Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Journal of Computational and Applied Mathematics 184 (2005) 10–49
JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS
www.elsevier.com/locate/cam

HIV dynamics: Modeling, data analysis, and optimal treatment protocols

B.M. Adams[a], H.T. Banks[a,*], M. Davidian[a], Hee-Dae Kwon[a,1], H.T. Tran[a], S.N. Wynne[a], E.S. Rosenberg[b]

[a] *Center for Research in Scientific Computation, Box 8205, North Carolina State University, Raleigh, NC 27695-8205, USA*
[b] *Massachusetts General Hospital and Harvard Medical School, I.D. Unit-Gray 5, Fruit Street, Boston, MA 02114, USA*

Received 22 February 2004; received in revised form 15 August 2004

Abstract

We present an overview of some concepts and methodologies we believe useful in modeling HIV pathogenesis. After a brief discussion of motivation for and previous efforts in the development of mathematical models for progression of HIV infection and treatment, we discuss mathematical and statistical ideas relevant to Structured Treatment Interruptions (STI). Among these are model development and validation procedures including parameter estimation, data reduction and representation, and optimal control relative to STI. Results from initial attempts in each of these areas by an interdisciplinary team of applied mathematicians, statisticians and clinicians are presented.

MSC: 34A56; 49J15; 34H05; 49L20; 65K10; 92B05; 92C50; 62H25

Keywords: HIV models; Parameter estimation; Data and model reduction; Structured treatment interruptions; Optimal control

1. Introduction—HIV modeling and STI

Although the correlates of immune protection in HIV infection remain largely unknown, our knowledge of viral replication dynamics and virus-specific immune responses has grown. Concurrent with these

* Corresponding author.
E-mail addresses: bmadams3@unity.ncsu.edu (B.M. Adams), htbanks@unity.ncsu.edu (H.T. Banks), davidian@unity.ncsu.edu (M. Davidian), tran@math.ncsu.edu (H.T. Tran), snwynne@unity.ncsu.edu (S.N. Wynne), erosenberg1@partners.org (E.S. Rosenberg).
[1] Department of Mathematics, Inha University, Inchon, South Korea.

0377-0427/$ - see front matter
doi:10.1016/j.cam.2005.02.004

advances, there has been an abundance of mathematical models that attempt to describe these phenomena. The models proposed have principally involved linear and nonlinear ordinary differential equations, both with and without delay terms. While data fitting problems motivated the development of some models, others have been proposed in a more abstract sense to suggest new experiments or possibilities for ways in which the body interacts with this pathogen. Mathematical models alone cannot answer questions about the pathogenesis of HIV infection or similar biological processes. But when used in conjunction with data as part of designed experiments, models can be a powerful tool in understanding mechanisms in complex systems. Moreover, data-oriented mathematical models can also stimulate further clinical and laboratory research. In any discussions of mathematical modeling of complex systems, it is appropriate to point out that while complex models may be needed to provide accurate descriptions of the underlying dynamics, the models are most useful when they can be compared to clinical and/or experimental data. In developing models for HIV infection and treatment, this requires that a balance be struck between complexity and utility.

We begin this paper with a brief summary of issues that have arisen in the development of models for HIV infection and motivate the model fitting problem. We also offer an introduction to STIs as a potentially improved treatment strategy and indicate how mathematical control theory can be helpful in finding treatment schemes. This background will set the stage for the model- and data-based examples that follow.

1.1. Models for HIV infection

Numerous factors have been considered in modeling HIV infection as one must typically choose only a critical subset of the many possible biological compartments and interactions. Moreover, scale is important in that one must decide whether to model at the micro level of viral RNA or more at the systemic level. Our focus is on compartmental models in which compartments each typically correspond to a type of cell population throughout the body. We do not attempt to provide a comprehensive survey of the extensive collection of mathematical models used with HIV infection. Rather, we refer the reader to one of the excellent survey articles already published; see, for example, [15,55]. We provide a brief overview of some important developments here.

Investigations of the kinetics of virus and CD4+ T-cell populations using mathematical models with data from patients undergoing highly active anti-retroviral therapy (HAART) support the theory of very rapid and constant turnover of the viral and infected cell populations in all individuals studied; see, for example [30,66,56]. This contrasts with researchers' previous assumptions that the stable viral and CD4+ T-cell concentrations seen during the period of clinical latency of chronic HIV infection were due to the absence of any significant viral replication. The studies by Ho, Wei and Perelson indicate that both the viral and infected cell populations are turning over rapidly and continuously. Further work in [54] revealed a second population of longer lived infected cells contributing to the population of viral RNA. Since these reports, numerous groups have used mathematical models to estimate decay rates for infected cell populations [42,44,46,47,70]. In Section 2 we present a model that can predict the observed persistent low-level replication of virus and includes multiple infected cell populations.

The early linear models developed in [66,30,56,54] are approximations to more realistic nonlinear models for viral and infected cell decay, and thus are applicable only over short periods of time, most likely on the order of days. While these linear models have been extremely useful in characterizing short-term dynamics of HIV infection after therapy, several researchers have attempted to use these models to

estimate time to eradication of virus from individuals. Such predictions involve periods of time which extend beyond that which is appropriate for approximation of the nonlinear dynamics by a linear model.

To model data over longer periods of time and make predictions about long term outcomes, nonlinear mathematical models are necessary. In addition to the unrealistic simplifying assumptions that make it difficult for linear models to accurately describe long-term HIV infection dynamics, factors that could play an important role in dynamic disease outcomes may be omitted in linear models. For example, several authors have raised the question as to whether or not these linear mathematical models have adequately described the decay of compartments relevant to HIV infection dynamics. The authors in [8] argue that more complex nonlinear models are needed to accurately describe long-term viral decay. In [42] the authors point out that the biphasic pattern which has been attributed to two populations of infected cells could be the result of exponential decay of a single population of infected cells with decreasing exponent over time. This phenomenon is well-known in population biology, and is often referred to as density-dependent decay.

Viral production by cells infected with HIV depends on the "age" (e.g., time since infection) of the infected cells and there are several different biological aspects of this age dependence. Intra-cellular delay due to viral reverse transcription, integration, transcription, and virion formation is described by Mittler et al. [45], extending the work in [54]. Mittler allows intra-cellular delay to vary across cells, and estimates these delays to be more significant than the pharmacological delays associated with drug absorption. Recent efforts [5] with in vitro data suggest the importance of modeling these distributed delays with some care. Incorporation of this variability of delays into models may lead to improved estimates of the half-life of free virus from short-term clinical data on patients undergoing HAART.

Since the qualitative behavior of a dynamical system is determined by its underlying parameters, knowledge of the bifurcation properties of the system is important for understanding the associated characteristics of the biological system described by the model. If the range of model parameters for a population is such that dramatically different outcomes are predicted for different individuals, bifurcation values for different parameters could suggest target interventions for continued successful treatment. For example, loss of stability of the zero steady state for viral load could be reversed by treatments affecting the parameters which influence this stability. In addition, variability in initial conditions, which one can consider as parameters in the model, can lead to trajectories of the system lying in different regions of attraction, i.e., different initial population sizes can lead to dramatically different qualitative outcomes in a nonlinear model. Such situations are described in [9,68] for structured treatment interruption (STI) studies, discussed below. The models in these reports and that discussed in Section 2 have multiple equilibria; different equilibria describe the success or failure of the immune system to control infection and the initial conditions and parameters of the system determine which equilibrium is realized. Careful qualitative analysis of mathematical models that describe HIV infection dynamics can contribute to understanding of fundamental qualitative features of infection, and possibly suggest targets for improved disease monitoring and/or treatment.

In using any of these complex models, one would like to understand the identifiability of parameters, i.e., which parameters can be successfully estimated and with what type of data. In Section 3, we examine the ability to estimate parameters in a typical nonlinear model of HIV infection dynamics. In doing so, we describe a relevant inverse problem methodology, including sensitivity analysis, the generalized least squares method, and estimation of standard errors in the estimation process.

The complexity of nonlinear HIV models and data sets may necessitate the extraction of only essential states or information in order to successfully fit models to data. In Section 4 we demonstrate the use of

the proper orthogonal decomposition (POD) method to capture essential elements of actual HIV patient time series data. This data reduction technique is similar to the singular value decomposition or principal component analysis often seen in statistical literature and could be used to reduce the dimensionality of ODE model solutions as well. While this tool may not be helpful for determining variability of individuals in a population, it can help determine the essential characteristics which pervade the entire population. One possibility is to fit ODE models to the obtained essential basis elements, rather than to the noisy aggregate population data.

1.2. STIs and control theory

In spite of the early hopes of a cure for HIV infection with potent antiretroviral therapy, it has become apparent that suppression, and not eradication, is most likely the best that can be achieved with this strategy. The current strategy of continuous combination therapy is difficult to maintain for long periods of time due to short- and long-term toxicities (metabolic abnormalities, body habitus changes, lipid abnormalities, mitochondrial toxicity and liver toxicity) as well as adherence challenges inherent in these complicated pill regimes, cost, and the development of resistance to medications.

Consequently, strategies have been sought that can decrease the side effects for and medication burden placed on patients while maintaining some control over HIV replication through an individual's immune system. Excitement for such strategies surfaced with the reporting of the "Berlin patient" [38], a patient initially treated during the course of seroconversion, who interrupted therapy temporarily twice within the first 6 months of treatment and then permanently stopped treatment and maintained control of HIV replication.

Numerous authors have discussed the effect on immune response as antigen declines [19,12,18,49,50,16,26,33,62,52,57,10,60], and these contributions suggest that viral replication is required to maintain an HIV-1 specific Cytotoxic T Lymphocyte (CTL) response. These observations are supported by data in [33,62] where it is observed that HIV-specific CTL activity increases in patients who experience viral rebound, either as a spike during ongoing suppression or as the result of changes in treatment regimen. These results, combined with the data on the "Berlin patient," have led a number of investigators to propose methods of immune stimulation through vaccine or structured treatment interruption (STI) once patients have successfully suppressed viral replication. Hence, an underlying premise in treatment interruption is that allowing virus to replicate stimulates the immune system and may even cause some patients to "self-vaccinate."

In an ideal scenario, individual immune response would be stimulated and augmented sufficiently through a number of interruptions to the point where patients control their infections without need for the cumbersome treatment regimens currently available. Initial investigations aimed at assessing immune response to viral replication through treatment interruptions offer promise. Control of viremia during therapy discontinuation has been associated with successful HIV-1 specific immune response [38,18,51]. A good overview of the concept of STI and its promise in several scenarios is offered in [37].

Many studies of treatment interruption assess changes in HIV-specific immune response [59,28,53,24]. In investigations of treatment interruption in acutely infected individuals [51,58], HIV-specific immune responses were associated with control of viremia. Ortiz et al. [51] note the role of broad and strong CTL response in tandem with neutralizing antibodies in association with control of viremia after drug discontinuation. Rosenberg et al. [58], report increased and sustained HIV-specific CTL response during one, two, or three treatment interruptions, and, once CTL response was established, eventual control

of viremia in all eight patients studied. These examples support the inclusion of immune response as a factor in any model of HIV infection to be used in STI studies. They also suggest that a realistic model should take into account the possibility of little or no immune response in some situations and a strong HIV-specific immune response in others. The model discussed in Section 2 below tries to capture these features.

The cited studies have involved small numbers of patients and most do not employ standardized protocols for interruption schedules (e.g., see [59]). Two of the primary questions associated with STI are: (i) Is there an optimal duration of therapy needed in order to benefit from structured treatment interruption? (shorter lengths of initial therapy may limit the number of toxicities and inconveniences experienced by patients but may not be sufficient to maximally suppress HIV in all body compartments to allow for controlled viral replication in a structured interruption setting); (ii) Can the host immune system be augmented through repeated exposure to viral antigens so as to control HIV infection without the use of anti-retroviral therapies? There is data to suggest that repeated exposure to antigen stimulates the immune system. Moreover, some individuals appear to control their infections after treatment interruptions with little or no need for further treatment.

In principle, these questions may be addressed using standard comparative statistical analysis techniques if sufficient numbers of subjects are randomized to different regimens (e.g., length of therapy). However, questions about the relationship between antigen level and immune response or factors which influence biological mechanisms associated with host immune control of infection may be more appropriately addressed through the use of mathematical modeling (including such models as those referenced above in Section 1.1) with formal estimation of and inference on biologically meaningful parameters combined with control theory.

Previous efforts on mathematical modeling of treatment interruption have been reported in [68,9]. Wodarz and Nowak [68] employ a mathematical model which represents uninfected CD4+ T-cells, infected CD4+ T-cells, CTL precursors (immune memory), and CTL effectors. Bonhoeffer et al. [9] use a model incorporating uninfected, actively infected, and latently infected T-cells, as well as immune response (CTL effectors). Both investigations offer important theoretical insights into immune control of virus based on treatment strategies. In what follows, we utilize another model that incorporates some of the ideas from these papers in the context of STI scenarios.

A number of researchers [11,14,21,32,36,67] have begun to consider theoretical control of HIV infection models in attempts to suggest optimal treatment strategies. Open loop control for differential equation models is considered in [14,36,21] to develop strategies for treatment in primary infection. Other authors [11,67] use feedback controls, i.e., control laws chosen in a real-time fashion based on current observations of the (full) state of the underlying dynamical systems. In these studies the "controls" range from chemotherapy to combination strategies, e.g., see [32], in which treatments enhancing the immune system component of the model are combined with those that directly delay progression of disease.

For example, to control the dynamics of the mathematical system that describes the viral and immune system dynamics during structured treatment interruption, one might consider a scaled periodic control input $u(t)$ (therapy) that is zero for some time Δt_1 and then rises with constant slope (linearly) to a value of one, remains there for time Δt_2, and then ramps linearly back to zero. This input affects states and parameters in the dynamic model that are related to infectivity. A cost functional, established in collaboration with clinicians, subject to the infection dynamics, could then be investigated. In a typical situation one might minimize the cost functional over permutations of treatment schedules; the free parameters would then be Δt_1 and Δt_2 in the control $u(t)$, yielding a treatment that is in some sense,

optimal. We demonstrate open loop control theoretic techniques for solving this problem with examples of continuously variable drug control (Section 5) and discrete (STI-like) control (Section 6).

Eventually, one should investigate these types of control problems in the context of patient data. Since patient data for immune response and antigen behavior are collected relatively infrequently and involve only partial state observations, a compensator theory combined with state feedback for nonlinear systems must be developed. Mathematical tools for closed-loop or feedback control of Structured Treatment Interruption for specific models are still needed. One must also determine the type and quantity of data on a patient's current health that would be required to successfully carry out an optimal or sub-optimal feedback control-based STI.

In the remainder of this paper, we discuss mathematical issues and tools relevant to STIs for HIV infection. Rather than purporting to provide a specific solution for HIV treatment, the nature of our presentation is more to demonstrate these tools with examples that may be pertinent to data available from clinicians. We suggest mathematical methods for solving some of the related problems and point out where further developments are needed.

2. A model for control and parameter estimation

In this section we discuss a number of important issues encountered in attempting to formulate a dynamic model of HIV progression in an individual. Desirable features in a model to be used with clinical observations of treated and untreated patients are presented. In addition to summarizing some of the difficulties in choosing a model, we ultimately must formulate a "typical" model to be employed in the subsequent discussions of mathematical and statistical methodologies presented in the remaining sections of this paper. Given this goal and given the existence of several recent survey papers on models noted in the Introduction, we do not include here discussions of all the important aspects of modeling of HIV. Nor do we represent that we have a good model of HIV progression and treatment; rather we have a model that contains some of the more desirable features one could expect!

A wide variety of mathematical models have been proposed to describe various aspects of in-host HIV infection dynamics. The most basic of these models typically include two or three of the key dynamic compartments: virus, uninfected target cells, and infected cells. A model used for estimation and control should include these compartments at a minimum. Infected and uninfected CD4+ T-cells are not usually censused separately, so they may need to aggregated for purposes of model fitting, i.e., the estimation of mechanisms and parameters in a model. However, current knowledge of the infection process justifies incorporation of three distinct physiologically relevant compartments. In addition, the documented importance of the immune system in responding to HIV infection (and especially its apparent crucial role during structured treatment interruptions) strongly motivates the inclusion of at least one model compartment representing immune response to the pathogen. We thus seek a model that includes some measure of cytotoxic T-lymphocyte (CD8) response to HIV infection. The incorporation of these compartments is important when relating the model to observable quantities in patient data.

A model used to suggest treatment strategies should be capable of incorporating the action of commonly used reverse transcriptase (RT) inhibitors and protease inhibitors (PIs). Inclusion of the latter usually implies inclusion of a compartment for virus rendered noninfectious by the PI. Given the predominance of HAART in the form of drug cocktails combining two or more drugs, the model should behave reasonably

when simulating multi-drug therapy. In spite of these comments, the model we will employ herein only includes RT inhibitors, but this aspect could be readily generalized.

As pointed out in [15] in the 2002 review paper, a reasonable model of HIV infection predicts a nonzero steady-state viral load in the presence of drug therapy. Patients who are subject to drug therapy often successfully suppress virus for a long time, potentially at undetectable levels. However, some reservoir exists which almost always causes the virus to grow out again upon removal of drug therapy. Hence we do not expect incorporation of drug therapy in the model, at a sensible efficacy, to drive the viral load to zero, but rather reduce it considerably, perhaps below the assay limits of detection. The authors of [15] note that many models typically employed to describe HIV infection do not exhibit a reasonable relationship between drug efficacy and the viral load. In these models, a very slight change in drug efficacy can yield a drastic change in the viral load equilibrium. This has important consequences for the control problem—a successful model must exhibit reasonable sensitivity of the viral load equilibrium to treatment efficacy.

The model we provide as an example is adapted from one proposed by Callaway and Perelson (denoted (5.3) in [15]). Our modified system of ODEs is given by

Type 1 target:

$$\dot{T}_1 = \lambda_1 - d_1 T_1 - (1-\varepsilon) k_1 V T_1,$$

Type 2 target:

$$\dot{T}_2 = \lambda_2 - d_2 T_2 - (1 - f\varepsilon) k_2 V T_2,$$

Type 1 infected:

$$\dot{T_1^*} = (1-\varepsilon) k_1 V T_1 - \delta T_1^* - m_1 E T_1^*,$$

Type 2 infected:

$$\dot{T_2^*} = (1 - f\varepsilon) k_2 V T_2 - \delta T_2^* - m_2 E T_2^*,$$

Free virions:

$$\dot{V} = N_T \delta (T_1^* + T_2^*) - cV - [(1-\varepsilon)\rho_1 k_1 T_1 + (1 - f\varepsilon)\rho_2 k_2 T_2] V,$$

Immune effectors:

$$\dot{E} = \lambda_E + \frac{b_E (T_1^* + T_2^*)}{(T_1^* + T_2^*) + K_b} E - \frac{d_E (T_1^* + T_2^*)}{(T_1^* + T_2^*) + K_d} E - \delta_E E. \tag{2.1}$$

This model entails two co-circulating populations of target cells, perhaps representing CD4 T-lymphocytes (T_1) and macrophages (T_2). The drug efficacy parameter ε models a reverse transcriptase (RT) inhibitor that blocks new infections and is potentially more effective in population 1 (T_1, T_1^*) than in population 2 (T_2, T_2^*), where the efficacy is $f\varepsilon$, with $f \in [0, 1]$. As in the Callaway–Perelson paper, we only model an RT inhibitor, though again, it is simple to incorporate a PI. The populations of uninfected target cells T_1 and T_2 may have different source rates λ_i and natural death rates d_i.

As is common in models of HIV infection, infected cells T_i^* result from encounters between uninfected target cells T_i and free virus V. The natural infection rate k_i may differ between the two populations, which could account for believed differences in activation rates between lymphocytes and macrophages. The differences in infection rates and treatment efficacy help create a low, but nonzero, infected cell steady

state for T_2^*, which is commensurate with the idea that macrophages may be an important source of virus after T-cell depletion. Infected cells may be removed from the system via either natural death or by the action of immune effector cells E described below. For our efforts here we assume that both target cell types have the same death rate δ, though this could be readily generalized as well.

To preserve simplicity in the model, we omit the chronically infected cell compartments proposed in the original Callaway–Perelson model. The important qualitative behaviors seem preserved in the model we propose and specifically modeling this feature is not essential to our present work. In particular, the existence of a low steady-state viral load equilibrium and sensitivity of the viral load equilibrium to the drug efficacy is obtained with or without such compartments. We note that while removing the chronically infected compartments does not affect the sensitivity to treatment, the addition of immune response terms does, as discussed below.

Free virus particles are produced by both types of infected cells, which we assume produce virus at the same rate (again this could be easily generalized to account for different productivity). In the Callaway–Perelson model, virus only leaves the V compartment via natural death at rate c; there is no removal term in the $\dot{V}$ equation representing loss of virus due to infection of a cell. One potential justification for this omission is offered by Nelson and Perelson in [55, p. 10] who suggest that this term can be omitted since the term $k_i T_i V$ is small in comparison to cV in the typical HIV-infected patient. They further assert that if T_i is approximately constant, one can absorb the loss of virus due to infection into the cV term, thus making it account for all clearance processes.

While the arguments offered by Nelson and Perelson could justify the exclusion of the virus removal term, we investigate situations where treatment is interrupted abruptly, potentially effecting drastic changes in all of the cell populations under consideration. We therefore include the term $[(1-\varepsilon)\rho_1 k_1 T_1 + (1-f\varepsilon)\rho_2 k_2 T_2]V$ in the $\dot{V}$ equation to account for the removal of free virus that takes place when free virions infect a T_1 or T_2 cell. We make the simplifying assumption $\rho_i = 1$, i.e., one free virus particle is responsible for each new infection. This could easily be adapted for multiple virus particles being responsible for each new infection by choosing $\rho_i > 1$.

Finally, the immune effectors E, or cytotoxic T-lymphocytes, are produced in response to the presence of infected cells and existing immune effectors. The immune response assumed here is similar to that suggested by Bonhoeffer, et al., in their 2000 paper [9], with a Michaelis–Menten type saturation nonlinearity. The infected cell dependent death term in the immune response represents immune system impairment "at high virus load". In [9], the authors demonstrate that a model with this structure of immune response and a latently infected cell compartment can exhibit transfer between "healthy" and "unhealthy" stable steady states via STI, making it a good candidate for our investigation. We also add a source term λ_E to create a nonzero off-treatment steady state for E. While immune effectors are not inherently present, some small persistence of them is to be expected during infection. We note that other immune responses models, such as those considered in [68] or [48] could be substituted, if desired. However, the latter does not appear to admit multiple stable off-treatment steady states.

The immune response we model is that of cytotoxic T-lymphocytes (CTL). CTL act by lysing infected cells, causing them to explode. Thus CTL remove infected cells from the system in the equations for $\dot{T}_1^*$ and $\dot{T}_2^*$, at rates m_1 and m_2, respectively. Unlike interferons, they do not directly target free virus, so there is no interaction term with the virus compartment. As with any immune system responders, we suspect that CTL sometimes mistarget or misidentify receptors and thus kill healthy cells or misidentify self versus antigen. But again for the sake of simplicity, we do not attempt to model that here.

Table 1
Parameters used in model (2.1)

Parameter	Value	Units	Description
λ_1	10,000	$\frac{cells}{ml \cdot day}$	Target cell type 1 production (source) rate
d_1	0.01**	$\frac{1}{day}$	Target cell type 1 death rate
ε	$\in [0, 1]$	–	Population 1 treatment efficacy
k_1	8.0×10^{-7}	$\frac{ml}{virions \cdot day}$	Population 1 infection rate
λ_2	31.98	$\frac{cells}{ml \cdot day}$	Target cell type 2 production (source) rate
d_2	0.01**	$\frac{1}{day}$	Target cell type 2 death rate
f	0.34 ($\in [0, 1]$)	–	Treatment efficacy reduction in population 2
k_2	1×10^{-4}	$\frac{ml}{virions \cdot day}$	Population 2 infection rate
δ	0.7*	$\frac{1}{day}$	Infected cell death rate
m_1	1.0×10^{-5}	$\frac{ml}{cells \cdot day}$	Immune-induced clearance rate for population 1
m_2	1.0×10^{-5}	$\frac{ml}{cells \cdot day}$	Immune-induced clearance rate for population 2
N_T	100*	$\frac{virions}{cell}$	Virions produced per infected cell
c	13*	$\frac{1}{day}$	Virus natural death rate
ρ_1	1	$\frac{virions}{cell}$	Average number virions infecting a type 1 cell
ρ_2	1	$\frac{virions}{cell}$	Average number virions infecting a type 2 cell
λ_E	1	$\frac{cells}{ml \cdot day}$	Immune effector production (source) rate
b_E	0.3	$\frac{1}{day}$	Maximum birth rate for immune effectors
K_b	100	$\frac{cells}{ml}$	Saturation constant for immune effector birth
d_E	0.25	$\frac{1}{day}$	Maximum death rate for immune effectors
K_d	500	$\frac{cells}{ml}$	Saturation constant for immune effector death
δ_E	0.1*	$\frac{1}{day}$	Natural death rate for immune effectors

Those in the top section of the table are taken directly from Callaway and Perelson. Parameters in the bottom section of the table are taken from Bonhoeffer et al., with K_b and K_d scaled to reflect the volumetric units used in our model and also adjusted. The superscripts * denote parameters the authors indicated were estimated from human data and ** denote those estimated from macqaque data.

2.1. Parameter specifications

The model (2.1) contains numerous parameters that must be assigned values before simulations can be carried out. In specifying model parameters, to the greatest extent possible we employ values similar to those reported or justified in the literature. The parameters indicated in Table 1 are principally extracted from the Callaway–Perelson [15] and Bonhoeffer et al. [9] papers.

Callaway and Perelson point out that several model parameters are not available from human or animal data. They choose the parameters $\lambda_1, k_1, \lambda_2,$ and k_2 such that several conditions on target cell and viral load equilibria are satisfied for their model. We note that these conditions are not precisely satisfied by our model which has no chronic cell compartment and an added immune response. However, the conditions are closely approximated by the model's behavior and we believe the parameters could be adjusted to obtain the same qualitative behavior. Further explanation of these criteria appears below.

Table 2
Off treatment ($\varepsilon = 0$) steady states (c/ml) for model (2.1). Nonphysical steady states have been omitted

	EQ_1	EQ_2	EQ_3	EQ_4
T_1	1 000 000	664 938	163 573	967 839
T_2	3198	50	5	621
T_1^*	0	1207	11945	76
T_2^*	0	11	46	6
V	0	6299	63919	415
E	10	207658	24	353108
Local stab.	Unstable	Unstable	Stable	Stable

In general, immune response parameters are not well known and are thus frequently chosen to demonstrate model behavior in simulations. The parameters m_1 and m_2 represent cytopathicity of the immune effectors. Their common value was taken from [15] where the authors note that the value was suggested originally in [48]. We scale the parameter K_b by a factor of 1000 since the Bonhoeffer paper appears to use units of microliters, whereas we consider infected cells per milliliter. The parameter K_d is adjusted by a factor of 100 in order to demonstrate the possibility of multiple stable equilibria for the model.

2.2. Model steady states and simulation

Given the specified parameters, in the absence of therapy, the model exhibits several steady states as shown in Table 2. Equilibrium EQ_1 represents the uninfected patient, with healthy T-cell counts and no virus or infected cells present. When the system is in this state, introduction of a small amount of virus causes the system to converge to EQ_3, where healthy target cells are substantially depleted and a dangerously high viral load is present. The system also exhibits a second stable equilibrium EQ_4, where a strong immune response has developed, successfully controlling the virus and consequently restoring target cell (T_1) help.

We turn to the criteria on the steady states suggested by Callaway and Perelson. When $\varepsilon = 1$ and $f = 0$, they desire a viral load of 100. In this scenario, our model exhibits a viral load of $V = 143\,\mathrm{c/ml}$, which is on that order. The equilibrium is:

$$T_1 = 1\,000\,000, \quad T_2 = 1314, \quad T_1^* = 0, \quad T_2^* = 27, \quad V = 143, \quad E = 20.$$

When $\varepsilon = 1$ and $f = 0.5$, they desire eradication of the virus. Our model does not precisely attain that, though the viral steady state ($V = 43\,\mathrm{c/ml}$) is below the limit of detection for most assays currently in use. The complete equilibrium is given by:

$$T_1 = 1\,000\,000, \quad T_2 = 2627, \quad T_1^* = 0, \quad T_2^* = 8, \quad V = 43, \quad E = 12.$$

Further, by increasing the treatment efficacy factor to $f = 0.6$, the virus is eradicated because the uninfected steady state becomes stable.

Fig. 1 depicts the sensitivity of the viral load equilibrium to drug efficacy ε. The introduction of the immune response terms causes a discontinuity in this curve as stability is exchanged between viral dominant and immune-dominant equilibria. This is in contrast to the sensitivity curve of the original Callaway–Perelson model which is continuous across the range of drug efficacies. However, this may be

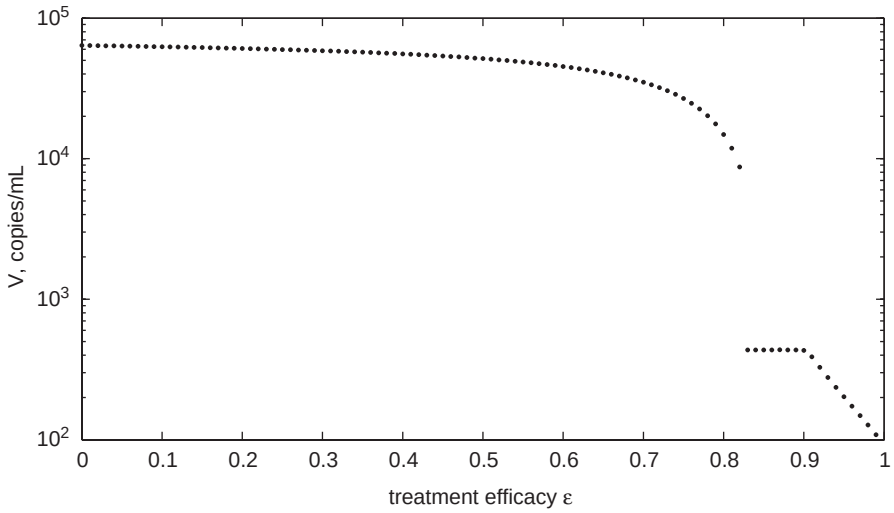

Fig. 1. Sensitivity of stable viral load equilibrium to drug.

reasonable—a certain amount of drug may be necessary to augment the immune system to counteract effects of the virus. The low (nonzero) steady state viral load is maintained across the entire range of reasonable drug efficacies. The model also exhibits approximately a 1–2 log drop in viral load across the range of drug efficacies, which is typical for monotherapy with RT inhibitors.

A simulation of early infection is shown in Fig. 2. Simulation is started near steady state EQ_1, with the addition of one (1) viral copy per ml ($V = 1\,\mathrm{c/ml}$). Upon infection, the virus replicates to a peak before converging in damped oscillations to equilibrium. There is a delayed initial immune response to the presence of the infected cells, but a sustained and vigorous immune response does not develop. The higher infection rate in the T_2 population is evident in its more dramatic decline than the T_1 cell population.

3. Sample inverse problem calculation and standard error estimates

As discussed above, if one wishes to use a mathematical model to make predictions about a particular individual or population, estimation of model parameters from data is crucial. In this section, we outline an inverse problem methodology for estimating parameters and associated uncertainty from single patient data. We use simulated data generated in a manner that respects conditions encountered in clinical studies. It is generally a standard (and wise) practice to test one's inverse algorithms on simulated data before attempting to use them with experimental or clinical data.

3.1. Simulated data generation

In generating simulated data for a single patient, we assume that measurement error is a dominant source of variation and that measurements of different system states (e.g., T_1 and V) are independent at any given time. This is valid for certain, but not all, assays, and a reasonable assumption given that the

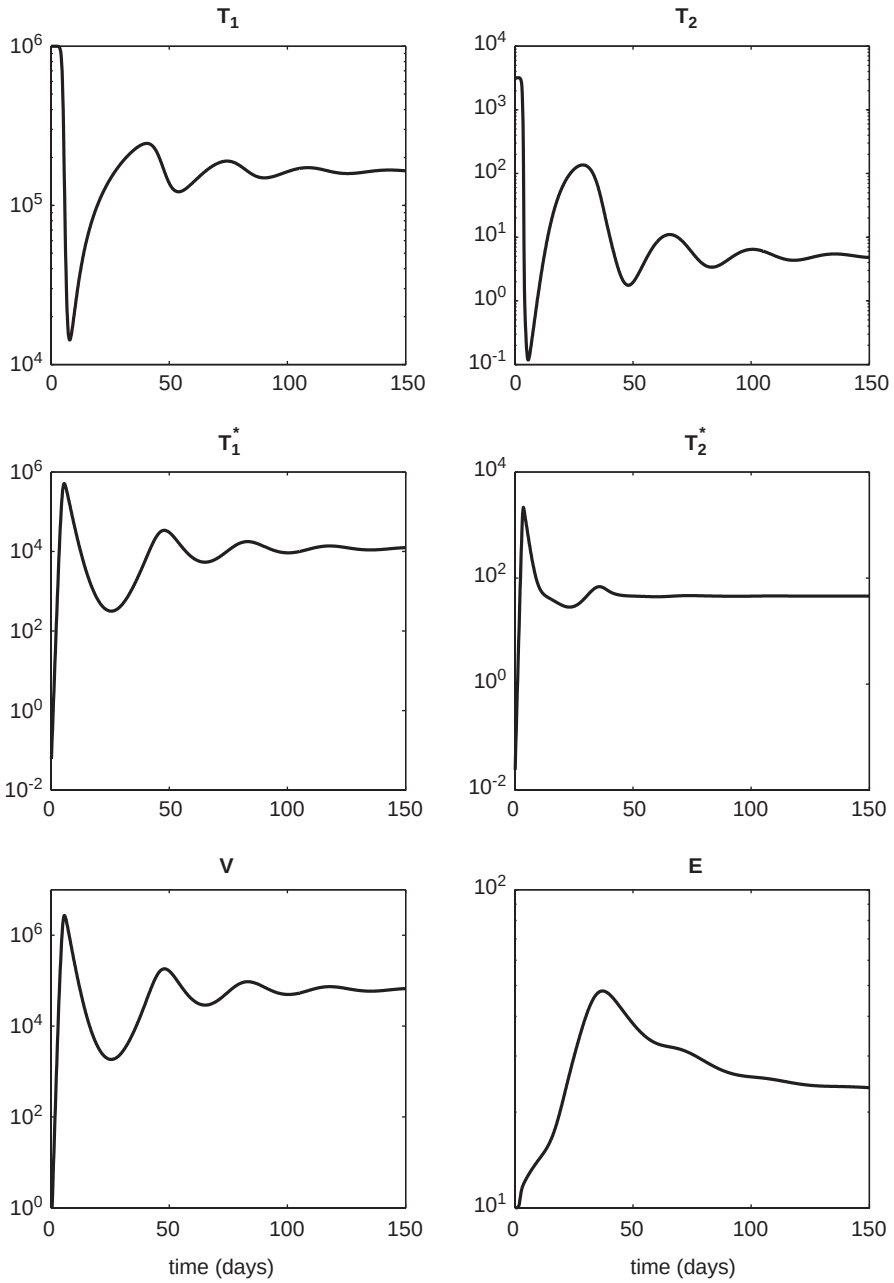

Fig. 2. Simulation of early infection scenario using model (2.1). Note varying scales in different subplots.

assays here are not performed sequentially but independently on distinct portions of the same specimen or on entirely different specimens. Measurements over time in clinical studies are generally taken on a given patient at intervals sufficiently separated in time that correlation over time may be reasonably

taken as negligible. We thus assume further that measurements over time are independent and take typical measurement intervals of 3–10 days.

Assume that the mean response for measurements on a single patient is given by the solution to the ODE model with some true parameter vector q_0. Specifically, let $\bar{x} = (T_1, T_2, T_1^*, T_2^*, V, E)$ be the state vector of the ODE model (2.1), and denote the model response for state i at time t^j by $\bar{x}_i^j(q_0) = \bar{x}_i(t^j; q_0)$, so that $i = 1, \ldots, 6$ indexes the state and $j = 1, \ldots, N$ indexes the observation times. Application of an observer matrix C to the model solution to observe only some states or combination of states results in values $x_i^j(q_0)$, $i = 1, \ldots, \rho$. This renders the model consistent with actual patient data where, for example, only the sum of T_1 and T_1^* may be measured. We then assume that y_i^j, observed measurements on x_i^j, are independent with mean $E(y_i^j) = x_i^j$.

Variation introduced by the assay and other sources is acknowledged to increase with the mean and is commonly represented by the variance model $\mathrm{Var}(y_i^j) = \sigma_i^2\{x_i^j\}^2$, where σ_i is the (constant) coefficient of variation corresponding to state i. Consistent with this, we take the y_i^j to be independently lognormally distributed, which is accomplished by generating independent z_i^j such that

$$z_i^j \sim \mathcal{N}(\log x_i^j - \log(\sigma_i^2 + 1)/2, \log(\sigma_i^2 + 1)) \tag{3.1}$$

and then setting $y_i^j = \exp(z_i^j)$. It may be verified using moment generating functions that y_i^j so generated have the desired mean and variance. An alternative probability model is the normal relative error model commonly used in the mathematical literature, where for $e_i^j \sim \mathcal{N}(0, \sigma_i^2)$, $y_i^j = x_i^j(1 + e_i^j)$. However, this model allows the possibility of generation of unrealistic negative measurements, making the lognormal a more natural choice for our systems and data.

We employ this method to generate simulated data with coefficient of variation $\sigma_i = 0.20$ for each state, which is in the range observed in practice for viral load and CD4 count assays (see, e.g., [69,63]). As above, we consider the early infection scenario, with parameters previously specified. We consider two scenarios—one where all six states are observed and one where there are only three observed quantities: $(T_1 + T_1^*)$, V, and E, as is typical in our clinical data sets.

3.2. Semi-relative sensitivity analysis

As a first step in testing the feasibility of inverse problem calculations, we attempt to estimate those parameters to which the model solution is most sensitive. These, in turn, are determined through a semi-relative sensitivity calculation.

To describe this methodology, suppose we wish to determine the relative sensitivity of the observed model quantities x (and consequently the model solution $\bar{x}$) to particular parameters q_k, $k = 1, \ldots, r$. The semi-relative sensitivity of the model solution to parameter q_k is given by

$$\frac{\partial \bar{x}(t; q)}{\partial q_k} \cdot q_k \tag{3.2}$$

(see [4]) and is computed by formally differentiating the ODE model

$$\frac{\mathrm{d}\bar{x}}{\mathrm{d}t} = f(t, \bar{x}; q)$$
$$\bar{x}(0) = \bar{x}_0,$$

with respect to q_k and interchanging the order of the time and parameter derivatives as outlined in [4,20,23]. In the case of r parameters and n state variables, we thus obtain an $(n \times r)$-dimensional system of differential equations for the sensitivities $\bar{x}_q(t; q) = (\partial\bar{x}/\partial q)(t; q)$, where q is the vector of parameters considered:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\bar{x}(t)}{\partial q}\right) = \frac{\partial f}{\partial\bar{x}}\,\frac{\partial\bar{x}(t)}{\partial q} + \frac{\partial f}{\partial q} \tag{3.3}$$

with initial condition

$$\frac{\partial\bar{x}(0)}{\partial q} = \frac{\partial\bar{x}_0}{\partial q} = 0. \tag{3.4}$$

The latter is simply the zero matrix whenever the initial condition is not a function of any of the parameters estimated. Here $\partial f/\partial\bar{x}$ is the Jacobian of the ODE system, and $\partial f/\partial q$ the derivative of the right side with respect to the parameters considered.

We solve {(3.3), (3.4)} for $\bar{x}_q(t; q)$ by coupling it with the original differential equation system to obtain an $(nr + n)$-dimensional system which we solve numerically with a Matlab ODE solver (the routine $ODE15s$). For another application of this methodology see [1]. Since our observation process on the model states is linear in these states, we can apply the observer C to these sensitivities to obtain the sensitivities $x_q(t; q)$ for the relevant observed model quantities.

This process yields sensitivity information as a function of time over the interval of interest. We wish to have some overall measure of the sensitivity of the solution to the parameters, so for each state/parameter combination, we take a norm (the L_2 norm) in time and then rank the resulting scalars to determine the most sensitive parameters. We find that these comparisons are similar using the two or sup norm—and report using the two norm here. Also, most of the interesting system dynamics occur in the first 50 days, so using a longer time span for these computations does not affect the results.

Fixing the parameters in (2.1) that have already been estimated from human or monkey data as known, we find that of the remaining parameters, the model solutions (especially states T_1 and V) are most sensitive to the parameters k_1, λ_1, k_2, and λ_2 (see Table 3). Several other parameters follow closely with less relative sensitivity, including b_E, d_E, and δ_E. Parameters near the end of the chart are likely more difficult to estimate from data.

Since the parameter ε=0 for these simulations, we find that as expected the model solution is completely insensitive to the treatment parameters ε and f. When considering a scenario with treatment $\varepsilon > 0$, the model solutions are highly sensitive to these parameters (results omitted). We also note that when only making partial state observations, the observables are most sensitive to these parameters in the same order as reported here.

There are alternate methods for computing the sensitivity matrix—see [6] for another efficient approach for some systems. For example, in [6], the authors determine the sensitivities x_q by taking forward difference quotient approximations near q. This replaces the sensitivity equation solution step with r extra solves of the ODE system. We note that for the initial conditions and parameters considered here it

Table 3
Ranked semi-relative sensitivity of various model states to parameters not previously estimated from human or monkey data

State x_i	Param q_k	$\left\| \frac{\partial x_i}{\partial q_k}(t) \right\|_2$	State x_i	Param q_k	$\left\| \frac{\partial x_i}{\partial q_k}(t) \right\|_2$
V	k_1	3.28488e+07	T_2^*	λ_1	1.33382e+03
T_1	k_1	1.04194e+07	E	λ_1	1.28026e+03
V	λ_1	6.51520e+06	T_1^*	δ_E	1.24200e+03
V	k_2	6.43367e+06	T_2	λ_2	1.08470e+03
T_1^*	k_1	6.21024e+06	E	k_1	1.05731e+03
T_1	λ_1	4.30398e+06	T_2^*	λ_2	1.03413e+03
T_1	k_2	1.70261e+06	T_1^*	m_1	6.32270e+02
T_1^*	λ_1	1.22538e+06	E	λ_E	6.08686e+02
T_1^*	k_2	1.21751e+06	E	K_d	6.04010e+02
V	λ_2	2.84012e+05	T_1^*	K_d	5.28876e+02
T_1	λ_2	2.16825e+05	T_1^*	λ_E	5.07648e+02
T_1^*	λ_2	5.35385e+04	E	K_b	3.69105e+02
T_2	k_1	1.94056e+04	T_1^*	K_b	3.20193e+02
V	b_E	1.76378e+04	V	m_2	2.41282e+02
T_2^*	k_1	1.55804e+04	T_1	m_2	1.21678e+02
T_1	b_E	1.43228e+04	E	λ_2	9.11780e+01
V	d_E	1.12343e+04	T_1^*	m_2	4.52933e+01
T_2	k_2	9.81344e+03	E	k_2	3.11716e+01
T_1	d_E	9.03250e+03	T_2	b_E	1.16408e+01
T_2^*	k_2	8.08379e+03	T_2	d_E	7.77859e+00
V	δ_E	6.63642e+03	T_2	δ_E	4.38937e+00
T_2	λ_1	6.27148e+03	T_2^*	b_E	3.13943e+00
T_1	δ_E	5.40935e+03	T_2	m_1	3.04099e+00
E	b_E	3.91597e+03	T_2	λ_E	2.53350e+00
V	m_1	3.37951e+03	T_2^*	d_E	2.00294e+00
T_1^*	b_E	3.30042e+03	T_2	K_d	1.55823e+00
V	K_d	2.82209e+03	T_2^*	δ_E	1.22305e+00
V	λ_E	2.71019e+03	T_2^*	m_2	1.13910e+00
T_1	m_1	2.55536e+03	T_2	K_b	9.81833e−01
E	d_E	2.52360e+03	T_2^*	m_1	6.96854e−01
T_1	K_d	2.31335e+03	T_2^*	λ_E	6.65082e−01
T_1	λ_E	2.27849e+03	E	m_1	6.28564e−01
T_1^*	d_E	2.10127e+03	T_2^*	K_d	5.07039e−01
V	K_b	1.70767e+03	T_2^*	K_b	3.34275e−01
E	δ_E	1.46946e+03	T_2	m_2	2.68566e−01
T_1	K_b	1.41645e+03	E	m_2	3.98107e−02

seems to make a considerable difference whether one uses one-sided (forward) or centered differences to approximate the derivatives. The latter provide much more accurate derivative information, at twice the computational cost.

3.3. Parameter estimation algorithm

We estimate the parameters from ρ observed quantities simultaneously based on data from N time points. The state variable observations are potentially different from each other by orders of magnitude,

so, intuitively, it is critical that the estimation scheme take this into account. One way to do this is by appropriately weighting the states in a least squares cost criterion. Formal statistical large sample theory implies that the optimal (in the sense of yielding the estimator for the true value q_0 making most efficient use of the data when N is large) least squares criterion uses weights equal to the inverse of the true model for variance [14]. Such weighting may be implemented via an iterative generalized least squares (GLS) algorithm based on the mean model

$$E\begin{bmatrix} y_1^j \\ y_2^j \\ \vdots \\ y_\rho^j \end{bmatrix} = \begin{bmatrix} x_1^j(q) \\ x_2^j(q) \\ \vdots \\ x_\rho^j(q) \end{bmatrix} \tag{3.5}$$

and covariance model

$$Var\begin{bmatrix} y_1^j \\ y_2^j \\ \vdots \\ y_\rho^j \end{bmatrix} = \begin{bmatrix} \sigma_1^2\{x_1^j(q)\}^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2\{x_2^j(q)\}^2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_\rho^2\{x_\rho^j(q)\}^2 \end{bmatrix} \tag{3.6}$$

for the data at each time point t^j.

To calculate initial weights for the GLS algorithm, we perform a preliminary weighted least squares estimation. We minimize a least squares objective functional over admissible parameters q to obtain the initial estimate $q^{(1)}$

$$q^{(1)} = \arg\min_q \sum_{i=1}^{\rho} \sum_{j=1}^{N} \frac{\{y_i^j - x_i^j(q)\}^2}{\{x_i^j(q^{(0)})\}^2}, \tag{3.7}$$

where $q^{(0)}$ is the initial iterate for the estimation process.

The estimation proceeds with the following iterative algorithm: Set $k = 1$.

1. Estimate coefficients of variation σ_i^2 for each state:

$$\sigma_i^2 = \frac{1}{N-r} \sum_{j=1}^{N} \frac{\{y_i^j - x_i^j(q^{(k)})\}^2}{\{x_i^j(q^{(k)})\}^2}, \tag{3.8}$$

   where we recall that r is the dimension of the free parameter vector q.
2. Form estimated weights

$$w_i^j = \frac{1}{\sigma_i^2\{x_i^j(q^{(k)})\}^2}. \tag{3.9}$$

3. Minimize the weighted least squares cost function to obtain the next estimate

$$q^{(k+1)} = \arg\min_{q} \sum_{i=1}^{\rho} \sum_{j=1}^{N} w_i^j \{y_i^j - x_i^j(q)\}^2. \tag{3.10}$$

4. Set $k = k + 1$ and repeat steps 1.–3. until convergence. One can use either a fixed number of GLS iterations, or iterate until the relative change in the parameter vector q is small (say 10^{-6}). The results that follow use the former criterion with a preliminary estimate followed by five GLS iterations.

We will denote the resulting optimal parameter vector estimate by q^*.

3.4. Standard error calculation

When estimating parameters using a procedure such as the one given above, it is vital to understand the uncertainty of the estimation *process*, that is, any estimate of mean model parameters from data should be accompanied by an estimate of uncertainty. Here we assess the variance in the estimated model parameters q^* by computing standard errors. The method here is a more general version of that described in [17] and uses sensitivity equations for derivative information, as described above. An example of this methodology in the single state case can be seen in [1].

We stack the model responses $X = (x_1^1, \ldots, x_1^N; x_2^1, \ldots, x_2^N; \ldots; x_\rho^1, \ldots, x_\rho^N)^{\mathrm{T}}$. Let X_q denote the matrix with entries $[X_q]_{ik} = \partial X_i / \partial q_k$ using the solution of the sensitivity equations (3.3). (Thus X_q is a $\rho N \times r$ matrix.) To obtain standard errors, this is computed at the optimal estimate for the parameters: $X_q(q^*)$.

Recall that we assume that the measurement process errors are independently distributed. With our assumed model for the distribution of the data, we expect the GLS estimate $\hat{q} = q^*$ to be approximately normally distributed (at least asymptotically). Specifically, for large samples,

$$\hat{q} = q^* \sim \mathcal{N}(q_0, \Sigma), \tag{3.11}$$

where q_0 is the true vector of parameters and the true covariance matrix is

$$\Sigma = \{X_q(q_0)^{\mathrm{T}} G^{-1} X_q(q_0)\}^{-1}. \tag{3.12}$$

Here G is the weighting matrix

$$G = \mathrm{diag}((\sigma_1^2\{x_1^1\}^2, \ldots, \sigma_1^2\{x_1^N\}^2; \sigma_2^2\{x_2^1\}^2, \ldots, \sigma_2^2\{x_2^N\}^2; \ldots; \sigma_\rho^2\{x_\rho^1\}^2, \ldots, \sigma_\rho^2\{x_\rho^N\}^2))$$

with σ_i^2 computed as in the GLS algorithm above. (See [17, Chapter 2] for an explanation of this idea.)

Having no better approximation to the true values q_0 available, we follow standard statistical practice and substitute the computed estimate q^* for q_0 in the above matrices in (3.12) to obtain standard errors for our estimates. In particular, taking Σ from (3.12) above (taken with the described substitutions), we then take $\sqrt{\Sigma_{kk}}$ to obtain the standard error for parameter component k.

3.5. Inverse problem results

To illustrate the type of information that can be obtained using the methodology discussed above, we present results for several inverse problems, considering first the case with full state observations

Table 4
Parameter estimates, standard errors, and cost function values at optimal parameters when estimating k_1 or k_2 with either full- or partial-state data

Parameter, obs. type	q^*	Std. Err.	Weighted cost	Unweighted cost
k_1, full state	8.0750e−07	1.1455e−09	120	4.121e+11
k_1, partial state	7.9388e−07	1.6828e−07	60	8.209e+10
k_2, full state	9.6730e−05	3.1889e−07	120	3.435e+11
k_2, partial state	9.1280e−05	5.5781e−04	60	7.390e+10

Results obtained with observations every 5 days ($N = 41$).

(T_1, T_2, T_1^*, T_2^*, V, E) as well as a case with partial state observations ($T_1 + T_1^*$, V, E). The latter case is representative of the type of clinical data discussed in Section 4. The data used simulated measurements every five days from 0 to 200 days, inclusive. We estimate three subsets of the model parameters: k_1 only, k_2 only, and the pair k_1, k_2 jointly. Recall that the simulated data are generated with true parameters $k_1 = 8.0 \times 10^{-7}$ and $k_2 = 1.0 \times 10^{-4}$. These infectivity parameters are important to estimate since they describe the viral infection rate for the two cell populations and the overall system dynamics are highly sensitive to them.

When estimating k_1 only, each of the initial iterates $q^{(0)} = 10^{-2}$, 10^{-4}, 10^{-6}, 10^{-8}, and 10^{-10} produces the same optimal parameter estimate. The same is true when estimating only k_2 for each of the initial iterates $q^{(0)} = 10^{-1}$, 10^{-2}, 10^{-4}, 10^{-6}, and 10^{-8}. This robustness of the estimation process to the choice of initial iterate would be particularly important when using clinical data. The resulting parameter estimates, standard errors, and cost function values are displayed in Table 4. In all four cases, we successfully recover reasonable estimates of the true parameters. Of course, we do not expect to precisely recover the true parameter values since we are fitting noisy simulated data (see Section 3.1 above). The standard errors for the full state data are acceptably small, although those for the partial state data are appreciably larger. These calculations are based on observing 6 states at each of 41 time points in the full state case (total 246 data points) and 3 states at the same time points in the partial state case (total 123 data points).

To estimate k_1 and k_2 jointly, we consider five initial iterates taken as ordered pairs from the iterates used for single parameter estimation above. When using full state observations, the algorithm returns the same parameter pair for each initial iterate pair: $q^* = (k_1, k_2) = (8.1065 \times 10^{-7}, 9.3961 \times 10^{-5})$ with corresponding standard errors $(5.0385 \times 10^{-9}, 1.5997 \times 10^{-6})$ and cost function values of 117 (weighted) and 3.7172e+11 (unweighted). The fit to data can be viewed in Fig. 3, where the solid lines are the simulations corresponding to the optimized parameters q^*.

For partial state observations we obtain similar results, although the initial iterate $(10^{-8}, 10^{-6})$ does not lead to convergence within the allotted five GLS iterations. (This is the only such exception among the results presented here.) Overall though, for the joint estimation case, the algorithm is relatively robust to the choice of initial iterates (see Table 5). Fig. 4 depicts a typical fit to the partial state data.

Even with only observations of $T_1 + T_1^*$, V, and E, the estimation process yields acceptable estimates of k_2, although the large standard errors do not lend a lot of confidence to their accuracy.

In a forthcoming paper, we will explore the ability to fit this model with greater numbers of free parameters and discuss the merits of different weighting schemes in the estimation process. As an alternative to the methodology presented here, one could use sampling-based methods, such as Monte Carlo Markov

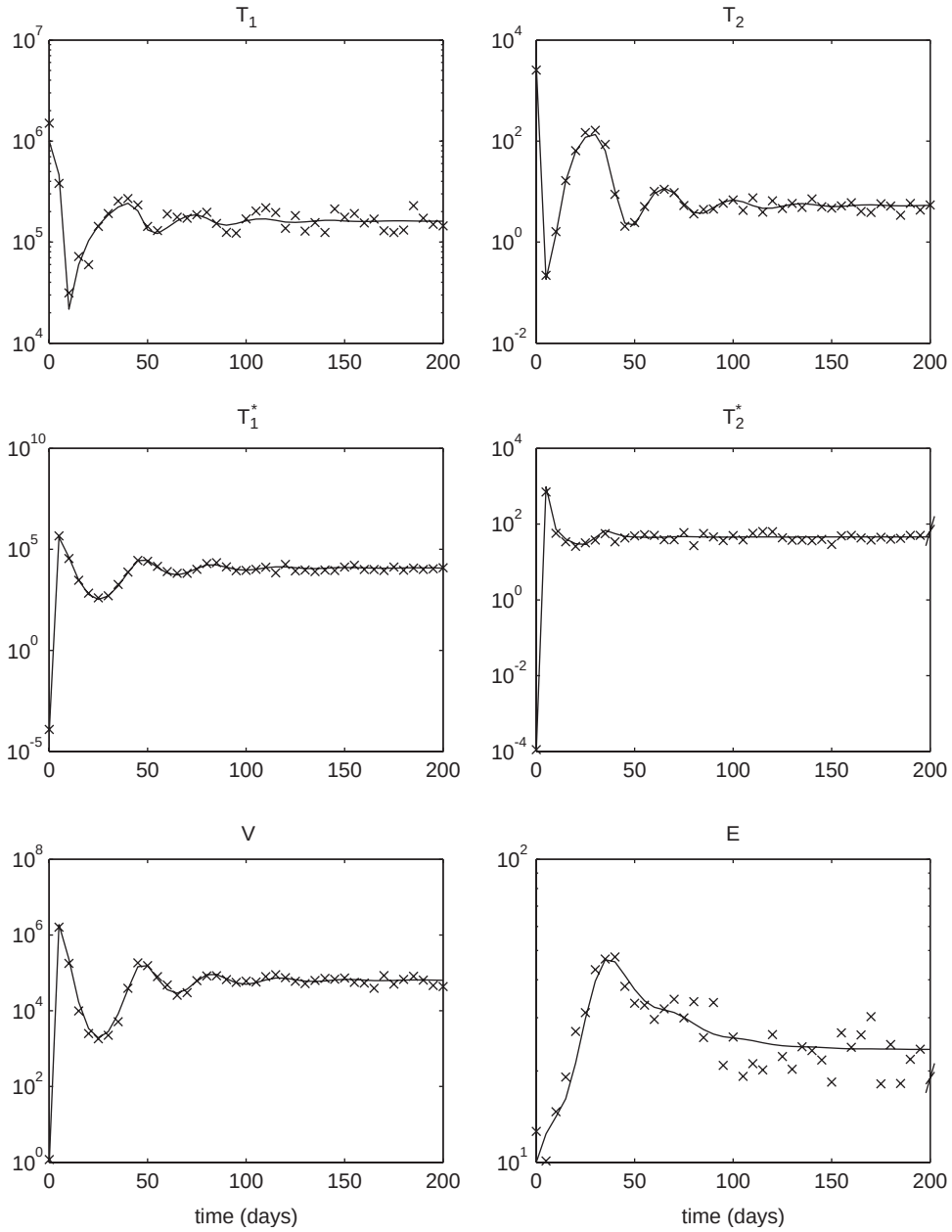

Fig. 3. Fit of model using estimated parameters (solid lines) to data ('x' markers) when jointly estimating k_1, k_2 using third initial iterate $(10^{-6}, 10^{-4})$. Estimates are based on full state observations. Note different scale on each subplot.

chain algorithms, to obtain parameter information from data. These methods yield an estimate for the distribution of a parameter (and hence estimates of the mean and variance similar to those we have reported here). Our group is currently comparing the efficiency of such methods to those discussed in this paper.

Table 5
Resulting parameter estimates, standard errors, and cost function values for partial state data, simultaneous estimation of k_1 and k_2

$q^{(0)}$	q^*	Std. Err	$J(q^*)$	
$[k_1, k_2]$	$[k_1, k_2]$	$[k_1, k_2]$	Weighted	Unweighted
$(10^{-2}, 10^{-1})$	(7.9420e−07, 9.9007e−05)	(3.0605e−07, 1.0957e−03)	58.5	7.9624e+10
$(10^{-4}, 10^{-2})$	(7.9420e−07, 9.9007e−05)	(3.0605e−07, 1.0957e−03)	58.5	7.9624e+10
$(10^{-6}, 10^{-4})$	(7.9420e−07, 9.9007e−05)	(3.0605e−07, 1.0957e−03)	58.5	7.9624e+10
$(10^{-8}, 10^{-6})$	(7.9373e−07, 1.0050e−04)	(3.0325e−07, 1.1074e−03)	58.0	8.3495e+10
$(10^{-10}, 10^{-8})$	(7.9420e−07, 9.8996e−05)	(3.0607e−07, 1.0956e−03)	58.5	7.9598e+10

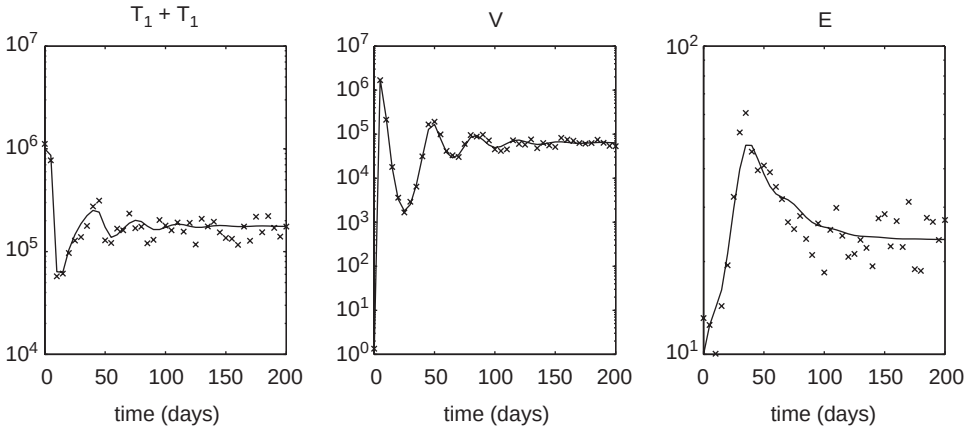

Fig. 4. Fit of model using estimated parameters (solid lines) to data ('x' markers) when jointly estimating k_1, k_2 using third initial iterate $(10^{-6}, 10^{-4})$. Estimates are based on partial state observations. Note different scale on each subplot.

4. Modeling HIV dynamics: a POD approach

There are many families of patterns occurring in both nature and in the sciences for which it is possible to obtain a useful systematic characterization. Often, the motivation is that the family is of low dimension, that is, in some sense, any given member of the family might be represented by a small number of parameters. Examples of these pattern families occur in turbulent flows [41,3,61,7], image processing [29,64], data compression [2], human speech [65], and human faces [35]. In these applications the technique used to represent such characteristics is known as the proper orthogonal decomposition (POD). In other disciplines it also called the Karhunen–Loève decomposition [39], principal component analysis [31], and the Hotelling transform [25].

Typically, the POD procedure is implemented in engineering or science applications where the physical or biological system may be described by a mathematical model, i.e., a nonlinear distributed parameter system with boundary and initial conditions. The pedagogy to obtain the governing equations begins with some simplifying assumptions about the physical or biological system and then applies first principles

to derive quantitative models. The POD procedure, applied together with the Galerkin, finite element, finite volume, or spectral method, results in a reduced order model which represents the underlying low dimensional system.

On the other hand, many problems in engineering and medicine, including HIV population level dynamics, are sufficiently complex that a mathematical description is either nearly impossible, or when a model is obtained, it cannot be implemented without undue complications. For example, the degree of freedom of distributed parameter systems is essentially infinite, and in the case of HIV population level models, the parameters are often formulated as random variables in order to accurately account for various sources of variability contained in aggregate clinical data. It is possible to apply the POD method to an ensemble of data (with or without an underlying model) to extract a basis that characterizes the salient features of the ensemble. In this section, we assume that the mathematical model for HIV dynamics is not yet known, but data to be used to guide and validate model development is available. Our primary goal is to show that characteristic features, extracted from collected data using the POD procedure, can be used to organize HIV data (RNA viral load values) more efficiently and accurately.

4.1. HIV STI clinical data

The data that we employ for our demonstration were collected by one of the authors (Rosenberg) from patients studied at Massachusetts General Hospital. The ensemble provides RNA viral load, CD4, and CD8 data for 102 patients. Approximately 40 patients have at least one treatment interruption, and some have as many as four treatment interruptions. The interruptions are either approved according to protocol, are due to illness or drug toxicity, or are patient initiated.

The data set contains a high degree of variability. For an individual patient viral load values may range from a censored value of fifty (c/ml) to several million (c/ml); the duration of the interruptions range from one week to several years. We have inter-patient variability (long-term nonprogressors versus rapid progressors) as well as intra-patient variability (on treatment, off treatment, poor adherence, short drug holidays, or structured treatment interruptions).

4.2. The proper orthogonal decomposition

In the following we provide a basic description of the POD method; a more detailed description is given in [43]. Let $U_i = U_i(t)$, $i = 1, 2, \ldots, N$, denote the set of N observations (also called snapshots) for $t \in \Omega$, Ω bounded. For example, U_i could represent the RNA values observed for patient i at times $t \in \Omega$. We wish to find an optimal compressed description of the sequence of data. One description of the process is a series expansion in terms of a set of basis functions. Intuitively, the basis functions should in some sense be representative of the members of the ensemble. Such a coordinate system is provided by the Karhunen–Loève expansion, where the basis functions Φ are, in fact, linear combinations of the snapshots and are given by

$$\Phi = \sum_{i=1}^{N} a_i U_i . \tag{4.1}$$

Here, the coefficients a_i are to be determined so that Φ given by (4.1) will most resemble the ensemble $\{U_i\}_{i=1}^{N}$. Indeed, it follows that (see, e.g., [43]) the coefficients a_i are the entries of the ith eigenvector

corresponding to the ith largest eigenvalue of the $N \times N$ covariance matrix

$$C_{\vec{u}} = \frac{1}{N} \int_{\Omega} U(t)U(t)^{\mathrm{T}} \, \mathrm{d}t, \tag{4.2}$$

where $U = (U_1, U_2, \ldots, U_N)^{\mathrm{T}}$. It is noted that the covariance matrix is a symmetric and nonnegative matrix and, hence, the eigenvalues, λ_i are real and nonnegative. We arrange the eigenvalues in decreasing order as $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \lambda_N \geqslant 0$. Thus Φ_1 is the basis function corresponding to the largest eigenvalue. Any snapshot in the ensemble can be represented as a linear combination of the basis functions as follows

$$U_i = \sum_{i=1}^{N} \langle \Phi_k, \, U_i \rangle_{L^2(\Omega)} \Phi_k. \tag{4.3}$$

Employing only the first M POD basis elements, we obtain the approximation U_i^M for U_i such that

$$U_i^M = \sum_{i=1}^{M} \langle \Phi_k, \, U_i \rangle_{L^2(\Omega)} \Phi_k. \tag{4.4}$$

It is well known that (see, e.g., [43]) expansion (4.4) is in some sense optimal. In particular, among all linear combinations, the POD is the most efficient, in the sense that, for a given number of modes M, the POD decomposition will capture the most possible kinetic energy. One way to quantify the error of the approximation is by computing the relative error

$$\mathrm{ER}(U_i^M) = \frac{\|U_i - U_i^M\|_{L^2(\Omega)}}{\|U_i\|_{L^2(\Omega)}}. \tag{4.5}$$

4.3. RNA modeling results and discussions

In this section, we demonstrate the effectiveness of using the POD expansion to model RNA values on an ensemble clinical data set collected at Massachusetts General Hospital. One could use the entire data set for the ensemble, using each patient as a snapshot. Indeed, initially we followed this approach with very limited success. This was due to the significant inter-patient and intra-patient variability observed in the data set. This contributed to the difficulty in representing the data by a reduced POD basis. Here, we first focus on a subset of the data ensemble—HIV patients that undergo treatment interruptions. We focus only on data points that correspond to ON treatment entries. Modeling of the OFF treatment entries will be addressed later.

Before applying the POD method to the restricted set of ON segments (snapshots), we first perform some preliminary data manipulations to obtain a more realistic "guess" of the OFF/ON treatment cycle. For example, in looking at the RNA data we often encountered segments similar to the following:

date	on/off	RNA value	comment
1/22/02	off	750,000	no medication
1/24/02	on		start medication
1/30/02	on	1000	

In this example, the 1/24/02 entry has no RNA value, and is ignored (deleted). As a result, the ON segment begins on 1/30/02 with a relatively low RNA value of 1000. For this example we would modify the data by copying the previous OFF treatment value as follows:

date	on/off	RNA value	comment
1/22/02	off	750,000	no medication
1/24/02	on	750,000	start medication
1/30/02	on	1000	

We feel this is a more realistic guess of the OFF/ON cycle. We did this for patient segments where a relatively recent OFF value was available, normally within one week of starting medications. If no recent OFF value was available, we left the data as is.

With these data modifications, we obtain 78 snapshots of RNA data for patients starting and continuing on treatment. The length of these segments vary from 13 days to 1421 days. The general ON profile starts with high viral load, it rapidly decreases due to medication, and then typically stays small (i.e., censored data).

To apply the POD method to the set of HIV ON snapshots, the snapshots should be defined on a similar time interval, i.e., $\Omega = [0, T]$. We start all snapshots at day zero, and then interpolate each snapshot onto a common (merged) set of time points. The ON snapshots are of variable length, i.e., $U_i(t)$ is defined on $[0, T_i]$ and not on $[0, T]$, except for $T = \min\{T_i\}$ small. We deal with the variable data length in one of two ways: we truncate a subset of patient data defined on $[0, T_{ON}]$ for a set of fixed length snapshots, or we compute the 'covariance' matrix $C_{\vec{u}}$ with variable length time series. For the later, $C_{\vec{u}}$ is computed as

$$C_{\vec{u}ki} = \frac{1}{N} \int_0^{T_{ki}} U_k(s) U_i(s)\, \mathrm{d}s, \quad T_{ki} = \min(T_k, T_i) \tag{4.6}$$

for $U_k(t)$ and $U_i(t)$ defined on $[0, T_k]$ and $[0, T_i]$, respectively.

To apply the POD method to fixed length vectors, we extract a subset of segments with length greater than T_{ON} and then truncate. For example, if $T_{\mathrm{ON}} = 482$ days, the subset includes 18 snapshots. A reduced order basis may be used to approximate the subset of truncated ON snapshots. The full POD basis may also be used to approximate the additional 60 ON segments, outside the subset of snapshots used to generate the POD basis, using approximation (4.4) computed on $[0, T_i]$ for $T_i < T_{\mathrm{ON}}$. In general, the relative error for the shorter snapshots is not great.

The POD method applied to the 78 variable length ON segments produces 78 POD basis elements defined on $T_{\max} = \max T_i$. We can reconstruct all 78 ON segments using approximation (4.4) with $M = 18$ POD basis elements to obtain a 5% average relative error. Nine snapshots have a relative error greater than 10%, i.e., $\mathrm{ER}(U_i^M) > 0.1$. In Fig. 5 we compare the relative errors of the fixed length POD approximations to the variable length POD approximations. In either case, the approximation is computed with $M = 18$ basis elements. The POD method based on the variable length snapshots obtains better overall approximations. The ability to successfully approximate most ON segments is likely due to the apparently homogeneous nature of patient profiles while on treatment. With good adherence to drugs, the RNA profile of most patients on treatment is relatively flat, i.e., the RNA values are $(<)50$ or $(<)400$, depending on the assay.

Given the variability of the data and the sometimes incomplete data records, we relax the notion of approximating all ON snapshots well. If we delete the nine snapshots with the worst POD approximation relative error from the ON treatment ensemble, 69 ON snapshots remain. Applying the POD

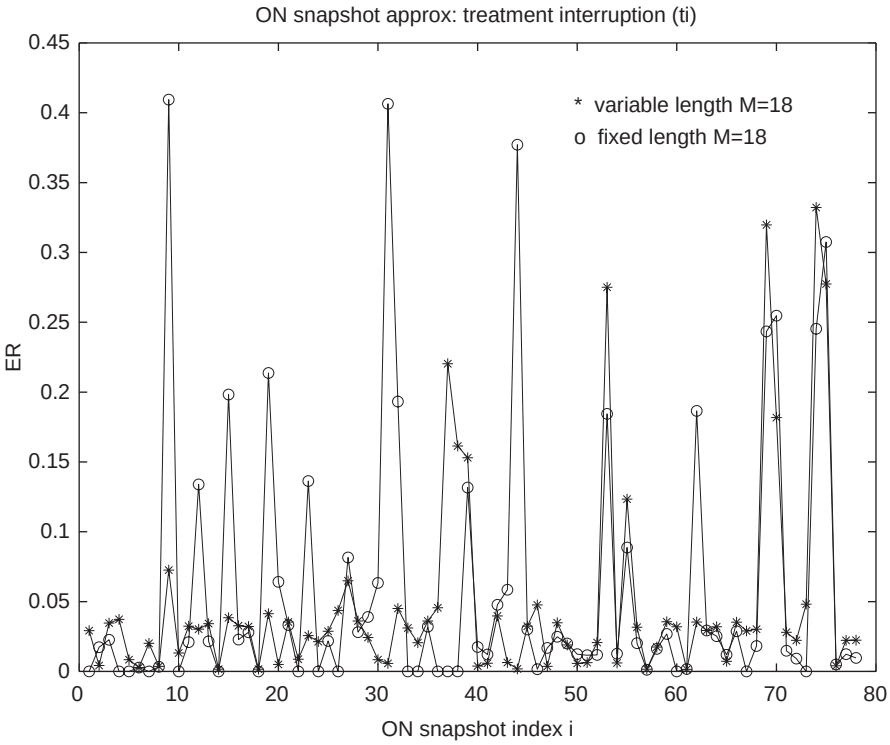

Fig. 5. Relative errors comparing fixed length (circle) POD approximations versus variable length (dot) POD approximation.

method to these remaining variable length ON segments, we again use $M = 18$ basis elements to reconstruct the snapshots. All remaining snapshots have a less than 10% approximation relative error, with $\max(\mathrm{ER}(U_i^M)) = 0.06$ and mean $(\mathrm{ER}(U_i^M)) = 0.02$ (see Fig. 6). Fewer basis vectors may be used, but the relative error increases rapidly for one segment in particular. See also Figs. 7 and 8 for the POD approximation of the 3 ON segments with the best and worst approximation relative error, respectively.

Next we turn to the original ensemble including all 102 patients, patients with treatment interruptions and with no treatment interruptions. We repeat the same process: manipulating the data to more accurately represent the ON/OFF treatment cycle using a slightly more rigorous protocol, and then extracting the ON segments. We obtain 138 ON segments with lengths varying from 4 days to 1833 days. We apply the POD method to the variable length ON snapshots, use $M = 33$ POD basis elements to approximate the snapshots, and identify and remove the 9 snapshots with more than 10% approximation relative error. If we rerun the POD method, again using $M = 33$ POD basis elements for approximation, we obtain good overall relative error (i.e., $\max(\mathrm{ER}(U_i^M)) = 0.084$ and mean $(\mathrm{ER}(U_i^M)) = 0.014$) (see Fig. 9).

Similarly, we have attempted to approximate the OFF treatment snapshots from patients with treatment interruptions. To date, the results are not as promising; the OFF segments achieve a 20% average approximation error. This is most likely due to the inhomogeneity of the OFF snapshots. Some patients maintain a low viral load after discontinuing treatment, some experience viral rebound and decline during treatment interruptions, while others exhibit a slow, steady increase in viral load. All of these scenarios

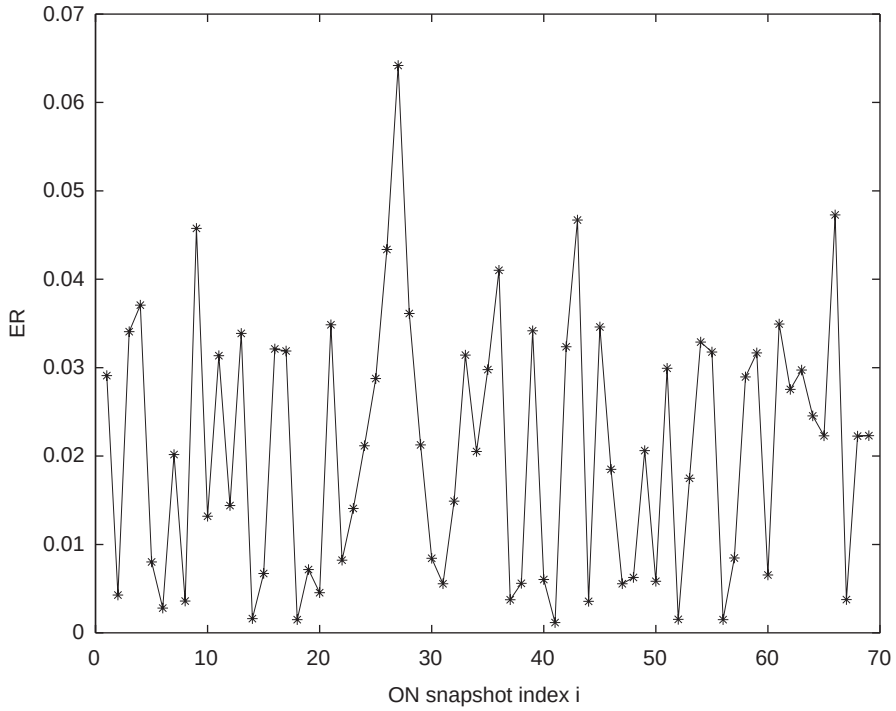

Fig. 6. Relative errors from reconstructing the 69 ON segments using only 18 POD basis elements.

are included in the OFF treatment ensemble. Future effort will include the application of the POD method to subgroups within the OFF treatment ensemble.

Procedures such as those described here should prove quite useful in selecting subsets of clinical data for use with the model fitting and validation methods presented in Section 3.

5. An optimal control problem

In this section, we consider optimal control methods to derive optimal drug treatments as functions of time. We attempt to control HIV populations in finite time intervals using a control function $\varepsilon(t)$ which represents the drug efficacy satisfying $0 \leqslant a \leqslant \varepsilon(t) \leqslant b < 1$. Here $\varepsilon(t) = b$ represents maximal efficacy, where we tacitly assume that we control efficacy by simply controlling the dose level.

Together with the model described by Eq. (2.1) for HIV infection we consider a cost functional given by

$$J(\varepsilon) = \int_{t_0}^{t_1} [QV(t) + R\varepsilon^2(t)]\,\mathrm{d}t, \tag{5.1}$$

where Q and R are weight constants of virus and control, respectively. The second term represents systemic costs of the drug treatment. Our goal is to minimize both the HIV virus population and the systemic cost

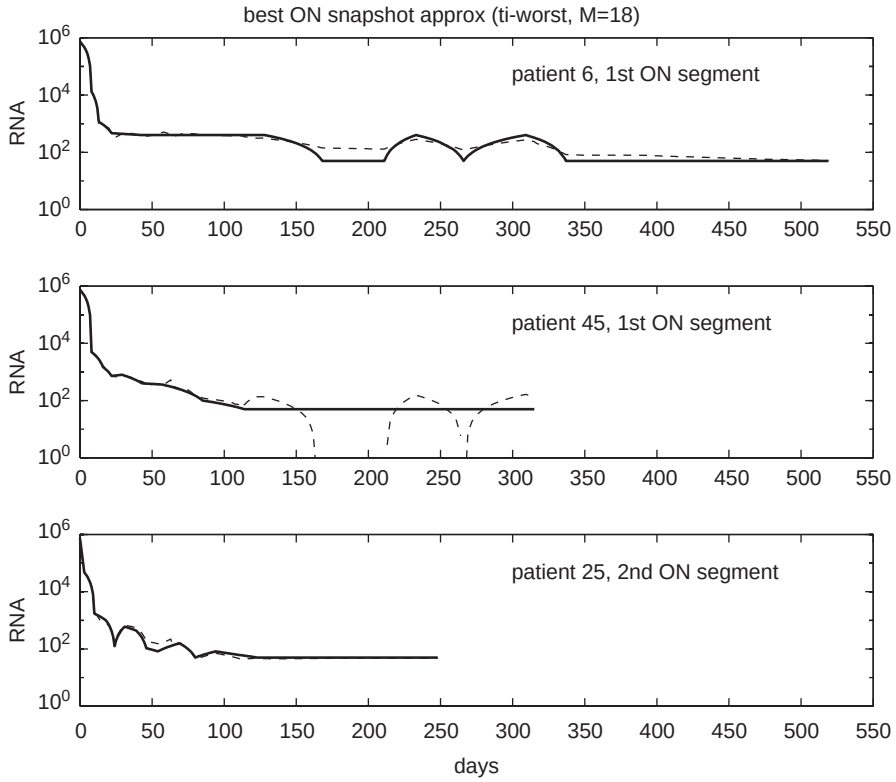

Fig. 7. Viral RNA (in c/ml)—POD approximation of 3 ON segments with "best" relative error.

to body. We seek an optimal control ε^* such that

$$J(\varepsilon^*) = \min\{J(\varepsilon)|\varepsilon \in U\},$$

where $U = \{\varepsilon(t)|\, \varepsilon \text{ is measurable}, a \leqslant \varepsilon \leqslant b,\ t \in [t_0, t_1]\}$ is the control set. Given the criterion (5.1) and the regularity of the system in our model equations (2.1), it is rather straightforward to establish existence of optimal controls by relying on standard results in control theory (e.g., see [22]). We turn therefore to methods for computation of optimal controls.

5.1. Optimality system

Since an optimal control exists for minimizing the cost functional (5.1) subject to (2.1), we present necessary conditions for optimality (see, e.g., [34,40] for details on these procedures). These can be used to compute candidates for optimal controls (which may or may not be uniquely determined).

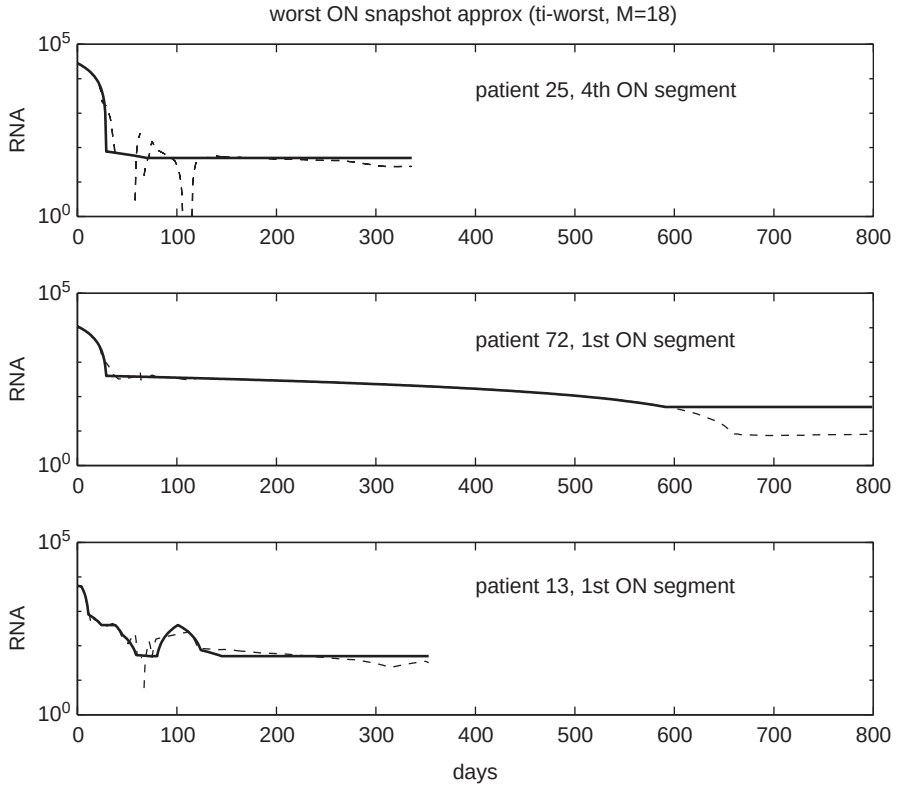

Fig. 8. POD approximation of 3 ON segments with "worst" relative error.

Given an optimal control ε^* and solutions of the corresponding state system (2.1), we can define adjoint variables ξ_i, $i = 1, \ldots, 6$, as follows:

$$
\begin{aligned}
\dot{\xi}_1 &= -\{\xi_1(-d_1 - (1-\varepsilon)k_1V) + \xi_3(1-\varepsilon)k_1V - \xi_5(1-\varepsilon)\rho_1 k_1 V\},\\
\dot{\xi}_2 &= -\{\xi_2(-d_2 - (1-f\varepsilon)k_2V) + \xi_4(1-f\varepsilon)k_2V - \xi_5(1-f\varepsilon)\rho_2 k_2 V\},\\
\dot{\xi}_3 &= -\left\{\xi_3(-\delta - m_1E) + \xi_5 N_T\delta + \xi_6\left(\frac{b_E E K_b}{(T_1^*+T_2^*+K_b)^2} - \frac{d_E E K_d}{(T_1^*+T_2^*+K_d)^2}\right)\right\},\\
\dot{\xi}_4 &= -\left\{\xi_4(-\delta - m_2E) + \xi_5 N_T\delta + \xi_6\left(\frac{b_E E K_b}{(T_1^*+T_2^*+K_b)^2} - \frac{d_E E K_d}{(T_1^*+T_2^*+K_d)^2}\right)\right\},\\
\dot{\xi}_5 &= -\{Q - \xi_1(1-\varepsilon)k_1T_1 - \xi_2(1-f\varepsilon)k_2T_2 + \xi_3(1-\varepsilon)k_1T_1 + \xi_4(1-f\varepsilon)k_2T_2,\\
&\quad + \xi_5(-c - (1-\varepsilon)\rho_1k_1T_1 - (1-f\varepsilon)\rho_2k_2T_2)\},\\
\dot{\xi}_6 &= -\left\{-\xi_3 m_1 T_1^* - \xi_4 m_2 T_2^* + \xi_6\left(\frac{b_E(T_1^*+T_2^*)}{T_1^*+T_2^*+K_b} - \frac{d_E(T_1^*+T_2^*)}{T_1^*+T_2^*+K_d} - \delta_E\right)\right\},\\
\xi_i(t_1) &= 0 \quad \text{for } i = 1, \ldots, 6.
\end{aligned}
\tag{5.2}
$$

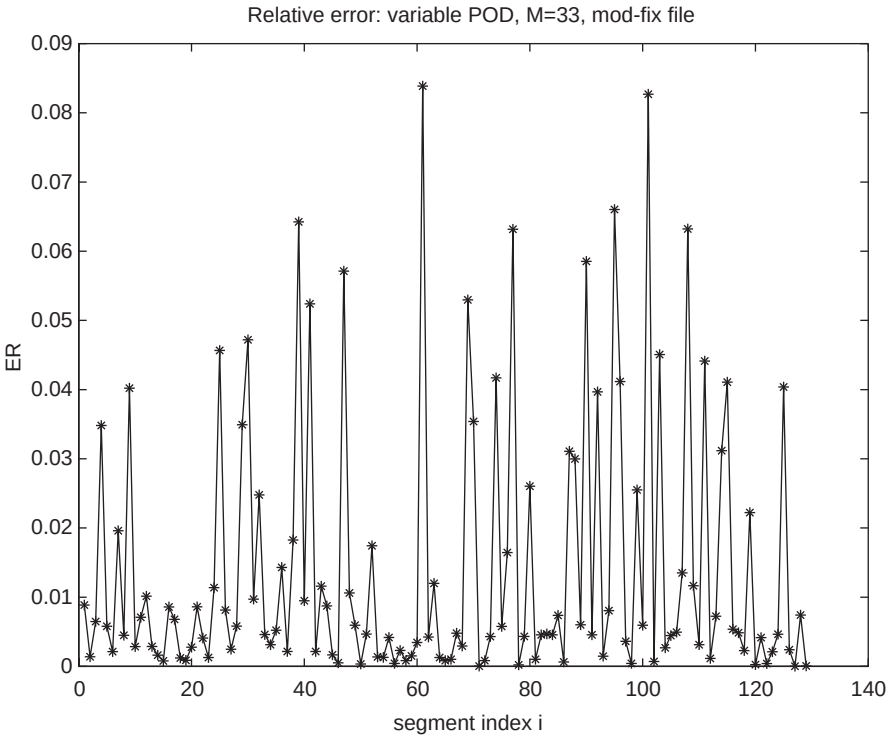

Fig. 9. Relative errors from reconstructing the ON segments from patients with and without treatment interruptions.

Using these we find that an optimal control ε^* is given by

$$\varepsilon^* = \max\left(a, \min\left(b, \frac{-(\xi_1 - \xi_3 + \rho_1\xi_5)k_1VT_1 - (\xi_2 - \xi_4 + \rho_2\xi_5)fk_2VT_2}{2R}\right)\right). \tag{5.3}$$

To see this, define the Lagrangian (which is the Hamiltonian augmented with penalty terms for the constraints) to be

$$\begin{aligned}
&L(T_1, T_2, T_1^*, T_2^*, V, E, \varepsilon, \xi_1, \xi_2, \xi_3, \xi_4, \xi_5, \xi_6)\\
&\quad = QV(t) + R\varepsilon^2 + \xi_1(\lambda_1 - d_1T_1 - (1-\varepsilon(t))k_1VT_1)\\
&\qquad + \xi_2(\lambda_2 - d_2T_2 - (1 - f\varepsilon(t))k_2VT_2) + \xi_3((1-\varepsilon(t))k_1VT_1 - \delta T_1^* - m_1ET_1^*)\\
&\qquad + \xi_4((1 - f\varepsilon(t))k_2VT_2 - \delta T_2^* - m_2ET_2^*)\\
&\qquad + \xi_5(N_T\delta(T_1^* + T_2^*) - cV - [(1-\varepsilon(t))\rho_1k_1T_1 + (1 - f\varepsilon(t))\rho_2k_2T_2]V)\\
&\qquad + \xi_6\left(\lambda_E + \frac{b_E(T_1^* + T_2^*)}{(T_1^* + T_2^*) + K_b}E - \frac{d_E(T_1^* + T_2^*)}{(T_1^* + T_2^*) + K_d}E - \delta_E E\right)\\
&\qquad - w_1(t)(\varepsilon(t) - a) - w_2(t)(b - \varepsilon(t)),
\end{aligned} \tag{5.4}$$

where $w_1(t) \geqslant 0$, $w_2(t) \geqslant 0$ are the penalty multipliers satisfying

$$w_1(t)(\varepsilon(t) - a) = 0 \quad \text{and} \quad w_2(t)(b - \varepsilon(t)) = 0 \quad \text{at optimal controls } \varepsilon^*.$$

We differentiate the Lagrangian with respect to states, T_1, T_2, T_1^*, T_2^*, V, and E, respectively, to obtain the following equations for the adjoint variables

$$\dot{\xi}_1 = -\frac{\partial L}{\partial T_1}, \quad \dot{\xi}_2 = -\frac{\partial L}{\partial T_2}, \quad \dot{\xi}_3 = -\frac{\partial L}{\partial T_1^*}, \quad \dot{\xi}_4 = -\frac{\partial L}{\partial T_2^*},$$
$$\dot{\xi}_5 = -\frac{\partial L}{\partial V} \quad \text{and} \quad \dot{\xi}_6 = -\frac{\partial L}{\partial E}.$$

By differentiating the Lagrangian L with respect to ε, we also have

$$\frac{\partial L}{\partial \varepsilon} = 2R\varepsilon + (\xi_1 - \xi_3 + \rho_1\xi_5)k_1VT_1 + (\xi_2 - \xi_4 + \rho_2\xi_5)fk_2VT_2 - w_1(t) + w_2(t) = 0.$$

Solving for the optimal control we find

$$\varepsilon^* = \frac{-(\xi_1 - \xi_3 + \rho_1\xi_5)k_1VT_1 - (\xi_2 - \xi_4 + \rho_2\xi_5)fk_2VT_2 + w_1(t) - w_2(t)}{2R}.$$

To determine an explicit expression for an optimal control without w_1 and w_2, we consider the following three cases:

(i) On the set $\{t | a < \varepsilon^*(t) < b\}$, we have $w_1(t) = w_2(t) = 0$. Hence the optimal control is

$$\varepsilon^* = \frac{-(\xi_1 - \xi_3 + \rho_1\xi_5)k_1VT_1 - (\xi_2 - \xi_4 + \rho_2\xi_5)fk_2VT_2}{2R}.$$

(ii) On the set $\{t | \varepsilon^*(t) = b\}$, we have $w_1(t) = 0$. Hence

$$b = \varepsilon^* = \frac{-(\xi_1 - \xi_3 + \rho_1\xi_5)k_1VT_1 - (\xi_2 - \xi_4 + \rho_2\xi_5)fk_2VT_2 - w_2(t)}{2R},$$

which implies that

$$\frac{-(\xi_1 - \xi_3 + \rho_1\xi_5)k_1VT_1 - (\xi_2 - \xi_4 + \rho_2\xi_5)fk_2VT_2}{2R} \geqslant b \quad \text{since } w_2(t) \geqslant 0.$$

(iii) On the set $\{t | \varepsilon^*(t) = a\}$, we have $w_2(t) = 0$. Hence

$$a = \varepsilon^* = \frac{-(\xi_1 - \xi_3 + \rho_1\xi_5)k_1VT_1 - (\xi_2 - \xi_4 + \rho_2\xi_5)fk_2VT_2 + w_1(t)}{2R}.$$

Since $w_1(t) \geqslant 0$, it follows that

$$\frac{-(\xi_1 - \xi_3 + \rho_1\xi_5)k_1VT_1 - (\xi_2 - \xi_4 + \rho_2\xi_5)fk_2VT_2}{2R} \leqslant a.$$

Combining these three cases, we obtain the optimal control given by (5.3).

The optimality system thus consists of the state system (2.1) coupled with the adjoint system (5.2) with the initial conditions and terminal conditions together with relationship (5.3). Any optimal controls must satisfy this system.

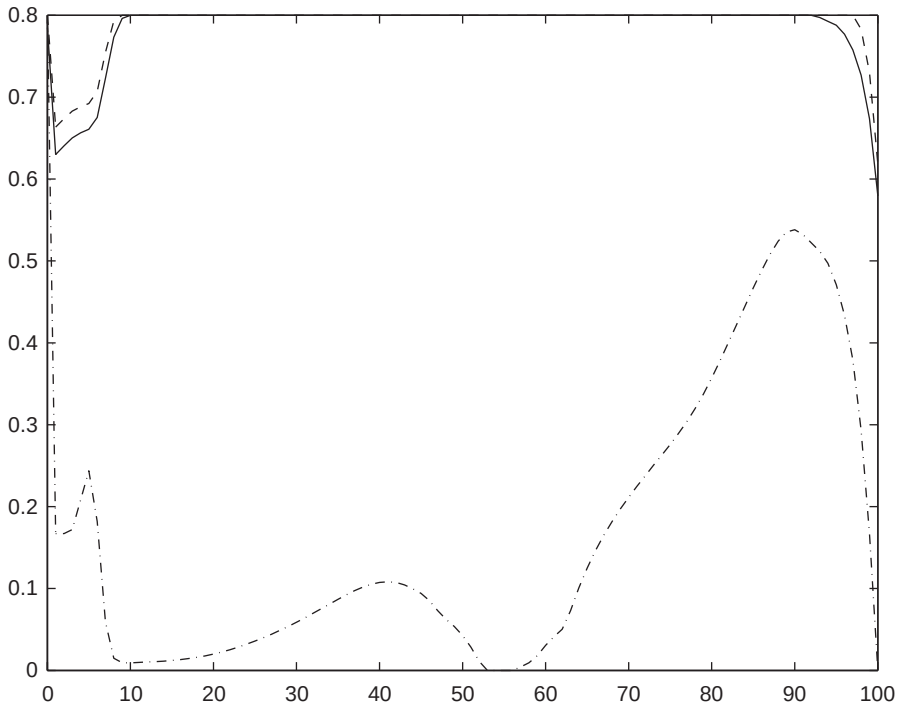

Fig. 10. Optimal controls with different weights R on control cost: $Q = 0.1$. $--$: $R = 5000$ $-$: $R = 10000$ $-\cdot$: $R = 15000$.

5.2. Continuous optimal treatment

The optimality system is a two-point boundary value problem since it includes initial conditions specified for the state equations and terminal conditions specified for the adjoint or costate system. We use a *gradient method* for solving the optimality system. The state system (2.1) with initial conditions is solved (once again we use standard Matlab routines) forward in time using an initial guess for the control and then the adjoint system (5.2) with terminal conditions is solved backward in time. The controls are updated in each iteration using the formula (5.3) for optimal controls. The iterations continue until convergence is achieved. For further discussion of this iterative method we refer the interested reader to [27]. The parameters used in solving the optimality system are summarized in Table 2.1. Treatment was simulated for 100 days.

We simulate early infection by introducing one virus particle per ml of blood plasma, i.e., $T_1(0) = 10^6$, $T_2(0) = 3198$, $T_1^*(0) = 10^{-4}$, $T_2^*(0) = 10^{-4}$, $V(0) = 1$ and $E(0) = 10$. In addition, we use $a = 0$ and $b = 0.8$. We ran simulations with three different values of weight factor R. The corresponding optimal control functions are presented in Fig. 10. As we increase R, thereby increasing the cost of therapy, the optimal control function is decreasing. In the other figures presented in this section, we present only the case $Q = 0.1$ and $R = 10000$ for brevity. Since the magnitude of virus population is much larger than the magnitude of the cost of the drug treatment in the objective function (5.1), this difference in magnitude is balanced by this choice of Q and R.

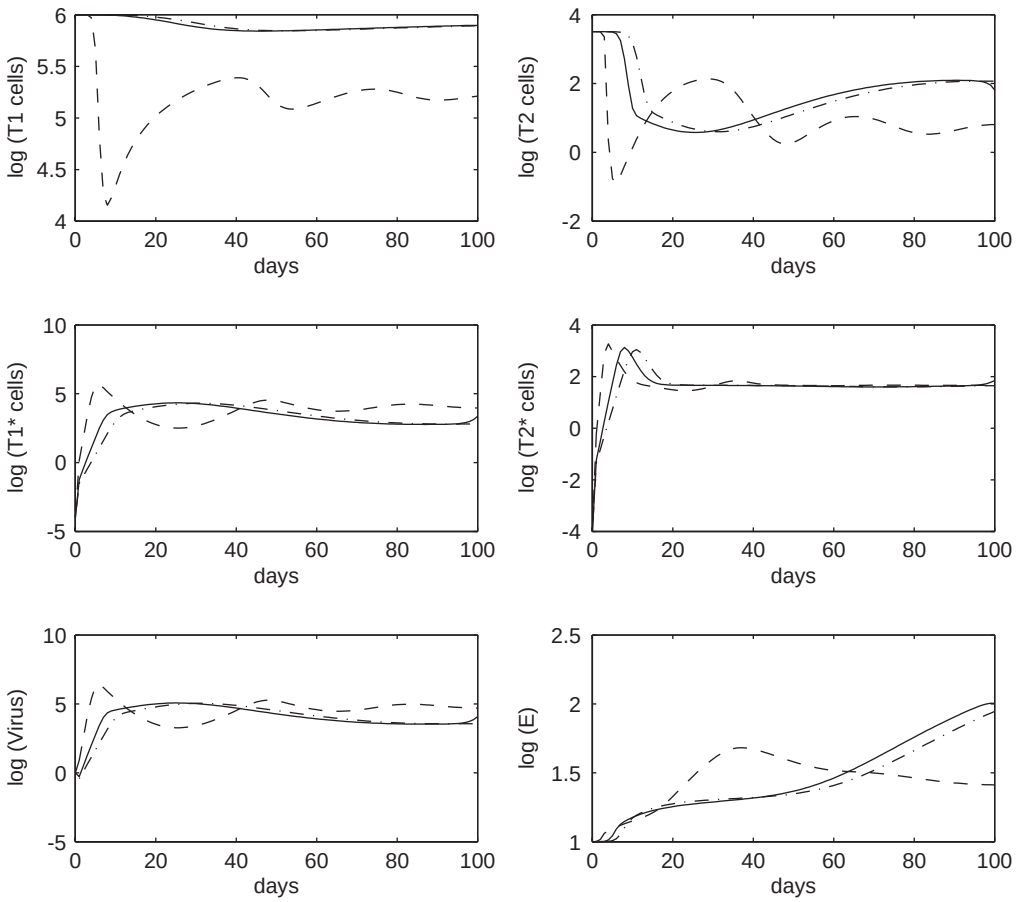

Fig. 11. Optimal solutions (−); solutions (−·) with full treatment (i.e., $\varepsilon = 0.8$); and solutions (−−) with no treatment (i.e., $\varepsilon = 0$) of early infection: $Q = 0.1$ and $R = 10000$.

Table 6
The values of objective functional

	Optimal control	$\varepsilon(t) = 0$	$\varepsilon(t) = 0.8$	STI control
J	9.4438e+05	1.2488e+06	9.6007e+05	1.1495e+06

We solve the HIV model (2.1) with no treatment ($\varepsilon \equiv 0$) and with fully efficacious treatment ($\varepsilon \equiv 0.8$). The numerical results for these two cases as well as the optimal case are presented in Fig. 11 for comparison. The corresponding values of the objective functional for all cases are given in Table 6. As depicted in Fig. 11 the optimal solutions are very close to the solutions corresponding to the fully efficacious treatment. However, the optimal cost is smaller than the cost functional corresponding to the fully efficacious treatment case.

6. An optimal STI control problem

In Section 5 we formulated and synthesized continuous optimal control solutions for a model-based infection treatment. However, this type of (continuous) treatment, especially, for long periods is difficult to maintain due to possible long-term toxicity of the drugs as well as possible development of resistance to medications. In this section we consider the optimal control of viremia through a number of drug structured treatment interruptions (STI).

To this end, we assume our control ε is a 1×100 vector which consists of only 0 or b in each vector element. If an element of a control vector is 0, it signifies drug treatment being off on that day and if an element of a control vector is b, it represents full drug treatment being on. Since we consider a drug treatment strategy over 100 days, the size of ε is 1×100. The set of all such control vectors is denoted by Λ. The goal is to seek the optimal control vector ε^* satisfying

$$J(\varepsilon^*) = \min_{\varepsilon \in \Lambda} J(\varepsilon),$$

subject to the state system (2.1) and where $J(\varepsilon)$ is defined by (5.1).

Since the number of elements of the set Λ is finite, existence of an optimal control solution is guaranteed. To find the discrete STI optimal solution, we could use a direct search approach and begin by selecting any element in the set Λ and then solving the state system using this element as a control. We next would select another element in the set Λ and solve again the state system using the element as a control. Then we compare the values of the objective functional, J, and select the element corresponding to the smaller cost functional value. If we iterate this strategy over all elements in the set Λ, we obtain the optimal control solution ε^*. However, this strategy to obtain the discrete optimal solution can lead to a large number of cost functional evaluations (and hence a large number of solutions to Eq. (2.1)). In our case, where the control vector ε is a 1×100 vector, the number of cost functional evaluations would be 2^{100}, rendering this approach computationally impractical.

There are several ideas that we considered to reduce the number of iterations. The simplest approach is to consider, for illustration, a 5 day segment (instead of a one day segment as above). This is reasonable from a practical point of view because it is not reasonable to have drug on one day and off the next day. For a 5 day segment, the size of the control vector is reduced to 1×20 from 1×100. Hence, the number of iterations to find the discrete optimal STI solution is now 2^{20} instead of 2^{100}.

Using the above iterative reduction technique, we obtained the optimal 5 day segment STI control vector

$$\varepsilon^* = (0\,0\,0\,0\,0\,b\,0\,0\,b\,0\,b\,b\,b\,0\,b\,b\,b\,b\,b\,0),$$

which is depicted in Fig. 12. The associated optimal state solutions are shown in Fig. 13.

As depicted in Fig. 12, the optimal 5 day segment STI strategy is that drug is off for the first 25 days, drug is on for the next 5 days, drug is off for the next 10 days, and so on. In addition, Fig. 13 reveals that solutions of the state system with the optimal STI control are similar to solutions with no treatment up to the 25th day. After that, the population of uninfected T_1 cells with STI approaches the population of uninfected T_1 cells with full treatment and the virus population is also maintained at a low level. Moreover, the immune effectors (E) with STI have larger values than in the other cases. That is, this optimal STI control boosts the immune effectors (E) while reducing virus load and increasing uninfected T_1 and T_2 cells over the treatment period.

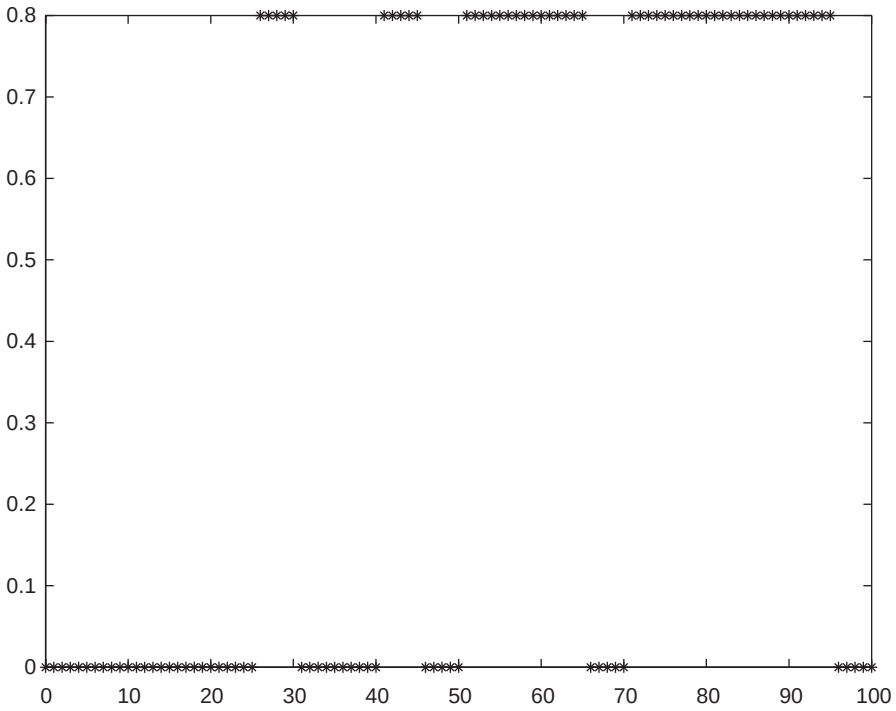

Fig. 12. Optimal STI control with $Q = 0.1$ and $R = 10000$ (5 day segments).

6.1. STI control problem with long time horizon

The above reduced iterative approach is still not computationally efficient for long time treatment regimens. For example, if one considers a treatment interruption strategy over 400 days duration of therapy, the resulting required number of cost functional evaluations is 2^{80} for a 5 day segment, which is still considerably large. One approach to mitigate this shortcoming is to consider subperiods of the given period such as [0, 50], [0, 100], [0, 150], . . . , [0, 400]. We then find an optimal STI control vector, ε_1^* over the first subperiod, [0, 50], using the reduced iteration technique as above. Since the size of ε_1^* is 1×10 (for a 5 day segment), its optimal solution can be obtained very quickly (with only $2^{10} = 1024$ iterations). In the second step, we consider our control vectors over the period [0, 100] as follows:

$$\varepsilon_2 = (\varepsilon_1^*, \star, \star, \star, \star, \star, \star, \star, \star, \star, \star) \quad \text{where } \star \text{ is 0 or b.}$$

That is, we fix the optimal STI ε_1^* as the first 10 elements of the control, and iterate ε_2 to find the last 10 elements of ε_2. In this case, the number of iterations is again just 2^{10}. We repeat this process to find an optimal STI control, ε_3^* over [0, 150] , ε_4^* over [0, 200], etc. The STI control obtained over the entire treatment period [0, 400] is $\varepsilon^* = \varepsilon_8^*$. It should be emphasized that ε^* is only suboptimal. However, our initial efforts suggest that it is a good approximation to the optimal STI control. In particular, the suboptimal STI control over [0, 100] that is obtained using the above subperiod method is almost identical (visually) to the optimal control using the reduced iterative method from above.

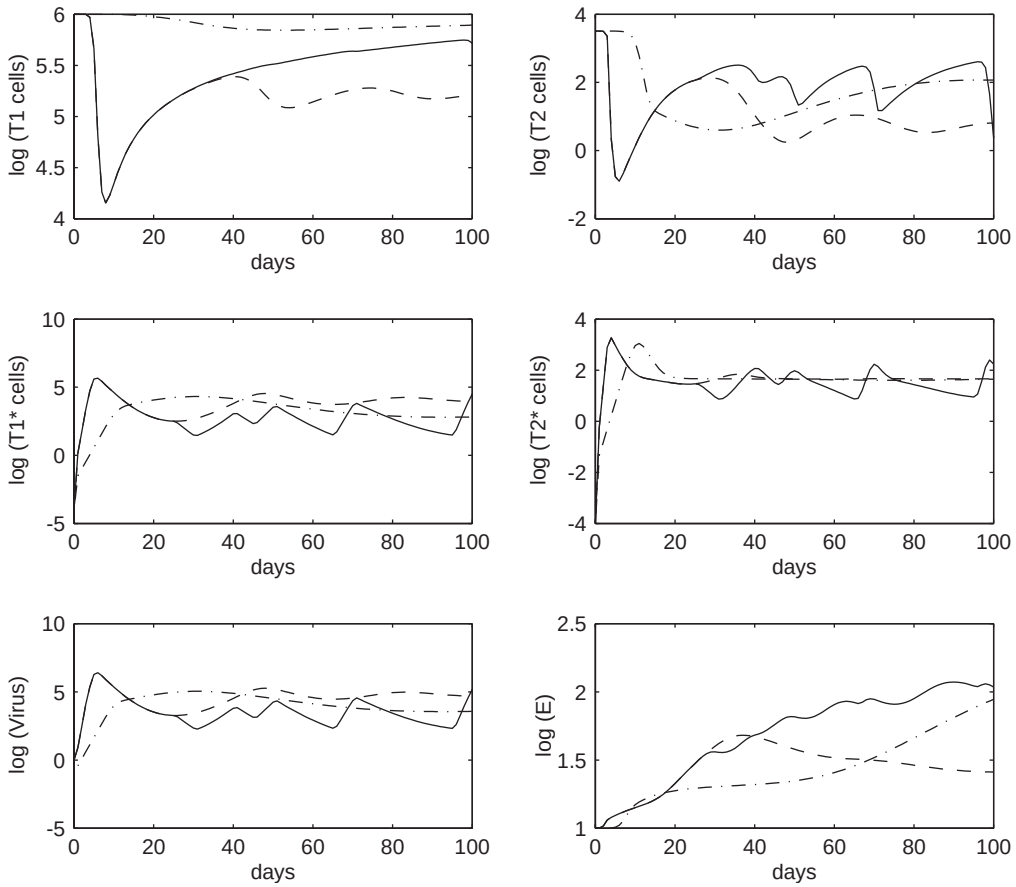

Fig. 13. Optimal 5 day segment STI solutions (−); solutions (−·) with full treatment (i.e., $\varepsilon = 0.8$); and solutions (−−) with no treatment (i.e., $\varepsilon = 0$) of early infection: $Q = 0.1$ and $R = 10000$.

Using the subperiod method (which might be viewed as a type of dynamic programming idea), the suboptimal STI control and its associated suboptimal solutions are depicted in Figs. 14 and 15, respectively. It is noted that the dynamics of uninfected T_1 cells with suboptimal STI control and full treatment are very close and the virus is also maintained at a low level. Moreover, the immune effectors with suboptimal STI control are maintained at a high level while the immune effectors with full treatment are at a lower level.

Finally, we consider a treatment interruption protocol where we apply suboptimal STI for the first treatment period [0, 300] then follow with full treatment over the final period [300, 400]. As depicted in Fig. 16 we see a substantial increase in the population of uninfected T_2 cells with suboptimal STI control followed by full treatment over [300, 400]. In addition, the immune effectors with the suboptimal STI and full treatment are also much higher on the 400th day. This increase in immune effectors allows the virus to maintain a small population. However, if we have full treatment for too long, (for example, if we consider a suboptimal STI over [0, 200] and follow with full treatment over [200, 400]), the immune effectors are increasing until around the 300th day and are decreasing after the 300th day.

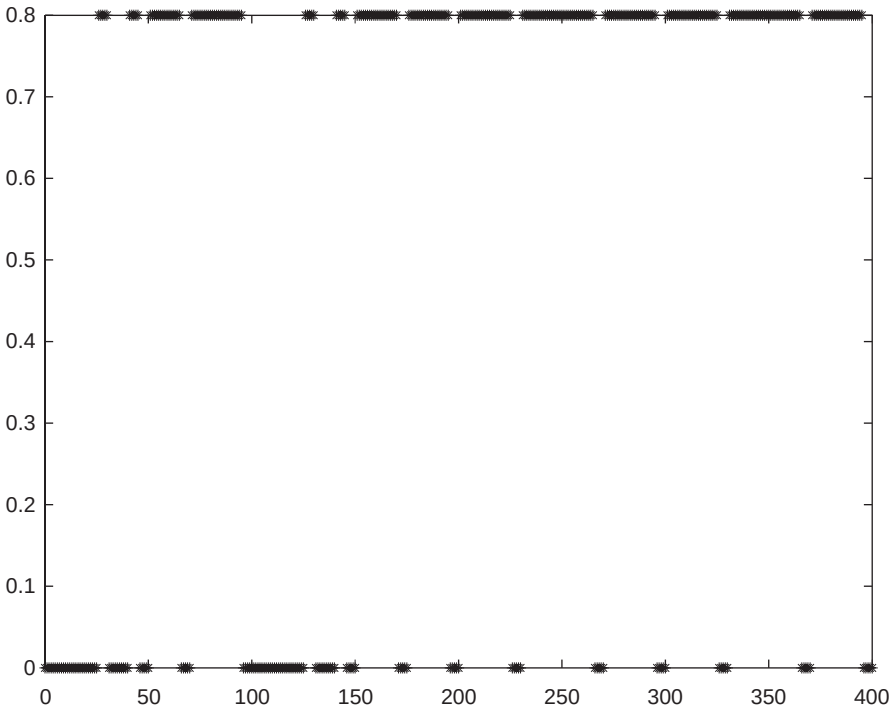

Fig. 14. Suboptimal STI control with $Q = 0.1$ and $R = 10000$ over [0, 400].

7. Concluding remarks

In this paper we have discussed a number of important issues relative to the investigation of HIV infection and treatment. The mathematical and statistical ideas we include are not meant to be exhaustive; rather they reflect some of those that our own experiences suggest are relevant in the investigation of models for the progression of diseases and possible therapeutic protocols. In particular, ideas and techniques in the important areas of model comparison and validation are not addressed here. While these are topics that need much attention, there are a few statistically-based model comparison methods currently available for use. These include the Akaike Information and related criteria (see [13] for extensive discussions and examples) as well as ANOVA type asymptotic statistical inverse problem methodology discussed in an HIV cellular level modeling context in [5].

The model we employed to illustrate the selected tools and techniques is not proposed as a penultimate model; it is simply one that we feel possesses features that we and other investigators believe are relevant in trying to understand certain quantitative and qualitative aspects of a very complex immunological process. Indeed, the immune response component of our model is naive in a number of ways. For example, it is so simple as to preclude investigations on a number of important immune response questions such as how early treatment during acute infection may affect and preserve certain aspects of the long term immune response in disease progression (e.g., the preservation of a memory cell based response).

In this regard we offer some philosophical thoughts about models and modeling in general. The very nature of modeling is that it is (or should be) an iterative process conductively (both collectively as a

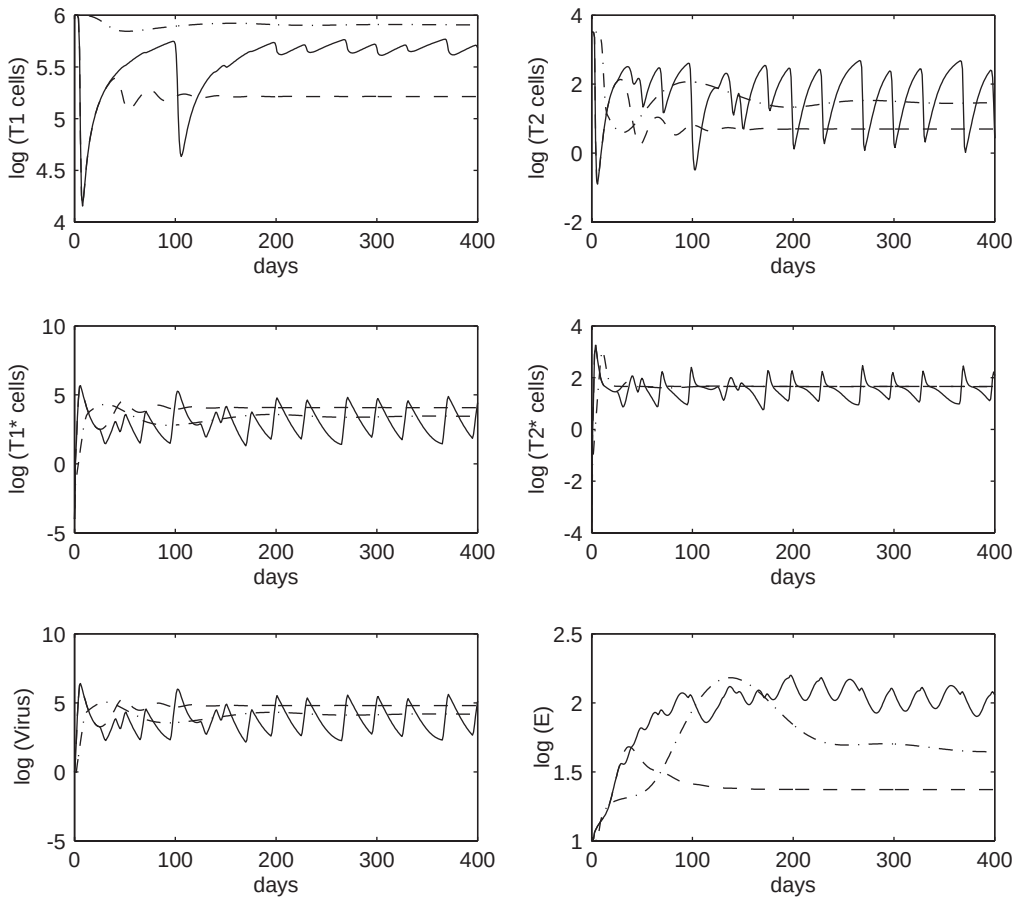

Fig. 15. Suboptimal STI solutions (−); solutions (−·) with full treatment (i.e., $\varepsilon = 0.8$); and solutions (−−) with no treatment (i.e., $\varepsilon = 0$) of early infection: $Q = 0.1$ and $R = 10000$.

community of HIV modelers and individually as an member of a focused team). Therefore it should be neither surprising nor bewildering to find (as discussed in Section 1.1) a substantial number of different types and levels of models for HIV progression. The "biologically correct" or "best" model (linear or nonlinear as may be required) is not the goal. Rather one models to gain insight and understanding into a biological or physical process (in our case, disease progression in the presence of therapy). This can result in a number of useful outcomes: the suggestion of new experiments (in vitro and/or in vivo) and clinical trials with new therapy regimes; the discovery of heretofore unelaborated mechanisms and/or pathways, etc. Model building and evaluation are challenging tasks and it is certainly true that one frequently faces difficult choices when deciding on which states (compartments) to include and which to omit. There are no clear, easy guidelines and procedures to follow other than some common sense points. For example, one cannot hope to ask and resolve questions about the possibility of immune system control of a disease unless some state(s) representing immune response levels are included. Moreover, the complex interplay between multiscale phenomena (e.g., cellular to system) always presents difficult modeling challenges;

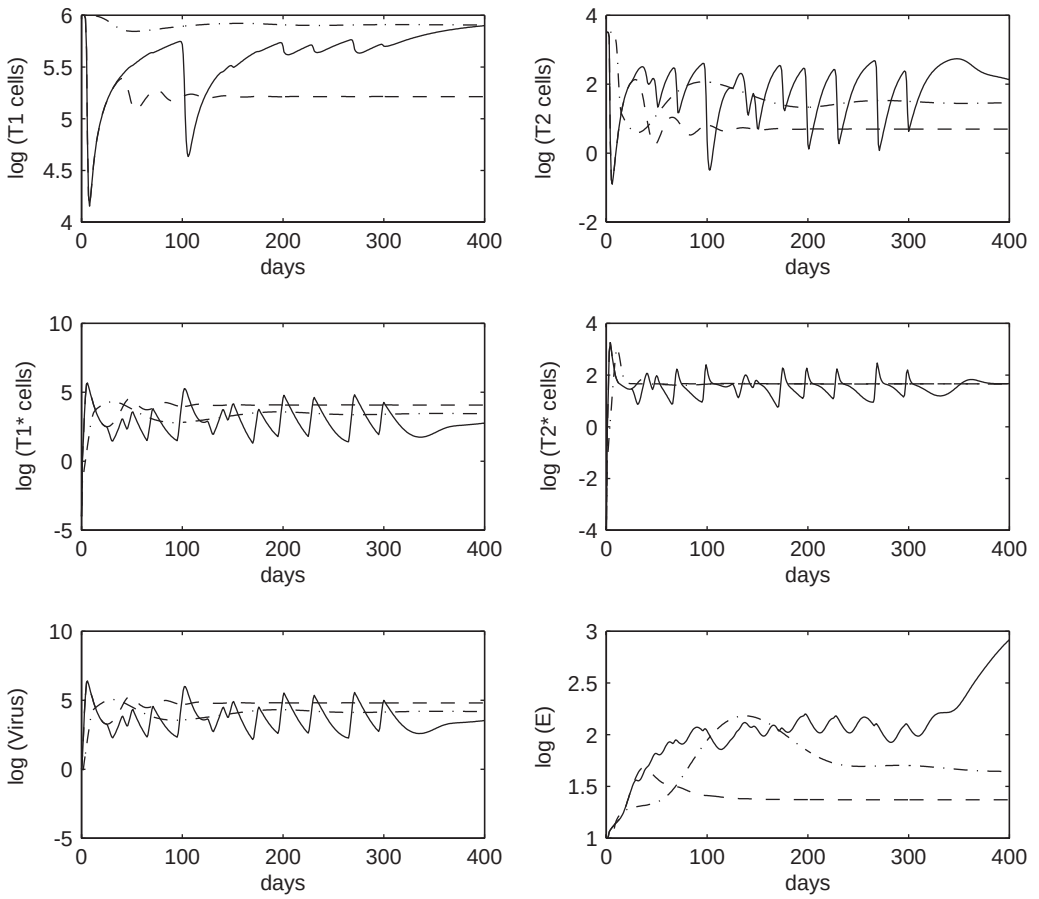

Fig. 16. STI solution (−) on [0, 300], coupled with full treatment over [300, 400]; solutions(−·) with full treatment (i.e., $\varepsilon = 0.8$) over [0, 400]; and solutions (−−) with no treatment (i.e., $\varepsilon = 0$) for early infection scenario: $Q = 0.1$ and $R = 10000$.

these challenges continue to grow as our knowledge of diseases at the genomic and cellular level becomes more extensive.

In spite of (or perhaps because of) these challenges, development of mathematical and statistical tools and our knowledge on how to use these in modeling efforts also continue to advance. While the human immune system and its related diseases are so complex as to often be perceived as overwhelming, it is precisely in such an environment that sophisticated quantitative modeling efforts should be pursued. The development of advanced mathematical, statistical and computational ideas for this endeavor offers fertile albeit challenging areas of scientific investigation for the foreseeable future and should be viewed as a source of great intellectual opportunity. We hope that this paper will help stimulate and promote such development with enthusiasm.

Acknowledgements

This research was supported in part by the Joint DMS/NIGMS Initiative to Support Research in the Area of Mathematical Biology under grant 1R01GM67299-01, and was facilitated through visits of the

authors to the Statistical and Applied Mathematical Sciences Institute (SAMSI), which is funded by NSF under grant DMS-0112069.

References

[1] B.M. Adams, H.T. Banks, J.E. Banks, J.D. Stark, Population dynamics models in plant-insect herbivore-pesticide interactions, CRSC Technical Report CRSC-TR03-12, NCSU, Raleigh, March, 2003 (revised August, 2003; Math. Biosci., to appear).
[2] C.A. Andrews, J.M. Davies, G.R. Schwarz, Adaptive data compression, Proc. IEEE 55 (1967) 267–277.
[3] N. Aubry, P. Holmes, J.L. Lumley, E. Stone, The dynamics of coherent structures in the wall region of a turbulent boundary layer, J. Fluid Mech. 192 (1988) 115–173.
[4] H.T. Banks, D.M. Bortz, A parameter sensitivity methodology in the context of HIV delay equation models, CRSC Technical Report CRSC-TR02-24, NCSU, Raleigh, August, 2002 (J. Math. Biol., to appear).
[5] H.T. Banks, D.M. Bortz, S.E. Holte, Incorporation of uncertainty in mathematical modeling of HIV infection dynamics, CRSC Technical Report CRSC-TR01-25, NCSU, Raleigh, September, 2001 (Math. Biosci. 183 (2003) 63–91).
[6] H.T. Banks, N.L. Gibson, W.P. Winfree, Electromagnetic crack detection inverse problems using terahertz interrogating signals, CRSC Technical Report CRSC-TR03-40, NCSU, Raleigh, October, 2003.
[7] G. Berkooz, P. Holmes, J.L. Lumley, J.C. Mattingly, Low-dimensional models of coherent structures in turbulence, Phys. Rep. 287 (1997) 338–384.
[8] S. Bonhoeffer, J.M. Coffin, M.A. Nowak, Human immunodeficiency virus drug therapy and virus load, J. Virol. 97 (1997) 3275–3278.
[9] S. Bonhoeffer, M. Rembiszewski, G.M. Ortiz, D.F. Nixon, Risks and benefits of structured antiretroviral drug therapy interruptions in HIV-1 infection, AIDS 14 (2000) 2313–2322.
[10] P. Borrow, H. Lewicki, B. Hahn, G.M. Shaw, M.B. Oldstone, Virus-specific CD8+ cytotoxic T-lymphocyte activity asssociated with control of viremia in primary human immunodeficiency virus type 1 infection, J. Virol. 68 (1994).
[11] M.E. Brandt, B. Chen, Feedback control of a biodynamical model of HIV-1, IEEE Trans. Biomed. Eng. 48 (2001) 754–759.
[12] R.P. Bucy, Immune clearance of HIV type 1 replication-active cells: a model of two patterns of steady state HIV infection, Aids Res. Hum. Retr. 15 (1999) 223–227.
[13] K.P. Burnham, D.R. Anderson, Model Selection and Multimodel Inference: a Practical Information-Theoretic Approach, Springer, New York, 2002.
[14] S. Butler, D. Kirschner, S. Lenhart, Optimal control of the chemotherapy affecting the infectivity of HIV, in: O. Arino, D. Axelrod, M. Kimmel (Eds.), Advances in Mathematical Population Dynamics-Molecules, Cells and Man, Word Scientific Press, Singapore, 1997, pp. 557–569.
[15] D.S. Callaway, A.S. Perelson, HIV-1 infection and low steady state viral loads, Bull. Math. Biol. 64 (2001) 29–64.
[16] M. Dalod, et al., Broad, intense anti-human immunodeficiency virus (HIV) ex vivo CD8+ responses in HIV type 1-infected patients: Comparison with anti-epstein-barr virus responses and changes during anti-retroviral therapy, J. Virol. 73 (1999) 7108–7116.
[17] M. Davidian, D.M. Giltinan, Nonlinear Models for Repeated Measurement Data, Chapman & Hall, London, 1995.
[18] R.T. Davey, N. Bhat, C. Yoder, et al., HIV-1 and T cell dynamics after interruption of highly active anti-retroviral therapy (HAART) in patients with a history of sustained viral suppression, Proc. Natl. Acad. Sci. 96 (1999) 15109–15114.
[19] C.A. Derdeyn, J.M. Kilby, G.D. Miralles, L.F. Li, G. Sfakianos, M.S. Saag, R.D. Hockett, R.P. Bucy, Evaluation of distinct blood lymphocyte populations in human immunodeficiency virus type-1-infected subjects in the absence or presence of efffective therapy, J. Inf. Dis. 180 (1999) 1851–1862.
[20] M. Eslami, Theory of Sensitivity in Dynamic Systems, Springer, Heidelberg, 1994.
[21] K.R. Fister, S. Lenhart, J.S. McNally, Optimizing chemotherapy in an HIV model, Electr. J. Differential Equations 32 (1998) 1–12.
[22] W.H. Fleming, R.W. Rishel, Deterministic and Stochastic Optimal Control, Springer, New York, 1975.
[23] P.M. Frank, Introduction to System Sensitivity Theory, Academic Press, New York, 1978.
[24] F. Garcia, M. Plana, C. Vidal, et al., Dynamics of viral load rebound and immunological changes after stopping effective antiretroviral therapy, AIDS 13 (1999) F79–F86.

[25] R.C. Gonzalez, P.A. Wintz, Digital Image Processing, Addison-Wesley, Reading, MA, 1987.
[26] C.M. Gray, et al., Frequency of class I HLA-restricted anti-HIV CD8+ T cells in individual receiving highly active anti-retroviral therapy (HAART), J. Immunol. 62 (1999) 1780–1788.
[27] M.D. Gunzburger, Perspectives in Flow Control and Optimization, SIAM, Philadelphia, 2003.
[28] P.A. Haslett, D.F. Nixon, Z. Shen, et al., Strong human immunodeficiency virus (HIV)-specific CD4+ T cell responses in a cohort of chronically infected patients are associated with interruptions in anti-HIV chemotherapy, J. Inf. Dis. 181 (2000) 1264–1272.
[29] R. Hilai, J. Rubinstein, Recognition of rotated images by invariant Karhunen–Loève expansion, J. Opt. Soc. Amer. A—Opt. Image Sci. Vis. 11 (1994) 1610–1618.
[30] D.D. Ho, A.U. Neumann, A.S. Perelson, et al., Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection, Nature 373 (1995) 123–126.
[31] I.T. Jolliffe, Principal Component Analysis, Springer, New York, 1986.
[32] H.R. Joshi, Optimal control of an HIV immunology model, Optim. Contr. Appl. Math. 23 (2002) 199–213.
[33] S.A. Kalams, et al., Levels of human immunodeficiency virus type 1-specific cytotoxic T-lymphocyte effector and memory responses decline after suppression of viremia with highly active anti-retroviral therapy, J. Virol. 73 (1999) 6721–6728.
[34] M.I. Kamien, N.L. Schwartz, Dynamic Optimization, North-Holland, Amsterdam, 1991.
[35] M. Kirby, L. Sirovich, Application of the Karhunen–Loève procedure for the characterization of human faces, IEEE Trans. Pattern Anal. Mach. Intell. 12 (1990) 103–108.
[36] D. Kirschner, S. Lenhart, S. Serbin, Optimal control of the chemotherapy of HIV, J. Math. Biol. 35 (1997) 775–792.
[37] J. Lisziewicz, F. Lori, Structured treatment interruptions in HIV/AIDS therapy, Microbes Infect. 4 (2002) 207–214.
[38] J. Lisziewicz, E. Rosenberg, J. Liebermann, Control of HIV despite the discontinuation of anti-retroviral therapy, New Engl. J. Med. 340 (1999) 1683–1684.
[39] M. Loève, Probability Theory, van Nostrand, Princeton, NJ, 1955.
[40] D.L. Lukes, Differential Equations: Classical to Controlled, Mathematics in Science and Engineering, Academic Press, New York, 1982.
[41] J.L. Lumley, The structure of inhomogeneous turbulent flows, in: A.M. Yaglom, V.I. Tatarski (Eds.), Atmospheric Turbulence and Radio Wave Propagation, Nauka, Moscow, 1967, pp. 166–178.
[42] K. Luzuriaga, et al., Dynamics of human immunodeficiency virus type 1 replication in vertically infected infants, J. Virol. 73 (1999) 362–367.
[43] H.V. Ly, H.T. Tran, Modeling and control of physical processes using proper orthogonal decomposition, Math. Comp. Model. 33 (2001) 223–236.
[44] A.J. Melvin, et al., HIV-1 dynamics in children, AIDS 20 (1999) 468–473.
[45] J.E. Mittler, B. Sulzer, A.U. Neumann, A.S. Perelson, Influence of delayed viral production on viral dynamics in HIV-1 infected patients, Math. Biosci. 152 (1998) 143–163.
[46] A.U. Neumann, R. Tubiana, V. Calvez, C. Robert, T.S. Li, H. Agut, B. Autran, HIV-1 rebound during interruption of highly active anti-retroviral therapy has no deleterious effect on re-initiated treatment, AIDS 13 (1999) 677–683.
[47] D.W. Notermans, J. Goudsmit, S.A. Danner, Rate of HIV-1 decline following anti-retroviral therapy is related to viral load at baseline and drug regimen, AIDS 12 (1998) 1483–1490.
[48] M.A. Nowak, C.R. Bangham, Population dynamics of immune responses to persistent viruses, Science 272 (1996) 74–79.
[49] G.S. Ogg, et al., Quantitation of HIV-1-specific cytotoxic T lymphocytes and plasma load of viral RNA, Science 279 (1998) 2103–2106.
[50] G.S. Ogg, et al., Decay kinetics of human immunodeficiency virus-specific effector cytotoxic T lymphocytes after combination anti-retroviral therapy, J. Virol. 73 (1999) 797–800.
[51] G. Ortiz, D. Nixon, A. Trkola, et al., HIV-1 specific immune responses in subjects who temporarily contain virus replication after discontinuation of highly active anti-retroviral therapy, J. Clin. Invest. 104 (1999) R13–R18.
[52] A. Oxenius, et al., Early highly active antiretroviral therapy for acute HIV-1 infection preserves immune function of CD8+ and CD4+ lymphocytes, Proc. Natl. Acad. Sci. 97 (2000) 3382–3387.
[53] E. Papasavvas, G.M. Ortiz, R. Gross, et al., Enhancement of human immunodeficiency virus type 1-specific CD4 and CD8 T cell responses in chronically infected persons after temporary treatment interruption, J. Inf. Dis. 182 (2000) 766–775.
[54] A.S. Perelson, A.U. Essunger, P.Y. Cao, et al., Decay characteristics of HIV-1 infected compartments during combination therapy, Nature 387 (1997) 188–191.
[55] A.S. Perelson, P.W. Nelson, Mathematical analysis of HIV-1 dynamics in vivo, SIAM Rev. 41 (1999) 3–44.

[56] A.S. Perelson, A.U. Neumann, M. Markowitz, et al., HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time, Science 271 (1996) 1582–1586.
[57] C.J. Pitcher, C. Quittner, D.M. Peterson, et al., HIV-1-specific CD4+ T cells are detectable in most individuals with active HIV-1 infection, but decline with prolonged viral suppression, Nature Med. 5 (1999) 518–525.
[58] E.S. Rosenberg, M. Altfeld, S.H. Poon, et al., Immune control of HIV-1 after early treatment of acute infection, Nature 407 (2000) 523–526.
[59] L. Ruiz, Martünez-Picado, J. Romeu, et al., Structured treatment interruption in chronically HIV-1 infected patients after long-term viral suppression, AIDS 14 (2000) 397–403.
[60] J.E. Schmitz, et al., Control of viremia in simian immunodeficiency virus infection by CD8+ lymphocytes, Science 283 (1999) 857–860.
[61] L. Sirovich, Turbulence and the dynamics of coherent structures, Part II: Symmetries and transformations, Quart. Appl. Math. XLV (1987) 573–582.
[62] H.M. Speigel, et al., Changes in frequency of HIV-1 specific cytotoxic T cell precursors and circulating effectors after combination anti-retroviral therapy in children, J. Inf. Dis. 180 (1999) 359–368.
[63] A.A. Tsiatis, V. DeGruttola, M.S. Wulfsohn, Modeling the relationship of survival to longitudinal data measured with error: Applications to survival and CD4 counts in patients with AIDS, J. Amer. Stat. Assoc. 90 (1995) 27–37.
[64] P.J.E. Vermeulen, D.P. Casasent, Karhunen–Loève techniques for optimal processing of time-sequential imagery, Opt. Eng. 30 (1991) 415–423.
[65] S. Watanabe, Karhunen–Loève expansion and factor analysis theoretical remarks and applications, in: Proceedings of the 4th Prague Conf. Inform. Theory, 1965.
[66] X. Wei, S.K. Ghosh, M.E. Taylor, et al., Viral dynamics of HIV-1 infection, Nature 373 (1995) 117–122.
[67] L.M. Wein, S.A. Zenios, M.A. Nowak, Dynamic multidrug therapies for HIV: a control theoretic approach, J. Theor. Biol. 185 (1997) 15–29.
[68] D. Wodarz, M.A. Nowak, Specific therapy regimes could lead to long-term immunological control of HIV, Proc. Natl. Acad. Sci. 96 (1999) 14464–14469.
[69] H. Wu, A.A. Ding, Population HIV-1 dynamics in vivo: applicable models and inferential tools for virological data from AIDS clinical trials, Biometrics 55 (1999) 410–418.
[70] H. Wu, et al., Characterization of viral dynamics in human immunodeficiency virus type 1-infected patients treated with combination anti-retroviral therapy: relationships to host factors, cellular restoration, and virologic end points, J. Inf. Dis. 179 (1999) 799–807.