on viral load and target cell equilibria are satisfied. These conditions are actually not reflected in our model since we have omitted the chronically infected state from the Callaway-Perelson model, though it is believed that small adjustments to the parameter values will obtain the same qualitative behavior (Adams et al., 2005).

The parameters used in the immune response state are not well known, and are chosen primarily to exhibit expected model behavior in the simulations. The parameters m_1, m_2 represent the effectiveness that the immune response E clears the infected cells of the target populations.

The commonly used value is taken from (Callaway and Perelson, 2002), who noted they obtained the value from (Nowak and Bangham, 1996).

We proceed by analyzing the three primary aspects of any dynamic system: existence, uniqueness, and stability. These three aspects have been discussed in depth for the nominal model parameter in (Adams, 2005; David, 2007) and we will summarize the key results here. In a later section we discuss the stability of the model equations for the patient specific parameter estimates.

Table 2.2: Nominal parameter values for the HIV model.

Parameter	Value	Units	Description
λ_1	10	cells/mm^3 day	Target cell type 1 source rate
λ_2	0.03198	cells/mm^3 day	Target cell type 2 source rate
d_1	0.01	1/day	Target cell type 1 death rate
d_2	0.01	1/day	Target cell type 2 death rate
k_1	8.0×10^{-4}	mm^3/virions day	Population 1 infection rate
k_2	0.1	mm^3/virions day	Population 2 infection rate
m_1	0.01	mm^3/cells day	Immune-induced clearance rate for population 1
m_2	0.01	mm^3/cells day	Immune-induced clearance rate for population 2
ρ_1	1	virions/cell	Average number virions infecting a type 1 cell
ρ_2	1	virions/cell	Average number virions infecting a type 2 cell
δ	0.7	1/day	Infected cell death rate
c	13	1/day	Virus natural death rate
f	$0.34 \in (0, 1)$		Treatment efficacy reduction in population 2
N_T	100	virions/cell	Virions produced per infected cell
λ_E	.001	cells/mm^3 day	Immune effector production (source) rate
δ_E	0.1	1/day	Natural death rate for immune effectors
b_E	0.3	1/day	Maximum birth rate for immune effectors
d_E	0.25	1/day	Maximum death rate for immune effectors
K_b	100	cells/ml	Saturation constant for immune effector birth
K_d	500	cells/ml	Saturation constant for immune effector death
ε_1	$.7 \in (0, 1)$		RTI efficacy
ε_2	$.3 \in (0, 1)$		PI efficacy

2.4.1 Existence and Uniqueness

Showing existence and uniqueness for a set of differential equations $\dot{x} = f(t, x; q)$ generally involves showing that the right hand side is Lipschitz continuous in x. By showing existence here, it allows us to justify the existence of sensitivity equations used later.

Due to product nonlinearities such as $k_1 T_1 V_I$, the right hand side of (2.3.1) does not satisfy Lipschitz continuity. To show existence, these nonlinear terms are replaced by piecewise differentiable terms that reach a saturation point once viral and cell populations reach a certain level. To incorporate this saturation, we consider a saturated model $\dot{x}^s = f^s(t, x^s)$, where product terms of the form $ax_i x_j$, $a \in \mathbb{R}$ are replaced with a saturation term

$$a_m^s(x_m) = \begin{cases} 0, & x_m < 0 \\ \sqrt{a} x_m, & x_m \in [0, x_m^M] \\ \sqrt{a} x_m^M, & x_m^M < x_m. \end{cases} \tag{2.4.1}$$

Introducing this term allows us to bound any product nonlinearity in the saturated model $\dot{x}^s = f^s(t, x^s)$ by $a x_i^M x_j^M$, and we note that this term is globally bounded and piecewise differentiable. Whenever the model states are below the saturation limit, the dynamics are determined by the original equations $\dot{x} = f(t, x; q)$. To complete the proof, we follow (Adams et al., 2005) and rewrite the system as

$$\dot{x} = S + \mathcal{L}(t)x + h(t, x), \tag{2.4.2}$$

where $S = [\lambda_1, \lambda_2, 0, 0, 0, 0, \lambda_E]$ are the source terms, $\mathcal{L}(t)x$ the terms that are linear in x, and $h(t, x)$ are the remaining nonlinearities, including the saturation dynamics. There is only one other nonlinear term in the original dynamics, $b\frac{T^*}{T^*+K}E$, where $T^* = T_1^* + T_2^*$. This term, however, is bounded above by

$$b\frac{T^*}{T^* + K}E \leq bE. \tag{2.4.3}$$

Its derivative with respect to the state is also bounded above by

$$b\frac{K}{(T^*+K)^2}E \leq \frac{bE}{K}, \tag{2.4.4}$$

for $T^* \geq 0$. These bounds directly imply that the derivative of the saturated nonlinear term is bounded above,

$$\|D_x h(t,x)\| < \infty, \tag{2.4.5}$$

and thus global Lipschitz continuity can be shown for the modified model

$$\dot{x}^s = f^s(t,x^s) = S + L(t)x^s + h(t,x^s) \tag{2.4.6}$$

as the right side satisfies a global Lipschitz condition:

$$\begin{aligned}
\|f^s(t,\varphi) - f^s(t,\psi)\| &= \|\mathcal{L}(t)(\varphi-\psi) + h(t,\varphi) - h(t,\psi)\| && (2.4.7)\\
&= \left\|\mathcal{L}(t)(\varphi-\psi) + \int_0^1 D_x h\big(t, y + s(\varphi-\psi)\big)(\varphi-\psi)ds\right\| && (2.4.8)\\
&\leq \|\mathcal{L}(t)\|\,\|\varphi-\psi\| + \|D_x h(t,x)\|\,\|\varphi-\psi\| && (2.4.9)\\
&= K_L\,\|\varphi-\psi\|\,. && (2.4.10)
\end{aligned}$$

Since the dynamics are Lipschitz continuous, the solutions exist and are unique (Perko, 2001). While the analysis has considered a saturated system, in simulation we use the original dynamics (2.3.1). Through all of the simulations a situation was never encountered where the state variables would grow without bound.

2.4.2 Stability Analysis

There are several steady states that exist for the HIV model under the nominal parameters. To find the equilibria, we set the treatment to off, $\varepsilon_1 \equiv 0 \equiv \varepsilon_2$ and set the $V_{NI} \equiv 0$ as it is uncoupled from the other model states. Equilibrium values are found by solving the system $f(x;q) = 0$, where $f(x;q)$ is the right hand side of our model equations. Solving this nonlinear system of equations symbolically is normally not possible, and numeric root finding methods must be employed to determine the equilibrium values.

Table 2.3: Steady state values under the nominal parameter with no treatment.

	EQ1	EQ2	EQ3
T_1	1000	163.57	967.84
T_2	3.198	0.005	0.6205
T_1^*	0	11.945	0.0760
T_2^*	0	0.0456	0.0061
V_I	0	63919	415.38
V_{NI}	0	0	0
E	0.01	0.0235	353.11
local stability	unstable	stable	stable

For the nominal parameters given in Table 2.2, three physically relevant steady states exist and are given in Table 2.3. There are other steady states, but are non-physical due to negative or complex equilibrium values. EQ1 is a healthy, uninfected patient with zero virus. This is an unstable steady state. The stability of each steady state is determined by a linear stability analysis, wherein we analyze the eigenvalues of the system Jacobian matrix evaluated at the steady state value.

When a small amount of virus is introduced, the system will converge to EQ2, where a high viral load is maintained over time and the immune response is depleted. The model under the nominal parameters also admits a third steady state EQ3, where an effective immune response

has developed and successfully suppresses the virus while regaining a healthy level of $CD4^+$ target cells. A more thorough treatment of the different steady states, as well as their sensitivities, is given in (Adams, 2005).

This third steady state is important because it represents a patient successfully controlling the virus in the absence of external HAART treatment. A central part of (David, 2007) is developing an optimal control methodology that would move a patient from EQ2 to EQ3 by way of a properly chosen control function. For each of the patients in our data set we perform a stability analysis using the resulting calibrated parameters and determine whether or not a treatment could be applicable in a clinical setting.

2.5 Numerical Solutions

Because the range of both the state variables and parameter values can vary over several orders of magnitude, the model is solved and calibrated using a $\log_{10}$ transformed system. We define the new state variable $x_i(t,q) = \log_{10} \overline{x}_i(t,q)$, where $\overline{x}_i$ represents the original, non-log transformed state. We also log-transform the parameters, $q \to \log_{10} q$, as they also varying widely in magnitude. The new system of equations are then given by

$$\dot{x}_i(t,q) = \frac{10^{-x_i}}{\ln 10} f(10^{x_i(q)}, 10^q) \tag{2.5.1}$$

where f is the right hand side of (2.3.1). Performing this substitution puts both the states and parameters within an order of (log) magnitude of each other, allowing for a more robust estimation procedure. This transformation also resolves issues of physically unrealistic negative solutions that occur due to round-off error. These solutions are more likely to happen during an optimization routine when the optimizer uses ill-chosen parameter values.

Model solutions are calculated with MATLAB's `ode15s` integrator with a relative error tolerance of 10^{-7}. The MATLAB integrator is a BDF-based stiff solver with variable tolerances (Shampine and Reichelt, 1997). A simulation of an early acute infection is shown in Figure

2.2. A small number of virus is introduced at $t = 0$, and the virus replicates to a peak before converging in a damped oscillation to a steady state. We see the delayed immune response E peak and damp out; a sustained and effective immune response does not develop.

2.6 Clinical Data Description

The data available has come from a clinical study at Massachusetts General Hospital of over 100 adults with acute HIV infection between 1996 and 2004. The central goal of the study was to determine the impact and consequences of early treatment initiation in the acute regime; the researchers wish to understand the role of early treatment in long term viral suppression. The data collected consists of $CD4^+$ T-lymphocyte count (cells/μl) and RNA viral load (RNA copies/ml). With respect to the model, the data represents $y_1 = T_1 + T_1^*$ for the $CD4^+$ count, and $y_2 = V_I + V_{NI}$ for the viral load. The viral load is quantified through different assay techniques; the standard assay can detect viral loads in the range of 400 to 750,000 copies/ml, while a high sensitivity assay has a detection range of 50 to 100,000 copies/ml. The high sensitivity assay is used when a measurement is below the 400 copies/ml lower detection limit for the standard assay, which can be the case when patients are successfully suppressing the virus.

Each patient in the study underwent combination HAART therapy consisting of three or more antiretroviral drugs; each specific patient protocol was determined by a qualified physician. Seven of the fourteen patients in our cohort underwent structured treatment interruptions according a study protocol. Other patients simply discontinued treatment on their own.

Due to the range sensitivity limits of the assays, the clinical viral load have upper and lower limits of quantification. The upper limit of quantification is handled through repeatedly diluting the sample until it falls within a detection range. The lower limit, or left censoring point, directly affects the observed data and a methodology is required to be able to use this data in the model calibration. When a data point is returned at a censor point, $L_1 = 400$ copies/ml or $L_2 = 50$ copies/ml, the only available knowledge is that the true measurement is between zero and the left quantification limit.

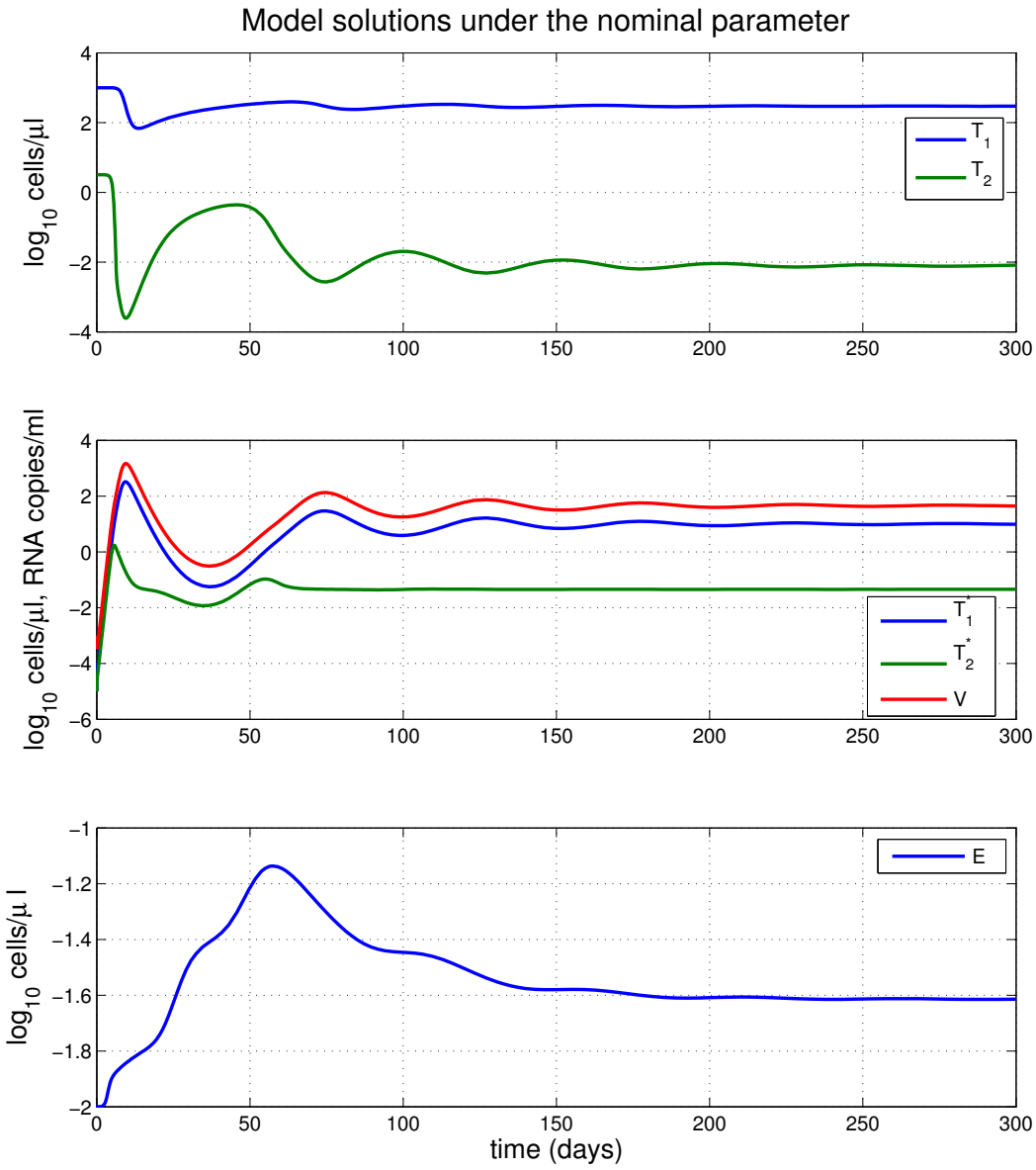

Figure 2.2: Model solutions under the nominal parameter given in Table 2.2. Note the differing logarithmic scales in each of the plots, and the quick transition to a steady state.

Data sets for patients 1 and 10 are shown in Figures 2.3, 2.4, respectively. The quantification limits have been highlighted by horizontal lines. The data points appearing directly on the horizontal reference lines are said to be censored. In the plots of collected data, the red bars on the bottom of the plot indicate non-adherence to the treatment regimen, whereas the green indicates the period of time the patient was on treatment. We collectively refer to this treatment as the "clinical control."

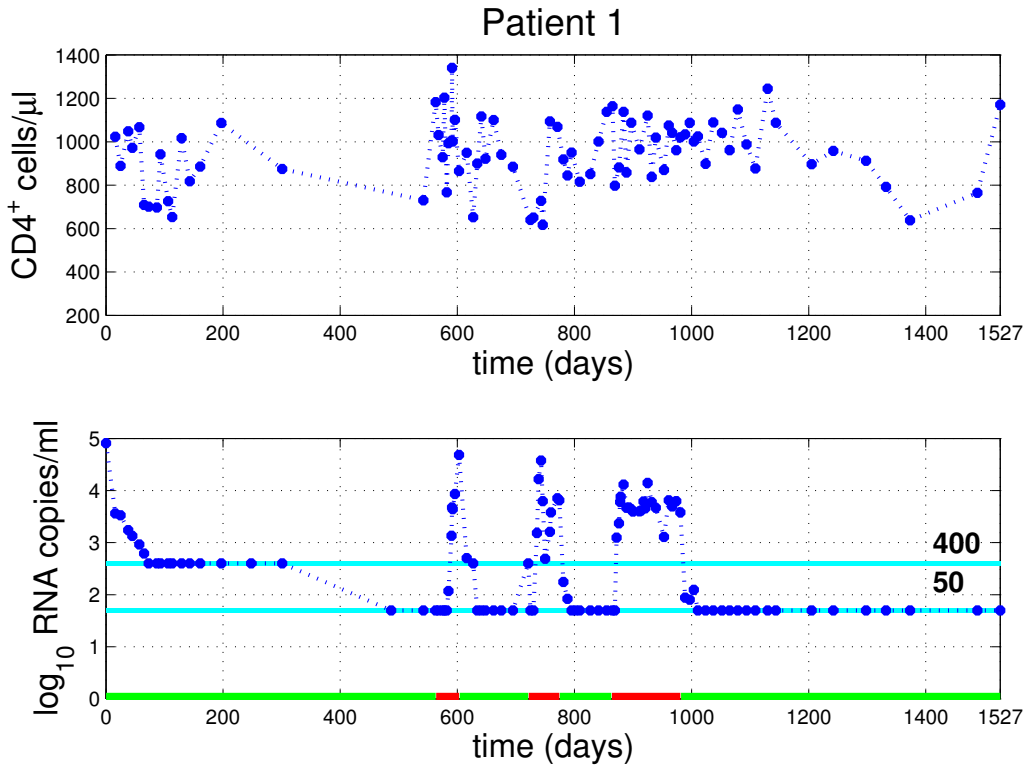

Figure 2.3: Patient 1 data set, with the censor points indicated by horizontal lines, $L_1 = \log_{10} 400, L_2 = \log_{10} 50$.

In addition to the data being censored at certain time values, the data are also collected at different time intervals. These intervals also differ between patients. This is visible in Figure 2.3 as well; the observations of viral load and $CD4^+$ may not have been made at the same

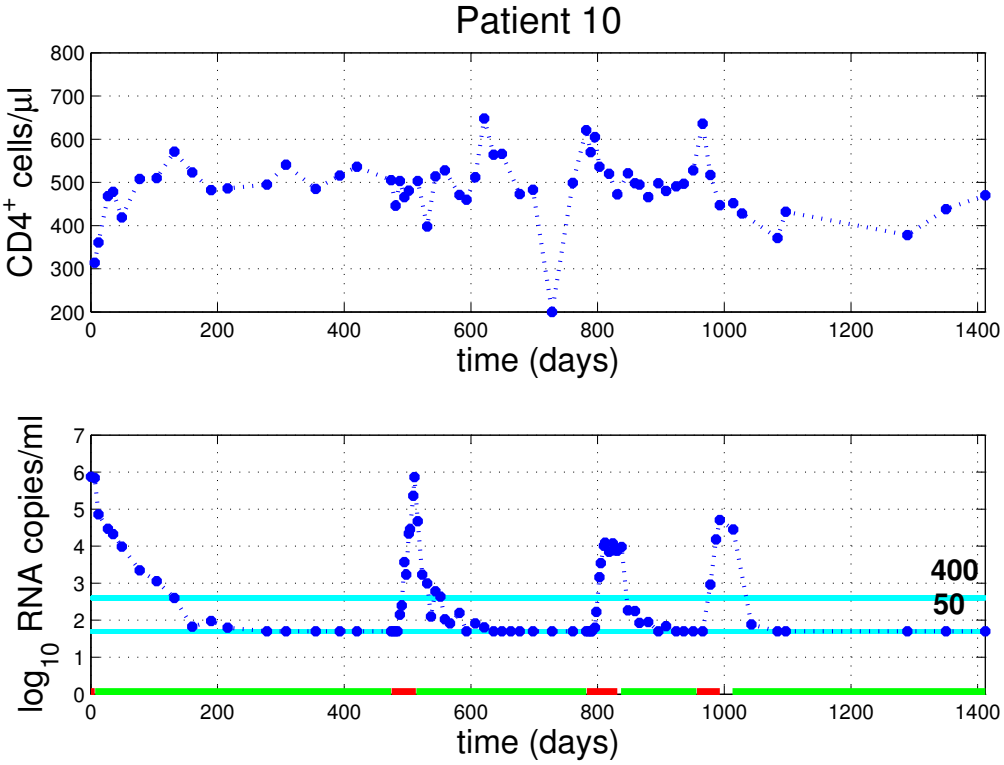

Figure 2.4: Patient 10 data set, with the censor points indicated by horizontal lines, $L_1 = \log_{10} 400, L_2 = \log_{10} 50$.

time points. In general for patient j (pt j), we have CD4$^+$ data pairs $(t_1^{ij}, y_1^{ij}), i = 1, \ldots, N_1^j$ and possibly different viral RNA pairs $(t_2^{ij}, y_2^{ij}), i = 1, \ldots, N_2^j$. Given the data available, we consequently have only reduced observations of the the model solutions.

2.7 Desired Research Outcomes

To this point we have described both the mathematical model and the clinical data at our disposal. Over the next few chapters, we combine the model and the data to form a calibrated, patient specific model. In doing so, we build a patient predictive model that can be used with a variety of different control methodologies, or treatment strategies, that can be used to investigate the best way to treat a patient to guarantee a certain outcome. In summary:

1. Calibrate the HIV model (2.3.1) parameters to the clinical datasets through an inverse problem based methodology.

2. Use the calibrated model in conjunction with optimal control theory to develop new, patient specific treatment strategies.

3. Compare the optimal strategies with the clinical control.

CHAPTER 3

Methodology of Model Calibration & Validation

Our first goal is model calibration and validation; this occurs when we fit the model parameters to the clinically collected data so that the resulting model trajectories match the data. Termed the inverse problem, model calibration is present in many areas of engineering and other physical problems. For a good overview of the inverse problem methodology and process, see (Banks and Tran, 2009) and the references therein.

In general, inverse problems are difficult problems to solve. Nonlinear and ill-posed problems are nearly the norm when confronting inverse problems. Additionally, not all parameters and states may be directly measurable or even observable, and the data available may be noisy and sparse. These problems in general may not have a unique solution, and require problem-specific expertise to solve. A summary of these uniqueness issues as well as stability considerations is given in (Aster et al., 2005).

Before rigorous parameter estimation techniques can be applied to any nonlinear ODE model, a large issue that must be overcome is the identification of which parameters, or subset of parameters, are identifiable given the available data. This question can be extended to investigate

how the data affects identifiability; e.g. determining at which time points having more data would lead to increased parameter identifiability. To address these issues, identifiability analysis is the first a priori step in model calibration and validation.

In this chapter we develop and discuss a few of these methods that are used in this work to identify the HIV model parameters to fit the clinical data, including calculus based sensitivity analysis, identifiability analysis, and methods for subset selection.

To develop these theories, we will motivate the analysis through the use of some low-dimension examples before addressing the general cases that we apply to this work.

3.1 Calculus Based Sensitivity Analysis

Sensitivity analysis is the study of the response of the system to small perturbations in data and system parameters. This analysis forms the basis of analysis of inverse problems, as it determines which parameters most affect system output. However we will find that even though a parameter is sensitive in some context, this does not mean that it may be uniquely identified from data or observations. In the past, sensitivity analysis was primarily used in the analysis of the forward, or simulation, problem when one needed to understand the effects of parameter perturbation on the output. In recent years sensitivity analysis has become a de-facto part of inverse problem analysis, partly because it directly aids in uncertainty analysis. When physiologically based mathematical models are developed, the parameters typically have biological relevance, and thus one may wish to estimate these parameters from data. Sensitivity analysis can also aid in the optimal design of experiments and improved data collection (Banks et al., 2007).

We have not addressed generalized sensitivity analyses, which provides an information-theoretical viewpoint of sensitivity analysis and can offer insight on the dependence of parameters on the model output. These functions also provides correlation information on which portions of data contain the most information for parameter estimations (Thomaseth and Cobelli, 1999; Batzel et al., 2007; Kiparissides et al., 2009; van Griensven et al., 2006).

Consider a function $y(q)$. The derivative $\frac{dy}{dq}$, defined as

$$y'(\hat{q}) = \lim_{h \to 0} \frac{y(\hat{q}+h) - y(\hat{q})}{h}, \tag{3.1.1}$$

at a point $\hat{q}$ determines how much y changes as q changes near $\hat{q}$. Large absolute values for $y'(\hat{q})$ indicate large changes in y near $\hat{q}$, whereas small values of $y'(\hat{q})$ indicate small changes in y, as presented in Figure 3.1.

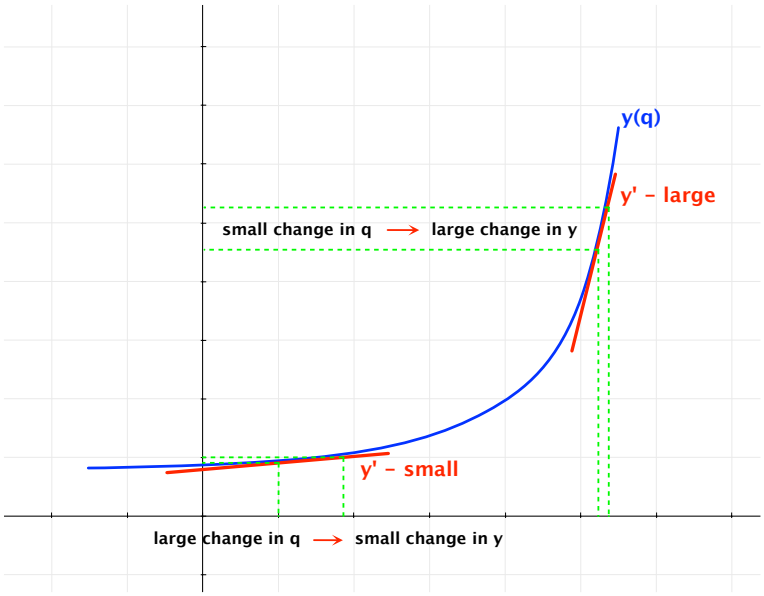

Figure 3.1: A function $y(q)$ and corresponding derivative $y'(q)$.

3.1.1 Direct Computation

Sensitivity functions can be computed in several different ways, each with varying degrees of accuracy, tedium, and computational time. Consider a simplified HIV model (Fink et al., 2007),

$$\begin{aligned}\dot{T} &= \lambda - dT - kVT \\ \dot{T}^* &= kVT - \delta T^* \\ \dot{V} &= N\delta T^* - cV,\end{aligned} \tag{3.1.2}$$

with observation function

$$\mathbf{y}(t) = \begin{bmatrix} T_1 + T_1^* \\ V_I + V_{NI} \end{bmatrix}. \tag{3.1.3}$$

This model neglects a second type of target cells and assumes no treatment nor immune response, and we use parameter values that are identical for the larger, seven state model. The state variables are the same as given in Table 2.1, and we take $\mathbf{q} = [\lambda, d, k, \delta, N, c]$. If we wanted to calculate

$$\frac{\partial \mathbf{y}}{\partial k} = \begin{bmatrix} \frac{\partial T_1}{\partial k} + \frac{\partial T_1^*}{\partial k} \\ \frac{\partial V_I}{\partial k} + \frac{\partial V_{NI}}{\partial k} \end{bmatrix}, \tag{3.1.4}$$

we could do so directly by taking derivatives,

$$\frac{\partial}{\partial k}\frac{dT}{dt} = \frac{\partial}{\partial k}(\lambda - dT - kVT) \tag{3.1.5}$$

$$\frac{d}{dt}\frac{\partial T}{\partial k} = -d\frac{\partial T}{\partial k} - VT - k\left(\frac{\partial V}{\partial K}T + V\frac{\partial T}{\partial k}\right). \tag{3.1.6}$$

In the above calculation, the quantity of interest is $\frac{\partial T}{\partial k}$, though we now need expressions for $\frac{\partial V}{\partial K}$ and $\frac{\partial T}{\partial k}$. We have also exchanged the order of differentiation on the left hand side, as we have implicitly assumed sufficient continuity requirements. Equation (3.1.6) is now a differential equation in $\frac{\partial T}{\partial k}$ that would be solved by standard numerical integration techniques. To integrate

(3.1.6) an initial condition is required, and we take

$$\frac{\partial T}{\partial \mathbf{q}}(0) = 0, \quad \frac{\partial T^*}{\partial \mathbf{q}}(0) = 0, \quad \frac{\partial V}{\partial \mathbf{q}}(0) = 0, \tag{3.1.7}$$

implying that the initial conditions for the states do not depend on the parameters.

Calculating sensitivities using this method is tedious and fraught with potential errors. Using a symbolic algebra system such as Maple could reduce the chance of an error being made, however the resulting set of equations would undoubtedly be unwieldily. In general, we are interested in the calculation of all sensitivity functions. For a system of n equations with m parameters, this leaves us with mn differential equations to be solved along with the original n equations. A more systematic method is required while still maintaining a high degree of accuracy.

The system (2.3.1) of differential equations can be written in the form

$$\frac{d\mathbf{x}(t)}{dt} = f\big(\mathbf{x}(\mathbf{q}), t; \mathbf{q}\big), \quad \mathbf{x}(t_0) = \mathbf{x}_0, \tag{3.1.8}$$

with observation function

$$\mathbf{y}(t) = g(\mathbf{x}(t)). \tag{3.1.9}$$

Here, $\mathbf{x} \in \mathbb{R}^n$ is a vector of the state variables $T_1, \ldots, E$ with $\mathbf{y} \in \mathbb{R}^p$. We assume the parameter vector $\mathbf{q} \in \mathbb{R}^m$ to be constant over time, though theory exists for cases when the parameters are considered functions rather than just constant values.

The sensitivity equations are formed by implicitly differentiating both sides of (3.1.9) with respect to $\mathbf{q}$ by way of the chain rule,

$$\frac{\partial \mathbf{y}}{\partial \mathbf{q}} = \frac{\partial g}{\partial \mathbf{x}} \frac{\partial \mathbf{x}}{\partial \mathbf{q}}. \tag{3.1.10}$$

These are the output sensitivities, and are in terms of the state sensitivities $\frac{\partial \mathbf{x}}{\partial \mathbf{q}}$. These equations

are obtained by differentiating (3.1.8) also with respect to $\mathbf{q}$. We again exchange the order of differentiation as we have implicitly made sufficient continuity assumptions. The mn sensitivity equations are then given by

$$\frac{d}{dt}\frac{\partial \mathbf{x}(t)}{\partial \mathbf{q}} = \frac{\partial f}{\partial \mathbf{x}}\frac{\partial \mathbf{x}}{\partial \mathbf{q}} + \frac{\partial f}{\partial \mathbf{q}}, \tag{3.1.11}$$

which is a linear, matrix differential equation in $\frac{\partial \mathbf{x}}{\partial \mathbf{q}}$. The calculation of the terms $\frac{\partial f}{\partial \mathbf{x}}$ and $\frac{\partial f}{\partial \mathbf{q}}$ must still be handled. These terms can either be calculated through manual techniques as discussed earlier and in (Valdez-Jasso et al., 2009), or by way of Automatic Differentiation (AD).

AD is essentially a systematic, numerical implementation of the chain rule of differentiation (Griewank, 1989; Verma, 2000). No matter how complicated a mathematical function is, it is still a combination of elementary functions such as trigonometric functions, addition, and multiplication. The derivatives of these functions are known and the total derivative is computed by table lookups and tracking all of the functional compositions. AD takes derivatives as people do: through product, quotient, and chain rules and provides a derivative that is numerically accurate to floating point (Carmichael et al., 1997). There is no sacrifice in accuracy as there is in finite difference methods, and AD can be relatively straight forward to implement. In our work, we have used a combination of AD and finite difference methods to compute the sensitivity functions.

3.1.2 Initial Condition Sensitivities

Because the initial model states are unknown for each patient, we must also estimate the initial conditions for the model. Doing so requires the formulation of initial condition sensitivity equations, or how the states respond to perturbations in initial condition. To construct these n^2 equations we differentiate (3.1.8) with respect to $\mathbf{x}_0$. After exchanging the order of integration,

we have

$$\frac{d}{dt}\frac{\partial \mathbf{x}(t)}{\partial \mathbf{x}_0} = \frac{\partial f}{\partial \mathbf{x}}\frac{\partial \mathbf{x}}{\partial \mathbf{x}_0}. \tag{3.1.12}$$

These equations are coupled with the original sensitivity and state equations yielding a system of $n + mn + n^2$ differential equations,

$$\begin{aligned} \frac{d\mathbf{x}(t)}{dt} &= f(\mathbf{x}(\mathbf{q}), \mathbf{q}) \\ \frac{d}{dt}\frac{\partial \mathbf{x}(t)}{\partial \mathbf{q}} &= \frac{\partial f}{\partial \mathbf{x}}\frac{\partial \mathbf{x}}{\partial \mathbf{q}} + \frac{\partial f}{\partial \mathbf{q}} \\ \frac{d}{dt}\frac{\partial \mathbf{x}(t)}{\partial \mathbf{x}_0} &= \frac{\partial f}{\partial \mathbf{x}}\frac{\partial \mathbf{x}}{\partial \mathbf{x}_0}, \end{aligned} \tag{3.1.13}$$

which would then be integrated in forward time giving the sensitivity functions. The Jacobian matrices $\frac{\partial f}{\partial \mathbf{x}}$ and $\frac{\partial f}{\partial \mathbf{q}}$ could either be computed and stored prior to integration, or computed on the fly with AD.

We assume a zero initial condition for the parameter sensitivity equations (3.1.11) at time $t = t_0$, as the initial conditions for model states are not parameter dependent and a unity initial condition for the initial condition sensitivities (3.1.12), since $\frac{\partial x(t)}{\partial x_0}\big|_{t=t_0} = \frac{\partial x_0}{\partial x_0} = 1$.

3.1.3 Finite Differencing

A final method of calculation of the sensitivity equations, and the method used in this research, is finite differencing. Consider the Taylor series of order n of a scalar valued function f,

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2}f''(x) + \cdots + \frac{h^n}{n!}f^{(n)}(x). \tag{3.1.14}$$

The derivative $f'(x)$ can then be written in terms of the higher order terms,

$$f'(x) = \frac{f(x+h) - f(x)}{h} + \mathcal{O}(h). \tag{3.1.15}$$

Equation (3.1.15) is referred to as the forward difference, whereas

$$f'(x) = \frac{f(x) - f(x-h)}{h} + \mathcal{O}(h) \tag{3.1.16}$$

is called the backwards difference. Combining these two we obtain the central finite difference,

$$f'(x) = \frac{f(x+h) - f(x-h)}{2h} + \mathcal{O}(h^2). \tag{3.1.17}$$

The central difference has the distinct advantage of being accurate on an order of h^2, versus order h for the forward and backward differences. This accuracy however comes at the cost of requiring an additional function evaluation. When we neglect the higher order terms, the difference approximations will be close to the true derivative value for sufficiently small steps h; as $h \to 0$, the finite difference approximations mathematically converge to the true derivative.

We are careful to note that this convergence occurs in a mathematical sense, however not in a numerical sense. Reducing the step size h to an arbitrary small value can introduce large amounts of floating point based error into the approximation, providing diminishing returns. Managing the step size h becomes nontrivial, especially because parameter values of interest can be on different orders of magnitude and the step size must scale accordingly. In (Dennis Jr. and Schnabel, 1996) it is shown that a good choice for h in forward and backward differencing schemes is $h = \sqrt{\text{macheps}} \cdot x$; for central differencing schemes a good value of $h = (\text{macheps})^{1/3} x$. These values are given in terms of macheps, which is the relative error in computing $f(x)$. These values minimize the round-off and machine error in the calculations.

For our calculations, the sensitivity functions are computed using a central finite difference approximation,

$$\frac{\partial \mathbf{y}}{\partial q_i} = \frac{\mathbf{y}(q_i + he_i) - \mathbf{y}(q_i - he_i)}{2h} + \mathcal{O}(h^2), \tag{3.1.18}$$

where $\mathbf{y}$ are the observations of the state solutions as computed from the integrator, and e_i is

an elementary basis vector which is all zeros except for the i^{th} slot, where the value is 1.

3.1.4 Sensitivity Rankings

After the sensitivity functions have been calculated, we need a way to compare the sensitivity of one parameter to another. Since parameters in a model are often of different units, we first non-dimensionalize the parameters by multiplying by the parameter value,

$$\frac{\partial y_i}{\partial q_j} q_j. \tag{3.1.19}$$

To improve the scaling of the functions, we also divide by the maximum value of the state variable. Finally, we take the L^2 norm, generating sensitivity coefficients,

$$C_{ij} = \left\| \frac{\partial y_i}{\partial q_j} \frac{q_j}{\max y_i} \right\|_2^2 = \int_{t_0}^{t_f} \left| \frac{\partial y_i}{\partial q_j} \frac{q_j}{\max y_i} \right|^2 \, dt \in \mathbb{R}. \tag{3.1.20}$$

Once these ranking values are computed, we can effectively compare the overall effect of one parameter on system output to another. Typically it is at this point that the parameters are ranked in descending order of sensitivity, beginning to give insight into the overall system identifiability. Some of the simplest identification techniques state simply that the larger the coefficients, then the more dramatic the system output with respect to that parameter and therefore a parameter is more likely to be identifiable when there is a greater sensitivity. As we will discuss in the next section, this is not always the case.

3.2 Identifiability Analysis

Sensitivity analysis as we have defined it so far is a local concept based around perturbations of parameters and initial state values. The sensitivity of a parameter is only one of many considerations that determines whether or not a parameter can be identified from data. Some parameters may have a very weak effect on the measured output, and successful estimation

is unlikely. However due to model structure and lack of measurements, parameters may also be unestimatable. The ideal estimation problem assumes an error free model and error free data. Since this is not often the case, we use identifiability analysis to shed light on how (the imperfect) model structure and available (noisy) data impact the ability to estimate parameters. In that sense, the identification problem is conditioned on the experiment. Identifiability is a basic system property, is theoretical in nature and is part of an a-priori analysis to parameter estimation (Li et al., 2004).

Identifiability in relation to HIV models have been considered in several works, (Xia and Moog, 2003; Xia, 2003; Wu et al., 2008; Guedj et al., 2007), though, in general, have been concerned with lower dimensional models and have not included clinical data. Indeed, many of the methods developed in these works simply do not easily scale to higher dimensional models such as ours. An excellent paper detailing identifiability issues in relation to viral dynamic models is (Miao et al., 2011). The authors define a system to be identifiable if the parameter $\mathbf{q}$ can be *uniquely* determined from the system input $u(t)$ and the measured system output $y(t)$. Otherwise, the system is said to be unidentifiable. In (Ljung and Glad, 1994), the definition distinguishing between globally identifiable, which states any two parameters $\mathbf{q}_1, \mathbf{q}_2$ are globally identifiable relative to system output $\mathbf{y}(\mathbf{q})$, we have $\mathbf{y}(\mathbf{q}_1) = \mathbf{y}(\mathbf{q}_2) \iff \mathbf{q}_1 = \mathbf{q}_2$. Local identifiability is similar, though the result holds only in a local neighborhood about some point $\mathbf{q}^*$.

We will follow a similar structure to (Miao et al., 2011), detailing the prior work in structural and practical identifiability before addressing the methodology used in this research, sensitivity-based identifiability analysis.

3.2.1 Structural Identifiability

Structural identifiability was first introduced in (Bellman and Åström, 1970), and explores how system structure, i.e., the model itself, affects parameter identifiability. These methods are not yet in widespread use, due to either computational complexity or the lack of effective evaluation

algorithms. To motivate how model structure can affect parameter identifiability, consider a system of equations describing radioactive marking of tumor cells. In this example, marked tumor cells are added in-vivo to the blood stream and the levels of cancer cells are measured over time. The system can be represented as

$$\dot{x}_1 = -\delta_1 x_1 - \delta_2 x_1 \tag{3.2.1}$$

$$\dot{x}_2 = \delta_2 x_1 - \beta_1 x_2, \tag{3.2.2}$$

where x_1, x_2 are the number of marked cells in two different environments such as capillaries and the lung. The parameters $\mathbf{q} = [\delta_1, \delta_2, \beta_1]$ are terms that regulate transfer rates between the two model states. Note that the first state can be written as $\dot{x}_1 = -(\delta_1 + \delta_2)x_1$, and so in estimation one is only truly able to estimate the sum term $\delta_1 + \delta_2$, rather than each term independently and uniquely. This model is structurally unidentifiable.

Consider a second model, for drug concentration:

$$\dot{x}(t) = -p_1 x(t) + p_2 u(t), \tag{3.2.3}$$

with observation $y(t) = p_3 x(t)$. Here, $x(t)$ is the drug concentration and the drug input is $u(t)$. The solution is given by, where we assume $x(0) = 0$,

$$y(t) = p_3 p_2 \int_0^t e^{-p_1(t-s)} u(s) ds, \tag{3.2.4}$$

and we are left with the unidentifiable product term $p_3 p_2$, despite not explicitly having a product term in the original dynamic equation. This model is also unidentifiable.

It is these kinds of analyses that structural identifiability addresses. For general nonlinear models, frameworks consisting of differential algebras have been developed (Ljung and Glad, 1994). Differential algebra methods are applied to biological systems in (Audoly et al., 2001). A more direct method of addressing structural issues is given in (Xia and Moog, 2003), and is

based on the implicit function theorem. Termed algebraic identifiability, the authors develop an identification function, proceed to show how parameters can be identified by solving algebraic equations based only on the initial conditions and the output variables. This function is derived by eliminating the unobserved state variables by taking higher order derivatives of the right hand side equations. A matrix, aptly named the identification matrix is formed by taking partial derivatives of these equations with respect to the unknown parameters. If this matrix is of full rank, then the parameters are identifiable. This method has been applied to models of up to dimension six (Miao et al., 2008; Xia and Moog, 2003). This method requires the construction of relatively large matrices and assumes a well conditioned rank test. As we discuss in a later section, these matrices are not singular, but rather nearly singular, and an informed decision is needed to declare the rank.

3.2.2 Practical Identifiability

Structural identifiability analysis is prefaced on two already mentioned tenants: that the model is perfect and uncorrupted, and that the data are exact and have no noise nor error. Given the undeniable fallacy of these assumptions, analysis is required beyond the structural level to determine parameter identifiability. It is in practical identifiable analysis where we begin to take the experiment vis a vis the data into account and build upon the structural analysis. We note that if a parameter is determined to be structurally unidentifiable, no amount of practical analysis will change that. Conversely, should a parameter be structurally identifiable, practical analysis may yield an unidentifiable result.

Practical identifiability analysis analyzes whether parameters identified as structurally identifiable can be estimated from the data with acceptable accuracy. In contrast to the structural analysis, this is a posterior (in relation to the experiment) analysis. The next subsections detail specific techniques to this end.

3.2.3 Sensitivity-Based Identifiability Analysis

In addition to parameters being insensitive and thus difficult to estimate, parameters may also be correlated. If there is a large correlation between parameters then the two parameters are likely to be indistinguishable. These dependencies can be analyzed and handled through analysis of the sensitivity function matrix.

This matrix is defined for outputs $y_1(t), \ldots, y_p(t)$ with respect to parameters $\mathbf{q} = [q_1, \ldots, q_m]^\top$ at the time points $t = (t_1, \ldots, t_n)$ by the $pn \times m$ matrix

$$S(\mathbf{q}, t) \equiv \begin{bmatrix} s_{1,1}(\mathbf{q}, t_1) & \cdots & s_{1,m}(\mathbf{q}, t_1) \\ s_{1,1}(\mathbf{q}, t_2) & \cdots & s_{1,m}(\mathbf{q}, t_2) \\ \vdots & \ddots & \vdots \\ s_{1,1}(\mathbf{q}, t_n) & \cdots & s_{1,m}(\mathbf{q}, t_n) \\ s_{2,1}(\mathbf{q}, t_1) & \cdots & s_{2,m}(\mathbf{q}, t_1) \\ \vdots & \ddots & \vdots \\ s_{p,1}(\mathbf{q}, t_n) & \cdots & s_{p,m}(\mathbf{q}, t_n) \end{bmatrix}, \quad s_{i,j}(\mathbf{q}, t_k) = \frac{\partial y_i(t_k)}{\partial q_j}. \tag{3.2.5}$$

The sensitivity matrix S represents the classical mechanism of sensitivity analysis (Frank, 1978; Eslami, 1994). A nominal parameter value is required for the evaluation of the sensitivity matrix, thus identifiability analysis based on sensitivity methods is evaluated with respect to a specific point in the parameter space. Consider the first order Taylor series expansion of the system response near a nominal parameter value $\mathbf{q}^*$,

$$\begin{aligned} \mathbf{y}(\mathbf{q}) &= \mathbf{y}(\mathbf{x}(t_k), \mathbf{q}) \\ &\approx \mathbf{y}(\mathbf{x}(t_k), \mathbf{q}^*) + \frac{\partial \mathbf{y}(\mathbf{x}(t_k), \mathbf{q})}{\partial \mathbf{q}}\bigg|_{\mathbf{q}=\mathbf{q}^*} (\mathbf{q} - \mathbf{q}^*), \end{aligned} \tag{3.2.6}$$

where in our notation $S = \frac{\partial \mathbf{y}(\mathbf{x}(t_k), \mathbf{q})}{\partial \mathbf{q}}\big|_{\mathbf{q}=\mathbf{q}^*}$. Let $\hat{\mathbf{y}}_k$ denote a model observation, e.g. a measurement at time t_k and $\triangle\mathbf{q} = \mathbf{q} - \mathbf{q}^*$. The residual sum of squares between the model observations and

the linear approximation (3.2.6) is

$$RSS(\triangle \mathbf{q}) = \sum_{k=1}^{N} \left[\hat{\mathbf{y}}_k - \mathbf{y}_k(\mathbf{q}^*) - \frac{\partial \mathbf{y}(\mathbf{x}(t_k), \mathbf{q})}{\partial \mathbf{q}} \bigg|_{\mathbf{q}=\mathbf{q}^*} \triangle \mathbf{q} \right]^2$$
$$= \sum_{k=1}^{N} \left[\frac{\partial \mathbf{y}(\mathbf{x}(t_k), \mathbf{q})}{\partial \mathbf{q}} \bigg|_{\mathbf{q}=\mathbf{q}^*} \triangle \mathbf{q} \right]^2, \tag{3.2.7}$$

where we have assumed that the residuals $\mathbf{r}_k = \hat{\mathbf{y}}_k - \mathbf{y}_k(\mathbf{q}^*)$ are small and negligible. Rewriting (3.2.7) in terms of the sensitivity function matrix S, we have

$$RSS(\triangle \mathbf{q}) = (S \triangle \mathbf{q})^\top S \triangle \mathbf{q}, \tag{3.2.8}$$

which has a minimum when $(S^\top S)\triangle \mathbf{q} = 0$. The matrix $S^\top S$ is referred to as the Fisher Information Matrix, and is a first order approximation to the Hessian of the cost function in an ordinary least squares sense. It follows that the parameters $\mathbf{q}$ are locally identifiable if and only if the column rank of the matrix S is equal to m, or equivalently $\det(S^\top S) \neq 0$. If $S^\top S$ is full rank, then $(S^\top S)\triangle \mathbf{q} = 0$ has a unique solution, $\mathbf{\Delta q} = 0$, and $\mathbf{q} = \mathbf{q}^*$ implying that $\mathbf{q}$ is locally identifiable at $\mathbf{q}^*$. If $S^\top S$ is singular, then there exists a nontrivial solution $\mathbf{q} \neq \mathbf{q}^*$, and the model parameters are not identifiable at $\mathbf{q}^*$.

It is clear that identifying $\mathbf{q}$ is contingent on F being full rank. It is typically the case, especially when working with complex compartment models the matrix F is not exactly singular. However F is often *nearly* singular, in the sense that its smallest singular value is much smaller than the largest singular value of F. The numerical rank of a matrix offers a measure of linear independence in the columns of a matrix. To compute the numerical rank of the matrix, we use the following measure with $\varepsilon = \varepsilon_{\|F\|}$, the floating point accuracy at $\|F\|$. This is a version of the equation presented in (Golub et al., 1976).

$$\text{rank}(A, \varepsilon) = \max \left\{ i \text{ s.t. } \frac{|\sigma_i|}{|\sigma_1|} > \varepsilon \, \|A\| \, m \right\}, \tag{3.2.9}$$

where σ_i are the singular values of the matrix, with $\sigma_1 \geq \sigma_i \ \forall i > 1$, and m is the number of columns. The calculation of the numerical rank is highly dependent on the tolerance used, especially when the spectrum is relatively smooth with no obvious breaks or gaps. The spectrum of the information matrix for two of the patients in the data set are given in Figures 3.2 and Figure 3.3. The obvious break in the spectrum continuity for patient 1 determines the numerical rank of the matrix to be 9. For patient 4, the spectrum is smoother and we must rely on only the error tolerance of the rank estimate; the sensitivity to the tolerance is higher. In this case, the rank of the matrix is given as 6.

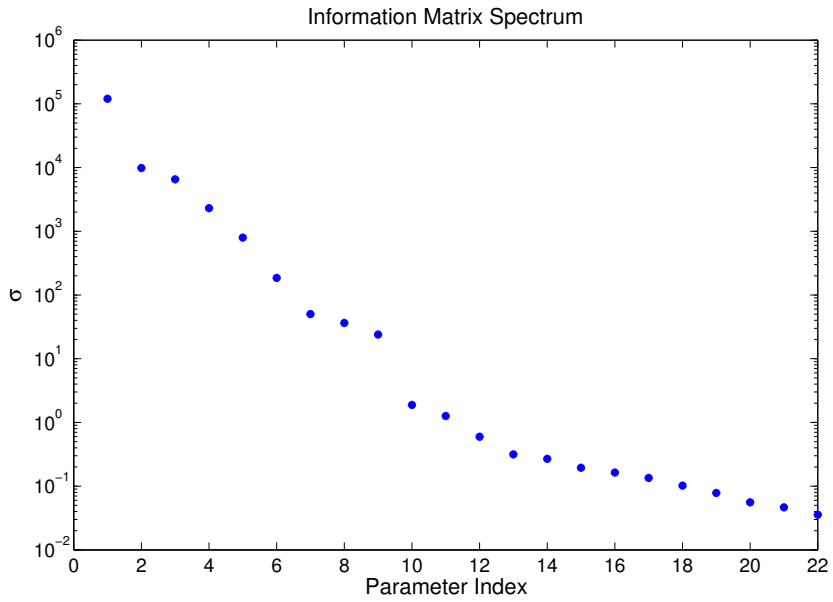

Figure 3.2: Spectrum of $S^\top S$ for pt 1. The gap in the continuity of the spectrum between parameter indices 9 and 10 determine the numerical rank of the matrix to be $k = 9$.

When the matrix F is of numerical rank k, we say that only k parameters are identifiable, as k parameters form a mostly linearly independent spanning set. The question then becomes which k parameters are identifiable, and this is answered through subset selection.

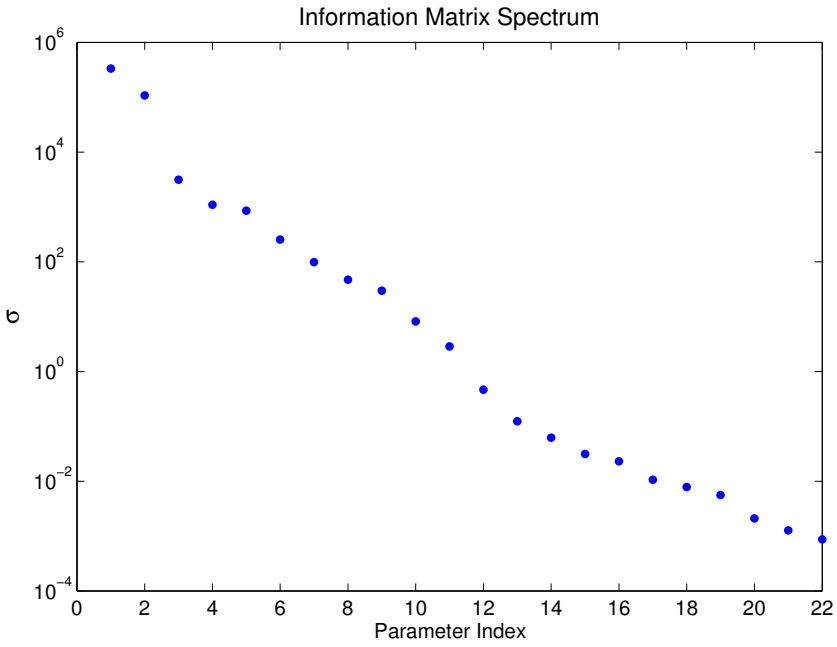

Figure 3.3: Spectrum of $S^\top S$ for pt 4. The gap in the continuity of the spectrum between parameter indices 6 and 7 determine the numerical rank of the matrix to be $k = 6$.

3.3 Subset Selection

A number of approaches have been developed and used to find an identifiable subset of parameters. In addition to parameters being insensitive, parameters may also be correlated, which also reduces identifiability (Thompson et al., 2009; Miao et al., 2011; Jacquez, 1972; Daun et al., 2008). Methods of subset selection seek to determine which grouping of parameters are identifiable in some sense, and are typically based on methods of maximizing linear independence in the sensitivity matrix. In this section we briefly discuss some common methods and algorithms for subset selection before settling on the method used in this work.

3.3.1 Structured Correlation Analysis

Subset selection based on correlation analysis relies on the analysis of the system correlation matrix. Assume that a parameter estimate $\hat{\mathbf{q}}$ has been computed by fitting the observations to

the model through a least squares process. The correlation matrix R has the form

$$C(\mathbf{q}) = \begin{bmatrix} r_{11}(\hat{q}_1, \hat{q}_1) & r_{12}(\hat{q}_1, \hat{q}_2) & \cdots & r_{1m}(\hat{q}_1, \hat{q}_m) \\ r_{21}(\hat{q}_2, \hat{q}_1) & r_{22}(\hat{q}_2, \hat{q}_2) & \cdots & r_{2m}(\hat{q}_2, \hat{q}_m) \\ \vdots & & \ddots & \\ r_{m1}(\hat{q}_m, \hat{q}_1) & r_{12}(\hat{q}_m, \hat{q}_2) & \cdots & r_{mm}(\hat{q}_m, \hat{q}_m) \end{bmatrix}, \tag{3.3.1}$$

where r_{ij}, $i, j = 1, \ldots, m$, $r_{ij} \in [-1, 1]$ are the correlation coefficients between parameter estimates $\hat{q}_i$ and $\hat{q}_j$. A strong correlation between two parameters is indicated if $|r_{ij}|$ is near 1. We use the term "near", as there is currently no theory nor empirical rule as to what constitutes a correlated parameters. The derivation for correlation matrix expression is given in (Daun et al., 2008; Rodriguez-Fernandez et al., 2006), where it is again used in a subset selection context.

This method uses at its base the Fisher information matrix, also known as the model Hessian. For a dynamic system of the form (3.1.8, 3.1.9), the Fisher matrix is formally given as

$$F = \sum_{k=1}^{N} \frac{\partial \mathbf{y}(\mathbf{x}(t_k), \mathbf{q})}{\partial \mathbf{q}}^{\top} \Sigma^{-1} \frac{\partial \mathbf{y}(\mathbf{x}(t_k), \mathbf{q})}{\partial \mathbf{q}} \tag{3.3.2}$$

$$= S^{\top} \Sigma^{-1} S, \tag{3.3.3}$$

where Σ is the measurement covariance. The Fisher matrix also provides the lower bound of variance of the parameter estimates according to Cramèr-Rao theory (Rao, 1945), and is equal to the inverse of the correlation matrix. That is,

$$C = F^{-1} \tag{3.3.4}$$

for $\det F \neq 0$. The correlation coefficients in matrix C are given by

$$r_{ij} = \frac{C_{ij}}{\sqrt{C_{ii} C_{jj}}}, \quad i \neq j \tag{3.3.5}$$

and 1's on the diagonal. The correlation matrix is also used to derive the standard errors, and this is discussed in section 3.4.1. In (Cintrón-Arias et al., 2009), the correlation matrix is used to compute what the authors term a vector of "coefficients of variation" $\nu(\mathbf{q}) \in \mathbb{R}^m$,

$$\nu_i(\mathbf{q}) = \frac{\sqrt{C(\mathbf{q})_{ii}}}{q_i}. \tag{3.3.6}$$

A selection score $\alpha(\mathbf{q})$ is then defined as $|\nu(\mathbf{q})|$ and parameters with high scores are suggested as having wide variability in at least some of their estimates, whereas $\alpha(\mathbf{q})$ near zero indicates lower uncertainty. One method of subset selection in (Olufsen and Ottesen, 2011) involves findings parameters with correlation values $|c_{ij}| > 0.9$, and sequentially removing parameters from estimation until the remaining parameters all have correlation values $|c_{ij}| < 0.9$. The correlation matrix for patient 8 is shown in Figure 3.4.

Using the correlation matrix to assess parameter identifiability assumes a well-conditioned Fisher matrix, which as discussed in section 3.2.3, may be difficult to assume.

3.3.2 Subset Selection Using Rank-Revealing QR

Subset selection through QR based methods follows on SVD based techniques put forth in (Golub and Loan, 1996). Subset selection with QR seeks to determine the *most* linearly independent columns of the Fisher matrix F. The parameters corresponding to these columns are termed the identifiable subset. This is a rank-revealing QR, and a strong algorithm is developed in (Gu and Eisenstat, 1996), and described and summarized in (Pope, 2009).

The rank revealing nature comes from a column pivoting process inside the QR algorithm. Given a matrix $A \in \mathbb{R}^{m\times n}$, $m \geq n$, the factorization with pivoting builds an orthogonal matrix $Q \in \mathbb{R}^{m\times m}$, an upper triangular matrix $R \in \mathbb{R}^{m\times n}$, and a permutation matrix $P \in \mathbb{R}^{n\times n}$ such

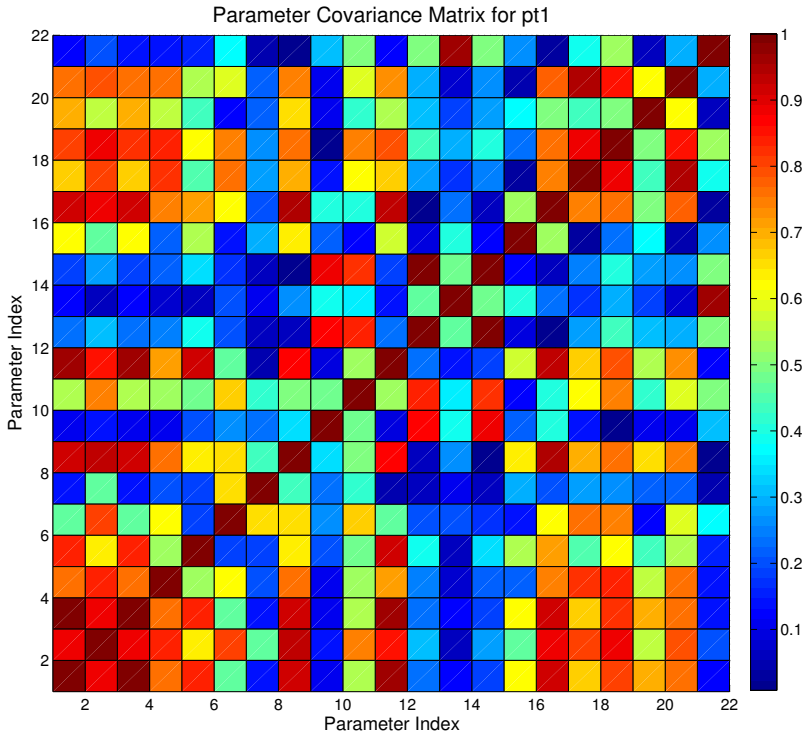

Figure 3.4: Correlation matrix for parameter estimates for patient 8 with no subset selection. The warmer colors indicate high parameter correlation.

that

$$AP = QR = \begin{bmatrix} Q_1 & Q_2 & Q_3 \end{bmatrix} \begin{bmatrix} R_{11} & R_{12} \\ 0 & R_{22} \\ 0 & 0 \end{bmatrix}. \tag{3.3.7}$$

The matrices are constructed in the following process:

1. Initialize: $i = 1, P = I, R = A = R_{22}, Q = I$.
2. Find j, the index of the column of R_{22} with greatest 2-norm.
3. Swap columns i and $i + j - 1$ in P and in R.
4. Use an orthogonal transformation Q_i that triangularizes the first i columns of R and assures non-negativity on the diagonal.
5. Update: $Q = QQ_i^\top, R = Q_i R$.
6. Repartition into
$$\begin{bmatrix} R_{11} & R_{12} \\ 0 & R_{22} \\ 0 & 0 \end{bmatrix},$$
where the first block R_{11} is moved left most at step i.
7. Update the counter i, and repeat steps 3 through 6 $n - 1$ times.

The orthogonal transformations at step five can be achieved through methods such as Gram-Schmidt, Householder transformations, or Givens rotations. The choice of norm used in step 2 is not unique. In (Kappel, 2009), different procedures of computing the QR factorization are outlined, and 1,2, and supremum norms are compared for the column selection step. It is shown that selection of the norm varies the permutation matrix P in some instances.

Another method of subset selection using QR utilizes a singular value decomposition prior to QR factorization to determine the numerical rank of the matrix F in a manner similar to

(3.2.9). A subset selection algorithm based on using SVD followed by QR with column pivoting is given next.

1. Given a Fisher matrix F, compute the numerical rank k.
2. Compute the eigen-decomposition of $F = V\Lambda V^\top$, with Λ a diagonal matrix of eigenvalues with corresponding eigenvectors as the columns of V.
3. Partition the eigenvector matrix V so that $V = [V_k \; V_{m-k}]$, where V_k are the first k eigenvectors.
4. Use QR with column pivoting on $V_k^\top$ so that $V_k^T P = QR$.
5. Use the partition matrix P to reorder the parameters in $\mathbf{q}$, $\mathbf{q}_\rho = P^\top \mathbf{q}$.

The identifiable parameters are then taken as the first k entries in $\mathbf{q}_\rho$. Using the model Hessian F in conjunction with SVD and QR is a process that is used in (Burth et al., 1999) in one of the earlier subset selection problems. This method is extended in numerous places, including (Miao et al., 2011; Olufsen and Ottesen, 2011; Vélez-Reyes, 1992; Heldt, 2004).

3.3.3 Subspace Selection

In this work we utilize a method based on minimizing distances between parameter subspaces and the eigenspace of the model Hessian $S^\top S$. This methodology is similar in principle to the orthogonal method described in (Yao et al., 2003), and seeks to examine the linear dependencies of the columns of F. Rather than use correlation based information, the method calculates the spanning distance of one column to the vector space spanned by the other columns. This methodology is first developed in (Kappel, 2009), and then expanded and described in detail in (Olufsen and Ottesen, 2011), and we reproduce the key results and notation here.

Subspace selection methods discussed thus far utilize the singular values of the Fisher matrix and the QR decomposition, where the parameter space is partitioned into an identifiable and non-identifiable subset. We begin this method by writing $F = V\Lambda V^\top$, where Λ is a

diagonal $m \times m$ matrix of eigenvalues, and the columns of V the eigenvectors. We first partition the sorted (from lowest magnitude to highest magnitude) eigenvalues and eigenvectors into two groups, $\{\lambda_1, \ldots, \lambda_k\}$ and $\{\lambda_{k+1}, \ldots, \lambda_m\}$ and the corresponding eigenvectors $\{v_1, \ldots, v_k\}$ and $\{v_{k+1}, \ldots, v_m\}$. These two groups correspond to the parameters which can and can not be identified, respectively. Then we denote the subspace spanned by the first k eigenvectors $\mathcal{W}_a = \text{span}\{v_1, \ldots, v_k\}$. The corresponding matrix $W_a = [v_1, \ldots, v_k]$ then has rank k and indicates the identifiable subspace. Therefore, W_a is the orthogonal projection onto the subspace $\mathcal{W}_a$.

The subspace, which when taken from all subspaces of dimension k spanned by canonical basis vectors is closest in distance to $\mathcal{W}_a$, is the subset of identifiable parameters. We compute all possible subspaces spanned by k canonical basis vectors, $\mathcal{E} = \text{span}\{e_{i_1}, \ldots, e_{i_k} \text{ s.t. } 1 \leq i_1 < \cdots < i_k \leq m\}$. The total number of subspaces is denoted $|\mathcal{E}|$, and is equal to $\binom{m}{k}$.

We move next to define a metric on all subspaces of dimension k so as to determine the minimal distance. Let S_1 and S_2 be subspaces of $\mathbb{R}^n$ with $\dim S_1 = \dim S_2 = k$ and let P_i be the orthogonal projection onto $S_i, i = 1, 2$. The distance between S_1 and S_2 is then defined as $\text{dist}(S_1, S_2) = \|P_1 - P_2\|_2$. Note that $\text{dist}(\cdot, \cdot)$ defines a metric on any set of subspaces of $\mathbb{R}^n$ with the same dimension. We then have the following key result.

Let W, Z be orthogonal matrices in $\mathbb{R}^{n \times n}$ and assume they are partitioned as $W = [W_1 \ W_2], Z = [Z_1 \ Z_2]$, where $W_1, Z_1 \in \mathbb{R}^{n \times k}$ and $W_2, Z_2 \in \mathbb{R}^{n \times n-k}$. Then for the subspaces $S_1 = \text{image } W_1$ and $S_2 = \text{image } Z_1$, we have $\text{dist}(S_1, S_2) = \left\|W_1^\top Z_2\right\|_2 = \left\|Z_1^\top W_2\right\|_2$. This result is provided by Theorem 2.6.1 in (Golub and Loan, 1996).

The subset of k identifiable parameters is then the subspace spanned by the canonical basis vectors which is closest in distance to $\mathcal{W}_a$. This is performed by minimizing the distance $\text{dist}(\mathcal{W}_a, E_{(i)_k}) \ \forall E_{(i)_k} \in \mathcal{E}$. Here the notation $(i)_k$ represents an ordered multi-index of length k, $(i)_k = (i_1, \ldots, i_k)$ with $1 \leq i_1 < \cdots < i_k \leq m$. This algorithm can be summarized as follows.

1. Given a parameter vector q, compute the sensitivity matrix S and the Fisher matrix $F = S^\top S$.

2. Compute the numerical rank k of matrix F.

3. Compute the eigenvalue decomposition $F = W\Lambda W^\top$, and partition the matrix of eigenvectors into $W = [W_k \; W_{m-k}]$

4. Calculate $\mathcal{W}_a = \text{span}\{W_1, \ldots, W_k\}$.

5. Generate all subspaces spanned by the canonical basis vectors $E_{(i)_k} = \text{span}\{e_1, \ldots, e_k\}$, and calculate the distance to $\mathcal{W}_a$ given by $\left\|(W_a^\top)E_{(i)_k}\right\|_2$.

6. Find the subspace with the smallest spanning distance over all $E_{(i)_k}$, and pick out the associated parameters.

This procedure will give you a subset of parameters whose span is closest to the eigenspace corresponding to the k most significant values. The subsets that are selected by the subspace selection algorithm, and the others, are not necessarily unique, and can (frequently) include non-sensitive parameters.

The subset of identifiable parameters is denoted ρ, with the subset of non-identifiable parameters $\overline{\rho}$. The parameters in $\overline{\rho}$ can be fixed to some nominal literature values. While this may not affect system output, fixing these parameters introduces a statistical bias into the parameter estimates. In this work, rather than fixing to a nominal a-priori value, we perform a limited (in iteration) global based optimization process and fix the parameters given in $\overline{\rho}$ to these results, reducing parameter bias. This process is explained in detail in section 4.2. In (Banks et al., 2008), the non-estimated parameters were fixed to population averages, rather than a patient specific estimate as in this work.

3.4 Combination Subset Selection and Sensitivity Analysis

The subset selection techniques up to this point may yield parameters that, while being linearly independent in some sense, may not be sensitive in the model output and thus (still) difficult to estimate. We would ultimately like to have a subset of sensitive and identifiable parameters,

and this can be accomplished by applying the subset selection methodology and comparing the results to the sensitivity analysis, and then successively reapplying the subset algorithms to a smaller, more sensitive group of parameters. Formally, this is accomplished as follows:

1. Perform subset selection sensitivity analysis, and numerical rank calculations on the entire set of parameters. For purposes of this algorithm, we will denote the k most sensitive parameters ρ^*, where k is the numerical rank of the Fisher matrix.

2. Remove the least sensitive parameter in the identifiable subset by fixing it to a literature or other value.

3. Reapply the subset algorithm to the reduced subset of parameters.

4. Continue steps 2 and 3 until the identifiable subset contains only parameters that are in the original set ρ^*.

This process will yield a subset of parameters that are linearly independent and sensitive. In practice, this involves regenerating a new, smaller, sensitivity matrix at each iteration. As a result, the numerical rank k can change over iterations, however when performing our simulations we keep k as the original numerical rank of the full sensitivity matrix.

3.4.1 Standard Error Analysis & Correlation Coefficients

One method of determining the effectiveness of different parameter estimation techniques is to look at the so-called standard errors of the resulting model predictions. To develop this idea mathematically we follow the notation in (Banks and Tran, 2009).

Consider a generic statistical model given by

$$Y_k \equiv f(t_k, q_0) + \mathcal{E}_k, \quad k = 1, \ldots, n, \tag{3.4.1}$$

where $f(t_k, q_0)$ is the dynamic model for the observations of the true states given the "truth" parameter $q_0 \in \mathbb{R}^p$. The observations are corrupted by errors $\mathcal{E}_k$ which are assumed to be

independently and identically distributed (i.i.d) random variables with zero mean and variance σ_0^2, where σ_0^2 is unknown. The observations Y_k are therefore random variables with mean $f(t_k, q_0)$ with variance σ_0^2.

To determine the unknown parameter value q, an ordinary least squares (OLS) process can be used. We use the realizations $\{y_k\}$ of the process $\{Y_k\}$ along with the model to compute q^* where

$$q^* = \arg\min J(q) = \sum_{k=1}^{n} \big(y_k - f(t_k, q)\big)^2. \tag{3.4.2}$$

Since Y_k is a random variable, so is q^*, with a distribution called the sampling distribution. This distribution characterizes all possible values that q^* could take on for a fixed data set of length n. The standard errors of q^* quantify the uncertainty in the estimation of q^* for a given data set and estimate of q^*.

The sample distribution can be estimated by a p-multivariate Gaussian distribution (with asymptotic convergence to distribution) with mean q_0 and covariance matrix $\Sigma_0 \approx \sigma_0^2\big(S^\top(q_0)S(q_0)\big)^{-1}$, where $S(q_0)$ is the local, regular sensitivity matrix evaluated at q_0. Therefore for n large, the sampling distribution approximates

$$q^*(Y) \sim \mathcal{N}_p(q_0, \Sigma_0) \approx \mathcal{N}_p\big(q_0, \sigma_0^2\big(S^\top(q_0)S(q_0)\big)^{-1}\big). \tag{3.4.3}$$

We have already discussed methods to compute the sensitivity matrix S, and to compute σ_0^2 we can use the simple approximation

$$\sigma_0^2 \approx \hat{\sigma}^2 = \frac{1}{n-p} \sum_{k=1}^{n} \big(y_k - f(t_k, q^*)\big)^2. \tag{3.4.4}$$

Standard errors are then given by the square-root of diagonal elements of the covariance matrix

(Casella and Berger, 2002),

$$e_k(q) = \sqrt{\Sigma_{kk}(q)}, k = 1, \ldots, p. \tag{3.4.5}$$

These errors can be used to construct confidence intervals at the normal $100(1-\alpha)\%$ level. The confidence measure is defined as

$$P\{q - t_{1-\alpha/2} e_k(q) < q_0 < q + 1_{t_{\alpha/2}} e_k(q)\} = 1 - \alpha, \tag{3.4.6}$$

where $\alpha \in [0, 1]$ and $t_{1-\alpha/2} \in \mathbb{R}^+$. For small values of α, such as 0.05, the values $t_{1-\alpha/2}$ are pulled from a student's t distribution with $n - p$ degrees of freedom.

3.5 Inverse Problem Formulation & Censored Data

We wish to use the HIV model given in (2.3.1) to describe clinically collected data and to make patient specific predictions. Before this can be done we must calibrate the model parameters to the clinical data through the inverse problem methodology.

The inverse problem is patient specific; we use data from a single patient j in order to estimate the patient specific parameter value q^j. Whereas in prior sections we have used $\hat{\mathbf{y}}(t_k)$ to denote model observations, we now switch notation, letting $z_s(t_s^{ij}), s = 1, 2$ denote the model observation for patient j, output s, at time index i. The goal function is given by

$$q^{\star j} = \arg\min J(q) = \sum_{s=1}^{2} \frac{1}{N_s^j} \sum_{i=1}^{N_s^j} \left| z_s(t_s^{ij}; q) - y_s^{ij} \right|^2, \tag{3.5.1}$$

over an admissible parameter space $\Omega \subseteq \mathbb{R}^\ell$, where ℓ is the number of parameters being identified by the least squares process.

Because the model solutions and collected data vary significantly in orders of magnitude, we fit and simulate the $\log_{10}$ scaled version of the model and data. From a statistical standpoint,

minimizing (3.5.1) is equivalent to computing a maximum likelihood estimate of q, assuming the log scaled measurements y_s^{ij} are normally distributed,

$$y_s^{ij} \sim \mathcal{N}(z_s(t^i; q^0), \sigma_s^2), \quad s = 1, 2,$$

for some true parameter value q^0 and variance σ_s^2. When viral load data are below the censor level, the observed values do not represent the truth value for the data value. The optimization process must be modified to account for this information, and is accomplished through standard techniques for censored data analysis.

In working with the censored data we exploit the fact that the log scaled data are normally distributed. Let $L_1 = \log_{10} 400$ and $L_2 = \log_{10} 50$, denoting the low and high sensitivity assay limits, respectively. For the censored data, the only knowledge available is that $y_2^i \leq L_i$. In this context, we observe pairs $(w^i, \chi^i), i = 1, \ldots, N$ where,

$$w^i = \begin{cases} y_2^i & \text{if } y_2^i > L_i \\ L_i & \text{if } y_2^i \leq L_i, \end{cases}$$

$$\chi^i = I_{\{y_2^i > L_i\}},$$

and I_A is the indicator function for the set A. Next we define the standard normal probability distribution function (pdf),

$$\varphi(\xi) = \frac{1}{\sqrt{2\pi}} \exp(-\xi^2/2) \tag{3.5.2}$$

with corresponding cumulative distribution function (cdf),

$$\Phi(\xi) = \int_{-\infty}^{\xi} \varphi(s) ds. \tag{3.5.3}$$

The viral load portion of the likelihood function given the observation w^i is

$$\overline{\mathcal{L}}(q,\sigma_2) = \prod_{i=1}^{N} \left[\frac{1}{\sigma_2}\varphi\left(\frac{w^i - z_2^i}{\sigma_2}\right)\right]^{\chi^i} \left[\Phi\left(\frac{w^i - z_2^i}{\sigma_2}\right)\right]^{1-\chi^i}, \tag{3.5.4}$$

where the first term accounts for the probability of observing w^i given it is not censored, and the second term accounting for the probability that the observation is censored and in the interval $(-\infty, L_i)$. Equation (3.5.4) is obtained under the assumptions of a truncated normal distribution for the censored observations. The corresponding log-likelihood is then

$$\begin{aligned}\mathcal{L}(q,\sigma_2) &= \sum_{i=1}^{N}\left(\chi^i\left[\log\varphi\left(\frac{w^i - z_2^i}{\sigma_2}\right) - \log\sigma_2\right] + (1-\chi^i)\left[\log\Phi\left(\frac{w^i - z_2^i}{\sigma_2}\right)\right]\right)\\ &= \sum_{i=1}^{N}\left(\chi^i\left[\log\varphi\left(\frac{y_2^i - z_2^i}{\sigma_2}\right) - \log\sigma_2\right] + (1-\chi^i)\left[\log\Phi\left(\frac{L_i - z_2^i}{\sigma_2}\right)\right]\right),\end{aligned} \tag{3.5.5}$$

which is then maximized to estimate q and σ_2. We use the Expectation Maximization (EM) (Dempster et al., 1977; McLachlan and Krishnan, 2008) algorithm to obtain the maximum value. EM works by iteratively updating the estimates for q, σ_2 until the maximum is reached. Defining $\xi^i = \frac{L_i - z_2^i}{\sigma_2}$ and $\Lambda(\xi^i) = \frac{\varphi(\xi^i)}{\Phi(\xi^i)}$ and using properties of the truncated normal distribution, we have

$$\begin{aligned}E[y_2^i \mid y_2^i \le L] &= z_2 - \sigma_2\Lambda(\xi^i),\\ E[(y_2^i)^2 \mid y_2^i \le L] &= (z_2^i)^2 - 2\sigma_2 z_2^i\Lambda(\xi^i) - \sigma_2^2\xi^i\Lambda(\xi^i) + \sigma_2^2.\end{aligned}$$

These are used to update the censored data and squared residuals through

$$\begin{aligned}\tilde{y}^i &= \chi^i y_2^i + (1-\chi^i)E[y_2^i \mid y_2^i \le L]\\ &= \chi^i y_2^i + (1-\chi^i)[z_2^i - \sigma_2\Lambda(\xi^i)],\end{aligned} \tag{3.5.6}$$

and for the residuals

$$\begin{aligned}\tilde{r}^i &= \chi^i E[(y_2^i - z_2^i)^2] + (1-\chi^i)E[(y_2^i - z_2^i)^2 \mid y_2^i \le L] \\ &= \chi^i (y_2^i - z_2^i)^2 + (1-\chi^i)\big[E[(y_2^i)^2 \mid y_2^i \le L_i] - 2z_2^i E[(y_2^i) \mid y_2^i \le L_i] + (z_2^i)^2\big] \\ &= \chi^i (y_2^i - z_2^i)^2 + (1-\chi^i)\sigma_2^2[1 - \xi^i \Lambda(\xi^i)]. \end{aligned} \tag{3.5.7}$$

The EM Algorithm is then given by the following.

1. Adjust the censored data down by half to $L_i/2$ and use ordinary least squares to estimate $\hat{q}^{(0)}$ using both the CD4$^+$ data y_1^i and viral load $\tilde{y}^i$, which includes the replaced censored data. Note that for censored data $\chi^i = 0$. Set $k = 0$ and compute an initial estimate for σ_2^2 from

$$(\hat{\sigma}_2^{(0)})^2 = \frac{1}{N_2}\sum_{i=1}^{N_2} \big|\tilde{y}^i - z_2(t_2^i; \hat{q}^{(0)})\big|^2.$$

2. Define $\hat{z}_2^{i(k)} = z_2(t^i; \hat{q}^{(k)})$ and $\hat{\xi}^{i(k)} = \frac{L_i - \hat{z}_2^{i(k)}}{\hat{\sigma}_2^{(k)}}$ and update the data and residuals by

$$\tilde{y}^{i(k)} = \chi^i y_2^i + (1-\chi^i)\big[\hat{z}_2^{i(k)} - \hat{\sigma}_2^{(k)}\Lambda(\hat{\xi}^{i(k)})\big] \tag{3.5.8}$$

$$\tilde{r}^{i(k)} = \chi^i (y_2^i - \hat{z}_2^{i(k)})^2 + (1-\chi^i)(\hat{\sigma}_2^{(k)})^2\big[1 - \hat{\xi}^{i(k)}\Lambda(\hat{\xi}^{i(k)})\big]. \tag{3.5.9}$$

3. Perform ordinary least squares to compute $\hat{q}^{(k+1)}, \hat{\sigma}_2^{(k+1)}$ by minimizing over q

$$\hat{q}^{(k+1)} = \arg\min \sum_{i=1}^{N_1} \big|y_1^i - z_1(t_1^i; q)\big|^2 + \sum_{i=1}^{N_2} \big|\tilde{y}^{i(k)} - z_2(t_2^i; q)\big|^2$$

and then computing and update to $\hat{\sigma}_2$,

$$\big(\hat{\sigma}_2^{(k+1)}\big)^2 = \frac{1}{N_2}\sum_{i=1}^{N_2} \tilde{r}^{i(k)}.$$

If the relative change in $\hat{q}, \sigma$ are small, terminate. Else set $k = k+1$ and continue.

The end result of EM yields estimates of the parameters, the expected value and variance of true values for the censored data. This information can then be used to compute standard errors and confidence measures for the parameter estimates, as discussed in section 3.4.1.

CHAPTER 4

Model Validation Results & Simulation

For each patient, we estimate a patient specific parameter set, including the seven initial conditions for the model. The following process is applied to each patient.

1. Solve the sensitivity equations (3.1.13) and build the sensitivity function matrix and corresponding Fisher matrix.

2. Compute the numerical rank k of the Fisher matrix $F = S^\top S$.

3. Determine the minimal spanning subset of parameters, determining the identifiable subset of parameters ρ_j by searching through all $\binom{n}{k}$ subsets.

4. Estimate the entire parameter and initial condition set. This involves a preliminary optimization with a sampling based algorithm. After only a few iterations of this algorithm we apply the EM algorithm in conjunction with a gradient based optimizer. The resulting parameter estimates are saved as $\overline{\rho}$, or the non-identifiable subset of parameters.

5. Operating only on the identifiable subset ρ, recompute the parameter estimates using the

results of the prior step as the initial optimization iterate.

4.1 Sensitivity Results

Because our sensitivity analysis is a local concept about the nominal patient parameter, the only distinguishing difference between the patients is adherence to the clinical control. It is the adherence patterns that directly effect which parameters come up as sensitive when performed as a local analysis.

Here we present representative results from two patients, patient 11 and patient 8. With the other patients, we see similar patterns in the sensitivity results, the clustering of parameters into distinct groups. Clustering of the parameters is evident in Figures 4.1, 4.2. In general the number of parameters in the most sensitive cluster nearly agrees with the numerical rank calculation.

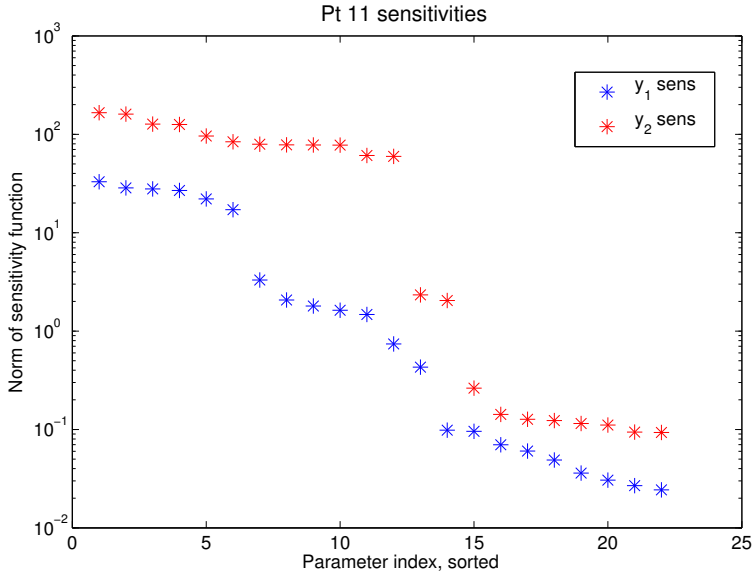

Figure 4.1: Pt 11 sensitivity values. Note that the clustering in the y_1 sensitivities agrees with the $k = 6$ rank calculation.

Table 4.1: Sensitivity results and values for patient 11. The line between parameters 6 and 7 indicate the numerical rank of the corresponding sensitivity matrix, $k = 6$.

Parameter	y_1 sens		Parameter	y_2 sens
λ_1	32.932		N_T	165.67
N_T	28.54		c	160.47
c	27.865		δ	127.04
k_1	26.885		ε_2	125.82
d_1	22.031		λ_2	95.972
δ	17.114		λ_1	83.854
λ_2	3.3046		k_1	79.207
d_2	2.0706		d_2	78.125
ε_2	1.795		ε_1	77.941
k_2	1.6258		k_2	77.697
ε_1	1.4738		d_1	60.849
f	0.74023		f	59.674
ρ_1	0.42985		ρ_2	2.3311
K_d	0.098362		ρ_1	2.0439
ρ_2	0.095749		K_d	0.26292
b_E	0.06998		d_E	0.14215
δ_E	0.060345		b_E	0.12699
λ_E	0.048854		m_1	0.12328
m_2	0.036065		δ_E	0.11497
d_E	0.030544		m_2	0.11106
m_1	0.026841		λ_E	0.093993
K_b	0.024296		K_b	0.093307

4.2 Patient Specific Subsets

For each patient, an a-priori sensitivity analysis is conducted at a nominal parameter value $\mathbf{q}^*$. The parameter values and descriptions are given in Table 2.2. As each patient model is evaluated at the same $\mathbf{q}^*$, the only factor that affects parameter identifiability is the patient adherence to treatment.

Before a subset specific estimate is formed, we first attempt to estimate all 22 parameters using the global optimization algorithm DIRECT with the EM algorithm. DIRECT is a bounded, sampling method based on dividing hyper-rectangles (Finkel and Kelley, 2004). The bounds

Table 4.2: Sensitivity results and values for patient 8. The line between parameters 6 and 7 indicate the numerical rank of the corresponding sensitivity matrix, $k = 6$.

Parameter	y_1 sens		Parameter	y_2 sens
d_1	26.479		N_T	267.82
λ_1	25.299		c	259.78
N_T	24.696		λ_1	246.66
k_1	24.11		d_1	232.4
c	23.982		k_1	230.25
δ	1.7784		δ	41.111
ρ_1	0.69431		λ_2	33.376
ε_2	0.63088		ε_2	30.394
λ_2	0.57694		ε_1	22.91
ε_1	0.57118		k_2	14.899
k_2	0.028845		ρ_1	7.5254
f	0.077068		f	3.8677
K_b	0.020505		ρ_2	1.5751
δ_E	0.019747		d_2	1.4192
ρ_2	0.01822		K_b	0.27754
m_1	0.15668		δ_E	0.21042
m_2	0.012461		m_2	0.16566
λ_E	0.012172		m_1	0.15609
d_2	0.011235		K_d	0.15279
K_d	0.009063		λ_E	0.14555
δ_E	0.0074273		δ_E	0.12937
b_E	0.004447		b_E	0.070737

for each parameter and initial state value are given in Table 4.3. The termination criterion for DIRECT is the number of iterations, and we only allow 3. DIRECT is sometimes used to determine a "better" local initial iterate for gradient based optimizers, and our approach is similar in this regard. The parameters that are not being estimated are fixed to the output of DIRECT optimization, and it is at this point we begin the gradient based method with EM to further refine the parameter estimates on the identifiable subset.

The number of identifiable parameters range from $k = 6$ to $k = 10$. In addition to these parameters, all seven initial conditions were estimated for each patient. Each patient subset is given in Table 4.4.

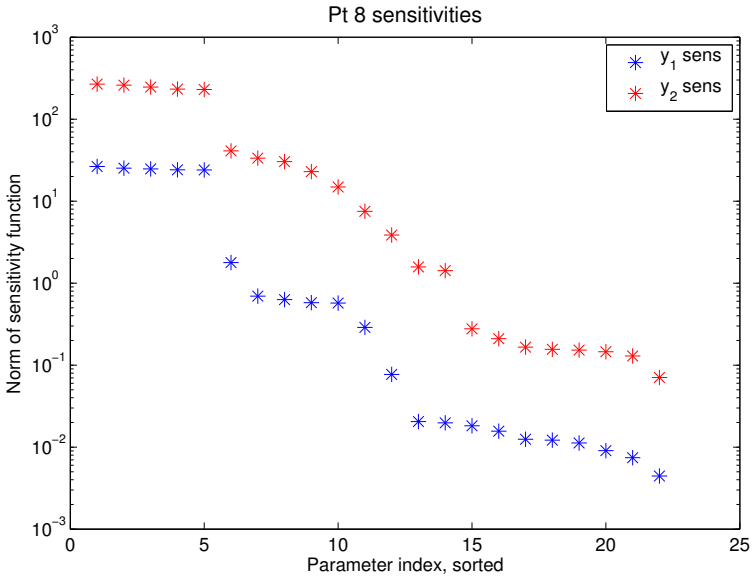

Figure 4.2: Pt 8 sensitivity values. Clustering of the parameters is evident.

Common in nearly all subsets are infection and immune induced clearance rates for both populations, k_i, m_i. Other identifiable parameters range in sensitivity, but their presence in the identifiable set indicates their mutual linear independence. Moreover, an identifiable subset is not unique.

4.3 Patient Specific Model Fits

In each figure the collected clinical data are given, as well as the adjusted data calculated through the expectation maximization algorithm. The forward model prediction is given in red. In general, the fits are good and the model prediction follows the clinical adherence pattern.

For some patients, namely patients 7, 9, and 12, the $CD4^+$ fits are not as good as they possibly could be. The parameter estimates are computed in batch in a parallel manner; estimates could be refined by restarting the optimizations in a patient specific manner. Additionally for these patients, there is a definite lack of data compared to some of the other patients. This is reflected in the goodness of fits in Figures 4.9, 4.10, and 4.14. There is no preference assigned to

Table 4.3: Bounding boxes used by DIRECT for initial estimation of model parameters and initial conditions. All bounds are $\log_{10}$ scaled from the values given below.

Parameter	Lower Bound	Upper Bound		Parameter	Lower Bound	Upper Bound
λ_1	0.1	100		λ_2	0.001	1
d_1	0.001	0.1		d_2	0.001	0.1
k_1	10^{-9}	10^{-4}		k_2	10^{-7}	0.01
m_1	10^{-8}	0.001		m_2	10^{-8}	0.001
ρ_1	0.99	1.01		ρ_2	0.99	1.01
δ	0	0.01		c	1	100
f	0	1		N_T	10	1000
λ_E	10^{-5}	.1		δ_E	0.01	0.9886
K_b	0.001	10		K_d	0.001	10
ε_1	10^{-6}	0.99		ε_2	10^{-6}	0.99
$T_1(0)$	$0.9 \cdot \min T_1(t)$	$1.1 \cdot \max T_1(t)$		$T_2(0)$	0.1	100
$T_1^*(0)$	0.01	$\max T_1(t)$		$T_2^*(0)$	10^{-5}	100
$V_I(0)$	100	10^6		$V_{NI}(0)$	1	10^4
$E(0)$	10^{-4}	10				

either the viral load or $CD4^+$ counts in the goal function (3.5.1), though by weighting towards the $CD4^+$ dataset, we perhaps could see an improved fit. This, however, may not be necessary, as there is in general a higher interest in the viral load progression than solely the $CD4^+$ count.

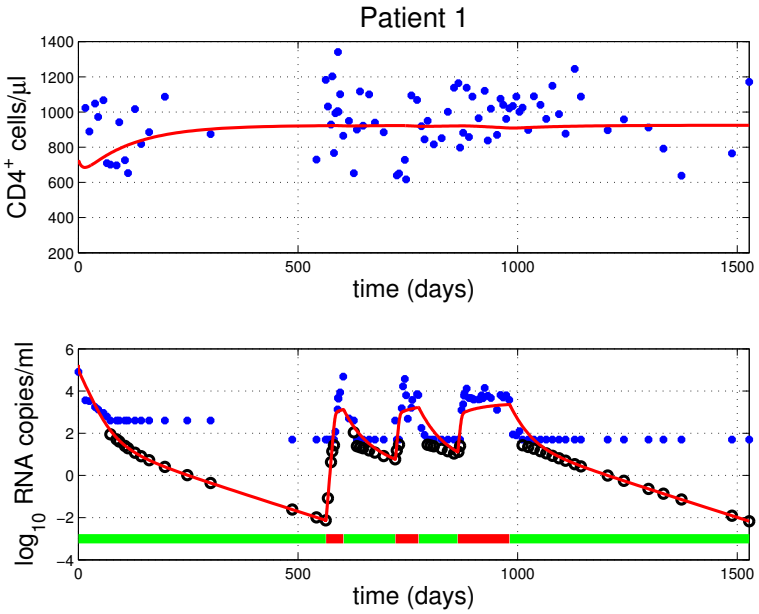

Figure 4.3: Collected and adjusted data with model prediction for patient 1.

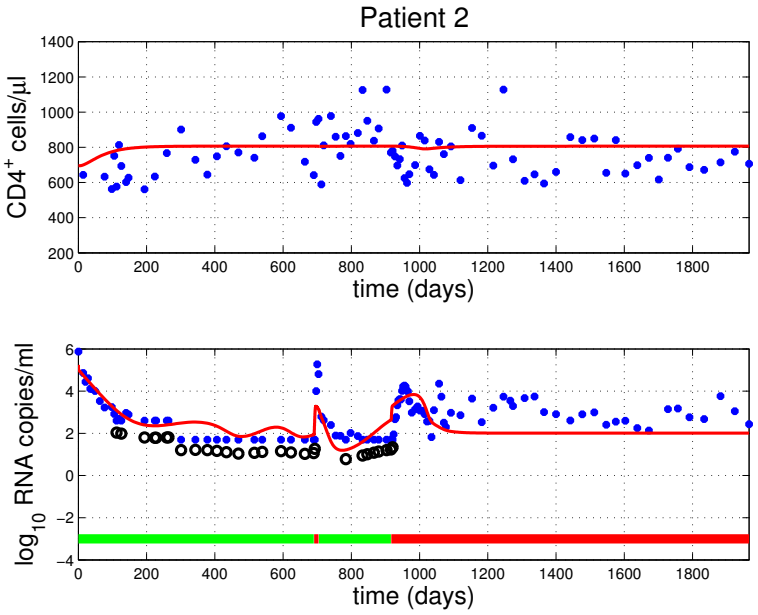

Figure 4.4: Collected and adjusted data with model prediction for patient 2.

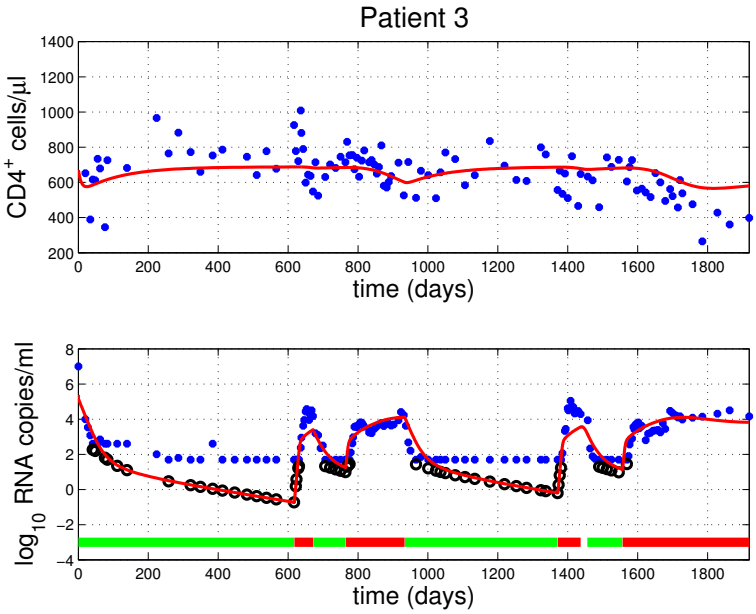

Figure 4.5: Collected and adjusted data with model prediction for patient 3.

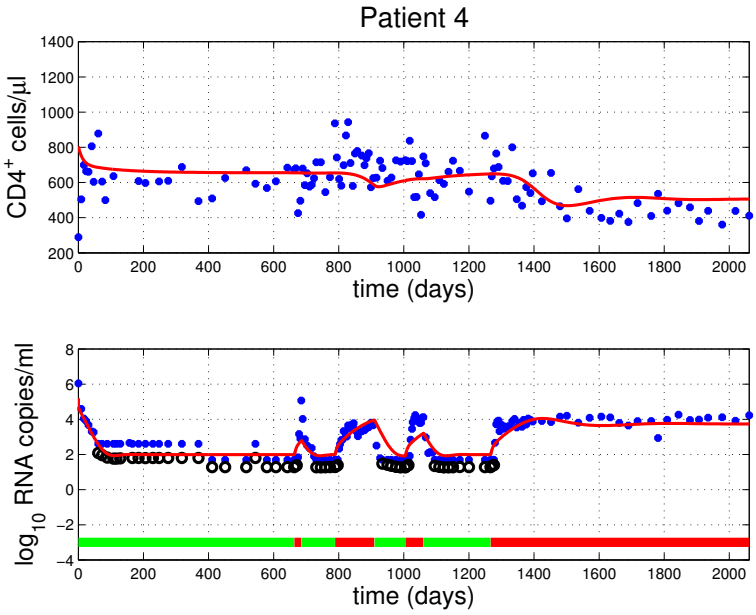

Figure 4.6: Collected and adjusted data with model prediction for patient 4.

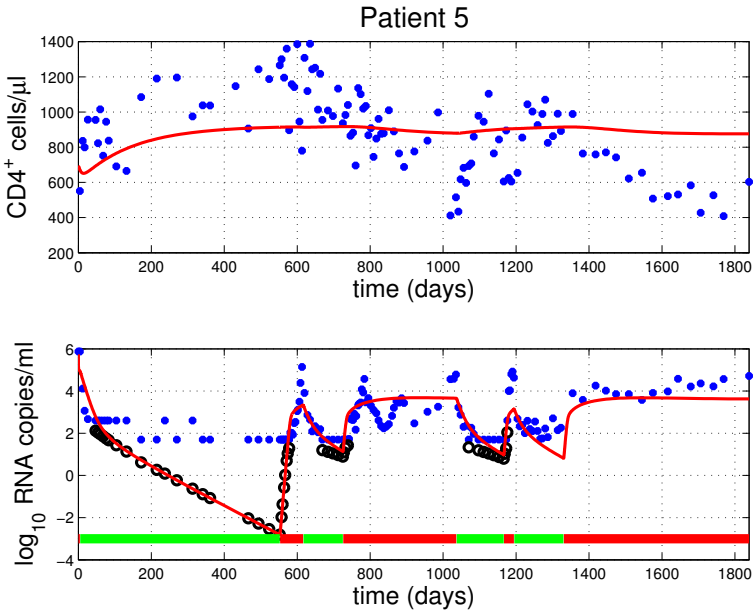

Figure 4.7: Collected and adjusted data with model prediction for patient 5.

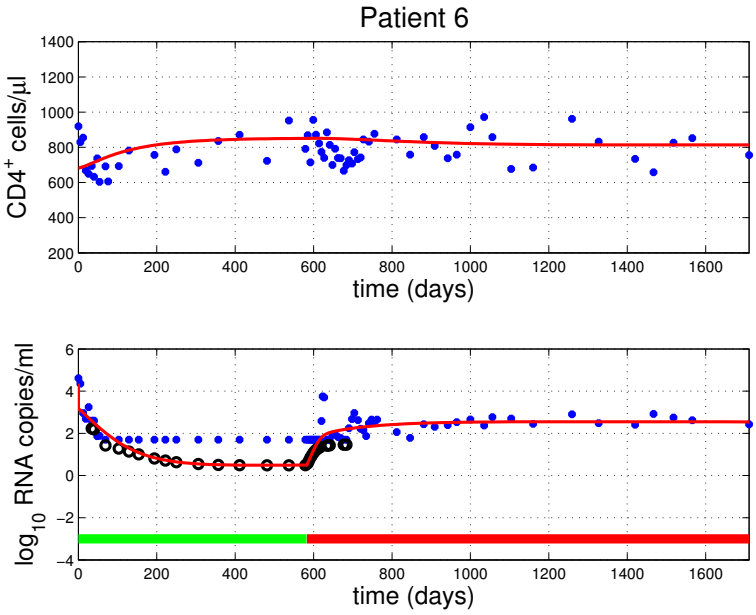

Figure 4.8: Collected and adjusted data with model prediction for patient 6.

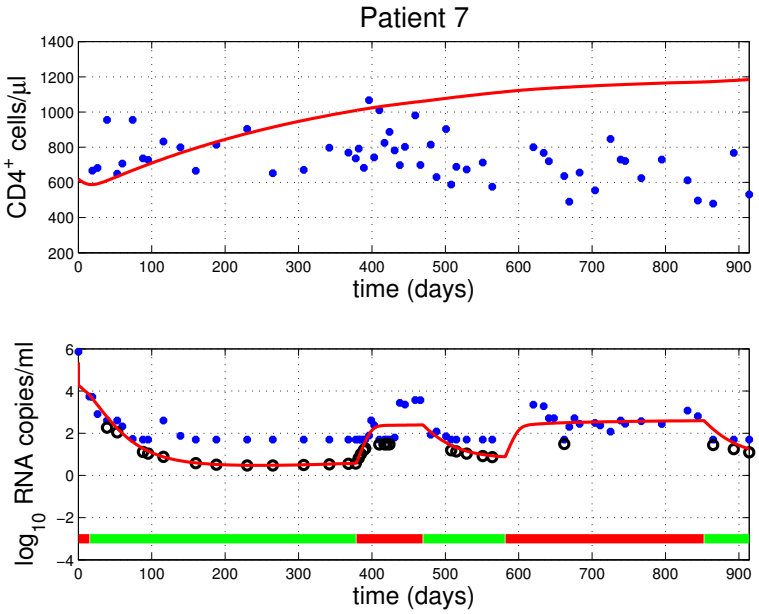

Figure 4.9: Collected and adjusted data with model prediction for patient 7.

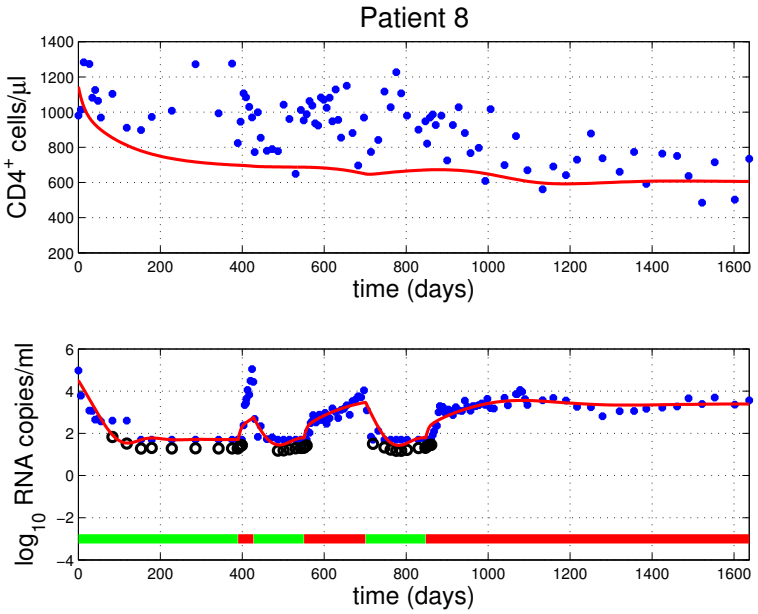

Figure 4.10: Collected and adjusted data with model prediction for patient 8.

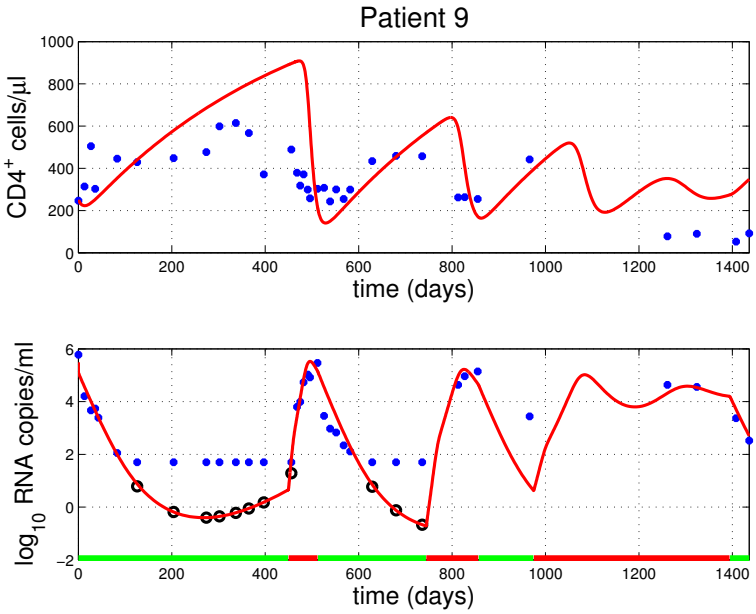

Figure 4.11: Collected and adjusted data with model prediction for patient 9.

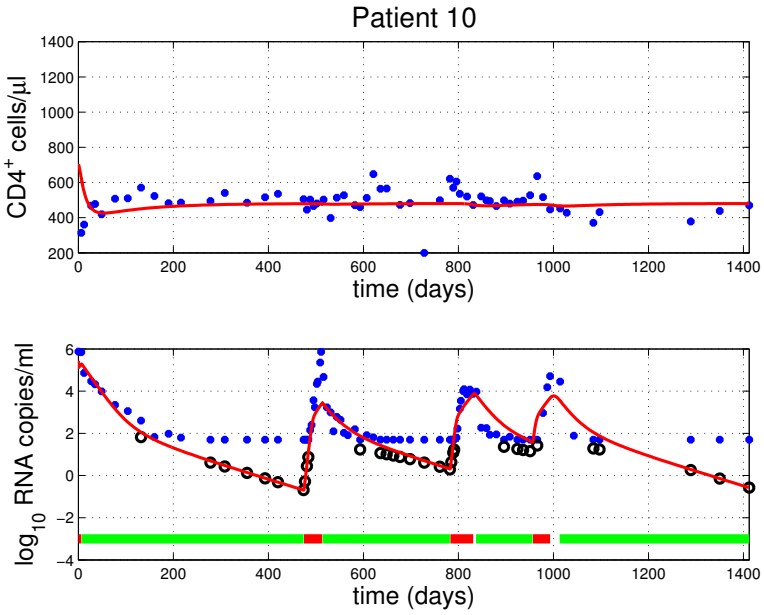

Figure 4.12: Collected and adjusted data with model prediction for patient 10.

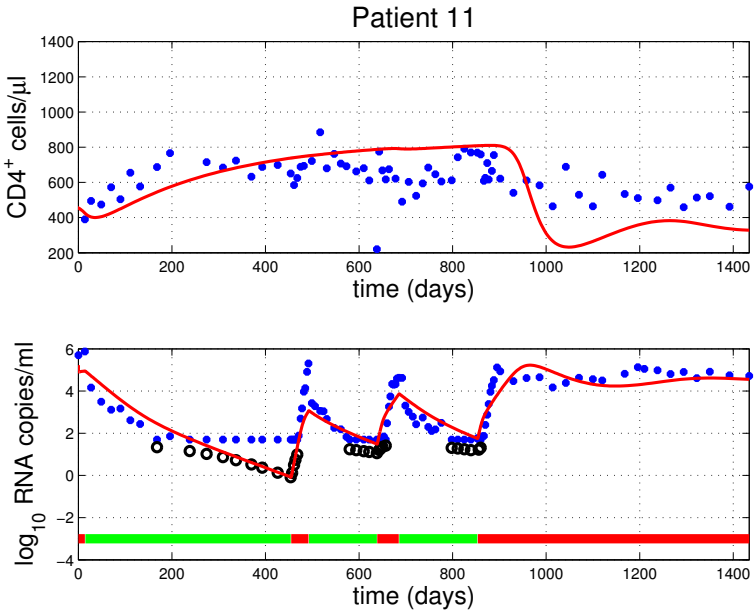

Figure 4.13: Collected and adjusted data with model prediction for patient 11.

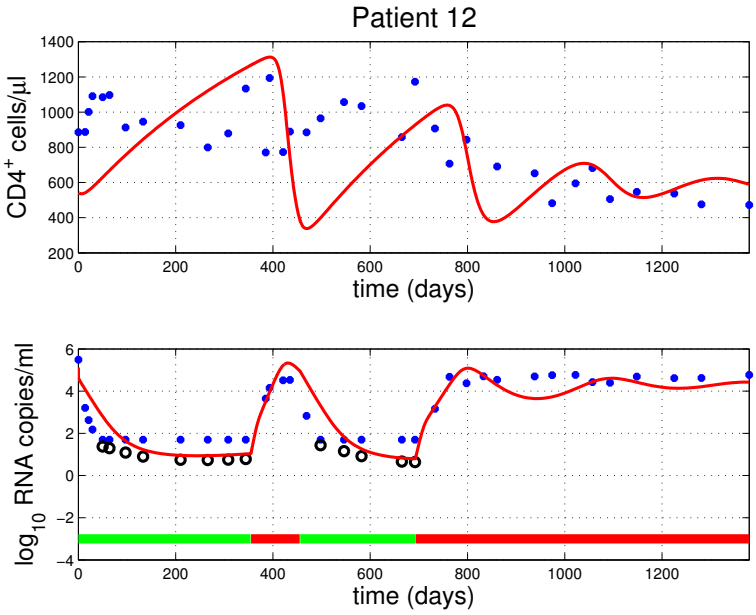

Figure 4.14: Collected and adjusted data with model prediction for patient 12.

Table 4.4: Patient specific identifiable subsets

Patient	Identifiable Subset ρ
1	$d_1, k_1, k_2, m_1, m_2, b_E, K_b, \varepsilon_2$
2	$d_1, k_1, k_2, m_1, m_2, f, d_E, \varepsilon_1$
3	$\lambda_2, k_1, k_2, m_1, m_2, d_E, K_b, \varepsilon_2$
4	$d_1, k_1, k_2, m_1, m_2, N_T$
5	$d_2, k_1, k_2, m_1, m_2, \delta, c, \varepsilon_2$
6	$\lambda_2, k_1, k_2, m_1, m_2, c, f, b_E, \varepsilon_1$
7	$d_2, k_1, k_2, m_1, m_2, \delta, \delta_E, \varepsilon_1$
8	$\lambda_2, k_1, k_2, m_1, m_2, N_T, d_E, K_d$
9	$\lambda_2, k_1, k_2, m_1, m_2, \delta, K_b, K_d$
10	$d_1, k_1, k_2, m_1, m_2, f, K_b, K_d$
11	$d_2, k_1, k_2, m_1, \rho_1, \rho_2, d_E, K_d$
12	$\lambda_2, k_1, m_2, m_1, m_2, \lambda_E, b_E, \varepsilon_2$
13	$\lambda_2, k_2, m_1, \rho_1, \delta, c, \varepsilon_1$
14	$\lambda_2, k_1, k_2, m_1, m_2, N_T, d_E, \varepsilon_2$

CHAPTER 5

Optimal Treatment Protocols

In this chapter we seek to compute optimal treatment strategies by calculating an optimal schedule for treatment. There have been various investigations into using optimal control based methodologies for the treatment of HIV through a variety of methods (Brandt and Chen, 2001; Ge et al., 2005; Joshi, 2002). Due to drug resistance, evolution of new viral strains, serious patient side effects, and imperfect patient compliance, the ability of HAART treatment to effectively control the virus long term fails in a high proportion of patients (Bajaria et al., 2004). Therefore concerns regarding the constant use of anti-retroviral therapies strongly motivate the investigation into their optimal use. In general, control of an HIV model is difficult due to the inherent nonlinearities in the mathematical model. In (Ge et al., 2005) a Lyapunov based design is used; in (Brandt and Chen, 2001) the idea of a feedback controller through a drug regimen is introduced onto a three dimensional HIV model. In the work done by (Adams et al., 2004; David, 2007; David et al., 2008), a calculus of variations approach was used to compute a control function. This was done either by a direct approach using the Pontryagin Minumum Principle, or an optimization approach as detailed in (Gunzburger, 2003).

These methods derive a continuous control function, which may be difficult to implement in a clinical setting, where treatment can only be altered at discrete intervals. To that end, we seek to compute treatment schedules that would be compatible with real world treatment scenarios.

5.1 Control Formulations

Together with the model equations (2.3.1), we consider a control problem defined by the objective function,

$$J\big(u(t)\big) = \int_0^{t_f} \left[QV(t) + R_1\varepsilon_1(t)^2 + R_2\varepsilon_2(t)^2 - SE(t)\right]\ dt, \tag{5.1.1}$$

where $\varepsilon_1(t) = \varepsilon_1 u(t)$, an RTI, and $\varepsilon_2(t) = \varepsilon_2 u(t)$, a PI. The values ε_1 and ε_2 are the patient specific values estimated in Chapter 4. The parameters Q, R_1, R_2 and S represent the control weights for the virus, control input, and immune response, respectively. By varying these weights, the optimal control can be customized in a patient specific sense, and exact treatments can be designed. By including the control terms in (5.1.1) we attempt to minimize the systematic cost of the drugs both in the sense of physiologic side effects as well as the monetary treatment cost. A similar optimal control based treatments are considered in (Culshaw et al., 2004; Kirschner et al., 1997), but with simpler dynamics with nominal parameter values, as well as simplified goal functions.

In (Adams et al., 2004), the authors use a direct search approach to find an optimal treatment by discretizing the control over a treatment interval. It is this technique that we further develop and implement. The control functions are represented as $\varepsilon_1(t) = \varepsilon_1 u(t)$ and $\varepsilon_2(t) = \varepsilon_2 u(t)$, where $u(t)$ is a binary treatment function, with a 1 representing a patient taking the drug during the given treatment interval, and a 0 meaning no treatment taken for that specific time period. The quality of control can depend on the treatment interval, which is the minimum of time necessary to change the treatment protocol. A sample control input is shown in Figure 5.1. Computing control functions of the form shown in Figure 5.1 is akin to calculating a series of

structured treatment interruptions, rather than a continuous treatment. These treatments are described further in the next section.

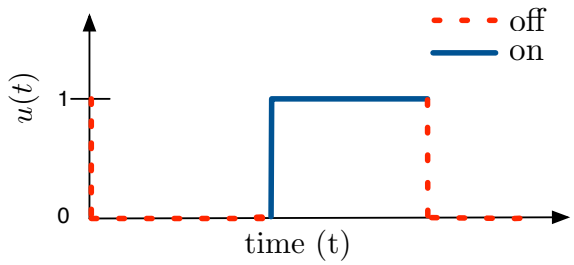

Figure 5.1: A sample control input for the HIV model. The switching of treatment represents a structured treatment interruption.

Mathematically, we represent the control numerically with a pair of vectors describing the control function $u(t)$ and the corresponding time intervals the control applies to. For example, for a treatment regimen with a treatment interval of $m = 15$ days, a possible control could be $u(t) = [1\ 1\ 0\ 0\ 1]$. This corresponds to an overall treatment length of 75 days with on treatment for the first 30 days, off for the next 30, and then back on for the final 15 days of treatment. For any treatment interval with n switchings of treatment, there exists a possible 2^n ways to define a treatment. Because this is a finite set, there exists an optimal control that minimizes (5.1.1). We denote the set of all control vectors $\mathcal{U}$, with $|\mathcal{U}| = 2^n$. We then seek the optimal control,

$$u^* \in \mathcal{U} = \arg\min J\big(\varepsilon_1(t), \varepsilon_2(t)\big). \tag{5.1.2}$$

On average, the data sets contain four years of longitudinal patient data, therefore $n \approx 1400$ days. Because n is relatively large, modeling daily changes in treatment is both impractical from a computational and clinical stand point. In practice, an interval of two to three weeks is clinically appropriate and is what we simulate.

If we assume a 30 day treatment interval, modeling the entire control interval could involve

the calculation of nearly 50 control values, or switches. As the problem is defined, $m = 30$ and $n \approx 1400$ would involve testing upwards of 2^{50} possible treatment protocols; clearly this is not computationally tractable. To mitigate this issue, we seek to reduce the number of iterations necessary to find the optimal control. One strategy is to break the treatment interval into bins and calculate independent controls on these intervals before combining all of these controls for the entire interval. For instance, for 30 day segments we first compute an optimal control over the set $[0, 30, 60, 90, 120, 150, 180]$, a calculation that involves the testing of $2^7 = 128$ distinct possible controls. Once this control is computed, the resulting trajectories are integrated to $t = 180$, and a control is computed over the set $[180, 210, 240, \ldots, 540]$. This control is then appended to the prior control vector, and this sequence is repeated until the final time t_f is reached.

Because we divide the treatment interval into several separate bins, the resulting control is only suboptimal; however in (Adams et al., 2004) they show that these results are reasonable approximations to a fully efficacious continuous therapy. In the simulations presented, we compare 10, 20, and 30 day treatment intervals.

5.2 Structured Treatment Interruptions

HAART has changed the course of HIV treatments since introduction. Under HAART, viral load decreases exponentially in the weeks after the start of treatment and is maintained at low levels so long as treatment is continued. Despite this exponential decline in viral load, these treatments do not completely clear the virus from the patient, and a life long therapy is necessary to maintain low viral load (Finzi et al., 1999). Even with this kind of success, the best way to treat acutely infected HIV patients is still an open question with many approaches under investigation (Bell et al., 2010). For many patients, long term continuous HAART is expensive and can include problems with drug toxicity and side effects, as well as increased drug resistance (Phillips, 2005). It has been shown that approximately 80% of patients who adhered to less than 70% of the prescribed treatment regimen experienced treatment failure, whereas patients of

high adherence (>95%) experienced the best results (Paterson et al., 2000). Long term use of protease inhibitors has been associated with insulin intolerance, cholesterol elevation, and the redistribution of body fat. Because of these reasons, some HIV infected patients will voluntarily terminate HAART. Some of these patients will also interrupt the continuous prescribed therapies for short or long periods. After discontinuing HAART, patients will usually experience a rapid increase in viral load coupled with a immediate decline in $CD4^+$ counts (Lisziewicz and Lori, 2002).

The canonical example of a patient undergoing unsupervised breaks in HAART is that of the "Berlin patient" (Lisziewicz et al., 1999). In this case, the patient was able to control viral load in the absence of treatment by cycling HAART on and off due to non-related infections. When treatment was fully stopped after day 176, the patient's viral load did not immediately rebound as is consistent with going off treatment. Due to this patient, interest in the use of structured treatment interruptions (STI) as a mechanism to regulate an HIV infection piqued. The investigators further note that the management of viral load coincided with the emergence and stimulation of a CD8 immune response, and speculate that it is this response that kept the viral load at bay. Indeed, in a randomized, controlled clinical trial with simian immunodeficiency virus infected macaques demonstrated that it was in fact the intermittent interruption of therapy that caused the boosted immune response (Lori et al., 2000). This immune response is further examined in (Carr et al., 1996), and optimal treatment protocols designed around the maximization of immune response is discussed in (Culshaw et al., 2004).

Structured treatment interruptions have been studied in three distinct contexts: the management of acutely infected patients, management of chronically drug suppressed patients, and "salvage treatment", in which patients have a severely depressed immune system. STI treatment is not in general without risk, and a main goal of the clinical trials is to assess how the risks can be weighed against the benefits. In Table 5.1 a summary of potential risks and benefits are listed. In (Lori and Lisziewicz, 2001), the authors survey many different STI based clinical trials and compare their results. Among their key findings are that STI treatments are most effective in

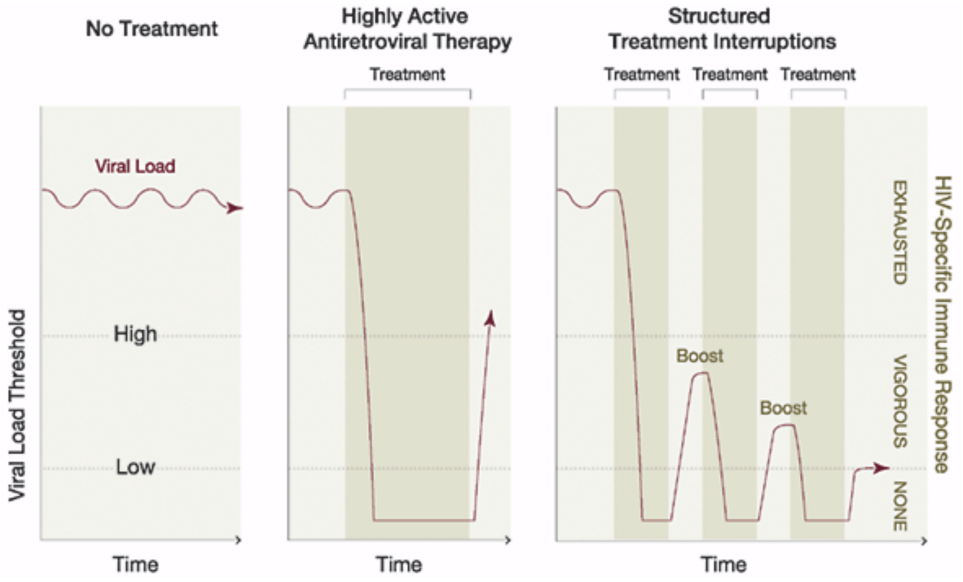

Figure 5.2: Qualitative effects on viral load given no treatment, continuous HAART treatment, and an STI type protocol. This figure is from (Lori and Lisziewicz, 2001).

the acute regime. In a clinical trial in Boston involving 8 subjects, treatment was discontinued and only turned back on when viral load exceeded 5000 copies/mL for three consecutive weeks. In this study, all of the patients achieved long-term control of HIV replication in the absence of treatment after the first treatment interruption (Rosenberg et al., 2000). Another STI success is discussed in (Palacios et al., 2010), where HIV infected children are put onto an STI therapy with twelve weeks on, four weeks off. Due to the STI, viral rebounds decreased in magnitude with each switch from on to off, along with an increased immune response.

For patients in later stages of HIV infection, extension of the positive acute results is more difficult because STI in the chronic case can not take advantage of a relatively healthy immune system which is present in the acute stage. Additionally HAART efficacy may be reduced in this phase, because of a high number of drug-resistant mutant viruses. The STI objective during chronic infection is to see if reliance on constant HAART can be reduced while still maintaining effective viral control, indicating a restoration of sensitivity to HAART. In a large Swiss-Spanish study, the patient group consisted of 128 chronically infected patients. Despite two weeks off

and eight weeks on HAART for four cycles, participants did not show an appreciable decline in viral load nor a boost in $CD4^+$ counts. Finally in a Philadelphia study which tracked five chronically infected, HAART suppressed patients, the patients experienced an increase in both $CD4^+$ and CD8 counts following STI. The results of this study suggest that HAART might be able to recover an appropriate HIV immune response during the chronic regime, and variations in the method of STI could have an effect on immune response.

One of the key variables in the design of any STI based protocol is the criteria in which patients are removed or placed back on treatment. In some studies the viral load was the primary indicator as to when treatment would resume, whereas in other studies off periods were of a fixed length, followed by a fixed length on period. The former approach is simpler, but harder to administer as it requires a higher level of patience scrutiny with more expensive and inconvenient blood tests. The latter strategy is easier to implement in larger patient groups, but the on/off periods are still up for debate and the subject of many clinical trials. Currently the period durations are determined either in an ad-hoc style, or by trying to adapt the methodologies of other trials.

The use of mathematical modeling to determine optimal STI regimes has flourished in recent years. The survey paper (Rosenberg et al., 2007) details the design of STI strategies using HIV models and emphasizes that the use of models may be beneficial in a clinical design scenario. In (Bajaria et al., 2004) the authors build upon the original work of (Ortiz et al., 2001) and use an in-host HIV model to simulate many possible different STI treatments. In contrast to our work however, there is no patient specific data or parameter estimates used and the STI treatments are not calculated, only evaluated; several insights into the effects of STI treatments are gained through the "virtual STI" trials.

5.3 Open Loop Control Formulation

There are two standard formulations of an optimal control. The first, and most common framework is an open loop control, or non feedback control. In an open loop control, the control

Table 5.1: Summary of potential benefits and risks associated with STI like therapies, adapted from (Lori and Lisziewicz, 2001).

	Risks	Benefits
Virological Control	Loss of viral control	Enhanced immune response
	Development of resistance	Delay of rebound
	Repopulation of viral reservoirs	Long-term treatment interruption
	Acute retroviral syndrome	Reduction in drug resistance
Immunity & Toxicity	$CD4^+$ cell decline	Increased tolerability of drug regimens
	Loss of immune response	Fewer toxic effects
	Clinical events	
Other	Recurrence of acute drug adverse effects	Increased access to treatment
	Pharmacokinetic Issues	Increased adherence levels
	Increased risk of transmission	Reduced monetary costs
	Social and psychological factors	

function is a function of only initial conditions and time,

$$u^*(t) = e(x(t_0), t). \tag{5.3.1}$$

A diagram of a generic open loop controller is given in Figure 5.3. Thus the optimal open-loop control is optimal only for a particular initial state value. While this is straightforward in a mathematical sense, it also poses several risks in an actual clinical setting. If the patient were to deviate from the optimal control, even in the slightest, then the resulting control is now suboptimal, and may not even be beneficial. Moreover, if the patient is under periodic clinical surveillance then an open loop control does not take advantage of new data that is collected. It is these reasons why we would like to consider a feedback based control methodology. A possible

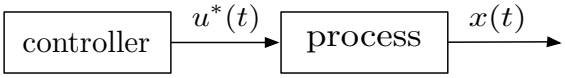

Figure 5.3: Diagram of a generic open loop controller. The model state does not affect the control beyond t_0.

procedure for finding an open loop control given our control problem is as follows.

1. Divide the treatment interval into $k = \lceil \frac{n}{m} \rceil$ bins.

2. Integrate model dynamics from t_0 to the end of the first bin. Determine the optimal control for the first bin by searching through all $u \in \mathcal{U}$.

3. For $k > 1$, integrate the model dynamics to the end of the k^{th} bin, using the optimal control determined in the prior steps for bins $1, \ldots k-1$. Calculate the optimal control for bin k by evaluating all $u \in \mathcal{U}$.

4. Repeat prior step until the end of the entire interval has been reached.

This is the process that is extended in the following sections.

5.4 Closed Loop (Feedback) Control

Feedback control methodologies include mechanisms to allow the control to be updated in real time according to progressing model dynamics. This allows for an adaptivity in the control to handle changing model dynamics, or even possibly external stimuli to the system. A generic feedback control system is shown in Figure 5.4. In (Brandt and Chen, 2001), a feedback system consisting of a therapeutic drug regimen parameterized by the model states is presented. However only a reduced order model was considered with nominal parameters with monotherapy.

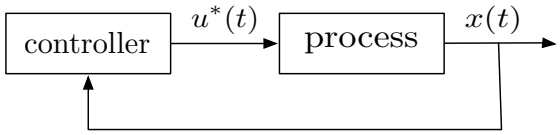

Figure 5.4: Diagram of a generic closed loop feedback controller. The control is updated with new information from the state beyond t_0.

5.4.1 Linear Feedback Laws

The HIV model being nonlinear poses a wide class of considerations to be made when formulating a feedback control. The control of a linear system has been extensively documented (Anderson and Moore, 2007). If the cost function is quadratic in both state and control, then the optimal control is a linear feedback law and the gains can be solved explicitly. These results are primarily due to R. E. Kalman. For the following derivation, we follow the process in (Kirk, 2004). In the linear case, the system is given by

$$\dot{x} = Ax(t) + Bu(t), \tag{5.4.1}$$

where the matrices A, B could be time varying. The performance measure, or goal function, for the class of linear quadratic regulator problems is given as

$$\min J(u) = \frac{1}{2}x^\top(t_f)Gx(t_f) + \frac{1}{2}\int_{t_0}^{t_f} x^\top(t)Qx(t) + u^\top(t)Ru(t)\ dt, \tag{5.4.2}$$

where the matrices G, Q are symmetric, semi-positive definite and R is symmetric, positive definite. The necessary conditions for optimality are derived through calculus of variations (Lewis and Syrmos, 1995), and are given by

$$\dot{x}^* = \frac{\partial H}{\partial p} = Ax^* + Bu^* \tag{5.4.3}$$

$$\dot{p}^* = -\frac{\partial H}{\partial x} = -Qx^* - A^T p^* \tag{5.4.4}$$

$$\frac{\partial H}{\partial u} = 0 = Ru^* + B^\top p^* \Rightarrow u^* = -R^{-1}B^\top p^*, \tag{5.4.5}$$

where $p(t)$ is a costate variable, otherwise known as a Lagrange multiplier or a "shadow" variable. We see that the optimal control $u^*(t)$ is in terms of the optimal costate $p^*(t)$. The existence of

R^{-1} in (5.4.5) is guaranteed since R is positive. Substituting (5.4.5) into (5.4.3) we obtain

$$\dot{x}^* = Ax^*(t) - BR^{-1}B^\top p^*(t), \tag{5.4.6}$$

resulting in a set of $2n$ linear, homogeneous differential equations

$$\begin{bmatrix} \dot{\mathbf{x}}^*(t) \\ \dot{\mathbf{p}}^*(t) \end{bmatrix} = \begin{bmatrix} A & -BR^{-1}B^\top \\ -Q & -A^\top \end{bmatrix} \begin{bmatrix} \mathbf{x}^*(t) \\ \mathbf{p}^*(t) \end{bmatrix}. \tag{5.4.7}$$

The solution to these equations has the form

$$\begin{bmatrix} \mathbf{x}^*(t_f) \\ \mathbf{p}^*(t_f) \end{bmatrix} = \Phi(t_f, t) \begin{bmatrix} \mathbf{x}^*(t) \\ \mathbf{p}^*(t) \end{bmatrix}, \tag{5.4.8}$$

where Φ is the state transition matrix of (5.4.7). Partitioning the matrix, we have

$$\begin{bmatrix} \mathbf{x}^*(t) \\ \mathbf{p}^*(t) \end{bmatrix} = \begin{bmatrix} \Phi_{11}(t_f, t) & \Phi_{12}(t_f, t) \\ \Phi_{21}(t_f, t) & \Phi_{22}(t_f, t) \end{bmatrix} \begin{bmatrix} \mathbf{x}^*(t_f) \\ \mathbf{p}^*(t_f) \end{bmatrix}, \tag{5.4.9}$$

where each partition of Φ is an $n \times n$ matrix. Boundary conditions for the solution are given by

$$x(t_0) = x_0 \tag{5.4.10}$$

$$\left.\frac{\partial h}{\partial x}\right|_{t_f} - p^*(t_f) \equiv 0 \Rightarrow p^*(t_f) = Gx^*(t_f), \tag{5.4.11}$$

where $h(x)$ is the final-time weighting function in (5.4.2). Substituting these results, namely for $\mathbf{p}^*(t_f)$ into (5.4.9),

$$\begin{aligned} \mathbf{x}^*(t_f) &= \Phi_{11}(t_f, t)\mathbf{x}^*(t) + \Phi_{12}(t_f, t)\mathbf{p}^*(t) \\ G\mathbf{x}^*(t_f) &= \Phi_{21}(t_f, t)\mathbf{x}^*(t) + \Phi_{22}(t_f, t)\mathbf{p}^*(t) \end{aligned} \tag{5.4.12}$$

Substituting the top into the bottom, and then solving for $\mathbf{p}^*(t)$ yields

$$\mathbf{p}^*(t) = \big(\Phi_{22}(t,t_f) - G\Phi_{12}(t,t_f)\big)^{-1}\big(G\Phi_{12}(t,t_f) - \Phi_{21}(t,t_f)\big)\mathbf{x}^*(t) \tag{5.4.13}$$

$$\triangleq \Pi(t)\mathbf{x}^*(t), \tag{5.4.14}$$

implying that $\mathbf{p}^*(t)$ is a linear function of the system state and $\Pi(t)$ is a time varying $n \times n$ matrix. To solve for $\Pi(t)$, differentiate both sides to obtain

$$\begin{aligned}
\dot{\mathbf{p}}^* &= \dot{\Pi}\mathbf{x}^* + \Pi\dot{\mathbf{x}}^* \\
&= \dot{\Pi}\mathbf{x}^* + \Pi(A\mathbf{x}^* + Bu) \\
&= \dot{\Pi}\mathbf{x}^* + \Pi\big(A\mathbf{x}^* + B(-R^{-1}B^\top\mathbf{p}^*)\big) \\
-Q\mathbf{x}^* - A^\top\mathbf{p}^* &= \dot{\Pi}\mathbf{x}^* + \Pi(Ax - BR^{-1}B^\top\Pi\mathbf{x}^*) \\
0 &= \dot{\Pi}\mathbf{x}^* + \Pi A\mathbf{x}^* - \Pi BR^{-1}B^\top\Pi\mathbf{x}^* + Q\mathbf{x}^* + A^\top\Pi\mathbf{x}^* \\
\dot{\Pi} &= -(\Pi A - \Pi BR^{-1}B^\top\Pi + A^\top\Pi + Q),
\end{aligned} \tag{5.4.15}$$

with boundary condition $\Pi(t_f) = G$. This matrix differential equation is of the Riccati type, and is referred to as the algebraic Riccati equation (ARE). The solutions are computed by integrating the final equation in (5.4.15) in backwards time, requiring the solution of $n(n+1)/2$ equations due to Π being symmetric. The matrix is then substituted back into (5.4.6) for the optimal state and feedback control.

5.4.2 Nonlinear Feedback Laws

For generic nonlinear systems, a different approach is required. For this derivation, we follow the process given in (Lewis and Syrmos, 1995). Assume we have a general nonlinear system model

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}, t), \qquad \mathbf{x}(t_0) = \mathbf{x}_0. \tag{5.4.16}$$

We seek to minimize the generic cost

$$J(\mathbf{x}(t_0), t_0) = \varphi\big(\mathbf{x}(t_f), t_f\big) + \int_{t_0}^{t_f} L(\mathbf{x}, \mathbf{u}, t)dt. \tag{5.4.17}$$

We proceed by first dividing the interval. If time t is the current time, then $t + \Delta t$ is a future time close to t. Then the cost corresponding to current time t, $J(\mathbf{x}(t), t)$ can be written as

$$J(\mathbf{x}, t) = \varphi\big(\mathbf{x}(t_f), t_f\big) + \int_{t+\Delta t}^{t_f} L(\mathbf{x}, \mathbf{u}, s)ds + \int_{t}^{t+\Delta t} L(\mathbf{x}, \mathbf{u}, s)ds. \tag{5.4.18}$$

This can be rewritten as

$$J(\mathbf{x}, t) = \int_{t}^{t+\Delta t} L(\mathbf{x}, \mathbf{u}, s)ds + J(\mathbf{x} + \Delta\mathbf{x}, t + \Delta t). \tag{5.4.19}$$

Note that in the sense of a Taylor expansion, the increment $\Delta\mathbf{x} = f(\mathbf{x}, \mathbf{u}, t)\Delta t$. Equation (5.4.19) describes all possible cost function values from time t to final time t_f. According to Bellman's principle of optimality, which states that no matter what the prior states of the system have been, all future decisions must constitute an optimal policy, the only candidates for the optimal cost $J^*(\mathbf{x}, t)$ are the costs $J(\mathbf{x}, t)$ that are optimal from $t + \Delta t$ to t_f. The optimality principle then requires

$$J^*(\mathbf{x}, t) = \min_{\substack{\mathbf{u}(s) \\ t \le s \le t+\Delta t}} \left[\int_{t}^{t+\Delta t} L(\mathbf{x}, \mathbf{u}, s)ds + J^*(\mathbf{x} + \Delta\mathbf{x}, t + \Delta t) \right]. \tag{5.4.20}$$

Equation (5.4.20) is referred to as the principle of optimality for continuous time systems, and does not provide a straightforward analytic process of determining the optimal control and cost. Assuming that the second partial derivatives of J^* exist and are bounded, we can expand

$J^*(\mathbf{x}+\Delta\mathbf{x}, t+\Delta t)$ in a Taylor series around the point $\mathbf{x}(t)$ to obtain

$$J^*(\mathbf{x},t) = \min_{\substack{\mathbf{u}(s) \\ t\le s\le t+\Delta t}} \left\{ \int_t^{t+\Delta t} L(\mathbf{x},\mathbf{u},s)ds + J^*(\mathbf{x},t) + \left[\frac{\partial J^*}{\partial t}(\mathbf{x},t)\right]\Delta t + \left[\frac{\partial J^*}{\partial \mathbf{x}}(\mathbf{x},t)\right]^\top \left[\mathbf{x}(t+\Delta t) - \mathbf{x}(t)\right] + \text{terms of higher order}\right\}. \tag{5.4.21}$$

For small values of Δt,

$$J^*(\mathbf{x},t) = \min_{\mathbf{u}(t)} \left\{ L(\mathbf{x},\mathbf{u},t)\Delta t + J^*(\mathbf{x},t) + \frac{\partial J^*}{\partial t}(\mathbf{x},t)\Delta t + \frac{\partial J^*}{\partial \mathbf{x}}^\top \mathbf{f}(\mathbf{x},\mathbf{u},t)\Delta t + \mathcal{O}(\Delta t)^2 \right\}. \tag{5.4.22}$$

To complete the derivation, we note that J^* and J_t^* are independent of the control $\mathbf{u}(\tau), t \le \tau \le t+\Delta t$ and can be removed from the optimization process. Hence,

$$\frac{\partial J^*}{\partial t}(\mathbf{x}(t),t)\Delta t + \min_{\mathbf{u}(t)} \left\{ L(\mathbf{x},\mathbf{u},t)\Delta t + \frac{\partial J^*}{\partial \mathbf{x}}^\top \big(\mathbf{f}(\mathbf{x},\mathbf{u},t)\big)\Delta t + \mathcal{O}(\Delta t)^2 \right\} = 0. \tag{5.4.23}$$

Dividing by Δt and taking the limit as $\Delta t \to 0$, we obtain

$$\frac{\partial J^*}{\partial t}(\mathbf{x}(t),t) + \min_{\mathbf{u}(t)} \left\{ L(\mathbf{x},\mathbf{u},t) + \frac{\partial J^*}{\partial \mathbf{x}}^\top (\mathbf{x},t)\mathbf{f}(\mathbf{x},\mathbf{u},t) \right\} = 0. \tag{5.4.24}$$

The boundary condition is given by

$$J^*(\mathbf{x}(t_f),t_f) = \varphi(\mathbf{x}(t_f),t_f). \tag{5.4.25}$$

The Hamiltonian $\mathcal{H}$ is defined as

$$\mathcal{H}(\mathbf{x},\mathbf{u},J_{\mathbf{x}}^*,t) \triangleq L(\mathbf{x},\mathbf{u},t) + \frac{\partial J}{\partial \mathbf{x}}^\top \mathbf{f}(\mathbf{x},\mathbf{u},t), \tag{5.4.26}$$

and

$$\mathcal{H}\big(\mathbf{x}, \mathbf{u}^*(\mathbf{x}, J^*_{\mathbf{x}}, t), J^*_{\mathbf{x}}, t\big) = \min_{\mathbf{u}(t)} \mathcal{H}(\mathbf{x}, \mathbf{u}, J^*_{\mathbf{x}}, t), \tag{5.4.27}$$

since the optimal control will depend only on $\mathbf{x}, J^*_{\mathbf{x}}$ and t. Using the Hamiltonian definitions above, we can rewrite (5.4.24) as

$$J^*_t(\mathbf{x}, t) + \mathcal{H}\big(\mathbf{x}, \mathbf{u}^*(\mathbf{x}, J^*_{\mathbf{x}}, t), J^*_{\mathbf{x}}, t\big) = 0. \tag{5.4.28}$$

Equation (5.4.28) is referred to as Halmilton-Jacobi-Bellman (HJB) equation. For nonlinear systems, the optimal feedback control in the case of quadratic cost is known to be of the form

$$\mathbf{u}^*(t) = -1/2R^{-1}BJ^*_x(t), \tag{5.4.29}$$

requiring the solution of (5.4.28), a partial differential equation. Solutions to the HJB equation exist analytically only for the most trivial of problems. Thus, computational methods have been developed to approximate the solution to the HJB, or to solve a related problem resulting in a suboptimal control. In general designing feedback control systems for nonlinear systems is a process with many areas to consider. A survey and comparison of these techniques is presented in (Beeler et al., 2000) and includes power series approximations, solving a state-dependent Riccati equation, and Galerkin solutions to the Halmilton-Jacobi-Bellman equations.

One approach that is used successfully in (Banks et al., 2006) building upon the algebraic Riccati equation derived in equations (5.4.15) by using a state dependent Riccati equation (SDRE). Using only partial state measurements that are similar to the clinical data sets, the authors are able to design multidrug strategies that reduce viral load, increase $CD4^+$ count while stimulating immune response. The idea behind the use of the SDRE is to rewrite the

right-hand-side equations in the form

$$\dot{x} = f(x) = A(x)x, \tag{5.4.30}$$

where clearly $A(x)$ is non unique. In (Banks et al., 2006), they show that $A(x) = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}$ is a very good choice for models comprised of a system of nonlinear ODEs. When f is written in this way, the SDRE is of the form

$$\Pi(x)A(x) + A^{\top}(x)\Pi(x) - \Pi(x)BR^{-1}B^{\top}\Pi(x) + Q = 0, \tag{5.4.31}$$

and the optimal feedback is given similarly as in the linear case,

$$u^{*}(x) = -R^{-1}B^{\top}\Pi(x)x. \tag{5.4.32}$$

Solving for the Riccati solution $\Pi(x)$ is more difficult in this state dependent case than in the constant coefficient case. A method discussed in (Beeler et al., 2000) involves a power series expansion to approximate the general solution. The matrix $A(x)$ is rewritten in terms of a constant matrix and the state dependent part,

$$A(x) = A_0 + \varepsilon \Delta A(x), \tag{5.4.33}$$

where ε is a just a temporary variable soon to be fixed to 1. Next $\Pi(x)$ is written out in a power series,

$$\Pi(x,\varepsilon) = \Pi(x) \mid_{\varepsilon=0} + \varepsilon \frac{\partial \Pi(x)}{\partial \varepsilon}\bigg|_{\varepsilon=0} + \frac{\varepsilon^2}{2}\frac{\partial^2 \Pi(x)}{\partial \varepsilon^2}\bigg|_{\varepsilon=0} + \cdots + \frac{\varepsilon^n}{n!}\frac{\partial^{(n)}\Pi(x)}{\partial \varepsilon^n} \tag{5.4.34}$$

$$= \sum_{n=1}^{\infty} \varepsilon^n L_n(x), \tag{5.4.35}$$

where each Π is symmetric, as is each L_n. By substituting the power series expression for $\Pi(x)$

back into (5.4.31) we obtain

$$\Big[\sum_{n=1}^{\infty}\varepsilon^n L_n(x)\Big]\big(A_0+\varepsilon\Delta A(x)\big)+\big(A_0^\top+\varepsilon\Delta A^\top(x)\big)\Big[\sum_{n=1}^{\infty}\varepsilon^n L_n(x)\Big]- \Big[\sum_{n=1}^{\infty}\varepsilon^n L_n(x)\Big]BR^{-1}B^\top\Big[\sum_{n=1}^{\infty}\varepsilon^n L_n(x)\Big]+Q=0. \tag{5.4.36}$$

Equation (5.4.36) can be expanded in ε. Equating coefficients for terms involving the same power of ε, we obtain a recurrence relation that can be used to determine L_n,

$$L_0A_0+A_0^\top L_0-L_0BR^{-1}B^\top L_0+Q=0, \tag{5.4.37}$$

$$L_1(A_0-BR^{-1}B^\top L_0)+(A_0^\top-L_0BR^{-1}B^\top)L_1+L_0\Delta A^\top L_0=0, \tag{5.4.38}$$

$$L_n(A_0-BR^{-1}B^\top L_0)+(A_0^\top-L_0BR^{-1}B^\top)L_N+L_{n-1}\Delta A+ \Delta A^\top L_{n-1}-\sum_{k=1}^{n-1}(L_kBR^{-1}B^\top L_{n-k})=0, \tag{5.4.39}$$

which is the standard ARE for the linear part A_0 coupled with higher order equations. These equations can be solved, but may be difficult if the expansion is complicated. For problems of certain structure, namely those in which ΔA has the same function of x in all of its elements, the coefficients L_n are much easier to compute. In problems of this form, including our problem, we can write

$$\Delta A(x)=g(x)\Delta A_c. \tag{5.4.40}$$

Similarly as before we can define

$$L_n(x)=g^n(x)(L_n)_C, \tag{5.4.41}$$

where $(L_n)_C$ is a constant matrix. Then, by (5.4.39), we obtain

$$(L_n)_C(A_0 - BR^{-1}B^\top L_0) + (A_0^\top - L_0BR^{-1}B^\top)(L_n)_C + (L_{n-1})_C\Delta A_C + \Delta A_C^\top(L_{n-1})_C - \sum_{k=1}^{n-1}((L_k)_CBR^{-1}B^\top(L_{n-k})_C) = 0, \tag{5.4.42}$$

which is a large constant valued matrix equation which can be solved with ease for a fixed n. Once as many L_n terms as desired are found, these terms can be substituted back into (5.4.35) with $\varepsilon = 1$. Given its numerical nature, this process approximates the solution to the SDRE, and so any resulting controls would be suboptimal in nature.

5.5 Receding Horizon Control

The control methodology for the HIV model described up to now has been an open loop control. We would like to install a feedback based control for several reasons. A feedback control could adjust for unexpected perturbations to the treatment schedule, e.g. the patient decides to go off-treatment for a period of time not otherwise prescribed. The feedback control could then adjust for this in future dosing. As established in the prior section, though several control paradigms exist for the control and simulation of nonlinear systems, not many nonlinear controllers can be applied to the discrete sample-and-hold methodology described in section 5.1. One method that is highly promising and is well suited to finding an optimal treatment amenable to HIV treatment is the receding horizon control (RHC) (García et al., 1989). Receding horizon control has also explicitly been used to compute optimal treatment schedules in (Zurakowski and Teel, 2006), though with a very different, immune response centric model. The authors were some of the first investigators to apply RHC methodologies to biological systems.

Receding horizon, or model predictive control seeks to gain the benefits of a feedback control while maintaining the overall existing control methodology (i.e. STI) and model structure. Developed in the late seventies, RHC was first applied to chemical reaction systems and has not seen wide spread application in biological systems (Cutler and Ramaker, 1980; Richalet

et al., 1978). The basic structure of the RHC methodology is shown in Figure 5.5. RHC is a feedback control system that operates by computing the current control by solving a finite time open loop control problem, using the current state of the system as the initial state. This implicit optimization then yields a control for future time, and the first control in the sequence is applied to the system for a specified control window. This is the primary difference with the open loop methodologies, where the controls are computed for the entire simulation interval before simulation.

This framework is uniquely suited for the computation of controls that fall within the STI paradigm under HAART. Since the number of elements in the control sequence if finite due to the on-off nature of the STI control, the existence of an optimal control on each window is guaranteed. Moreover, this approach is model invariant, in that the HIV model can be updated or swapped out with a model with minimal modifications to the control setup. Lastly the RHC parameters such as horizon and window length, discussed later, can be easily modified based on clinical needs and observations.

There are several basic elements to the RHC system, outlined in (Camacho and Bordons, 2004):

1. Model equations which govern the system dynamics.

2. The calculation of a sequence of optimal control laws, subject to the cost function.

3. A "receding horizon" strategy, so that on each interval or horizon, that the control is computed is shifted forward in time. The control that is computed is then only used for a portion of the horizon length.

To solidify the RHC methodology, we present the following mathematical framework.

Let $[t_i \ t_{i+1}]$ be a sequence of time intervals. Let $t_{hor,i}$ such that $t_{hor,i} \geq t_{i+1} - t_i$ be the control horizon. We denote $t_{i+1} - t_i$ to be t_{win}, the current control window. Consider a sequence

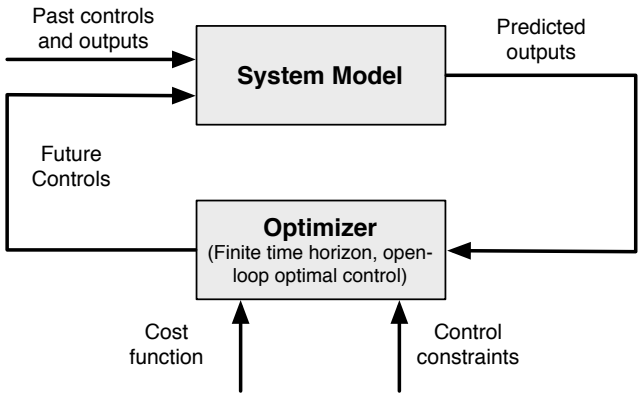

Figure 5.5: Schematic diagram of the receding horizon control.

of control problems P_i,

$$u^* \in \mathcal{U} \text{ s.t. } u^* = \arg\min J(u) = \int_{t_i}^{t_i + t_{hor,i}} L(x, u, t)dt, \tag{5.5.1}$$

subject to

$$\dot{x}(t) = f(x, u), \quad x(t_i) = x_i, \tag{5.5.2}$$

on the interval $[t_i \ t_{i+1}]$. The solution to each control problem P_i is computed by following the direct search routine described in the prior section. This process is extended by RHC, outlined as follows:

1. Given an initial condition $x(t_i)$ solve the optimal control problem P_i on the horizon interval $[t_i \ t_{hor,i}]$.

2. Use the control calculated in the prior step to compute the trajectory over the current control window $[t_i \ t_{i+1}]$.

3. Repeat this process by extending to the next control horizon to compute the next control. Terminate when the next $t_i > t_f$.

A schematic overview of this process is given in Figure 5.6. Note that in this system, each successive control computed on each new window is now implicitly a function of the prior window's final state – now the current window's initial condition. The window length and horizon length vary per simulation given different patient parameters, and are listed with each simulation presented.

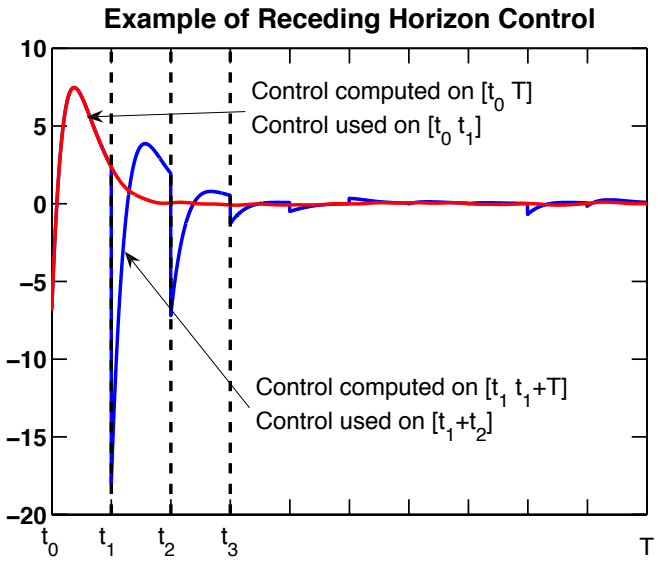

Figure 5.6: An example of receding horizon control. Future controls are calculated by computing controls on successive intervals.

Receding control depends on having full state knowledge in the observations of the system. Our control functions are computed after the completion of the inverse problem, so we assume our model to be accurate and predictive for each patient, therefore yielding full state knowledge at each time step.

For problems where there exists only a reduced observation of the system, a (typically) nonlinear estimation of the true state values is needed for the RHC to operate successfully. In cases such as this, the estimator is installed between the system model and the optimizer, as shown in Figure 5.7.

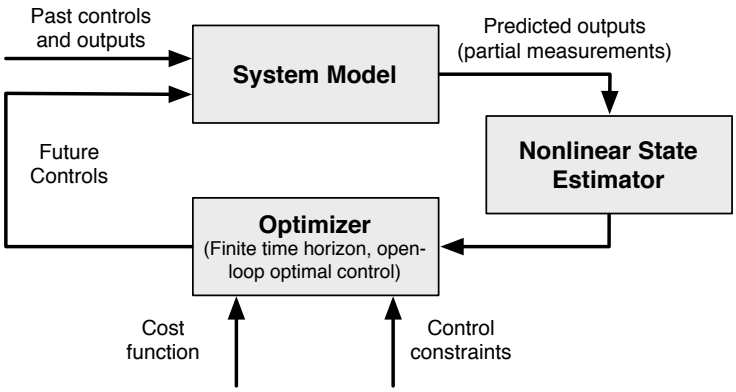

Figure 5.7: Schematic diagram of the receding horizon control with a nonlinear state estimator.

5.6 Simulations

In this section we present optimal controls and the resulting state solutions for several different patients under utilizing differing control weights and treatment intervals. We wish to illustrate that treatment protocols can be swayed and adjusted on a patient specific basis using the control weights. We also show cases with an "unexpected" perturbation to the optimal treatment, e.g. forced off treatment during an interview to demonstrate the feedback capability of the RHC. The control weights are presented as the vector $[Q\ S\ R_1\ R_2]$. In all of the following plots, the green and red bars at the bottom of the plot indicate clinical adherence to control, whereas the cyan and magenta indicate the computed control of on treatment and off treatment, respectively.

5.6.1 Varying Treatment Intervals

We first investigate the impact of different treatment intervals has on viral load. Figure 5.8 and Figure 5.9 differ only in the treatment interval. We first see a 20 day schedule, which results in an average viral of load of less than 1000 RNA copies/ml. The control is on for alternating periods of 60 days, followed by 40 days off, then 60 days on again.

When the interval is increased to 30 days in Figure 5.9 the average viral load increases to over 1500 RNA copies/ml. The control is on for the first 60 days, followed by an off treatment for

30 days, repeating periodically. Despite having a greater number of switching times, the control is off for fewer day, 30 off, versus 40 off with $m = 20$ in Figure 5.8. With 30 day treatment intervals, the viral load peaks to a level less than that of using 20 day intervals. The on/off treatment doesn't decrease the viral load to points achieved due to clinical adherence, which is constant on for nearly two years. However, when the patient goes off treatment even for a matter of days, the viral load spikes to nearly 40,000 RNA copies/ml. This kind of spike does not occur with the STI style regime produced by the RHC.

Compare these results with Figure 5.10, where we use a 15 day treatment interval. Treatment begins with 45 days of on treatment, before entering into a periodic switching of 30 days off, 15 days on, 15 days off, 30 days on. Due to the greater amount of time spent off control, the lowest viral load is greater than that of other treatment intervals. Through these results we see that having access to a greater number of switchings, or shorter treatment intervals, does not necessarily produce better results overall. Further, we see in general a greater amount of control used in the first 100 or so days, indicating the need to control HIV greater in the acute infection regimes.

5.6.2 Varying control weights

One advantage of the control methodology presented is being able to customize a treatment for a given patient. This can be done by varying the control weights to emphasize or demphasize specific terms in the goal function (5.1.1). Between Figures 5.8 and 5.11 we see the effect of changing the weight on viral load reduction by a factor or 10. Nearly twice as much control is used when $Q = 0.01$ versus $Q = 0.001$. There are currently ongoing clinical trials that are exploring STI treatments by using more frequent on-treatments compared to prior studies.

5.6.3 Unexpected Treatment Perturbations

In Figure 5.12 we force an off treatment protocol for 100 days early in the simulation. The control returns the viral load to the same level in Figure 5.11, where treatment is not modified.

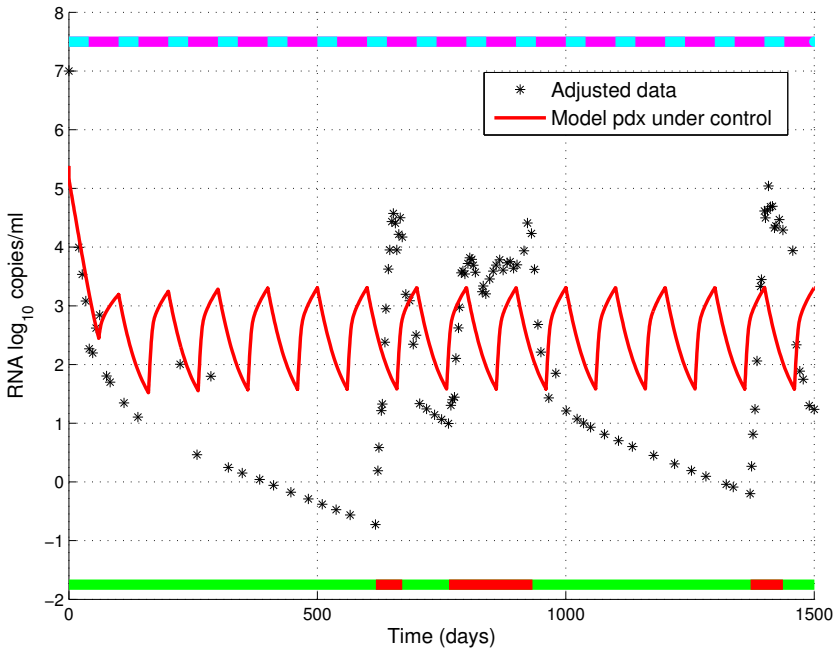

Figure 5.8: Patient 3, with weights [.01 1 .01 .1], with 20 day treatment intervals, 100 day control window, and 720 day control horizon.

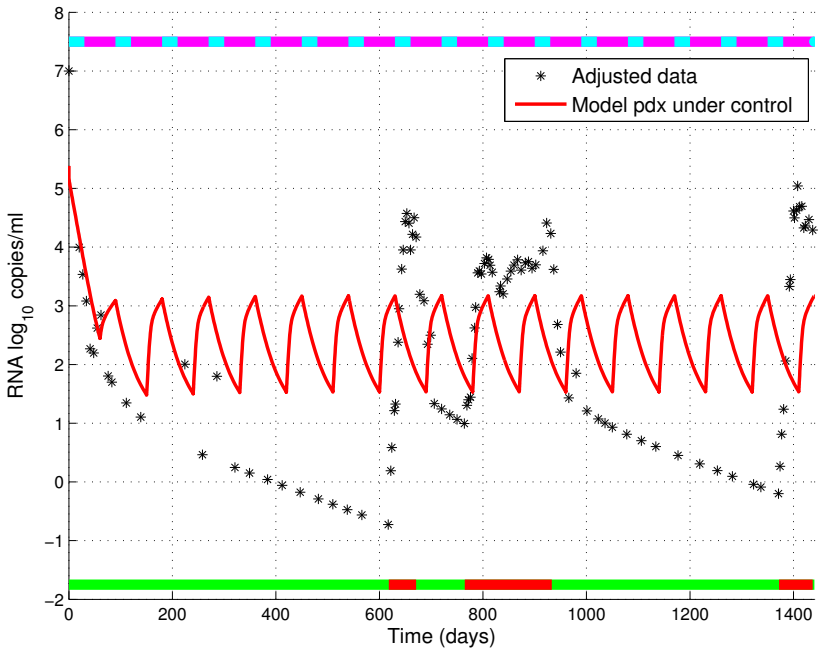

Figure 5.9: Patient 3, with weights [.01 1 .01 .1], with 30 day treatment intervals, 90 day control window, and 720 day control horizon.

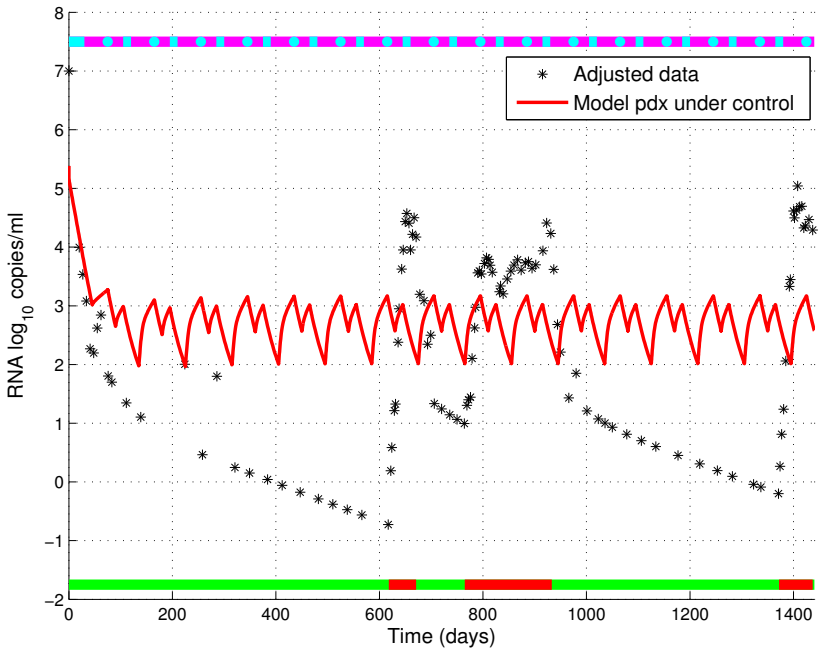

Figure 5.10: Patient 3, with weights [.01 1 .01 .1], with 15 day treatment intervals, 90 day control window, and 720 day control horizon.

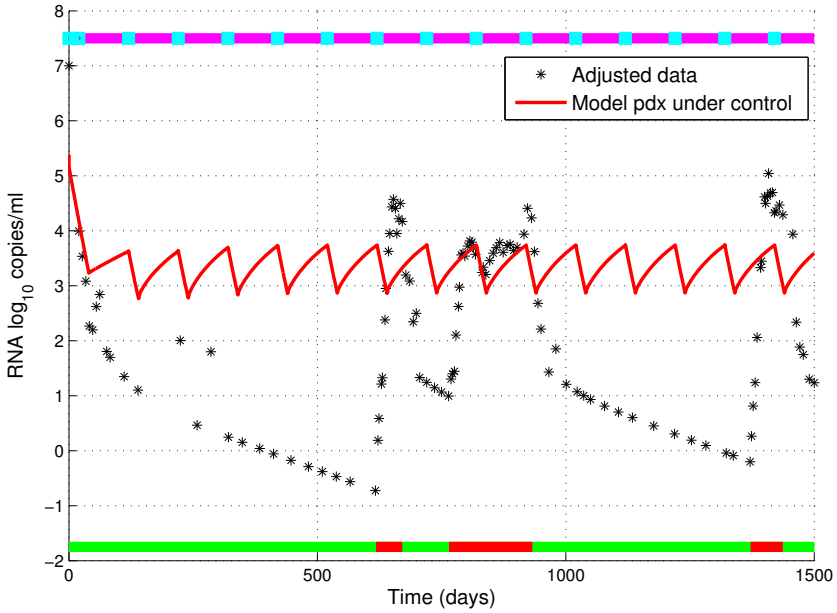

Figure 5.11: Patient 3, with weights [.001 1 .01 .01], with 20 day treatment intervals, 100 day control window, and 720 day control horizon.

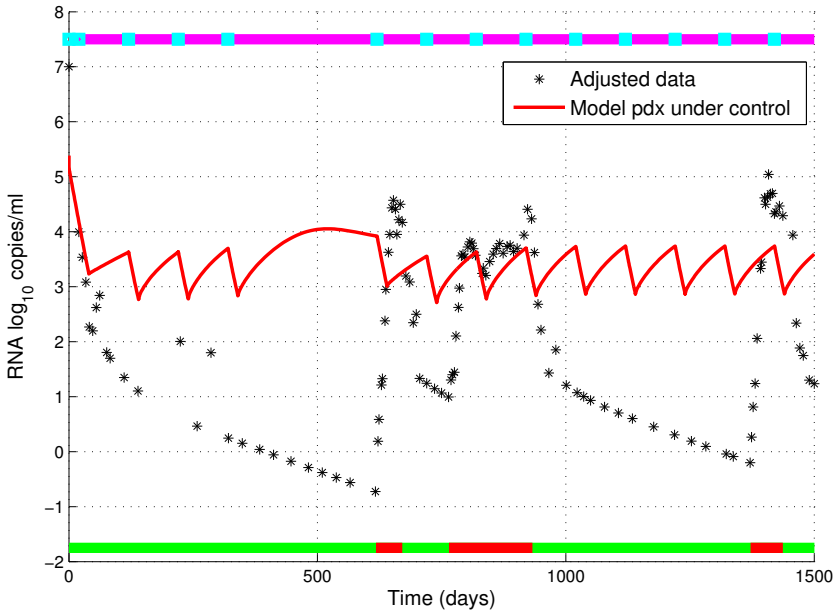

Figure 5.12: Patient 3, with weights [.001 1 .01 .01], with 20 day treatment intervals, 100 day control window, and 720 day control horizon. The control was fixed to zero for 100 days early on in the simulation.

CHAPTER 6

Conclusion and Future Work

In this work we have examined an in-vivo model describes the acute and chronic infection regimes of HIV. With this model, we have performed a patient specific model calibration to clinically collected data sets consisting of $CD4^+$ count as well as viral load, which are censored. Because the dynamic model contains over 20 biologically relevant parameters, we must first determine which parameters are to be estimated. Traditionally, this process has relied solely on a local relative sensitivity analysis, which identifies parameters that have a significant effect on the output of the system given small changes. However we have shown that simply being sensitive does not constitute identifiability, and other methods are needed to identify the parameters that can be estimated in a least-squares sense. To that end, we utilize subset selection algorithms that identify the source of rank-dependence in the Fisher matrix. While the parameters that are identified through this process are identifiable in some sense, they may not be sensitive. We therefore propose an algorithm that would lead to a sensitive, and identifiable set of parameters.

To use the censored data within the inverse problem methodology, statistical methods such as expectation maximization are utilized to provide estimates to the true data that is unobserved.

We initialize the methods using a global sampling algorithm to provide a good initial iterate for optimization. Using these statistical estimates provides better parameter estimates than just having simply ignored the censored clinical data.

In concluding this work, we take the first steps in investigating the construction of optimal treatment strategies, noting the lack of consensus in the best standards of treatment for acutely infected patients. To construct the treatment protocols, we first describe in detail linear and non-linear feedback control systems in a mathematical context. For a feedback law in our model, we use receding horizon control which computes several successive controls over a shifting interval. The broader controls are then used on short time intervals, allowing for the control to adjust to unexpected physical perturbations to the system. These controls mimic the form of structured treatment interruptions, and there are several ongoing clinical studies to assess their effectiveness against standard HAART regimes. In our control implementation, we have assumed full state knowledge of the dynamic system. This is an unrealistic assumption and the use of nonlinear estimators would be perfunctory.

Going forward, there are several areas that can be addressed. Calibration and validation of more advanced HIV models could lead greater insight into the replication processes during acute infection. Parameter estimates can be improved, and this would rely on richer data being available. Additionally, the development of new immune response models, an area that is currently poorly understood, would shed light into the complicated immune processes that occur post-infection. There is CD8 immune response data available, but we do not use it in this work, as the data does not match the type of immune response that we have modeled.

In performing the identifiability analysis and subset selection, it is only the clinical adherence to therapy that determines the identifiability of parameters on a patient specific basis as a nominal patient parameter is used in the subset selection algorithm. Therefore, identifiability of parameters is directly affected by experiment design and execution. Constructing optimal experiments (Banks et al., 2010) could lead to enhanced parameter identification and in turn, improved model analysis and calibration. ■

REFERENCES

Adams, B.M.: Non-parametric parameter estimation and clinical data fitting with a model of HIV infection. Ph.D. thesis, North Carolina State University (2005)

Adams, B.M., Banks, H., Davidian, M., Rosenberg, E.: Estimation and prediction with HIV-treatment interruption data. Bull Math Bio **69**, 563–584 (2007). URL http://dx.doi.org/10.1007/s11538-006-9140-6. 10.1007/s11538-006-9140-6

Adams, B.M., Banks, H.T., Davidian, M., dae Kwon, H., Tran, H.T., Wynne, S.N., Rosenberg, E.S.: HIV dynamics: modeling, data analysis, and optimal treatment protocols. J. Comput. Appl. Math **184**, 10–49 (2005)

Adams, B.M., Banks, H.T., Kwon, H.D., Tran, H.T.: Dynamic multidrug therapies for hiv: optimal and sti control approaches. Math Biosci Eng **1**(2), 223–241 (2004). URL http://www.ncbi.nlm.nih.gov/pubmed/20369969

Anderson, B., Moore, J.: Optimal Control: Linear Quadratic Methods. Dover Books on Engineering Series. Dover Publications (2007). URL http://books.google.com/books?id=PZToAAAACAAJ

Andriote, J.: Victory Deferred: How AIDS Changed Gay Life in America. University of Chicago Press (1999). URL http://books.google.com/books?id=dAI7nOe1KfYC

Aster, R., Thurber, C., Borchers, B.: Parameter Estimation And Inverse Problems. International Geophysics Series. Elsevier Academic Press (2005). URL http://books.google.com/books?id=LXgSA6YEke4C

Audoly, S., Bellu, G., D'Angio, L., Saccomani, M., Cobelli, C.: Global identifiability of nonlinear models of biological systems. IEEE Trans Biomed Eng **48**(1), 55 –65 (2001). DOI 10.1109/10.900248

Bailey, J.J., Fletcher, J.E., Chuck, E.T., Shrager, R.I.: A kinetic model of cd4+ lymphocytes with the human immunodeficiency virus (HIV). Biosystems **26**(3), 177 – 183 (1992). DOI 10.1016/0303-2647(92)90077-C. URL http://www.sciencedirect.com/science/article/pii/030326479290077C

Bajaria, S.H., Webb, G., Kirschner, D.E.: Predicting differential responses to structured treatment interruptions during HAART. Bulletin of Mathematical Biology **66**(5), 1093 – 1118 (2004). DOI 10.1016/j.bulm.2003.11.003

Banks, H., Davidian, M., Hu, S., Kepler, G., Rosenberg, E.: Modeling HIV immune response and validation with clinical data. J Biol Dyn **2**(4), 357–385 (2008)

Banks, H., Tran, H.: Mathematical and Experimental Modeling of Physical and Biological Processes. Textbooks in Mathematics. CRC Press (2009). URL http://books.google.com/books?id=SSRapIe8p3QC

Banks, H.T., Dediu, S., Ernstberger, S.L., Kappel, F.: Generalized sensitivities and optimal experimental design. Journal of Inverse and III-posed Problems **18**(1), 25–83 (2010). DOI 10.1515/jiip.2010.002

Banks, H.T., Dedui, S., Ernstberger, S.L.: Sensitivity functions and their uses in inverse problems. J. Inverse and Ill-posed Problems **15**, 1–26 (2007)

Banks, H.T., Kwon, H.D., Toivanen, J.A., Tran, H.T.: A state-dependent riccati equation-based estimator approach for HIV feedback control. Optimal Control Applications and Methods **27**(2), 93–121 (2006). DOI 10.1002/oca.773. URL http://dx.doi.org/10.1002/oca.773

Batzel, J.J., Kappel, F., Schneditz, D., Tran, H.T.: Cardiovascular and Respiratory Systems. Society for Industrial and Applied Mathematics, Philadephia, PA (2007). DOI DOI:10.1137/1.9780898717457. URL http://dx.doi.org/10.1137/1.9780898717457

Beeler, S.C., Tran, H.T., Banks, H.T.: Feedback control methodologies for nonlinear systems. Journal of Optimization Theory and Applications **107**, 1–33 (2000). URL http://dx.doi.org/10.1023/A:1004607114958. 10.1023/A:1004607114958

Bell, S.K., Little, S.J., Rosenberg, E.S.: Clinical management of acute HIV infection: Best practice remains unknown. Journal of Infectious Diseases **202**(Supplement 2), S278–S288 (2010). DOI 10.1086/655655. URL http://jid.oxfordjournals.org/content/202/Supplement_2/S278.abstract

Bellman, R., Åström, K.: On structural identifiability. Mathematical Biosciences **7**(3â“4), 329 – 339 (1970). DOI 10.1016/0025-5564(70)90132-X. URL http://www.sciencedirect.com/science/article/pii/002555647090132X

Bonhoeffer, S., Coffin, J.M., Nowak, M.A.: Human immunodeficiency virus drug therapy and virus load. J Virol. **71**(4), 3275–3278 (1997)

Bonhoeffer, S., Rembiszewski, M., Ortiz, G.M., Nixon, D.F.: Risks and benefits of structured antiretroviral drug therapy interruptions in HIV-1 infection. AIDS **14**(15) (2000). URL http://journals.lww.com/aidsonline/Fulltext/2000/10200/Risks_and_benefits_of_structured_antiretroviral.12.aspx

Brandt, M., Chen, G.: Feedback control of a biodynamical model of HIV-1. IEEE Trans Biomed Eng **48**(7), 754 –759 (2001). DOI 10.1109/10.930900

Burth, M., Verghese, G., Velez-Reyes, M.: Subset selection for improved parameter estimation in on-line identification of a synchronous generator. IEEE Trans Power Sys (1999). URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=744536

Callaway, D., Perelson, A.: HIV-1 infection and low steady state viral loads. Bull Math Bio **64**, 29–64 (2002). 10.1006/bulm.2001.0266

Camacho, E., Bordons, C.: Model Predictive Control. Advanced Textbooks in Control and Signal Processing. Springer (2004). URL http://books.google.com/books?id=Sc1H3f3E8CQC

Carmichael, G.R., Sandu, A., florian A. Potra: Sensitivity analysis for atmospheric chemistry models via automatic differentiation. Atmos Environ **31**(3), 475 – 489 (1997). DOI DOI: 10.1016/S1352-2310(96)00168-9. URL http://www.sciencedirect.com/science/article/pii/S1352231096001689

Carr, A., Emery, S., Kelleher, A., Law, M., Cooper, D.A.: CD8 lymphocyte responses to antiretroviral therapy of HIV infection. JAIDS Journal of Acquired Immune Deficiency Syndromes (4), 320–326 (1996)

Casella, G., Berger, R.: Statistical inference. Duxbury advanced series in statistics and decision sciences. Thomson Learning (2002)

Cintrón-Arias, A., Banks, H.T., Capaldi, A., Lloyd, A.L.: A sensitivity matrix based methodology for inverse problem formulation. J. Inverse and Ill-posed Problems **17**, 545–564 (2009). DOI 10.1515/JIIP.2009.034

Culshaw, R.V., Ruan, S.: A delay-differential equation model of HIV infection of CD4+ t-cells. Mathematical Biosciences **165**(1), 27 – 39 (2000). DOI 10.1016/S0025-5564(00)00006-7. URL http://www.sciencedirect.com/science/article/pii/S0025556400000067

Culshaw, R.V., Ruan, S., Spiteri, R.J.: Optimal HIV treatment by maximising immune response. Journal of Mathematical Biology **48**, 545–562 (2004). URL http://dx.doi.org/10.1007/s00285-003-0245-3. 10.1007/s00285-003-0245-3

Cutler, C., Ramaker, B.: Dynamic matrix control - a computer control algorithm. Proc. Joint Aut. Control Conf. (1980)

Daun, S., Rubin, J., Vodovotz, Y., Roy, A., Parker, R., Clermont, G.: An ensemble of models of the acute inflammatory response to bacterial lipopolysaccharide in rats: Results from parameter space reduction. J Theor Biol **253**(4), 843 – 853 (2008). DOI DOI:10.1016/j.jtbi.2008.04.033. URL http://www.sciencedirect.com/science/article/B6WMD-4SF30B6-6/2/8dc55e06ddfa9539af973484fedba6ad

David, J.A.: Optimal control, estimation, and shape design: Analysis and applications. Ph.D. thesis, North Carolina State University (2007)

David, J.A., Banks, H.T., Tran, H.T.: HIV model analysis under optimal control based treatment strategies. Tech. Rep. TR08-07, North Carolina State University (2008)

Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum likelihood from incomplete data via the em algorithm. J. Roy. Stat. Soc, Ser. B **39**(1), 1–38 (1977)

Dennis Jr., J.E., Schnabel, R.B.: Numerical Methods for Unconstrained Optimization and Nonlinear Equations (Classics in Applied Mathematics, 16). Soc for Industrial & Applied Math (1996)

Eslami, M.: Theory of Sensitivity in Dynamic Systems: An Introduction. Springer (1994)

Fink, M., Attarian, A., Tran, H.: Subset selection for parameter estimation in an HIV model. PAMM **7**(1), 1121,501–1121,502 (2007). DOI 10.1002/pamm.200700319. URL http://dx.doi.org/10.1002/pamm.200700319

Finkel, D.E., Kelley, C.: Convergence analysis of the direct algorithm. Opt Online **14**(2), 1–10 (2004). URL http://www.optimization-online.org/DB_HTML/2004/08/934.html

Finzi, D., Blankson, J., Siliciano, J., Margolick, J., Chadwick, K., Pierson, T., Smith, K., Lisziewicz, J., Lori, F., Flexner, C., Quinn, T., Chaisson, R., Rosenberg, E., Walker, B., Gange, S., Gallant, J., Siliciano, R.: Latent infection of CD4 t cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy. Nature Medicine **5**(5), 512–517 (1999). URL http://www.scopus.com/inward/record.url?eid=2-s2.0-0032953920&partnerID=40&md5=aaaca2836647600ddd16322194dd6482. Cited By (since 1996) 923

Frank, P.: Introduction to System Sensitivity Theory. Academic Press Inc (1978)

García, C.E., Prett, D.M., Morari, M.: Model predictive control: Theory and practice—a survey. Automatica **25**(3), 335–348 (1989). URL http://www.sciencedirect.com/science/article/pii/0005109889900022

Ge, S., Tian, Z., Lee, T.H.: Nonlinear control of a dynamic model of HIV-1. IEEE Trans Biomed Eng **52**(3), 353 –361 (2005). DOI 10.1109/TBME.2004.840463

Golub, G., Klema, V., Stewart, G.W.: Rank degeneracy and least squares problems. Tech. rep. (1976)

Golub, G., Loan, C.F.V.: Matrix Computations. John Hopkins University Press (1996)

van Griensven, A., Meixner, T., Grunwald, S., Bishop, T., Diluzio, M., Srinivasan, R.: A global sensitivity analysis tool for the parameters of multi-variable catchment models. Journal of Hydrology **324**(1â“4), 10 – 23 (2006). DOI 10.1016/j.jhydrol.2005.09.008. URL http://www.sciencedirect.com/science/article/pii/S0022169405004488

Griewank, A.: On automatic differentiation. Tech. Rep. CRPC-TR89003, Rice University (1989)

Gu, M., Eisenstat, S.C.: Efficient algorithms for computing a strong rank-revealing qr factorization. SIAM Journal on Scientific Computing **17**(4), 848–869 (1996). DOI 10.1137/0917055. URL http://link.aip.org/link/?SCE/17/848/1

Guedj, J., Thiébaut, R., Commenges, D.: Practical identifiability of HIV dynamics models. Bull Math Bio **69**, 2493–2513 (2007). URL http://dx.doi.org/10.1007/s11538-007-9228-7. 10.1007/s11538-007-9228-7

Gunzburger, M.: Perspectives in Flow Control and Optimization. Advances in Design and Control. Society for Industrial and Applied Mathematics (2003). URL http://books.google.com/books?id=dWKcdjHmMvYC

Heldt, T.: Computational models of cardiovascular response to orthostatic stress. Ph.D. thesis, Massachusetts Institute of Technology (2004)

Ho, D.D., Neumann, A.U., Perelson, A.S., Chen, W., Leonard, J.M., Markowitz, M.: Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. Nature **373**(6510), 123–126 (1995). URL http://dx.doi.org/10.1038/373123a0

Jacquez, J.: Compartmental analysis in biology and medicine: Kinetics of distribution of tracer-labeled materials. Elsevier Pub. Co. (1972). URL http://books.google.com/books?id=Us1qAAAAMAAJ

Joshi, H.R.: Optimal control of an HIV immunology model. Optimal Control Appl. Methods **23**, 199–213 (2002)

Kappel, F.: Subset selection. Tech. rep., University of Graz (2009)

Kiparissides, A., Kucherenko, S.S., Mantalaris, A., Pistikopoulos, E.N.: Global sensitivity analysis challenges in biological systems modeling. Industrial and Engineering Chemistry Research **48**(15), 7168–7180 (2009). DOI 10.1021/ie900139x. URL http://pubs.acs.org/doi/abs/10.1021/ie900139x

Kirk, D.: Optimal Control Theory: An Introduction. Dover books on engineering. Dover Publications (2004). URL http://books.google.com/books?id=fCh2SAtWIdwC

Kirschner, D., Lenhart, S., Serbin, S.: Optimal control of the chemotherapy of HIV. Journal of Mathematical Biology **35**, 775–792 (1997). URL http://dx.doi.org/10.1007/s002850050076. 10.1007/s002850050076

Lewis, F., Syrmos, V.: Optimal Control. A Wiley-Interscience publication. J. Wiley (1995). URL http://books.google.com/books?id=jkD37elP6NIC

Li, R., Henson, M., Kurtz, M.: Selection of model parameters for off-line parameter estimation. IEEE Trans Contr Sys Tech **12**(3), 402 – 412 (2004). DOI 10.1109/TCST.2004.824799

Lisziewicz, J., Lori, F.: Structured treatment interruptions in HIV/AIDS therapy. Microbes and Infection **4**(2), 207 – 214 (2002). DOI 10.1016/S1286-4579(01)01529-5. URL http://www.sciencedirect.com/science/article/pii/S1286457901015295

Lisziewicz, J., Rosenberg, E., Lieberman, J., Jessen, H., Lopalco, L., Siliciano, R., Walker, B., Lori, F.: Control of HIV despite the discontinuation of antiretroviral therapy. New England Journal of Medicine **340**(21), 1683–1683 (1999). DOI 10.1056/NEJM199905273402114

Ljung, L., Glad, T.: On global identifiability for arbitrary model parametrizations. Automatica **30**(2), 265 – 276 (1994). DOI 10.1016/0005-1098(94)90029-9. URL http://www.sciencedirect.com/science/article/pii/0005109894900299

Lori, F., Lewis, M.G., Xu, J., Varga, G., Zinn, D.E., Crabbs, C., Wagner, W., Greenhouse, J., Silvera, P., Yalley-Ogunro, J., Tinelli, C., Lisziewicz, J.: Control of SIV rebound through

structured treatment interruptions during early infection. Science **290**(5496), 1591–1593 (2000). DOI 10.1126/science.290.5496.1591. URL http://www.sciencemag.org/content/290/5496/1591.abstract

Lori, F., Lisziewicz, J.: Structured treatment interruptions for the management of HIV infection. JAMA: The Journal of the American Medical Association **286**(23), 2981–2987 (2001). DOI 10.1001/jama.286.23.2981. URL http://jama.ama-assn.org/content/286/23/2981.abstract

McLachlan, G., Krishnan, T.: The EM algorithm and extensions. Wiley series in probability and statistics. Wiley-Interscience (2008). URL http://books.google.com/books?id=BW4ZAQAAIAAJ

Miao, H., Dykes, C., Demeter, L., Cavenaugh, J., Park, S., Perelson, A., Wu, H.: Modeling and estimation of kinetic parameters and replicative fitness of HIV-1 from flow-cytometry-based growth competition experiments. Bulletin of Mathematical Biology **70**, 1749–1771 (2008). URL http://dx.doi.org/10.1007/s11538-008-9323-4. 10.1007/s11538-008-9323-4

Miao, H., Xia, X., Perelson, A.S., Wu, H.: On identifiability of nonlinear ode models and applications in viral dynamics. SIAM Rev. **53**(1), 3–39 (2011). DOI DOI:10.1137/090757009. URL http://dx.doi.org/doi/10.1137/090757009

Nelson, P.W., Gilchrist, M.A., Coombs, D., Hyman, J.M., Perelson, A.S.: An age-structured model of hiv infection that allows for variations in the production rate of viral particles and the death rate of productively infected cells. Math Biosci Eng **1**(2), 267–288 (2004)

Nelson, P.W., Perelson, A.S.: Mathematical analysis of delay differential equation models of HIV-1 infection. Mathematical Biosciences **179**(1), 73 – 94 (2002). DOI 10.1016/S0025-5564(02)00099-8. URL http://www.sciencedirect.com/science/article/pii/S0025556402000998

Nowak, M.A., Bangham, C.R.M.: Population dynamics of immune responses to persistent viruses. Science **272**(5258), 74–79 (1996). DOI 10.1126/science.272.5258.74. URL http://www.sciencemag.org/content/272/5258/74.abstract

Olufsen, M.S., Ottesen, J.T.: A practical approach to parameter estimation applied to model predicting heart rate regulation. J Math Biol (2011). DOI 10.1007/s00285-012-0535-8

Ortiz, G.M., Wellons, M., Brancato, J., Vo, H.T.T., Zinn, R.L., Clarkson, D.E., Loon, K.V., Bonhoeffer, S., Miralles, G.D., Montefiori, D., Bartlett, J.A., Nixon, D.F.: Structured antiretroviral treatment interruptions in chronically HIV-1-infected subjects. Proceedings of the National Academy of Sciences of the United States of America **98**(23), pp. 13,288–13,293 (2001). URL http://www.jstor.org/stable/3057079

Palacios, G.C., Sanchez, L.M., Briones, E., Ramirez, T.J., Castillo, H., Rivera, L.G., Vazquez, C.A., Rodriguez-Padilla, C., Holodniy, M.: Structured interruptions of highly active antiretroviral therapy in cycles of 4 weeks off/12 weeks on therapy in children having a chronically undetectable viral load cause progressively smaller viral rebounds. International

Journal of Infectious Diseases **14**(1), e34 – e40 (2010). DOI 10.1016/j.ijid.2009.03.003. URL http://www.sciencedirect.com/science/article/pii/S1201971209001374

Paterson, D.L., Swindells, S., Mohr, J., Brester, M., Vergis, E.N., Squier, C., Wagener, M.M., Singh, N.: Adherence to protease inhibitor therapy and outcomes in patients with HIV infection. Annals of Internal Medicine **133**(1), 21–30 (2000). URL http://www.annals.org/content/133/1/21.1.abstract

Perelson, A.S., Kirschner, D.E., Boer, R.D.: Dynamics of hiv infection of cd4+ t cells. Mathematical Biosciences **114**(1), 81 – 125 (1993). DOI 10.1016/0025-5564(93)90043-A. URL http://www.sciencedirect.com/science/article/pii/002555649390043A

Perelson, A.S., Nelson, P.W.: Mathematical analysis of HIV-1 dynamics in vivo. SIAM Rev. **41**, 3–44 (1999). DOI 10.1137/S0036144598335107

Perelson, A.S., Neumann, A.U., Markowitz, M., Leonard, J.M., Ho, D.D.: HIV-1 dynamics in vivo: Virion clearance rate, infected cell life-span, and viral generation time. Science **271**(5255), 1582–1586 (1996). DOI 10.1126/science.271.5255.1582

Perko, L.: Differential Equations and Dynamical Systems. Texts in Applied Mathematics. Springer (2001). URL http://books.google.com/books?id=A7fvvz9Puf8C

Phillips, A.: Long term probability of detection of HIV-1 drug resistance after starting antiretroviral therapy in routine clinical practice. AIDS **19**(5), 487–494 (2005). URL http://www.scopus.com/inward/record.url?eid=2-s2.0-17144369901&partnerID=40&md5=853e8a357b6f0987ad4011bd1b370d74. Cited By (since 1996) 89

Pope, S.R.: Parameter identification in lumped compartment cardiorespiratory models. Ph.D. thesis, North Carolina State University (2009)

Rao, R.C.: Information and the accuracy attainable in the estimation of statistical parameters. Bull. Calcutta Math. Soc. **37**, 81–91 (1945). URL http://www.ams.org/mathscinet-getitem?mr=0015748

Richalet, J., Rault, A., Testud, J.L., Papon, J.: Model predictive heuristic control: Applications to industrial processes. Automatica **14**(5), 413–428 (1978). URL http://www.sciencedirect.com/science/article/pii/0005109878900018

Rodriguez-Fernandez, M., Mendes, P., Banga, J.R.: A hybrid approach for efficient and robust parameter estimation in biochemical pathways. Biosys **83**(2-3), 248 – 265 (2006). DOI DOI:10.1016/j.biosystems.2005.06.016. URL http://www.sciencedirect.com/science/article/B6T2K-4HC776X-4/2/2a48c31a0d9aa413bc616023689e55c8. 5th International Conference on Systems Biology - ICSB 2004

Rong, L., Feng, Z., Perelson, A.S.: Mathematical analysis of age-structured HIV-1 dynamics with combination antiretroviral therapy. SIAM J. on Appl Math **67**(3), 731–756 (2007)

Rong, L., Perelson, A.S.: Modeling HIV persistence, the latent reservoir, and viral blips. Journal of Theoretical Biology **260**(2), 308 – 331 (2009). DOI 10.1016/j.jtbi.2009.06.011. URL http://www.sciencedirect.com/science/article/pii/S0022519309002665

Rosenberg, E.S., Altfeld, M., Poon, S.H., Phillips, M.N., Wilkes, B.M., Eldridge, R.L., Robbins, G.K., D'Aquila, R.T., Goulder, P.J.R., Walker, B.D.: Immune control of hiv-1 after early treatment of acute infection. Nature **407**(6803), 523–526 (2000). URL http://dx.doi.org/10.1038/35035103

Rosenberg, E.S., Davidian, M., Banks, H.T.: Using mathematical modeling and control to develop structured treatment interruption strategies for HIV infection. Drug and Alcohol Dependence **88, Supplement 2**(0), S41 – S51 (2007). DOI 10.1016/j.drugalcdep.2006.12.024. URL http://www.sciencedirect.com/science/article/pii/S0376871606004765. ¡ce:title¿Customizing Treatment to the Patient: Adaptive Treatment Strategies¡/ce:title¿

Shampine, L.F., Reichelt, M.W.: The matlab ode suite. SIAM Journal on Scientific Computing **18**(1), 1–22 (1997). DOI 10.1137/S1064827594276424. URL http://link.aip.org/link/?SCE/18/1/1

Menezes Campello de Souza, F.: Modeling the dynamics of HIV-1 and CD4 and CD8 lymphocytes. Engineering in Medicine and Biology Magazine, IEEE **18**(1), 21 –24 (1999). DOI 10.1109/51.740960

Tan, W.Y., Wu, H.: Stochastic modeling of the dynamics of CD4+ t-cell infection by HIV and some monte carlo studies. Mathematical Biosciences **147**(2), 173 – 205 (1998). DOI 10.1016/S0025-5564(97)00094-1. URL http://www.sciencedirect.com/science/article/pii/S0025556497000941

Thomaseth, K., Cobelli, C.: Generalized sensitivity functions in physiological system identification. Annals of Biomedical Engineering **27**, 607–616 (1999). URL http://dx.doi.org/10.1114/1.207. 10.1114/1.207

Thompson, D.E., McAuley, K.B., McLellan, P.J.: Parameter estimation in a simplified mwd model for hdpe produced by a ziegler-natta catalyst. Macromolecular Reaction Engineering **3**(4), 160–177 (2009). DOI 10.1002/mren.200800052. URL http://dx.doi.org/10.1002/mren.200800052

Valdez-Jasso, D., Haider, M., Banks, H., Santana, D., German, Y., Armentano, R., Olufsen, M.: Analysis of viscoelastic wall properties in ovine arteries. IEEE Trans Biomed Eng **56**(2), 210 –219 (2009). DOI 10.1109/TBME.2008.2003093

Vélez-Reyes, M.: Decomposed algorithms for parameter estimation. Ph.D. thesis, Massachusetts Institute of Technology (1992). URL http://hdl.handle.net/1721.1/12845

Verma, A.: An introduction to automatic differentiation. Current Science **78**(7), 804–807 (2000)

Wei, X., Ghosh, S.K., Taylor, M.E., Johnson, V.A., Emini, E.A., Deutsch, P., Lifson, J.D., Bonhoeffer, S., Nowak, M.A., Hahn, B.H., Saag, M.S., Shaw, G.M.: Viral dynamics in

human immunodeficiency virus type 1 infection. Nature **373**(6510), 117–122 (1995). URL http://dx.doi.org/10.1038/373117a0

WHO: Global HIV/AIDS response, epidemic update and health sector progress towards universal access. progress report 2011. Tech. rep., World Health Organization (2011)

Wodarz, D.: Helper-dependent vs. helper-independent CTL responses in HIV infection: Implications for drug therapy and resistance. Journal of Theoretical Biology **213**(3), 447 – 459 (2001). DOI 10.1006/jtbi.2001.2426. URL http://www.sciencedirect.com/science/article/pii/S0022519301924269

Wodarz, D., Nowak, M.A.: Specific therapy regimes could lead to long-term immunological control of HIV. Proc Natl Acad Sci USA **96**(25), 14,464–14,469 (1999)

Wu, H., Zhu, H., Miao, H., Perelson, A.: Parameter identifiability and estimation of HIV/AIDS dynamic models. Bulletin of Mathematical Biology **70**, 785–799 (2008). URL http://dx.doi.org/10.1007/s11538-007-9279-9. 10.1007/s11538-007-9279-9

Xia, X.: Estimation of HIV/aids parameters. Automatica **39**(11), 1983 – 1988 (2003). DOI 10.1016/S0005-1098(03)00220-6

Xia, X., Moog, C.: Identifiability of nonlinear systems with application to HIV/aids models. IEEE Trans Auto Contr. **48**(2), 330 – 336 (2003). DOI 10.1109/TAC.2002.808494

Yao, K.Z., Shaw, B.M., Kou, B., McAuley, K.B., Bacon, D.W.: Modeling ethylene/butene copolymerization with multisite catalysts: Parameter estimability and experimental design. Polymer Reaction Engineering **11**(3), 563–588 (2003). DOI 10.1081/PRE-120024426. URL http://www.tandfonline.com/doi/abs/10.1081/PRE-120024426

Yuan, Y., Allen, L.J.: Stochastic models for virus and immune system dynamics. Mathematical Biosciences **234**(2), 84 – 94 (2011). DOI 10.1016/j.mbs.2011.08.007. URL http://www.sciencedirect.com/science/article/pii/S0025556411001313

Zurakowski, R., Teel, A.R.: A model predictive control based scheduling method for HIV therapy. Journal of Theoretical Biology **238**(2), 368 – 382 (2006). DOI 10.1016/j.jtbi.2005.05.004. URL http://www.sciencedirect.com/science/article/pii/S0022519305002067