Deterministic and Stochastic Models of AIDS Epidemics and HIV Infections with Intervention

WAI-YUAN TAN
University of Memphis, USA

HULIN WU
University of Rochester, USA

World Scientific

CONTENTS

CHAPTER 11

IDENTIFIABILITY OF HIV/AIDS MODELS

Annah M. Jeffrey and Xiaohua Xia

There is a need to accurately estimate all viral, host cell and immune response specific parameters. Before such an extensive parameter estimation exercise can be carried out, an identifiability analysis needs to be done to investigate whether or not it is possible to determine all the parameters. If it is found to be possible, then the conditions or restrictions that apply need to be known beforehand. The issues to address are the variables to be measured, the minimal number of measurements for a complete determination of all parameters, the frequency and when, during the course of the viral infection, such measurements can be taken. It is important to know this in advance, especially where budgets are concerned. The measured variable combination and conditions that result in the least number of measurements or cost less could then be selected. All the foregoing will be investigated in this article using a well-established nonlinear system identifiability theory, where identifiability is a basic system property of whether all model parameters can be calculated from the measured system outputs and known inputs.

Keywords: HIV models, identifiability, minimal measurements, algebraic framework.

1. Introduction

1.1. *A control system perspective*

It is needless to say that the modelling approach has paved the way of theoretical research. In order to use the HIV/AIDS models as a tool for treatment decisions, it is necessary to determine all parameters of the models for individuals. Even though there are general observations that can be made from the model and its structure, it is only when the model is tailored to each patient's individual parameters that clear benefits in the treatment arise. After casting the HIV/AIDS models into a control system

framework, many well-established control system techniques can be used to study various issues of HIV/AIDS modelling and control.

The available works on the engineering and related aspects of HIV/AIDS management generally involve sensor design (Grant and Xu, 2002), modelling the spread of infection in high risk populations, computer aided health care support systems (Beerenwinkel *et al.*, 2001), cellular level modelling of the interaction between the virus and the immune system (various), controllability and timing the initiation of therapy (Jeffrey, Xia and Craig, 2003a), feedback control of the viral load (Alvarez–Ramirez, Meraz and Velasco–Hermandez, 2000; Brandt and Chen, 2001), as well as optimal control and multi-drug therapy scheduling (Kirschner, Lenhart and Serbin, 1997; Wein, Zenios and Nowak, 1997; de Souza, Caetano and Yoneyama, 2000). Other model analysis works such as viral load time response under therapy and bifurcation are also available (Gumel, Twizell and Yu, 2000; Jeffrey and Xia, 2002). However, in total, the number of papers with a control system perspective is limited. There is scope and potential for more work to be done in order to further assist or help alleviate the burden imposed on health care givers by this wide spreading viral infection. To this end, this paper presents an identifiability analysis of some HIV/AIDS models. The intention is to provide more insight into the properties of these models, as well as help to formulate guidelines for and aid clinicians in measurement and test scheduling for the eventual estimation of the model parameters. The actual parameter estimation techniques or algorithms will however, not be dealt with here. For that, the reader is referred to Xia (2003), and Filter and Xia (2003).

1.2. *Use for parameter estimates*

If model parameter estimates can be obtained during the early stages of the HIV infection, they can be used to predict viral load set points, which are an important indicator of disease progression (Dewhurst, da Cruz and Whetter, 2000). In this case, if such estimates can be obtained, then it must be within a reasonably short period of time. From an HIV vaccination point of view, such estimates can be used to determine the vaccine efficacy. Estimates for model parameters are available in for example (Kirschner, Lenhart and Serbin, 1997; Nowak and May, 2000; Perelson *et al.*, 1996a; Ramratnam *et al.*, 1999). Not much effort however, has been put into *simultaneously* estimating all of these parameters. The available

parameter estimates, especially for the compartmental models, are sparse and incomplete. There are indications though, that accuracy of estimates for some of the model parameters is increasing as more innovative approaches for their estimation are employed.

Another reason for the need for accurate parameter estimates is the fact that there are inter-individual variations in parameters (Muller, Maree and De Boer, 2001a; Little *et al.*, 1999). Furthermore, parameters are thought to vary from one stage to the next as the infection progresses (Kramer, 1999; Nowak and May, 2000). There is a possibility though, that what seems to be changes in model parameters with time could just be the effect of unmodelled dynamics or external disturbances. Obtaining an individuals parameters at different stages of the viral infection could therefore, settle this issue.

HIV drug pharmacokinetics, pharmacodynamics and adverse reactions are genetically predisposed (Pirmohamed and Back, 2001; Roden and George, 2002). Furthermore, the response to therapy, for example, the time to effectively suppress the viral load, are parameter dependent (Jeffrey, Xia and Craig, 2003b). Given this parameter dependence of the response to therapy, one can therefore consider exploiting inter-individual variations in parameters to individualize treatment and enhance the benefits of anti-retroviral therapy (Becker and Hoetelmans, 2002; Lindpainter, 2001). There will be a need then, for test measurements to be done over a short period of time.

1.3. *Layout of paper*

An identifiability analysis of some HIV/AIDS models will be presented in this paper. The intention is to provide more insight into the properties of these models, as well as help to formulate guidelines for and aid clinicians in measurement and test scheduling for the eventual estimation of the model parameters. The issues to address are the variables to be measured, the minimal number of measurements for a complete determination of all parameters, the frequency and when, during the course of the viral infection, such measurements can be taken.

The layout of the paper is as follows: Section 2 presents the models to be analyzed. Section 3 introduces the identifiability concept and the procedure to follow when applying this concept. Section 4 presents and discusses the analysis results of the HIV/AIDS models presented in Section 2, while Section 5 has the conclusions.

2. HIV/AIDS Models

Mathematical models that describe the host-pathogen interaction between the immune system and HIV should be able to explain the initial high rise in plasma viral load, its decline and settling to levels that are much lower than the peak viral load. The subsequent dramatic increase in the viral load during the later stage of the infection and timing of this increase should also be explained. Models have been developed and explained in, for example, Kirschner (1996), Perelson and Nelson (1999), Nowak and May (2000), Hraba, Dolezal and Celikovsky (1990), Hraba and Dolezal (1996), Tan and Wu (1998), Gumel, Shivakumar and Sahai (2001), de Souza (1999), Culshaw and Ruan (2000), Murray *et al* (1998). Most of these models are deterministic and based on balancing population dynamics, while some are stochastic and take into account the random variations in HIV dynamics. Some of these models are single compartment and address the population dynamics of the virus and $CD4^+$ T cells in plasma, while others are multi compartment and take into consideration other cells in the body like macrophages that are as susceptible to the virus.

None of these models however, can completely exhibit all that is observed clinically and account for the full course of the infection. The main reason for the models' limitation is lack of a good understanding

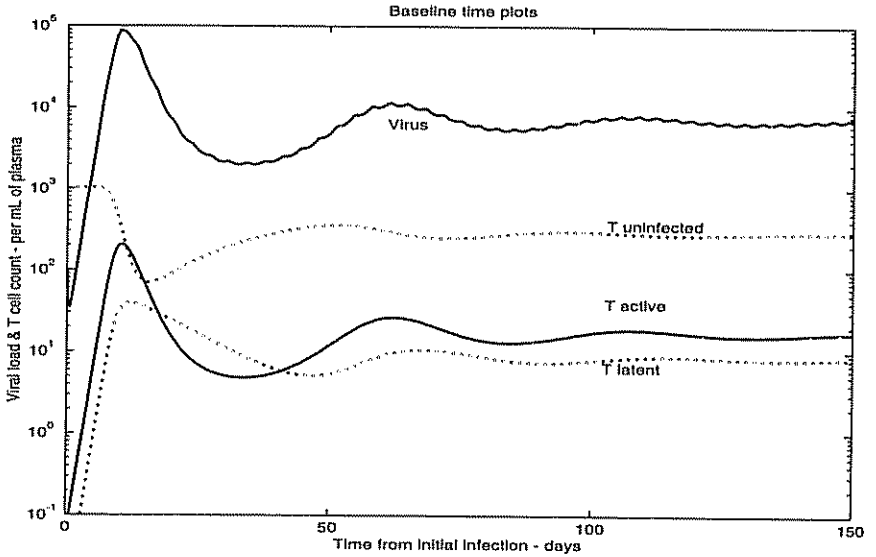


Fig 1 Simulated viral load and $CD4^+$ T cell variation as the HIV infection progresses

of the immunology of the human body against HIV. Biological systems tend to exhibit multi-compartmental interactions that are usually not well understood and as a result, cannot be accurately modelled mathematically. Another point to consider is that these models do not take into account other extenuating environmental, social and welfare factors that may affect the progression of the infection.

These models, however, do adequately explain the interaction of the virus and the immune system up to the clinical latency stage as illustrated in Fig. 1. In an attempt to account for the later or advanced stage of the infection, some model parameters are assumed to change as the infection progresses. These assumptions, though not clinically validated, do give a virus and $CD4^+$ cell profile that complies with clinical observations. Other suggestions are that the prolonged production and destruction of the $CD4^+$ lymphocytes ultimately results in an immune collapse. More recent studies have however, challenged these paradigms (Dewhurst, da Cruz and Whetter, 2000).

The following is a selected presentation of a class of HIV/AIDS models in Perelson and Nelson (1999).

2.1. *The basic 3-dimensional model*

The basic 3-dimensional (3D) model presented below in Eq. (1) is single compartment and shows the interaction between the virus and the $CD4^+$ T cells in blood.

$$\sum_{3D} = \begin{cases} \dot{T} &= s - dT - \beta TV, \\ \dot{T}^* &= \beta TV - \mu T^*, \\ \dot{V} &= pT^* - cV. \end{cases} \tag{1}$$

The state variables T, T^* and V are the plasma concentrations of the uninfected $CD4^+$ T cells, the actively infected $CD4^+$ T cells and the free virus particles, respectively. The first equation of (1) describes the population dynamics of the uninfected $CD4^+$ T cells. It shows that they are produced from a source at a rate s and die with a rate constant d.

Some authors assume that the source term s is constant. Other authors suggest that the source term depends on other variables of the system to best fit the clinical data, even though there is no direct medical evidence. Since HIV may be able to infect cells in the thymus and bone marrow and thus lead to a reduced production of new immunocompetent T cells, it is believed that the source term s is a decreasing function of the viral load. The following forms exist in literature: $s(v) = se^{(-\theta v)}$ in Perelson (1989);

$s(v) = \theta s/(\theta + v)$ in Perelson, Kirschner and De Boer (1993); Alvarez–Ramirez, Meraz and Velasco–Hernandez (2000); $s(v) = \theta_1 s + \theta_2 s/(B_s + v)$ in Kirschner and Webb (1996).

A proliferation rate term is sometimes added to the uninfected $CD4^+$ T cell dynamics. Since there are suggestions that the proliferation rate is density dependent with the rate of proliferation slowing as the T cell count gets high (Ho *et al.*, 1995), the most common form for proliferation is taken as the following logistic function (Perelson, Kirschner and De Boer, 1993):

$$\phi(T) = rT\left(1 - \frac{T}{T_{max}}\right), \tag{2}$$

with r as the proliferation rate constant and T_{max} is the T cell population density at which proliferation shuts off.

Uninfected $CD4^+$ T cells are infected by the virus at a rate that is proportional to the product of their abundance and the amount of free virus particles. The proportionality constant β is an indication of the effectiveness of the infection process and includes the rate at which virus particles find target cells, the rate of virus entry and probability of successful infection. The second equation of (1) describes the population dynamics of the actively infected $CD4^+$ T cells and shows that μ is their death rate constant. The third equation similarly describes the population dynamics of the free virus particles and it can be seen that an actively infected $CD4^+$ T cell produces infectious free virus particles with a rate constant p and c is the death rate constant for these virus particles.

2.2. *The latently infected 4-dimensional model*

There are additional reservoirs of virus. There is a pool of latently infected, resting $CD4^+$ T cells that is established early during primary HIV infection (Chun *et al.*, 1998). A quantitative image analysis technique was used to reveal viral burden in lymphoid tissue (Haase *et al.*, 1996), particularly on the surface of folicular dendritic cells (FDC). Perelson *et al.* observed that, after the rapid first phase of decay during the initial 1–2 weeks of antiretroviral therapy, plasma virus levels declined at a considerably slower rate (Perelson *et al.*, 1997). This second phase of viral decay was attributed to the turnover of a longer lived virus reservoir of infected cell population, which was determined to have a half-life of 1–4 weeks. This means that on average, it would take between $\frac{1}{2}$ and 3 years of perfectly effective antiretroviral therapy to eradicate the virus. This estimate generated a lot of enthusiasm and optimism in 1996 (Perelson *et al.*, 1996).

The so-called latently infected cell model presented in (3) takes note of the fact that not all CD4$^+$ T cells actively produce virus upon successful infection. This is reflected by dividing the infected cell pool into actively and latently infected cells.

$$\sum_{4D} = \begin{cases} \dot{T} &= s - dT - \beta TV, \\ \dot{T}_1 &= q_1\beta TV - kT_1 - \mu_1 T_1, \\ \dot{T}_2 &= q_2\beta TV + kT_1 - \mu_2 T_2, \\ \dot{V} &= pT_2 - cV. \end{cases} \tag{3}$$

In this 4-dimensional (4D) model, state variable T_2 represents the cells that actively produce virus, while the latently infected cells, denoted by T_1, harbor HIV proviral DNA and do not produce virus until they are activated. Parameter k is the activation rate constant. Parameters q_1 and q_2 are the probabilities that upon infection, a CD4$^+$ T cell will become either latent or actively produce virus.

2.3. *The extended 6-dimensional model*

While the latently infected cell model fits the patient data, it is not the only reasonable biological model. The source underlying the second phase kinetics might be the release of virus trapped in the lymphoid tissue. It could be linked to infected macrophage (Perelson *et al.*, 1997). Continued follow-up of persons who have remained on HAART for extended periods of time has provided strong evidence for the existence of a possible reservoir in long-lived CD4$^+$ memory T lymphocytes, a third phase of HIV decay observed during HAART. The kinetics of decay are extremely slow, and the half-life of the memory cell reservoir has been estimated at between 6 and 44 months. As a consequence, the predicted time for effective therapy required to fully eradicate the virus from the body ranges from 9 to 72 years. This suggested that a true virologic cure is unattainable using the conventional antiretroviral drugs. Table 1 is a summary of these important findings.

Table 1 Virus reservoir and life span

Infected cell	Size	Half-life	Eradication
Active CD4	3×10^7	1 day	25 days
FDC	3×10^8–10^{11}	1–4 wks	0.5–2.8 yrs
Macrophage	?	1–4 wks	?
Memory CD4	10^5–10^6	6–44 months	9–72 yrs

Most of the figures can be found or deduced from Dewhurst, da Cruz and Whetter (2000).

Another totally different theory proposed by Grossman and colleagues says that the very slow rates of viral decay which occur following the initiation of HAART can be best explained if most of the virus is produced by cells infected after the commencement of treatment (Grossman *et al.*, 1999). These developments have brought to us challenges to face.

An upgrade of the 4D model is the extended 6-dimensional (6D) model given in (4), as it is apparent that there are other cells in the body besides CD4$^+$ T cells that are as susceptible to the virus. The release of the virus from these cells and other infected compartments has been shown to affect the virus kinetics in plasma (Muller, Maree and De Boer, 2001b). The particular cells of interest here are macrophages. These are large cells that live longer than the CD4$^+$ T cells and are chronic virus producers.

$$\sum_{6D} = \begin{cases} \dot{T} &= s_T - d_T T - \beta_T TV, \\ \dot{T}_1 &= q_1 \beta_T TV - kT_1 - \mu_1 T_1, \\ \dot{T}_2 &= q_2 \beta_T TV + kT_1 - \mu_2 T_2, \\ \dot{M} &= s_M - d_M M - \beta_M MV, \\ \dot{M}^* &= \beta_M MV - \delta M^*, \\ \dot{V} &= p_T T_2 + p_M M^* - cV. \end{cases} \tag{4}$$

M and M^* are the concentrations of the uninfected and infected macrophages respectively. The fourth equation of (4) shows that uninfected macrophages are produced from a source at a rate s_M, die with a rate constant d_M and are infected by the virus at a rate that is proportional to their abundance. Parameter β_M is an indication of the efficiency of the infection process. The fifth and sixth equations of (4) show that infected macrophages die with a rate constant δ and an infected macrophage cell produces virus with rate constant p_M. Table 2 summarizes the parameters that are in each model.

3. Identifiability Concepts

Observability (identifiability) is a basic system property of whether all state variables (all parameters) can be calculated from the measured output. The precise meaning of observability and identifiability is defined in (Conte, Moog and Perdon, 1999; Glad, 1997). Various works on nonlinear system identification can be found for example, in Tunali and Tarn (1987), Diop and Fliess (1991) and Ljung and Glad (1994).

Table 2. Parameters in each model

Parameter and description		3D	4D	6D
s, s_T	Source rate for uninfected CD4$^+$T cells	$\star$	$\star$	$\star$
d, d_T	Death rate for uninfected CD4$^+$ T cells	$\star$	$\star$	$\star$
β, β_T	Infection rate for CD4$^+$ T cells by virus	$\star$	$\star$	$\star$
μ, μ_2	Death rate for actively infected CD4$^+$ T cells	$\star$	$\star$	$\star$
μ_1	Death rate for latently infected CD4$^+$ T cells		$\star$	$\star$
k	Activation rate for latently infected CD4$^+$ T cells		$\star$	$\star$
q_1	Fraction of infected CD4$^+$ T cells that become latent		$\star$	$\star$
q_2	Fraction of infected CD4$^+$ T cells that become active		$\star$	$\star$
s_M	Source rate for uninfected macrophages			$\star$
d_M	Death rate for uninfected macrophages			$\star$
β_M	Infection rate for macrophages by virus			$\star$
δ	Death rate for infected macrophages			$\star$
p, p_T	Virus particle production rate per infected CD4$^+$ T cell	$\star$	$\star$	$\star$
p_M	Virus particle production rate per infected macrophage			$\star$
c	Death rate for virus particle	$\star$	$\star$	$\star$
Total number of parameters in model		6	10	15

*indicates the applicable parameter.
All parameters except s, s_T, and s_M are rate constants

The identifiability concept applied here was presented and analyzed by Xia and Moog (2003). The concept is based on the practical requirement that if parameters can be expressed as functions of known quantities of the model, then it is possible to work in the algebraic framework. In essence, if as many such function expressions as there are unknown parameters can be generated, then it is possible to solve for all the unknown parameters. The identifiability characterizations presented here lend themselves to isolate the conditions under which the system parameters cannot be determined.

The following is the identifiability concept and the calculation procedure to follow when applying this concept to test a system's identifiability properties.

3.1. *Concepts*

Consider a nonlinear system,

$$\sum_{\theta}: \begin{cases} \dot{x} = f(x, \theta, u), & x(0, \theta) = x_0, \\ y = h(x, \theta, u), \end{cases} \tag{5}$$

where $x \in R^n$, $u \in R^m$ and $y \in R^p$ are the state, input and output variables of the system. Assume that

$$\operatorname{rank} \partial h(x, \theta, u)/\partial x = p. \tag{6}$$

θ is the parameter to be identified, and is assumed to belong to $\mathcal{P}$ which is an open subset of R^q. The functions $f(x, \theta, u)$ and $h(x, \theta, u)$ are meromorphic functions on a connected open subset $\mathcal{M}$ of R^n. Moreover, without loss of generality, x_0 is assumed to be independent of θ and not an equilibrium point of the system.

An input function $u(t)$: $[0, T] \to \mathcal{U}$, where $\mathcal{U}$ is an open subset of R^m, is called an admissible input (on $[0, T]$). For any initial condition x_0 and an admissible input $u(t)$ on $[0, T]$, there exists, on a possibly smaller time interval, $[0, \bar{T}]$, $\bar{T} \leq T$, a parameterized solution $x(t, \theta, x_0, u)$. The corresponding output is denoted by $y(t, \theta, x_0, u)$.

A classical definition of identifiability can be found in Tunali and Tarn (1987).

Definition 1: The system Σ_θ is said to be x_0-identifiable at θ through an admissible input u(on $[0, T]$) if there exists an open set $\mathcal{P}^0 \subset \mathcal{P}$ containing θ such that for any two θ_1, $\theta_2 \in \mathcal{P}^0$, $\theta_1 \neq \theta_2$, the solutions $x(t, \theta_1, x_0, u)$ and $x(t, \theta_2, x_0, u)$ exist on $[0, \epsilon]$, $0 < \epsilon \leq T$, and their corresponding outputs satisfy, on $t \in [0, \epsilon]$,

$$y(t, \theta_1, x_0, u) \neq y(t, \theta_2, x_0, u).$$

This property was termed in Tunali and Tarn (1987) as (instantaneously) locally strongly identifiability. We are more interested in a generic property of identifiability. This property was studied in Denis–Vidal, Joly–Blanchard and Noiret (2001) for polynomial systems. To introduce such a concept, we need a topology for the input function space. For any $T > 0$ and a positive integer N, the space $C^N[0, T]$ is the space of all functions on $[0, T]$ which have continuous differentiations up to the order N. A topology of the space $C^N[0, T]$ is the one associated with following well-defined norm: for $r(t) \in C^N[0, T]$,

$$\|r(t)\| = \sum_{i=0}^{N} \max_{t \in [0,T]} |r^{(i)}(t)|.$$

For any $T > 0$ and positive integer N, denote $C_{\mathcal{U}}^N[0, T]$ the set of all admissible inputs (on $[0, T]$) that have continuous derivatives up to the order N. The topology of $C_{\mathcal{U}}^N[0, T]$ is defined to be the m-fold product

topology of $C^N[0,T]$. The topology of $C^N_{\mathcal{U}}[0,T] \times C^N_{\mathcal{U}}[0,T]$ is defined to be the product topology of $C^N_{\mathcal{U}}[0,T]$. The M-fold product of $C^N_{\mathcal{U}}[0,T]$ is denoted as $(C^N_{\mathcal{U}}[0,T])^M$.

Definition 2: The system Σ_θ is said to be structurally identifiable if there exist a $T > 0$, and a positive integer N, and open and dense subsets $\mathcal{M}^0 \subset \mathcal{M}$, $\mathcal{P}^0 \subset \mathcal{P}$, $\mathcal{U}^0 \subset C^N_{\mathcal{U}}[0,T]$ such that the system Σ_θ is x_0-identifiable at θ through u, for every $x_0 \in \mathcal{M}^0$, $\theta \in \mathcal{P}^0$ and $u \in \mathcal{U}^0$.

The structural identifiability is also interchangeably called geometrical identifiability in this paper, because Definition 2 is the generic version of the definition of Tunali and Tarn (1987). The structural identifiability is used to characterize the one-to-one property of the map from the parameters to the system output. The algebraic identifiability is about construction of parameters from algebraic equations of the system input and output. This concept was first employed in Glad (1997), and Ljung and Glad (1994) and later formally defined in Diop and Fliess (1991) in the differential algebraic framework.

Definition 3: The system Σ_θ is said to be algebraically identifiable if there exist a $T > 0$, positive integers N and k, and a meromorphic function Φ: $R^q \times R^{(k+1)m} \times R^{(k+1)p} \to R^q$ such that

$$\det \frac{\partial \Phi}{\partial \theta} \not\equiv 0 \tag{7}$$

and

$$\Phi(\theta, u, \dot{u}, \ldots, u^{(k)}, y, \dot{y}, \ldots, y^{(k)}) = 0 \tag{8}$$

hold, on $[0,T]$, for all $(\theta, u, \dot{u}, \ldots, u^{(k)}, y, \dot{y}, \ldots, y^{(k)})$ where (θ, x_0, u) belong to an open and dense subset of $\mathcal{P} \times \mathcal{M} \times C^N_{\mathcal{U}}[0,T]$, and $\dot{u}, \ldots, u^{(k)}$ are the corresponding derivatives of u, and $y, \dot{y}, \ldots, y^{(k)}$ are the derivatives of the corresponding output $y(t, \theta, x_0, u)$.

Algebraic identifiability enables one to construct the parameters from solving algebraic equations depending only on the information of the input and output. As a matter of fact, under condition (7), one can solve locally an expression for θ from Eq. (8) by using the Implicit Function Theorem.

Sometimes an initial condition is known for a system. The information of the known initial state may provide additional help in determining the parameters. This phenomenon was recognized in Tunali and Tarn (1987), Glad (1997), and Ljung and Glad (1994).

Definition 4: The system Σ_θ is said to be identifiable with known initial conditions if there exist a positive integer k and a meromorphic function Φ: $R^q \times R^n \times R^{(k+1)m} \times R^{(k+1)p} \to R^q$ such that

$$\det \frac{\partial \Phi}{\partial \theta} \not\equiv 0 \tag{9}$$

and

$$\Phi(\theta, x_0, u(0^+), \dot{u}(0^+), \ldots, u^{(k)}(0^+), y(0^+), \dot{y}(0^+), \ldots, y^{(k)}(0^+)) = 0 \tag{10}$$

hold for all $(\theta, x_0, u(0^+), \dot{u}(0^+), \ldots, u^{(k)}(0^+), y(0^+), \dot{y}(0^+), \ldots, y^{(k)}(0^+))$, where $(\theta, x_0, u(0^+), \dot{u}(0^+), \ldots, u^{(k)}(0^+)$ belong to an open and dense subset of $\mathcal{P} \times \mathcal{M} \times \mathcal{U}^{(k+1)m}$, and $y(0^+), \dot{y}(0^+), \ldots, y^{(k)}(0^+))$ are the derivatives of the corresponding output $y(t, \theta, x_0, u)$ evaluated at $t = 0^+$.

Again, under condition (9), one can solve locally an expression for θ from Eq. (10) by using the Implicit Function Theorem. The expression will depend on the known initial state, the input and the output.

3.2. *Characterizations*

We will give characterizations of the algebraic identifiability and structural identifiability in the linear algebraic framework of nonlinear systems (Conte, Moog and Perdon, 1999). The characterizations also lend themselves to isolate the initial conditions and inputs that are not persistently exciting, i.e. where the system parameters cannot be determined.

To recall the linear algebraic framework, let $\mathcal{K}$ be the field consisting of meromorphic functions of x, θ, u and finite derivatives of u, and define

$$E = \mathrm{span}_{\mathcal{K}}\{\mathrm{d}\mathcal{K}\},$$

that is, a vector in E is a linear combination of a finite number of one-forms from $\mathrm{d}x, \mathrm{d}\theta, \mathrm{d}u, \mathrm{d}\dot{u}, \ldots, \mathrm{d}u^{(k)}, \ldots$, with coefficients in $\mathcal{K}$. The vectors in E are called one-forms.

The differentiation of a function $\phi(x, \theta, u, \ldots, u^{(k)})$ along the dynamics of the system (5) is defined as

$$\dot{\phi} = \frac{\partial \phi}{\partial x} f(x, \theta, u) + \sum_{i=0}^{k} \frac{\partial \phi}{\partial u^{(i)}} u^{(i+1)},$$

and this operation can be extended to differential one-forms $\omega = \kappa_x \mathrm{d}x + \kappa_\theta \mathrm{d}\theta + \sum \eta_i \mathrm{d}u^{(i)} \in E$ as the following:

$$\dot{\omega} = \dot{\kappa}_x \mathrm{d}x + \dot{\kappa}_\theta \mathrm{d}\theta + \sum \dot{\eta}_i \mathrm{d}u^{(i)} + \kappa_x \mathrm{d}f(x, \theta, u) + \sum \eta_i \mathrm{d}u^{(i+1)}.$$

Table 3. Characterization of identifiability notions.

Algebraic identifiability	:	$\Theta \subset \mathcal{Y} + \mathcal{U}$
Identifiability with known initial conditions	:	$\Theta \subset \mathcal{X} + \mathcal{Y} + \mathcal{U}$
Geometric identifiability	:	$\Theta \subset \mathcal{X} + \mathcal{Y} + \mathcal{U}$

Note that $\dot{\omega} \in E$.

Denote

$$\mathcal{Y} = \bigcup_{k=0}^{\infty} \operatorname{span}\{\mathrm{d}y, \mathrm{d}\dot{y}, \ldots, \mathrm{d}y^{(k)}\},$$
$$\mathcal{X} = \operatorname{span}\{\mathrm{d}x\},$$
$$\mathcal{U} = \bigcup_{k=0}^{\infty} \operatorname{span}\{\mathrm{d}u, \mathrm{d}\dot{u}, \ldots, \mathrm{d}u^{(k)}\},$$
$$\Theta = \operatorname{span}\{\mathrm{d}\theta\},$$

then the characterizations of the different notions of identifiability can be summarized in Table 3. For proof of these characterizations, the reader is referred to Xia and Moog (2003). It can be seen from the table that geometric identifiability is equivalent to identifiability with known initial conditions.

3.3. *Calculation*

Let us now investigate the computational issues for the determination of the parameters.

Note that $\mathcal{X} \cap (\mathcal{Y} + \mathcal{U})$ is the observation cospace of system (5) (Conte, Moog and Perdon, 1999), while $\mathcal{X} \cap (\mathcal{Y} + \Theta + \mathcal{U})$ can be interpreted as the observation cospace with parameter.

One checks the identifiability through calculating the input-output relations of the system. One way of doing this is first to eliminate x through observability properties of x.

Define, for system (5), the so-called observability indices. Let

$$\mathcal{F}_k := \mathcal{X} \cap \left(\operatorname{span}\{\mathrm{d}y, \mathrm{d}\dot{y}, \ldots, \mathrm{d}y^{(k-1)} + \mathcal{U} + \Theta\}\right)$$

for $k = 1, \ldots, n$. Consider the filtration

$$\mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots \subset \mathcal{F}_n.$$

Then, as done in Krener and Respondek (1985) for nonlinear systems without parameters and which are linearizable by output injections, define

$d_1 := \dim \mathcal{F}_1$ and $d_k := \dim \mathcal{F}_k - \dim \mathcal{F}_{k-1}$, for $k = 2, \ldots, n$. Let

$$k_i := \max\{k \mid d_k \geq i\}.$$

Then the list $\{k_1, k_2, \ldots, k_p\}$ is the list of observability indices and d_k represents the number of observability indices which are greater than or equal to k, for $k = 1, \ldots, n$. Reorder, if necessary, the output components such that

$$\begin{aligned} \operatorname{rank} \frac{\partial(y, \dot{y}, \ldots, y^{(n-1)}}{\partial x} &= \operatorname{rank} \partial(y_1, \dot{y}_1, \ldots, y_1^{(k_1-1)}, y_2, \ldots, y_2^{(k_2-1)}, \ldots, \\ &\qquad y_p, \dot{y}_p, \ldots, y_p^{(k_p-1)})/\partial x \qquad (11) \\ &= k_1 + k_2 + \cdots + k_p. \qquad (12) \end{aligned}$$

Thanks to the assumption (6), the p observability indices are well defined. Compute

$$\begin{aligned} \mathrm{d}y_1 &= \xi_{11}\mathrm{d}x + \gamma_{11}\mathrm{d}\theta (\bmod \mathcal{U}), \\ &\vdots \\ \mathrm{d}y_p &= \xi_{p1}\mathrm{d}x + \gamma_{p1}\mathrm{d}\theta (\bmod \mathcal{U}). \end{aligned}$$

By assumption, $\operatorname{rank} \begin{bmatrix} \xi_{11} \\ \vdots \\ \xi_{p1} \end{bmatrix} = p$. And more generally, compute

$$\mathrm{d}y_i^{(j-1)} = \xi_{ij}\mathrm{d}x + \gamma_{ij}\mathrm{d}\theta (\bmod \mathcal{U}),$$

for $i = 1, \ldots, p$ and $j = 1, \ldots, k_i$. From (11), any ξ_{ij} can be written as a linear combination of $\{\xi_{11}, \ldots, \xi_{1k_1}, \ldots, \xi_{p1}, \ldots, \xi_{pk_p}\}$.

Higher order time derivatives $\mathrm{d}y_i^{(j)}$ can be computed and, from the implicit function theorem, $\mathrm{d}x$ can be substituted to obtain

$$\mathrm{d}y_i^{(j)} = \left(\sum_{r=1}^{p} \sum_{s=1}^{k-r} \eta_{rs} \mathrm{d}y_r^{(s-1)} \right) + \gamma_{i,j+1}\mathrm{d}\theta (\bmod \mathcal{U}).$$

Then, the system is geometrically identifiable if and only if there are integers k_i^*, for $i = 1, \ldots, p$, such that

$$\operatorname{rank} \begin{bmatrix} \gamma_{11} \\ \vdots \\ \gamma_{1k_1^*} \\ \gamma_{21} \\ \vdots \\ \gamma_{pk_p^*} \end{bmatrix} = q. \qquad (13)$$

The system is algebraically identifiable if and only if there exist p integers l_i^*, for $i = 1, \ldots, p$, such that

$$\operatorname{rank} \begin{bmatrix} \gamma_{1(k_1+1)} \\ \vdots \\ \gamma_{1l_1^*} \\ \gamma_{2(k_2+1)} \\ \vdots \\ \gamma_{pl_p^*} \end{bmatrix} = q. \tag{14}$$

A trajectory such that (13) (or (14)) holds is called geometrically (or algebraically) persistently exciting.

4. Model Analysis

The success of any control system depends very much on the measurement subsystem. This is especially true in the modelling of HIV/AIDS dynamics. As a matter of fact, "the impetus for further modelling came with the development of rapid, sensitive and accurate methods of measuring the number of virus particles in the blood" (Perelson and Nelson, 1999). All variables in the models presented in this and Section 2 can essentially be measured, though some with less accuracy and high cost. However, variables that are routinely measured for deciding when to initiated therapy and for monitoring of patients on antiretroviral therapy are the viral load and the total $CD4^+$ T cell count T_{tot}, which is the sum of the uninfected and infected $CD4^+$ T cells. In settings where all variable measurements are obtainable, this will improve the identifiability properties of the system.

4.1. *Identifiability properties of the 3D model*

The identifiability properties of the basic 3D model (1) will be analyzed. Consider the viral load and uninfected $CD4^+$ T cell count as the measured outputs. Taking outputs as

$$y_1 = T, \quad y_2 = V; \quad y = (y_1, y_2),$$

compute

$$\dot{y}_1 = s - dy_1 - \beta y_1 y_2. \tag{15}$$

Thus, output y_1 has an observability index equal to 1 and three parameters that may be identified. Higher order derivatives yield

$$\ddot{y}_1 = -d\dot{y}_1 - \beta(y_1 y_2)^{(1)}, \tag{16}$$

$$y_1^{(3)} = -d\ddot{y}_1 - \beta(y_1 y_2)^{(2)}. \tag{17}$$

Now we have three equations (15), (16) and (17) with three unknown parameters. Parameters s, d and β can therefore, be computed from any persistently exciting trajectory $y(t)$ such that rank $\partial(\dot{y}_1, \ddot{y}_1, y_1^{(3)})/\partial(s, d, \beta) = 3$. That is, if

$$\operatorname{rank} \begin{bmatrix} 1 & -y_1 & -y_1 y_2 \\ 0 & -\dot{y}_1 & -(y_1 y_2)^{(1)} \\ 0 & -\ddot{y}_1 & -(y_1 y_2)^{(2)} \end{bmatrix} = 3.$$

The three equations (15), (16) and (17) can be solved to get a unique solution for s, d and β as functions of $y(t_o)$, $\dot{y}(t_o)$, $\ddot{y}(t_o)$ and $y_1^{(3)}(t_o)$ at some instant t_o. To cope with these order of derivatives, one can conclude that at least four measurements of the uninfected CD4$^+$ T cell count y_1 and at least three measurements of the viral load y_2, are needed for a complete first determination of these three parameters.

For the remaining 3 parameters, consider output y_2 and compute

$$\dot{y}_2 = pT^* - cy_2, \tag{18}$$

$$\ddot{y}_2 = \theta_1 \dot{y}_2 + \theta_2 c y_2 + \theta_3 y_1 y_2, \tag{19}$$

where

$$\theta_1 = -(\mu + c), \quad \theta_2 = -\mu c, \quad \theta_3 = p\beta.$$

Output y_2 has an observability index equal to 2 and higher order derivatives yield

$$y_2^{(3)} = \theta_1 \ddot{y}_2 + \theta_2 \dot{y}_2 + \theta_3 (y_1 y_2)^{(1)}, \tag{20}$$

$$y_2^{(4)} = \theta_1 y_2^{(3)} + \theta_2 \ddot{y}_2 + \theta_3 (y_1 y_2)^{(2)}. \tag{21}$$

Again, we have three equations (19), (20) and (21) and three unknown parameters. "Parameters" θ_1, θ_2 and θ_3 can be similarly computed from any persistently exciting trajectory $y(t)$ such that rank $\partial(\ddot{y}_2, y_2^{(3)}, y_2^{(4)})/\partial(\theta_1, \theta_2, \theta_3) = 3$. That is, if

$$\operatorname{rank} \begin{bmatrix} \dot{y}_2 & y_2 & y_1 y_2 \\ \ddot{y}_2 & \dot{y}_2 & (y_1 y_2)^{(1)} \\ y_2^{(3)} & \ddot{y}_2 & (y_1 y_2)^{(2)} \end{bmatrix} = 3.$$

Table 4 Minimum measurements for 3D model

Measured	Property	T_{tot}	T	T^*	V	Total
T, V	Algebraic	—	4	0	5	9
T_{tot}, V	Algebraic	4/5	—	—	5/4	9

— indicates that measurement is not applicable

The map $\Theta = (\theta_1, \theta_2, \theta_3) = (-(\mu + c), -\mu c, p\beta)$ is one-to-one, hence the identifiability of $\theta_1, \theta_2, \theta_3$ is equivalent to that of μ, c, p. The three equations (19), (20) and (21) can be solved to get a unique solution for μ, c and p as functions of $y(t_o)$, $\dot{y}(t_o)$, $\ddot{y}(t_o)$, $y_2^{(3)}(t_o)$ and $y_2^{(4)}(t_o)$ at some instant t_o, provided $\beta \neq 0$ and $c \neq \mu$. For the determination of these three parameters, at least three measurements of the uninfected CD4$^+$ T cell count y_1 and five measurements of the viral load y_2 are necessary. The system Σ_{3D} (1) is therefore algebraically identifiable and all six parameters can be computed from the measurements of the viral load and uninfected CD4$^+$ T cells. Overall, at least four measurements of the uninfected CD4$^+$ T cell count and five viral load measurements are required for a complete first estimate of all six parameters. Results are summarized in Table 4.

In current practice, discriminatory CD_4^+ T cell measurements are not readily attainable and the cell count that is actually measured is the total of the infected and uninfected CD4$^+$ T cells. Identifiability of this 3D model (1) has been analyzed with the viral load and the total CD4$^+$ T cell count as the measured outputs. Even though this results in the same total number of measurements, as indicated in Table 4, the eventual extraction of the parameters from the measured outputs will be more mathematically intensive.

It has been observed that for individuals in the asymptomatic stage of the infection and those on antiretroviral drugs, the infected CD4$^+$ T cell pool makes a very small percentage of the total CD4$^+$ T cell count (Embretson *et al.*, 1993; Janeway and Travers, 1997; Chun *et al.*, 1997). This means that in this case, the assumption that the uninfected CD4$^+$ T cell count is approximately equal to the total CD4$^+$ T cell count can be made. We have therefore opted to take the uninfected CD4$^+$ T cell count as the measured output instead of the total CD4$^+$ T cell count.

4.2. *When to take measurements*

If either one of the measured outputs y_1 or y_2 is constant, the higher order derivatives (16), (17), (20) and (21) in the previous section will be zero

and parameter extraction from the measured outputs will not be possible. When one therefore, considers the variation of the model variables with time as depicted in Fig. 1, then one can see that the measured outputs are constant for the asymptomatic stage. It has been observed that for HIV-infected individuals, the viral load remains relatively constant during this long asymptomatic stage, while the $CD4^+$ T cell count slowly declines. This means that measurements should be taken during the acute infection stage and during the advanced stage of the HIV infection.

For individuals in the asymptomatic stage of the infection, one then needs to use antiretroviral drugs to perturb the quasi-steady state. When measurements are taken during the acute infection stage of the infection, then the assumption that the measured total $CD4^+$ T cell count is representative of the uninfected cell population does not hold. This will necessitate for discriminatory $CD4^+$ T cell measurements to be the standard practice, unless antiretroviral agents are again used, but in this case to reduce the proportion of infected cells in the total cell count.

4.3. *Identifiability with the use of antiretroviral agents*

The two classes of commonly used anti-retroviral agents are Reverse Transcriptase Inhibitors (RTI) and Protease Inhibitors (PI). Both agents work within the $CD4^+$ T cell because they do not prevent the virus from entering the cell. Model parameters that are affected by antiretroviral agents have been identified (McLeod and Hammer, 1992; Perelson and Nelson, 1999). RTIs reduce successful infection by reducing the value of β. Perfect inhibition, therefore, occurs when $\beta = 0$. In practice however, perfect inhibition is not attainable. PIs on the other hand, inhibit the protease enzyme so that the virus particles that are produced are mostly noninfectious. There are therefore two types of virus particles when protease inhibitors are used. The first type are the infectious virus particles that still continue to infect $CD4^+$ T cells and the other is the noninfectious type. Similarly, perfect inhibition occurs when all virus particles that are produced are non-infectious. Current therapies use a combination of Reverse Transcriptase and Protease Inhibitors and the combined therapy model can be presented as

$$\sum_{3Da} = \begin{cases} \dot{T} &= s - dT - u_{RT}\beta T V_I, \\ \dot{T}^* &= u_{RT}\beta T V_I - \mu T^*, \\ \dot{V}_I &= u_{PI} p T^* - c V_I, \\ \dot{V}_N &= (1 - u_{PI}) p T^* - c V_N. \end{cases} \tag{22}$$

$u_{RT} = 1 - \eta_{RT}$ and $u_{PI} = 1 - \eta_{PI}$ are the respective control efforts for the reverse transcriptase and protease inhibitors. η_{RT}, $0 \leq \eta_{RT} < 1$ is the combined effectiveness of all the reverse transcriptase inhibitors used and η_{PI}, $0 \leq \eta_{PI} < 1$ is the combined effectiveness of all the protease inhibitors used. State variables V_I and V_N are the plasma concentrations of the infectious and noninfectious virus particles respectively. It is assumed that both types of virus particles have the same death rate constant c.

For parameter identifiability under the use of antiretroviral agents, the effect of the drugs will affect the identification of the affected model parameters. In fact, for the model Σ_{3Da}, the identifiable parameters will be $u_{RT}\beta$ instead of β and $u_{PI}p$ instead of p. This means that the drug control effort or efficacy cannot be separated from the parameter it affects. Another point worth noting is that, current assays do not differentiate between infectious and noninfectious virus particles. That is, $V_{tot} = V_I + V_N$ is the measured viral load.

4.4. *Identifiability properties of the 4D model*

In this section, consider the latently infected cell model Σ_{4D} in Eq. (3). Taking the uninfected $CD4^+$ T cell count and the viral load as the measured outputs, then again,

$$y_1 = T, \quad y_2 = V.$$

Compute

$$\dot{y}_1 = s - dy_1 - \beta y_1 y_2. \tag{23}$$

The identifiability of parameters s, d and β has been presented in Section 4.1.

For the other parameters $(q_1, q_2, k, \mu_1, \mu_2, p, c)$, compute

$$\dot{y}_2 = pT_2 - cy_2, \tag{24}$$

$$\ddot{y}_2 = pq_2\beta y_1 y_2 + k_1 pT_1 - (\mu_2 + c)\dot{y}_2 - \mu_2 c y_2, \tag{25}$$

$$y_2^{(3)} = \theta_1(y_1 y_2)^{(1)} + \theta_2\ddot{y}_2 + \theta_3\dot{y}_2 + \theta_4 y_1 y_2 + \theta_5 y_2, \tag{26}$$

where

$$\begin{aligned}
\theta_1 &= pq_2\beta, \\
\theta_2 &= -(\mu_1 + \mu_2 + c + k), \\
\theta_3 &= -(\mu_2 c + \mu_1\mu_2 + \mu_1 c + ck + k\mu_2), \\
\theta_4 &= kpq_1\beta + pq_2\beta(\mu_1 + k), \\
\theta_5 &= -c\mu_2(\mu_1 + k).
\end{aligned}$$

Thus output y_2 has an observability index equal to 3, and higher order derivatives of y_2 will just read as

$$y_2^{(i)} = \theta_1 (y_1 y_2)^{(i-2)} + \theta_2 y_2^{(i-1)} + \theta_3 y_2^{(i-2)} + \theta_4 (y_1 y_2)^{(i-3)} + \theta_5 y_2^{(i-3)}. \quad (27)$$

The five "parameters" $\theta_1, \ldots, \theta_5$ can be computed from any persistent exciting trajectory $y(t)$ such that rank $\partial(y_2^{(3)}, \ldots, y_2^{(7)})/\partial(\theta_1, \ldots, \theta_5) = 5$. That is, if

$$\operatorname{rank} \begin{bmatrix} (y_1y_2)^{(1)} & \ddot{y}_2 & \dot{y}_2 & y_1y_2 & y_2 \\ (y_1y_2)^{(2)} & y_2^{(3)} & \ddot{y}_2 & (y_1y_2)^{(1)} & \dot{y}_2 \\ (y_1y_2)^{(3)} & y_2^{(4)} & y_2^{(3)} & (y_1y_2)^{(2)} & \ddot{y}_2 \\ (y_1y_2)^{(4)} & y_2^{(5)} & y_2^{(4)} & (y_1y_2)^{(3)} & y_2^{(3)} \\ (y_1y_2)^{(5)} & y_2^{(6)} & y_2^{(5)} & (y_1y_2)^{(4)} & y_2^{(4)} \end{bmatrix} = 5.$$

Thus, the system Σ_{4D} is not algebraically identifiable with the outputs taken as the uninfected CD4$^+$ T cell count and the viral load. Besides the identification of s, d and β, five of the seven parameters can be computed in terms of the measurements and two remaining parameters, if

$$q_1 q_2 \beta p (\mu_1 + k - \mu_2)(\mu_2 - c)(\mu_1 + k - c) \neq 0. \quad (28)$$

Some interesting conclusions can be drawn from (28):

(i) If $q_1 = 0$, then μ_1 and k cannot be separated from $\mu_1 + k$. Medically, this would mean that if the channel is cut off for infected cells to become latent, then one will not be able to tell from the viral load and the CD4$^+$ T cell count, how many latently infected cells die naturally and how many of them turn to actively infected cells.

(ii) If $p = 0$, then the actively infected CD4$^+$ T cells do not produce any virus. No information about these infected cells can therefore be extracted from measurements of the viral load.

(iii) If $\mu_1 + k = \mu_2$, then the latently infected cells and the actively infected cells disappear from their corresponding pools at the same rate. The dynamics of these two cell types will therefore be indistinguishable.

(iv) If $\mu_2 = c$, then the actively infected cells and the virus disappear at the same rate from their respective pools, hence it will not be possible to distinguish their dynamics.

(v) If $\mu_1 + k = c$, then similarly, the latently infected cells and the virus dynamics will be indistinguishable if they are cleared at the same rate from their respective pools.

Table 5 Identifiability of system Σ_{4D} with some known parameters.

Known parameters	Identifiability of remaining parameters
q_1, q_2	identifiable
q_1, μ_1	identifiable
q_1, k	identifiable
q_1, μ_2	not identifiable
q_1, p	identifiable
q_1, c	not identifiable
q_2, μ_1	identifiable
q_2, k	identifiable
q_2, μ_2	not identifiable
q_2, p	identifiable
q_2, c	not identifiable
μ_1, k	not identifiable
μ_1, μ_2	not identifiable
μ_1, p	identifiable
μ_1, c	not identifiable
k, μ_2	not identifiable
k, p	identifiable
k, c	not identifiable
μ_2, p	not identifiable
μ_2, c	not identifiable

It is noted that some of the parameters can be determined by other methods. For example, through the experiment of Chun *et al.* (1997), q_1 was found to be $q_1 = 0.01$. Janeway and Travers (1997) found that $q_2 = 0.02$. An exhaustive list of all cases where two of the model parameters are known is presented in Table 5.

The identifiability property of system Σ_{4D} (3) can be improved if the initial values of some other variables are known. In identifiability theory, this property is characterized by the so-called geometric identifiability, or equivalently, identifiability with known initial conditions.

To inspect the geometric identifiability of the seven remaining parameters $\{q_1, q_2, k, \mu_1, \mu_2, p, c\}$, introduce the notation $\Theta_1 = pq_2\beta$ and $\Theta_2 = kpq_1\beta$. Consider the new list of parameters $\{\Theta_1, \Theta_2, \mu_1, \mu_2, k, p, c\}$ and compute $\ddot{y}_2$, $y_2^{(3)}$ as

$$\ddot{y}_2 = \Theta_1 y_1 y_2 + kpT_1 - p\mu_2 T_2 - c\dot{y}_2, \tag{29}$$

$$\begin{aligned} y_2^{(3)} &= \Theta_1 (y_1 y_2)^{(1)} + \Theta_2 y_1 y_2 - \mu_2(\ddot{y}_2 + c\dot{y}_2) - c\ddot{y}_2 \\ &\quad - (\mu_1 + k)[\ddot{y}_2 - \Theta_1 y_1 y_2 + \mu_2(\dot{y}_2 + cy_2) - c\dot{y}_2]. \end{aligned} \tag{30}$$

Higher order derivatives are obtained by differentiating (30).

Introduce the notation

$$
\begin{aligned}
A &= y_1 y_2, \\
B &= (y_1 y_2)^{(1)} + (\mu_1 + k)(y_1 y_2), \\
C &= -[\ddot{y}_2 - \Theta_1 y_1 y_2 + \mu_2(\dot{y}_2 + c y_2) - c\dot{y}_2], \\
D &= -[(\mu_1 + k)(\dot{y}_2 + c y_2) + \ddot{y}_2 + c\dot{y}_2], \\
E &= -[y_2^{(2)} - \Theta_1 y_1 y_2 + \mu_2(\dot{y}_2 + c y_2) + c\dot{y}_2], \\
F &= -[(\mu_1 + k)(\mu_2 y_2 + \dot{y}_2) + \mu_2 \dot{y}_2 + \ddot{y}_2],
\end{aligned}
$$

and compute Γ_g as
$\partial(\dot{y}_2, \ddot{y}_2, y_2^{(3)}, y_2^{(4)}, y_2^{(5)}, y_2^{(6)}, y_2^{(7)})/\partial(\Theta_1, \Theta_2, \mu_1, \mu_2, k, p, c)$.
That is,

$$
\Gamma_g = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & T_2 & -y_2 \\
A & 0 & 0 & -pT_2 & pT_1 & [kT_1 - \mu_2 T_2] & -\dot{y}_2 \\
B & A & C & D & E & 0 & F \\
\dot{B} & \dot{A} & \dot{C} & \dot{D} & \dot{E} & 0 & \dot{F} \\
\ddot{B} & \ddot{A} & \ddot{C} & \ddot{D} & \ddot{E} & 0 & \ddot{F} \\
B^{(3)} & A^{(3)} & C^{(3)} & D^{(3)} & E^{(3)} & 0 & F^{(3)} \\
B^{(4)} & A^{(4)} & C^{(4)} & D^{(4)} & E^{(4)} & 0 & F^{(4)}
\end{bmatrix}.
$$

Then rank $\Gamma_g = 7$ if (28) holds. That is if,

$$
q_1 q_2 \beta p(\mu_1 + k - \mu_2)(\mu_2 - c)(\mu_1 + k - c) \neq 0.
$$

Thus, the system is geometrically identifiable and all the original parameters of Σ_{4D} are identifiable from the measurements of the viral load and uninfected CD4$^+$ T cell count if the initial values for both the actively and latently infected CD4$^+$ T cells for the individual are known. One therefore, needs to have a *comprehensive* test before measurements are taken. With the advancement of faster and cheaper measuring devices, it is envisaged that discriminatory CD4$^+$ T cell count measurements will soon be routine and available cost-effectively. The required minimum number of measurements for a complete first determination of all ten parameters are as summarized in Table 6.

Table 6 Minimum measurements for 4D model

Measured	Property	T	T_1	T_2	V	Total
T, V	Geometric	6	1	1	8	16

4.5. *Identifiability properties of the 6D model*

An analysis of the 6D model Σ_{6D} in (4) will be presented for when different combinations of model variables are the measured outputs.

4.5.1. *Identifiability with viral load and uninfected CD4$^+$ T cell count measurements*

In the first instance, take outputs as the uninfected CD4$^+$ T cell count, T and the viral load V

$$y_1 = T, \quad y_2 = V.$$

Output y_1 has observability index equal to 1, while that of y_2 can be shown to be equal to 5. The resulting equations for outputs y_1 and y_2 are as given in (31) and (32), respectively.

$$\dot{y}_1 = s_T - d_T y_1 - \beta_T y_1 y_2, \tag{31}$$

$$\begin{aligned} y_2^{(5)} &= p_T q_2 \beta_T (y_1 y_2)^{(3)} + p_T \beta_T (\theta_2 + \delta q_2)(y_1 y_2)^{(2)} \\ &\quad + p_T \beta_T \delta \theta_2 (y_1 y_2)^{(1)} - (\theta_1 + c + \mu_2) y_2^{(4)} - (\theta_4 + (k+\mu_1)\theta_3) y_2^{(3)} \\ &\quad + (\theta_5 + s_M - (k+\mu_1)\theta_4)\ddot{y}_2 + (k+\mu_1)\theta_5 \dot{y}_2 \\ &\quad + s_M(\Sigma + (\psi_1 - \beta_M y_2)\ddot{y}_2) + \frac{\dot{\Sigma}}{\Sigma} - (d_M - \beta_M y_2)\Lambda, \end{aligned} \tag{32}$$

where,

$$\begin{aligned} \theta_1 &= k + \mu_1 + \delta, \\ \theta_2 &= kq_1 + (k+\mu_1)q_2, \\ \theta_3 &= \delta + c + \mu_2, \\ \theta_4 &= \delta c + \mu_2(\delta + c), \\ \theta_5 &= p_M \beta_M s_M - \delta c \mu_2, \\ \psi_1 &= \mu_2 - d_M - \beta_M y_2, \\ \Sigma &= \ddot{y}_2 + (\psi_1 - \beta_M y_2)\dot{y}_2, \\ \Lambda &= y_2^{(4)} - p_T q_2 \beta_T (y_1 y_2)^{(2)} - p_T \beta_T (\theta_2 + \delta q_2)(y_1 y_2)^{(1)} \\ &\quad - p_T \beta_T \delta \theta_2 y_1 y_2 + (k + \mu_1 + \theta_3) y_2^{(3)} + (\theta_4 + \theta_3(k+\mu_1))\ddot{y}_2 \\ &\quad - (\theta_5 + s_M - \theta_4(k+\mu_1))\dot{y}_2 - \theta_5(k+\mu_1)y_2 - \psi_1 s_M y_2. \end{aligned}$$

However, the system is not algebraically identifiable. Besides the identification of s_T, d_T and β_T, only eight of the remaining twelve parameters can be computed in terms of the measured outputs and the other four parameters. The system can be shown to be geometrically identifiable, or identifiable with known initial conditions. One therefore, needs a comprehensive test to obtain initial measurements for both the actively and latently infected

$CD4^+$ T cells, as well as for the infected and uninfected macrophages. Even though this system is geometrically identifiable with measured outputs taken as the viral load and uninfected $CD4^+$ T cells, the required minimum number of measurements of these outputs is too high as outlined in Table 7. Attempts to obtain estimates of all 15 parameters of this model, with the viral load and uninfected $CD4^+$ T cell counts as the measured outputs will therefore, not be a practical approach when cost and patient discomfort are taken into consideration. One therefore needs to consider measuring something else or increasing the number of measured outputs.

4.5.2. *Identifiability with viral load and uninfected $CD4^+$ T cell count and macrophage measurements*

The identifiability property of system Σ_{6D} (4) can be improved by also measuring the uninfected macrophages. Then, taking the outputs as

$$y_1 = T, \quad y_2 = V, \quad y_3 = M, \quad \text{and} \quad \text{let } x_1 = y_1 y_2, \quad x_2 = y_2 y_3,$$

and compute

$$\dot{y}_1 = s_T - d_T y_1 - \beta_T x_1. \tag{33}$$

Then, s_T, d_T and β_T can therefore be computed from any persistently exciting trajectory $y(t)$ such that rank $\partial(\dot{y}_1, \ddot{y}_1, y_1^{(3)})/\partial(s_T, d_T, \beta_T) = 3$. That is, if

$$\operatorname{rank} \begin{bmatrix} 1 & -y_1 & -x_1 \\ 0 & -\dot{y}_1 & -\dot{x}_1 \\ 0 & -\ddot{y}_1 & -\ddot{x}_1 \end{bmatrix} = 3.$$

For output y_3, compute

$$\dot{y}_3 = s_M - d_M y_3 - \beta_T x_2. \tag{34}$$

Similarly, s_M, d_M and β_M can be computed from any persistently exciting trajectory $y(t)$ such that rank $\partial(\dot{y}_3, \ddot{y}_3, y_3^{(3)})/\partial(s_M, d_M, \beta_M) = 3$. That is,

if

$$\operatorname{rank}\begin{bmatrix} 1 & -y_3 & -x_2 \\ 0 & -\dot{y}_3 & -\dot{x}_2 \\ 0 & -\ddot{y}_3 & -\ddot{x}_2 \end{bmatrix} = 3.$$

For the remaining nine parameters $(\mu_1, \mu_2, q_1, q_2, k, p_T, p_M, \delta, c)$, define

$$\begin{aligned}
\theta_1 &= k + \mu_1 + \delta, \\
\theta_2 &= kq_1 + (k + \mu_1)q_2, \\
\theta_3 &= \delta + c + \mu_2, \\
\theta_4 &= \delta c + \mu_2(\delta + c), \\
\theta_5 &= \delta c(\delta + c)(k + \mu_1), \\
\theta_6 &= c\mu_2(c + \mu_2)(k + \mu_1), \\
\theta_7 &= \mu_2\delta(\mu_2 + \delta)(k + \mu_1).
\end{aligned}$$

Then compute

$$\dot{y}_2 = p_T T_2 + p_M M^* - c y_2, \tag{35}$$

$$\begin{aligned}
\ddot{y}_2 = {} & q_2 p_T \beta_T x_1 + p_M \beta_M x_2 + p_T k T_1 \\
& - (\delta + c)\dot{y}_2 - \delta c y_2 - p_T(\mu_2 - \delta)T_2,
\end{aligned} \tag{36}$$

$$\begin{aligned}
y_2^{(3)} = {} & q_2 p_T \beta_T \dot{x}_1 + p_M \beta_M \dot{x}_2 + p_T \beta_T (k q_1 + \delta q_2) x_1 \\
& + p_M \beta_M \mu_2 x_2 - \theta_3 \ddot{y}_2 - \theta_4 \dot{y}_2 - \mu_2 \delta c y_2 \\
& - p_T k (k + \mu_1 - \delta) T_1,
\end{aligned} \tag{37}$$

$$\begin{aligned}
y_2^{(4)} = {} & q_2 p_T \beta_T \ddot{x}_1 + p_M \beta_M \ddot{x}_2 + p_T \beta_T (\theta_2 + \delta q_2)\dot{x}_1 \\
& + p_M \beta_M (k + \mu_1 + \mu_2)\dot{x}_2 + p_T \beta_T \delta \theta_2 x_1 \\
& + p_M \beta_M \mu_2 (k + \mu_1) x_2 - (\theta_3 + k + \mu_1) y_2^{(3)} \\
& - (\theta_4 + \theta_3(k + \mu_1))\ddot{y}_2 - (\mu_2 \delta c + \theta_4(k + \mu_1))\dot{y}_2 \\
& - (k + \mu_1)\mu_2 \delta c y_2.
\end{aligned} \tag{38}$$

The remaining nine parameters can therefore be computed from any persistently exciting trajectory $y(t)$ such that $\operatorname{rank} \partial(y_2^{(4)}, \ldots, y_2^{(12)})/\partial(\mu_1, \mu_2, q_1, q_2, k, p_T, p_M, \delta, c) = 9$. That is, if

$$\operatorname{rank}\begin{bmatrix} \psi_1 & \psi_2 & \psi_3 & \cdots & \psi_8 & \psi_9 \\ \dot{\psi}_1 & \dot{\psi}_2 & \dot{\psi}_3 & \cdots & \dot{\psi}_8 & \dot{\psi}_9 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ \psi_1^{(8)} & \psi_2^{(8)} & \psi_3^{(8)} & \cdots & \psi_8^{(8)} & \psi_9^{(8)} \end{bmatrix} = 9,$$

where,

$$\begin{aligned}
\psi_1 &= p_T\beta_T q_2(\dot{x}_1 + \delta x_1) + p_M\beta_M(\dot{x}_2 + \mu_2 x_2) \\
&\quad - y_2^{(3)} - \theta_3\ddot{y}_2 - \theta_4\dot{y}_2 - \mu_2\delta c y_2, \\
\psi_2 &= p_M\beta_M(\dot{x}_2 + (k+\mu_1)x_2) - y_2^{(3)} \\
&\quad - (\theta_4 + c)\ddot{y}_2 - \theta_5\dot{y}_2 - (k+\mu_1)\delta c y_2, \\
\psi_3 &= p_T\beta_T k(\dot{x}_1 + \delta x_1), \\
\psi_4 &= p_T\beta_T(\ddot{x}_1 + \theta_1\dot{x}_1 + \delta(k+\mu_1)x_1), \\
\psi_5 &= p_T\beta_T(q_1+q_2)(\dot{x}_1 + \delta x_1) + p_M\beta_M(\dot{x}_2 + \mu_2 x_2) \\
&\quad - y_2^{(3)} - \theta_3\ddot{y}_2 - \theta_4\dot{y}_2 - \mu_2\delta c y_2, \\
\psi_6 &= \beta_T(q_2\ddot{x}_1 + (\theta_3 + \delta q_2)\dot{x}_1 + \delta\theta_3 x_1), \\
\psi_7 &= \beta_M(\ddot{x}_2 + \mu_2(k+\mu_1)x_2), \\
\psi_8 &= p_T\beta_T(q_2\dot{x}_1 + \theta_3 x_1) - y_2^{(3)} \\
&\quad - (k + \mu_1 + \delta + \mu_2)\ddot{y}_2 - \theta_6\ddot{y}_2 - (k+\mu_1)\mu_2 c y_2, \\
\psi_9 &= -y_2^{(3)} - (\mu_2 + \theta_1)\ddot{y}_2 - \theta_7\dot{y}_2 - (k+\mu_1)\mu_2\delta y_2.
\end{aligned}$$

System Σ_{6D} can be shown to be algebraically identifiable, and all fifteen parameters can be computed from measurements of the uninfected CD4$^+$ T cells, the viral load and uninfected macrophages, if

$$p_T\beta_T\beta_M k q_1 q_2(k - q_1)(k + \mu_1 - \mu_2) \neq 0,$$

$$\mu_2 \neq c \quad \text{and} \quad c \neq \delta.$$

Again, bearing in mind that the infected cells are a small portion of the total cell count, then in most instances where it is not possible to obtain discriminatory macrophage and CD4$^+$ T cell count measurements, one can take the total cell count as representative of the uninfected cells.

4.5.3. *Identifiability with viral load and discriminatory CD4$^+$ T cell count measurements*

Considering that macrophage measurements are currently more difficult to obtained compared to discriminatory CD4$^+$ T cell count measurements, one option would be to measure the actively infected CD4$^+$ T cells instead of the uninfected macrophages. Setting

$$y_1 = T, \quad y_2 = V, \quad y_4 = T_2, \quad \text{and} \quad x_1 = y_1 y_2,$$

then outputs y_1, y_2 and y_4 have observability indices $r_1 = 1$, $r_2 = 3$ and $r_4 = 2$, respectively. Identifiability of parameters s_T, d_T and β_T is as

presented in Section 4.5.2. For identifiability of the remaining twelve parameters, define

$$\begin{aligned}\triangle_1 &= \frac{\dot{y}_2}{y_2} - d_M - \beta_M y_2,\\ \triangle_2 &= \ddot{y}_2 - p_T\ddot{y}_4 + (\delta + c)\dot{y}_2 - p_T\delta y_4 + \delta c y_2,\end{aligned}$$

and compute

$$\dot{y}_2 = p_T y_4 + p_M M^* - c y_2, \tag{39}$$

$$\ddot{y}_2 = p_T\ddot{y}_4 + p_T\delta y_4 - (\delta + c)\dot{y}_2 - \delta c y_2 + p_M\beta_M y_2 M, \tag{40}$$

$$\begin{aligned}y_2^{(3)} &= p_T\dddot{y}_4 - (\delta + c + d_M)\ddot{y}_2 + p_T(\delta + d_M)\ddot{y}_4\\ &\quad - (\delta c + d_M(\delta + c))\dot{y}_2 + p_T d_M\delta y_4\\ &\quad - (d_M\delta c - p_M\beta_M s_M)y_2 + \left(\frac{\dot{y}_2}{y_2} - \beta_M y_2\right)\triangle_2.\end{aligned} \tag{41}$$

Then the seven parameters $(p_T, p_M, \delta, c, s_M, d_M, \beta_M)$ in $y_2^{(3)}$ can be computed from any persistently exciting trajectory $y(t)$ such that rank $\partial(y_2^{(3)}, \ldots, y_2^{(9)})/\partial(p_T, p_M, \delta, c, s_M, d_M, \beta_M) = 7$. That is, if

$$\operatorname{rank}\begin{bmatrix} \upsilon_1 & s_M\upsilon_6 & \upsilon_2 & \upsilon_3 & p_M\upsilon_6 & \upsilon_4 & \upsilon_5\\ \dot{\upsilon}_1 & s_M\dot{\upsilon}_6 & \dot{\upsilon}_2 & \dot{\upsilon}_3 & p_M\dot{\upsilon}_6 & \dot{\upsilon}_4 & \dot{\upsilon}_5\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ \upsilon_1^{(6)} & s_M\upsilon_6^{(6)} & \upsilon_2^{(6)} & \upsilon_3^{(6)} & p_M\upsilon_6^{(6)} & \upsilon_4^{(6)} & \upsilon_5^{(6)}\end{bmatrix} = 7,$$

where,

$$\begin{aligned}\upsilon_1 &= \dddot{y}_4 + (\delta - \triangle_1)\ddot{y}_4 - \delta\triangle_1 y_4,\\ \upsilon_2 &= -\ddot{y}_2 + p_T\dddot{y}_4 - (c - \triangle_1)\dot{y}_2 - p_T\triangle_1 y_4 + \triangle_1 c y_2,\\ \upsilon_3 &= -\ddot{y}_2 - (\delta - \triangle_1)\dot{y}_2 + \triangle_1\delta y_2,\\ \upsilon_4 &= -\ddot{y}_2 + p_T\dddot{y}_4 - (\delta + c)\dot{y}_2 + p_T\delta y_4 - \delta c y_2,\\ \upsilon_5 &= (p_M s_M + \triangle_2)y_2,\\ \upsilon_6 &= \beta_M y_2.\end{aligned}$$

However, the above matrix only has rank = 6, and therefore not all the seven parameters can be estimated. It can be shown that one needs prior knowledge of either s_M or p_M in order to determine the other five parameters.

For the still remaining parameters $(\mu_1, \mu_2, q_1, q_2, k)$, compute

$$\dot{y}_4 = q_2\beta_T x_1 + kT_1 - \mu_2 y_4, \tag{42}$$

$$\ddot{y}_4 = q_2\beta_T \dot{x}_1 + \beta_T(kq_1 + (k+\mu_1)q_2)x_1 - (k+\mu_1+\mu_2)\dot{y}_4 - \mu_2(k+\mu_1)y_4. \tag{43}$$

The five parameters can be computed from any persistently exciting trajectory $y(t)$ such that rank $\partial(\ddot{y}_4, \ldots, y_4^{(6)})/\partial(\mu_1, \mu_2, k, q_1, q_2) = 5$. That is, if

$$\text{rank}\begin{bmatrix} \phi_1 & \phi_2 & \phi_1 + \beta_T q_1 x_1 & \beta_T k x_1 & \phi_3 \\ \dot{\phi}_1 & \dot{\phi}_2 & \dot{\phi}_1 + \beta_T q_1 \dot{x}_1 & \beta_T k \dot{x}_1 & \dot{\phi}_3 \\ \ddot{\phi}_1 & \ddot{\phi}_2 & \ddot{\phi}_1 + \beta_T q_1 \ddot{x}_1 & \beta_T k \ddot{x}_1 & \ddot{\phi}_3 \\ \phi_1^{(3)} & \phi_2^{(3)} & \phi_1^{(3)} + \beta_T q_1 x_1^{(3)} & \beta_T k x_1^{(3)} & \phi_3^{(3)} \\ \phi_1^{(4)} & \phi_2^{(4)} & \phi_1^{(4)} + \beta_T q_1 x_1^{(4)} & \beta_T k x_1^{(4)} & \phi_3^{(4)} \end{bmatrix} = 5,$$

where,

$$\phi_1 = -\dot{y}_4 + \beta_T q_2 x_1 - \mu_2 y_4,$$

$$\phi_1 = -\dot{y}_4 - (k+\mu_1)y_4,$$

$$\phi_3 = \beta_T(\dot{x}_1 + (k+\mu_1)x_1.$$

However, the above matrix has rank < 5, and therefore not all five remaining parameters can be estimated. One needs prior knowledge of either μ_1, q_1 or k in order to determine the other four parameters. The system is therefore not algebraically identifiable if the viral load, uninfected and actively infected CD4$^+$ T cell counts are the measured outputs.

To test for geometric identifiability of the remaining parameters, use (40) and (42) to generate a 7th equation for the y_2 dynamics and a 5th and 6th equation for the y_4 dynamics.

The system is geometrically identifiable if the viral load, the uninfected and actively infected CD4$^+$ T cells counts are the measured outputs. The initial measurements for the latently infected CD4$^+$ T cells, T_1 and uninfected macrophages, M will also be required.

For the output options considered in this section, measuring the actively infected CD4$^+$ T cells instead of the uninfected macrophage cells significantly reduces the number of required measurements, even though the system is no longer algebraically identifiable. This illustrates that improving the identifiability property of a system does not necessarily imply a reduction in the required number of measurements. More importantly, the point that is being illustrated here is that careful consideration of what needs to

Table 7 Minimum measurements for 6D model.

Measured	Property	T	T_1	T_2	M	M^*	V	Total
T, V	Geometric	15	1	1	1	1	17	36
T, M, V	Algebraic	11	0	0	11	0	13	35
T, T_2, V	Geometric	5	1	7	1	1	8	23

be measured is necessary. Table 7 summarizes the results for when various model variables are the measured outputs.

5. Conclusions

In this paper, a nonlinear system identifiability theory has been applied to analyze the identifiability properties of some HIV/AIDS models. The intention was to investigate the possibility of simultaneously estimating all the model parameters from measured system outputs, such as the viral load and $CD4^+$ T cell count. Other issues addressed include the minimum number of measurements for a complete first approximation of all parameters, the timing and conditions under which such measurements can be taken.

This information will be useful for the eventual parameter estimation, as well as for formulating guidelines for clinical practice.

References

Alvarez–Ramirez J, Meraz M and Velasco–Hermandez JX (2000) Feedback control of the chemotherapy of HIV, *Int J Bifurcat Chaos* **10**: 2207–2219.

Becker SL and Hoetelmans RM (2002) Exploiting pharmacokinetics to optimize antiretroviral therapy, http://www.medscape.com/viewprogram/703

Beerenwinkel N, Lengauer T, Selbig J, Schmidt B *et al.* (2001) Geno2pheno: interpreting genotypic HIV drug resistance tests, *IEEE Intell Syst Bio* Nov/Dec: 35–41.

Brandt ME and Chen G (2001) Feedback control of a biodynamical model of HIV-1, *IEEE Trans BME* **48**: 754–759.

Chun TW, Carruth L, Finzi D, Xhen X *et al.* (1997) Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection, *Nature* **387**: 183–188.

Chun TW, Engel D, Berrey MM, Shea T, Corey L, Fauci AS (1998) Early establishment of a pool of latently infected, resting $CD4^+$ T cells during primary HIV-1 infection, *Proc Natl Acad Sci USA* **95**: 8869–8873.

Conte G, Moog CH and Perdon AM (1999) *Nonlinear Control Systems: An Algebraic Setting*, Springer Verlag, London.

Culshaw RV and Ruan S (2000) A delay-differential equation model of HIV infection of $CD4^+$ T-cells, *Math Biosci* **165**: 27–39.

de Souza FMC (1999) Modeling the dynamics of HIV-1 and CD4 and CD8 lymphocytes, *IEEE Eng Med Bio* Jan/Feb: 21–24.

de Souza JAMF, Caetano MAL and Yoneyama T (2000) Optimal control theory applied to the anti-viral treatment of AIDS, in *Proc IEEE Conference on Decision and Control*, pp. 4839–4844, Sydney, Australia.

Denis-Vidal L, Joly-Blanchard G and Noiret C (2001) Some effective approaches to check the identifiability of uncontrolled nonlinear systems, *Math Comput Simulat* **57**: 35–44.

Dewhurst S, da Cruz RLW and Whetter L (2000) Pathogenesis and treatment of HIV-1 infection: recent developments (Y2K update), *Front Biosci* **5**: 30–49.

Diop S and Fliess M (1991) Nonlinear observability, identifiability and persistent trajectories, *Proc 30th CDC*, Brighton, UK: 714–719.

Embretson J, Zupancic M, Beneke J, Till M, Wolinski S, Ribas JL *et al.* (1993) Analysis of HIV infected tissues by amplification and *in situ* hybridization reveals latent and permissive infections at single-cell resolution, *Proc Acad Sci* **90**: 357–361.

Filter R. and Xia X (2003) A penalty function approach to HIV/AIDS model parameter estimation, 13th *IFAC Symp Syst Ident* 27–29 Aug. Rotterdam, Netherlands.

Glad ST (1997) Solvability of differential algebraic equations and inequalities: an algorithm, *Proc 4th ECC*, Bruxelles.

Grant SA and Xu J (2002) Investigation of an optical dual receptor method to detect HIV, *IEEE Sensors* **2**(5): 409–415.

Grossman Z, Polis M, Feinberg MB, Levi I, Jankelevich S, Yarchoan R *et al.* (1999) Ongoing dissemination during HAART, *Nat Med* **5**: 1099–1104.

Gumel AB, Shivakumar PN and Sahai BM (2001) A mathematical model for the dynamics of HIV-1 during the typical course of infection, *Nonlin Anal* **47**: 1773–1783.

Gumel AB, Twizell EH and Yu P (2000) Numerical and bifurcation analyses for a population model of HIV chemotherapy, *Math Comp Sim* **54**: 169–181.

Haase AT, Henry K, Zupancic M, Sedgewick G, Faust R *et al.* (1996) Quantitative image analysis of HIV-1 infection in lymphoid tissue, *Science* **274**: 985–989.

Ho DD, Newman AU, Perelson AS *et al.* (1995) Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection, *Nature* **273**: 123–126.

Iraba T, Dolezal J and Celikovsky S (1990) Model based analysis of $CD4^+$ lymphocyte dynamics in HIV infected individuals, *Immunbiol* **181**: 108–118.

Iraba T and Dolezal J (1996) A mathematical model and $CD4^+$ lymphocyte dynamics in HIV infection, *Emerg Infect Dis* **12**(4): 299–305.

Janeway CA and Travers P (1997) *Immunology: The Immune System in Health and Disease*, Garland, New York.

Jeffrey AM and Xia X (2002) Estimating the viral load response time after HIV chemotherapy, *IEEE Africon*, pp. 77–80, George, South Africa, 2–4 October.

Jeffrey AM, Xia X and Craig IK (2003a) When to initiate HIV therapy: a control theoretic approach, *IEEE Trans BME* **50**(11): 1213–1220.

Jeffrey AM, Xia X and Craig IK (2003b) On attaining maximal and durable suppresion of the viral load, *1st African Control Conference AFCON*, Cape Town, South Africa, 3–5 Dec.

Kirschner D (1996) Using mathematics to understand HIV immune dynamics, *Notices AMS* **43**(2): 191–202.

Kirschner D, Lenhart S and Serbin S (1997) Optimal control of the chemotherapy of HIV, *J Math Biol* **35**: 775–792

Kirschner D and Webb GF (1996) A model for treatment strategy in the chemotherapy of AIDS, *Bull Math Biol* **58**: 367–390.

Kramer I (1999) Modeling the dynamical impact of HIV on the immune system: viral clearance, infection and AIDS, *Math Comp Mod* **29**: 95–112.

Krener AJ and Respondek W (1985) Nonlinear observers with linearizable error dynamics, *SIAM J Contr Optimiz* **23**: 197–216.

Lindpainter K (2001) Pharmacogenetics and the future of medical practice: conceptual considerations, *Pharmacogen J* **1**: 23–26.

Little SJ, McLean AR, Spina CA, Richman DD and Hvalir DV (1999) Viral dynamics of acute HIV-1 infection, *J Exp Med* **190**: 841–850.

Ljung L and Glad ST (1994) On global identifiability for arbitrary model parameterizations, *Automatica* **30**: 265–276.

McLeod GX and Hammer SM (1992) Zidovudine: 5 years later, *Ann Intern Med* **117**: 487–501.

Muller V, Maree AFM and De Boer RJ (2001a) Small variations in multiple parameters account for large variations in HIV-1 set-points: a novel modelling approach, *Proc R Soc Lond B Biol Sci* **268**: 235–242.

Muller V, Maree AFM and De Boer RJ (2001b) Release of virus from lymphoid tissue affects human immunodeficiency virus type 1 and hepatitis C virus kinetics in the blood, *J Virol* **75**: 2597–2603.

Murray JM, Kaufmann G, Kelleher AD and Cooper DA (1998) A model of primary HIV-1 infection, *Math Biosci* **154**: 57–85.

Nowak MA and May RM (2000) *Virus Dynamics: Mathematical Principles of Immunology and Virology*, Oxford University Press, New York.

Perelson AS (1989) Modelling the interaction of the immune system with HIV, in Castillo–Chavez C (ed), *Mathematical and Statistical Approaches to AIDS Epidemiology*, Lecture Notes in Biomathematics, Vol 83, pp 350–370, Springer Verlag, New York.

Perelson AS, Kirschner D and De Boer R (1993) Dynamics of HIV infection of $CD4^+$ T cells, *Math Biosci* **114**: 81–125.

Perelson AS and Nelson PW (1999) Mathematical analysis of HIV-1 dynamics *in vivo*, *SIAM Rev* **41**: 3–44.

Perelson AS, Neumann AW, Markowitz M *et al.* (1996) HIV-1 dynamics *in vivo*: virion clearance rate, infected cell life-span and viral generation time, *Science* **271**: 1582–1586.

Perelson AS, Essunger P, Markowitz M and Ho DD (1996) How long should treatment be given if we had an antiretroviral regimen that completely blocked HIV replication? *XIth Intl Conf AIDS Abstracts*.

Perelson AS, Essunger P, Cao Y, Vesanen M, Hurley A, Saksela K, Markowitz M and Ho DD (1997) Decay characteristics of HIV-1 infected compartments during combination therapy, *Nature* **387**: 188–191.

Pirmohamed M and Back DJ (2001) The pharmacogenetics of HIV therapy, *Pharmacogen J Nature* **1**: 243–253.

Ramratnam B, Bonhoeffer S, Binley J, Hurley A *et al.* (1999) Rapid production and clearance of HIV-1 and hepatitis C virus assessed by large volume plasma apheresis, *Lancet* **354**: 1782–1785.

Roden DM and George AL Jr. (2002) The genetic basis of variability in drug responses, *Nat Rev Drug Disc* **1**: 37–44.

Tan W-Y and Wu H (1998) Stochastic modeling of the dynamics of $CD4^+$ T cell infection by HIV and some Monte Carlo studies, *Math Biosci* **147**: 173–205.

Tunali ET and Tarn TJ (1987) New results for identifiability of nonlinear systems, *IEEE Trans Auto Contr* **32**: 146–154.

Wein LM, Zenios SA and Nowak MA (1997) Dynamic multidrug therapies for HIV: a control theoretic approach, *J Theor Biol* **185**: 15–29.

Xia X (2003) Estimation of HIV/AIDS parameters, *Automatica* **39**: 1983–1988.

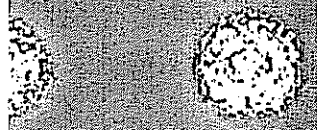
Xia X and Moog CH (2003) Identifiability of nonlinear systems with application to HIV/AIDS models, *IEEE Tran Auto Contr* **48**: 330–336.

With contributions from an international team of leading researchers, the book pulls together updated research results in the area of HIV/AIDS modeling to provide readers with the latest information in the field. Topics covered include: AIDS epidemic models; vaccine models; models for HIV/cell dynamics and interactions; cellular kinetics; viral dynamics with antiviral treatments; modeling of drug resistance and quasispecies.

Extensive deterministic models, statistical models, stochastic models and state space models on treating AIDS patients with anti-retroviral drugs are provided, as well as an in-depth discussion of these models. The book also contains updated reviews on mathematical models for assessing effects of AIDS vaccines, statistical methods for analyzing clinical trial data on AIDS vaccines, and overviews of models and statistical methods for assessing drug resistance of HIV to anti-retroviral drugs. Some important statistical methods specific to the intervention and prevention of HIV epidemic are also discussed.

This will be a useful reference source for graduate students and researchers in biomathematics and biostatistics, as well as for HIV/AIDS epidemiologists and clinical investigators learning quantitative methods to study AIDS epidemics and HIV infection.

Deterministic and Stochastic Models of AIDS Epidemics and HIV Infections with Intervention



World Scientific

ISBN 981-256-139-0

With contributions from an international team of leading researchers, the book pulls together updated research results in the area of HIV/AIDS modeling to provide readers with the latest information in the field. Topics covered include AIDS epidemic models, vaccine models, models for HIV/cell dynamics and interactions, cellular kinetics, viral dynamics with antiviral treatments, modeling of drug resistance and quasispecies.

Extensive deterministic models, statistical models, stochastic models and state space models on treating AIDS patients with antiretroviral drugs are provided, as well as an in-depth discussion of these models. The book also contains updated reviews on mathematical models for assessing effects of AIDS vaccines, statistical methods for analyzing clinical trial data on AIDS vaccines, and overviews of models and statistical methods for assessing drug resistance of HIV to antiretroviral drugs. Some important statistical methods specific to the intervention and prevention of HIV epidemic are also discussed.

This will be a useful reference source for graduate students and researchers in biomathematics and biostatistics, as well as for HIV/AIDS epidemiologists and clinical investigators learning quantitative methods to study AIDS epidemics and HIV infection.

Deterministic and Stochastic Models of AIDS Epidemics and HIV Infections with Intervention

World Scientific

ISBN 981-256-139-0