When to Initiate HIV Therapy: A Control Theoretic Approach

Annah M. Jeffrey*, Xiaohua Xia, *Senior Member, IEEE,* and Ian K. Craig, *Senior Member, IEEE*

***Abstract*—This paper shows an application of control theory to human immunodeficiency virus (HIV)/AIDS models. Minimum singular value decomposition is applied to HIV/AIDS models to measure the extent to which the different stages in the progression of HIV/AIDS disease are controllable and, consequently, when best to initiate therapy such that the general objectives of therapy are satisfied. Simulations will be used to demonstrate the effect of treatment at various stages. Comparisons will be made between mono-class and combination therapies and between when therapy is initiated at the acute infection, asymptomatic and the advanced stages.**

***Index Terms*—Chemotherapy, controllability, HIV, initiate therapy, singular value.**

I. Introduction

ERADICATION of human immunodeficiency virus (HIV) infection does not seem possible with currently available anti-retroviral drugs. This is due primarily to the establishment of a pool of latently infected $CD4^+$ T cells during the very earliest stages of acute HIV infection that persists with an extremely long half-life, even with prolonged suppression of the plasma viral load using highly active anti-retroviral therapy (HAART) [6], [11]. There is no doubt, however, that anti-retroviral drugs reduce the viral load, can maintain an acceptable $CD4^+$ cell count and, consequently, prolong the life of the infected person. What is not clear though, is when, during the disease progression is it best to initiate therapy.

Guidelines are available for HIV therapies and are in conflict as to when to start therapy. Two large medical organizations have published treatment guidelines that make general recommendations regarding when people living with HIV should start anti-retroviral therapy. These organizations are the United States Public Health Service (USPHS) and the International AIDS Society-USA (IAS-USA). In general, the IAS-USA guidelines recommend starting therapy earlier than the USPHS guidelines. The optimal time to initiate therapy in the course of HIV infection still remains unclear.

HIV chemotherapy is initiated either in the asymptomatic stage, termed early therapy or in the advanced stage and referred to as late therapy. There are instances where therapy is initiated in the acute infection stage. Therapy in most cases entails the use of anti-retroviral drugs that inhibit the replication rate of the virus. Other therapies that reconstruct the immune system are also available. The issues of when to initiate therapy and the necessary dosage have been analyzed in [7], [20], [25]. Some authors believe that early therapy, when the $CD4^+$ T cell count is still high is best, while some believe that late therapy during the final decline of the $CD4^+$ T cells is best. The clinical benefits of initiating therapy very early in the acute infection stage are not very clear.

The reason for the lack of consensus is that the chemotherapy of HIV has multiple objectives and the studies that have been done and used to formulate these guidelines have different objectives. There is a general consensus though, from these guidelines and conference presentations that the viral load prior to therapy does not appear to play an important role in clinical response to HAART [30].

It is apparent that HIV therapy has to be administered indefinitely or at least till such a time when alternative drugs are available. The objectives of HIV chemotherapy according to the USPHS guidelines are in general, "maximal and durable suppression of viral load, restoration and/or preservation of immunologic function, improvement of quality of life, and reduction of HIV-related morbidity and mortality" [32]. On an individual basis, the objectives range from suppressing the viral load to below detectable levels, to maintaining $CD4^+$ T cell counts at levels that are just sufficient to delay the onset of AIDS and opportunistic infections. Generally, the objectives call for moderate or low-dose therapy schedules where possible that can effectively suppress the viral load while avoiding the emergence of resistant virus strains. Furthermore, therapy should be initiated at a time when the immune system is still functional or at least repairable.

Maximal suppression of the viral load to below detectable levels has been achieved using HAART. Durable suppression has, however, proven to be difficult using HAART because the high drug doses used have adverse side effects that make adherence to therapy very difficult. It has also been shown that high drug doses may lead to early emergence of resistance [3]. Studies are and have been carried out on the clinical benefits of initiating therapy on patients at different stages of the disease. Some studies [13], [17], [24] have shown that there is a higher mortality rate for patients who start therapy in the advanced stages of the disease as opposed to those who started earlier.

This paper adopts a control theoretic approach to when best to initiate therapy. The application of control theory to HIV dynamics has been explored from an optimal control point of view

Manuscript received June 25 2002; revised February 27, 2003. *Asterisk indicates corresponding author.*

*A. M. Jeffrey is with the Department of Electrical, Electronic, and Computer Engineering Engineering, University of Pretoria, Pretoria 0002, South Africa (e-mail: mjeffrey@tuks.co.za)

X. Xia and I. K. Craig are with the Department of Electrical, Electronic, and Computer Engineering Engineering, University of Pretoria, Pretoria 0002, South Africa.

Digital Object Identifier 10.1109/TBME.2003.818465

in [5], [9], [19], and [33]. The authors of [1] and [4] adopted a feedback control point of view. Most of these HIV control works are aimed at suppressing the plasma viral load of infected persons, while some focus on maintaining the $CD4^+$ T cell count within a given range or above a specified level. There is a general view though, that treatment should vary with time and depend on the individual patients' response to treatment. [20] has addressed this issue of when to initiate therapy by using simulations. This paper gives a control theoretic base and simulations will be used to back up control theoretical derivations.

Motivated by the lack of consensus on when best to initiate therapy, the aim of this paper, therefore, is to shed light, from a control theoretic point of view, on this issue. The argument being put forth in this paper is that therapy is best initiated at the time when the viral load is easiest to control. The reasoning is that; an easier to control viral load implies less control effort in the form of lower drug doses. Lower drug doses in turn imply that therapy can be administered and be effective with milder side effects over a longer period of time. This in light of the objectives of therapy translates to maximal and durable suppression of the viral load and improvement of the quality of life.

Minimum singular value decomposition (SVD) will be used as a measure of the extent to which the various stages of the disease progression are controllable. SVD is a valid method to use here because the same variables are used to measure controllability and compared at different times as the disease progresses. The easier to control stages will, therefore, be identified and simulations will be used to demonstrate the effect on the viral load of initiating various types of therapies at different stages as the disease progresses. Comparison will be made between when therapy is initiated in the acute infection stage, the asymptomatic stage, and the later stages of the disease. Comparisons will also be made between no therapy, mono-class, or class-sparing therapies using either protease inhibitor(s) (PI) only, reverse transcriptase (RT) inhibitor(s) only, and multidrug therapy using a combination of PI and RT inhibitors.

The layout of the paper is as follows: Section II presents the working model under mono- and multitherapies. Section III introduces the structure of the control input and minimum SVD of the working model. Section IV has simulations of how the virus responds to therapy at different stages of the disease and Section V has the conclusions that are drawn from this study.

II. The Working Model

Mathematical models that describe the host-pathogen interaction between the immune system and HIV should be able to explain the initial high increase in plasma viral load, its decline and settling to levels that are much lower than the peak viral load. The subsequent dramatic increase of the viral load during the later stage of the disease and timing of this increase should also be explained. Models have been developed and explained in, for example, [14], [18]–[20], [23], [26], and [31]. Most of these models are deterministic and based on balancing population dynamics, while some are stochastic and take into account the random variations in the HIV dynamics.

None of these models, however, can completely exhibit all that is observed clinically and account for the full course of the disease. The main reason for the models' limitation is lack of a good understanding of the immunology of the human body against HIV. Biological systems exhibit multicompartmental interactions that are usually not well understood and as a result, can not be accurately modeled mathematically. The accuracy of the models though, is increasing with new medical discoveries [21]. 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 disease.

TABLE I
Parameter Estimates

s	$10mm^{-3}day^{-1}$	β	$7.5\times10^{-6}mm^{-3}day^{-1}$
p	$0.03day^{-1}$	δ	$0.5day^{-1}$
T_m	$1500mm^{-3}$	N	2000 virions cell^{-1}
d_T	$0.02day^{-1}$	c	$5day^{-1}$

The mathematical model as presented by [26] is adopted for this paper and the following is a summary of that model:

$$\dot{x}_1 = s + px_1\left(\frac{1-x_1}{T_m}\right) - d_T x_1 - \beta x_1 x_3 \quad (1)$$

$$\dot{x}_2 = \beta x_1 x_3 - \delta x_2 \quad (2)$$

$$\dot{x}_3 = N\delta x_2 - cx_3. \quad (3)$$

The state variables x_1, x_2, and x_3 are the concentrations of the uninfected $CD4^+$ T cells, the infected $CD4^+$ T cells, and the free virus particles, respectively. Equation (1) describes the population dynamics of the uninfected $CD4^+$ T cells. It shows that they are produced from a source at a constant rate s and proliferate to a maximum given by T_m, at a rate that is proportional to their abundance, with p as the proliferation rate constant. Uninfected $CD4^+$ T cells die with a rate constant d_T and 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. Equation (2) shows that infection of healthy $CD4^+$ T cells produces infected $CD4^+$ T cells that die with a rate constant δ. Equation (3) similarly shows that an infected $CD4^+$ T cell produces N free virus particles during its lifetime. These free virus particles, which are also known as virions, die with a rate constant c.

Estimates for model parameters are available in [19], [20], [23], [26]–[28], and [33] and some parameters may vary with time [8], [23]. A method for obtaining online estimates by using control techniques can be found in [35]. For this paper, the estimates used are in Table I and are sourced from [1], [23], and [26].

RT inhibitors block infection by reducing the value of β. Refer to, for example, [22] and [26], on which model parameters are affected by therapy. Perfect inhibition occurs when $\beta = 0$. In practice, however, perfect inhibition is not attainable.

Equations (1)–(3) under imperfect RT inhibitor(s) can be modified to

$$\dot{x}_1 = s + px_1\left(\frac{1-x_1}{T_m}\right) - d_T x_1 - u_1\beta x_1 x_3 \quad (4)$$

$$\dot{x}_2 = u_1\beta x_1 x_3 - \delta x_2 \quad (5)$$

$$\dot{x}_3 = N\delta x_2 - cx_3 \quad (6)$$

where $u_1 = 1 - \eta_{\rm RT}$ is taken as a control input, and $\eta_{\rm RT}$, $0 \leq \eta_{\rm RT} \leq 1$ is the effectiveness or combined effectiveness of the RT inhibitor(s) used. Perfect inhibition occurs, therefore, when $\eta_{\rm RT} = 1$ and there is no inhibition when $\eta_{\rm RT} = 0$.

PIs on the other hand, do not block infection, but rather block the protease enzyme so that the virus particles that are produced are noninfectious. There are, therefore, two types of virus particles when PIs 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.

Equations (1)–(3) under PI(s) can be modified to

$$\dot{x}_1 = s + px_1\left(\frac{1-x_1}{T_m}\right) - d_T x_1 - \beta x_1 x_3 \tag{7}$$

$$\dot{x}_2 = \beta x_1 x_3 - \delta x_2 \tag{8}$$

$$\dot{x}_3 = u_2 N\delta x_2 - cx_3 \tag{9}$$

$$\dot{x}_4 = (1-u_2)N\delta x_2 - cx_4 \tag{10}$$

where $u_2 = 1 - \eta_{\rm PI}$ is taken as the control input, and $\eta_{\rm PI}$, $0 \leq \eta_{\rm PI} \leq 1$ is the combined effectiveness of the PIs used. Similarly, perfect inhibition occurs when $\eta_{\rm PI} = 1$ and there is no inhibition when $\eta_{\rm PI} = 0$. Equation (9) describes the population dynamics of the infectious virus particles, while (10) is for the noninfectious virus particles. It is assumed that both types of virus particles have the same death rate constant c.

Current therapies use a combination of RT and PIs and (7)–(10) can be modified to reflect the effect of both inhibitors.

III. Controllability Analysis

Medically, it seems that the viral load in infected persons is controllable by the prevailing anti-retroviral drugs. Controllability analysis for HIV models using only RT inhibitors has been addressed by [15] and is included in this paper for completeness. Analysis will, therefore, be extended to models using only PIs and combinations of protease and RT inhibitors.

A. Anti-Retroviral Drugs as Control Inputs

Administering an anti-retroviral drug is in fact equivalent to introducing an input signal that perturbs the HIV dynamics. As a control input, the model parameters that are affected by these drugs have been identified [22], [26].

An important point to consider is that these drugs are administered periodically. In this paper, it is assumed that the interval between doses is constant and that the drug potency from the dosing time rises then decreases exponentially with time. This assumption is based on drug pharmacokinetics as explained in [2]. The increase in plasma drug concentration is determined by the rate at which the drug enters the plasma by absorption and simultaneously removed from plasma by either distribution to other body compartments or by elimination. A maximum concentration is reached when there is a balance between the entry and removal rates of the drug into and out of the plasma. Thereafter, the drug plasma concentration declines because the combined rate at which it is distributed and eliminated exceeds the rate at which it is absorbed. This plasma drug concentration decline rate is usually biphasic if the rate at which the drug is distributed exceeds the rate at which it is eliminated. If one assumes, therefore, that the plasma concentration of a drug at a particular point in time is an indication of the efficacy or instantaneous inhibition [10] of virus replication by the drug, then one can model the drug's efficacy as illustrated in Fig. 1. Refer to, for example, [12] and [16] for the relationship between plasma drug concentration and efficacy.

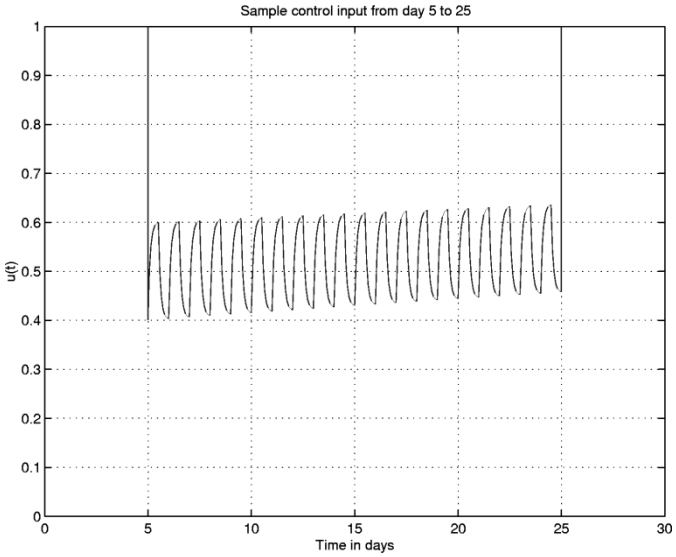

Fig. 1. Sample control input. One cycle represents one day.

The drug is usually effective until the time when resistance emerges and the effectiveness η ($\eta_{\rm RT}$ or $\eta_{\rm PI}$) can also be assumed to decline exponentially with time. The decay time constant for the effectiveness η are drug and dosage dependent. The admissible control for the chemotherapy of HIV in this paper is, therefore, an oscillatory and decaying function of time that can be represented as

$$u(t) = \begin{cases} 1, & t < t_i, t > t_f \\ 1 - \eta(t), & t_i \leq t \leq t_f \end{cases} \tag{11}$$

where t_i is the time when therapy is initiated and t_f is when therapy is ended. The expression for $\eta(t)$ over one cycle is as shown in (12) at the bottom of the page, where a_w and a_v are the width and average values, τ_e is the rise and decay time constant after each dose and τ_d is the decay time constant of the drugs effectiveness.

Fig. 1 shows a sample control variable $u(t)$ for when therapy is on for 20 days from day 5 to day 25. What will be done then is to incorporate these constraints into the control input when doing analysis and test simulations. For simplicity, the drug is

$$\eta(t) = \begin{cases} (a_v - a_w(1 - e^{-(t-t_i)/\tau_e}))e^{-(t-t_i)/\tau_d}, & t_i \leq t \leq t_i + 0.5 \\ (a_v - a_w e^{-(t-t_i)/\tau_e})e^{-(t-t_i)/\tau_d}, & t_i + 0.5 \leq t \leq t_i + 1 \end{cases} \tag{12}$$

assumed to be administered daily and, therefore, one cycle represents a day.

When therapy is on, the steady states for the infectious and noninfectious virus particles are given by

$$x_{3ss} = \frac{a_{v2}Ns}{c} + \frac{p - d_T}{a_{v1}\beta} - \frac{pc}{a_{v2}N(a_{v1}\beta)^2 T_m} \tag{13}$$

$$x_{4ss} = \frac{1 - a_{v2}}{a_{v2}} x_{3ss} \tag{14}$$

where a_{v1} and a_{v2} are the control input averages for the RT and PIs, respectively. The viral load steady states without treatment are as given below

$$x_{3ss} = \frac{Ns}{c} + \frac{p - d_T}{\beta} - \frac{pc}{N\beta^2 T_m} \tag{15}$$

$$x_{4ss} = 0. \tag{16}$$

It is apparent, therefore, that when treatment is initiated, the viral load will settle at a new on treatment steady state that is determined by the drug(s) efficacy. Conversely, the drug(s) efficacy required can be determined, given the desired treatment steady state. Therapy, therefore, moves the states from one point to another. Initiating therapy when the viral load is below this treatment steady-state will result in an increasing viral load, which is interpreted as failure to control the viral load. A higher dose will be needed in order to obtain a reduction in viral load. Initiating therapy when the viral load is higher than this treatment steady state will obviously result in some degree of viral load control even though the viral load will eventually settle to the same steady state.

B. Minimum Singular Value Decomposition

The variable to be controlled is the viral load. Medically, the viral load is considered to be controllable if the control law in use can reduce it by 90% in eight weeks from the time treatment is initiated and continue to suppress it to below 50 copies per milliliter of plasma in six months [32].

Under mono-class therapy using one or more RT inhibitors, an approximate analysis by linearizing the nonlinear equations in (4)–(6) at the operating points can be obtained. The Jacobians for (4)–(6) when evaluated at an operating point $\left(x^0, u_1^0\right)$ are given by

$$A_{\mathrm{RT}} = \begin{bmatrix} \kappa_1 & 0 & -u_1^0\beta x_1^0 \\ u_1^0\beta x_3^0 & -\delta & u_1^0\beta x_1^0 \\ 0 & N\delta & -c \end{bmatrix}$$

$$B_{\mathrm{RT}} = \beta x_1^0 x_3^0 \begin{bmatrix} -1 & 1 & 0 \end{bmatrix}^T$$

where $\kappa_1 = p(1 - 2x_1^0/T_m) - d_T - u_1^0\beta x_3^0$.

The controllability matrix [29] $[B_{\mathrm{RT}}\ A_{\mathrm{RT}}B_{\mathrm{RT}}\ A_{\mathrm{RT}}^2 B_{\mathrm{RT}}]$ is given by

$$M_{\mathrm{RT}} = \beta x_1^0 x_3^0 \begin{bmatrix} -1 & -\kappa_1 & -\kappa_1^2 - \kappa_2 \\ 1 & -\kappa_3 & -u_1^0\beta x_3^0\kappa_1 + \delta\kappa_3 + \kappa_2 \\ 0 & N\delta & -N\delta(\kappa_3 + c) \end{bmatrix}$$

where $\kappa_2 = u_1^0\beta x_1^0 N\delta$ and $\kappa_3 = u_1^0\beta x_3^0 + \delta$.

Under mono-class therapy using one or more PIs, the Jacobians for (7)–(10) when evaluated at an operating point $\left(x^0, u_2^0\right)$ are given by

$$A_{\mathrm{PI}} = \begin{bmatrix} \kappa_4 & 0 & -\beta x_1^0 & 0 \\ \beta x_3^0 & -\delta & \beta x_1^0 & 0 \\ 0 & u_2^0 N\delta & -c & 0 \\ 0 & \left(1 - u_2^0\right) N\delta & 0 & -c \end{bmatrix}$$

$$B_{\mathrm{PI}} = N\delta x_2^0 \begin{bmatrix} 0 & 0 & 1 & -1 \end{bmatrix}^T$$

where $\kappa_4 = p(1 - 2x_1^0/T_m) - d_T - \beta x_3^0$.

The controllability matrix $[B_{\mathrm{PI}}\ A_{\mathrm{PI}}B_{\mathrm{PI}}\ A_{\mathrm{PI}}^2 B_{\mathrm{PI}}\ A_{\mathrm{PI}}^3 B_{\mathrm{PI}}]$ is given by the equation shown at the bottom of the page, where $\kappa_5 = \beta x_1^0 u_2^0 N\delta + c^2$ and $\kappa_6 = \beta x_3^0 + \delta + c$.

Similarly, under multiple therapy using a combination of RT and PIs, the Jacobians when evaluated at an operating point $\left(x^0, u_1^0, u_2^0\right)$ are given by

$$A_{\mathrm{CO}} = \begin{bmatrix} \kappa_1 & 0 & -u_1^0\beta x_1^0 & 0 \\ u_1^0\beta x_3^0 & -\delta & u_1^0\beta x_1^0 & 0 \\ 0 & u_2^0 N\delta & -c & 0 \\ 0 & (1 - u_2^0)N\delta & 0 & -c \end{bmatrix}$$

$$B_{CO} = \begin{bmatrix} -\beta x_1^0 x_3^0 & 0 \\ \beta x_1^0 x_3^0 & 0 \\ 0 & N\delta x_2^0 \\ 0 & -N\delta x_2^0 \end{bmatrix}.$$

The controllability matrix $M_{\mathrm{CO}} = [B_{\mathrm{CO}}\ \ A_{\mathrm{CO}}B_{\mathrm{CO}}\ A_{\mathrm{CO}}^2 B_{\mathrm{CO}}\ A_{\mathrm{CO}}^3 B_{\mathrm{CO}}]$ can be determined in a similar manner.

The matrices M_{RT}, M_{PI}, and M_{CO} are not of full row rank only when the viral load is zero and when the T cell count is zero. A zero viral load is invalid because the patient is assumed to be actively infected. When the T cell count is zero then the immune system is completely damaged. In this case, there is no point trying to control the virus. All other states apart from when the T cell count or the viral load is zero are, therefore, controllable. When minimum SVD is applied to the controllability matrices, an estimate measure of how controllable the system is at a particular time during the progression of the disease can be obtained.

$$M_{\mathrm{PI}} = N\delta x_2^0 \begin{bmatrix} 0 & -\beta x_1^0 & -\beta x_1^0(c - \kappa_4) & \beta x_1^0(\kappa_4^2 - c\kappa_4 - \kappa_5) \\ 0 & \beta x_1^0 & -\beta x_1^0\kappa_6 & \beta x_1^0(\beta x_3^0(\kappa_4 - c) + \delta\kappa_6 + \kappa_5) \\ 1 & -c & \kappa_5 & -u_2^0 N\delta\beta x_1^0\kappa_6 - \kappa_5 c \\ -1 & c & \beta x_1^0 N\delta - \kappa_5 & -N\delta\beta x_1^0((u_2^0 - 1)\kappa_6 - c) + \kappa_5 c \end{bmatrix}$$

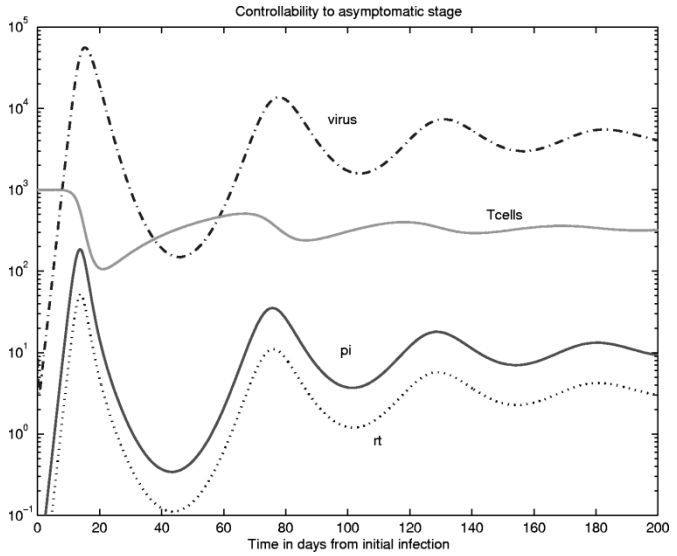

Fig. 2. Controllability to asymptomatic stage. Graphs pi and rt generated using matrices $M_{\rm pi}$ and $M_{\rm rt}$. $V_o = 10$, $T_o = 1000$. Parameters in Table I.

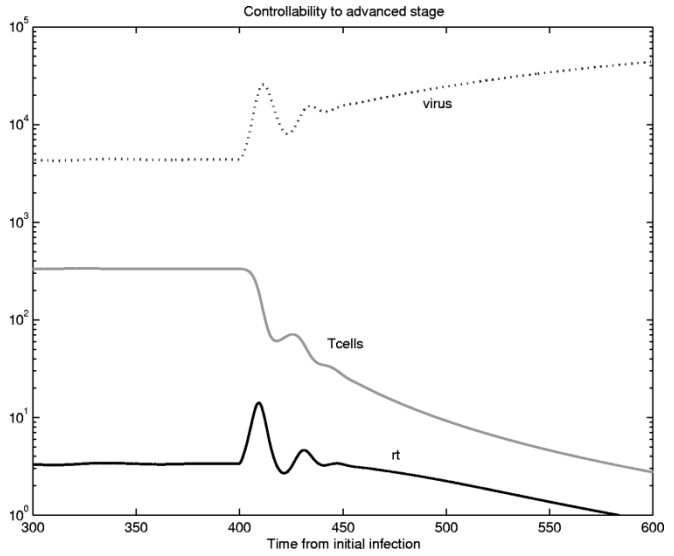

Fig. 3. N and β changing at rates ($r_N = 0.05$, and $r_\beta = 0.05$). Graph rt generated using matrix M_{rt}.

Fig. 2 shows how the concentrations of the uninfected $\rm CD4^+$ T cells, free virus particles, and the minimum singular value vary with time from initial infection to the asymptomatic stage. The stage from initial infection to before the viral load settles is known as the acute infection or preasymptomatic stage and the stage when the viral load has settled is known as the asymptomatic or latent stage and can last for up to 10 years with the infected person showing no symptoms of the disease.

The graphs labeled $\rm pi$ and $\rm rt$ in Fig. 2 show how the minimum singular value varies with time for matrices $M_{\rm PI}$ and $M_{\rm RT}$, respectively. The controllability profile for the RT and PIs are similar in that where one controls most effectively, the other one does also. For a particular drug, a higher singular value indicates an easier to control viral load in the sense that the transition to the treatment steady state is faster. A lower singular value indicates a more difficult to control viral load characterized by a slow transition to the final state, given the same control effort. Graphs in Fig. 2 show that the very early stages of the acute infection stage are relatively more difficult to control as compared to the asymptomatic stage. The section of the acute infection stage where the viral load is much higher than the steady-state viral load is relatively the easiest to control. It can also be seen that up to the asymptomatic stage, controllability and viral load are correlated, whereas there is no obvious correlation between the T cell count and controllability.

Model parameters are thought to vary with time as has been shown in [8], [23]. For this paper, in an attempt to account for the rapid increase of the viral load at the later stages of the disease, parameters N and β are assumed to increase linearly with time as given by (17)

$$y(t) = \begin{cases} y_o, & 0 \le t \le t_b \\ y_o(1 + r_y(t - t_b)), & t > t_b \end{cases} \tag{17}$$

where the variable y represents the concerned parameter (N or β), y_o is the original value, and r_y is the rate at which the parameter changes. All parameters are taken to be constant from the time of infection until a time t_b when the immune system breaks down. These assumptions, though not clinically validated, do give virus and $\rm CD4^+$ T cell profiles that comply with clinical observations. For illustrative purposes, the immune system is taken to break down 400 days from initial infection.

Fig. 3 shows the minimum singular value plots to the advanced stage with β and N changing at $r_\beta = 0.05$ and $r_N = 0.05$, respectively. It can be seen that the very advanced stage is not as controllable as the asymptomatic stage and controllability correlates with the $\rm CD4^+$ T cells. There is a period of time from just when the virus rebounds and the $\rm CD4^+$ T cells decline, when the viral load is more controllable than during the asymptomatic stage.

IV. Simulation

Incorporating the nonlinearities of the models, computer simulations will be used to support the theoretical deductions of this analysis.

The control variable $u(t)$ is as presented in Section III-A. Therapy is assumed to be administered once daily as it has been shown by [20] that it is just as effective as when administered thrice daily. Simulation results are presented here for when therapy is initiated from the acute infection stage to the steady-state stage. In particular, simulation is performed from days 5 and 10 before the viral load reaches its natural peak value. These simulations are compared with when therapy is initiated at day 300 from initial infection.

Using the asymptomatic stage at day 300 under RT inhibitors as a reference, "moderate" (mod) therapy is taken as the dosage that is required at the asymptomatic stage to suppress the viral load to below 50 copies before it rebounds, when initiated at day 300. This corresponds to a control input that fluctuated between 0.3 and 0.45 and has an average value of 0.375 when therapy is on. Given that the relationship between control u effectiveness η is $u(t) = 1 - \eta(t)$, moderate therapy then has an average effectiveness of 62.5%. "Low" therapy is selected such that it can reduce the viral load by 90% or 1 log scale and fluctuated between 0.45 and 0.65 with 40% average effectiveness. "Strong"

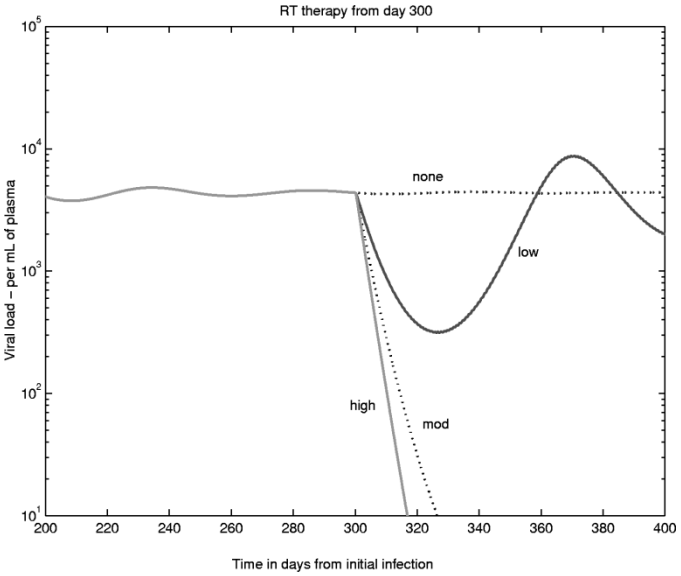

Fig. 4. Viral load response to varying RT efficacy when initiated at day 300: low = 40%, mod = 62.5%, and high = 80% effectiveness.

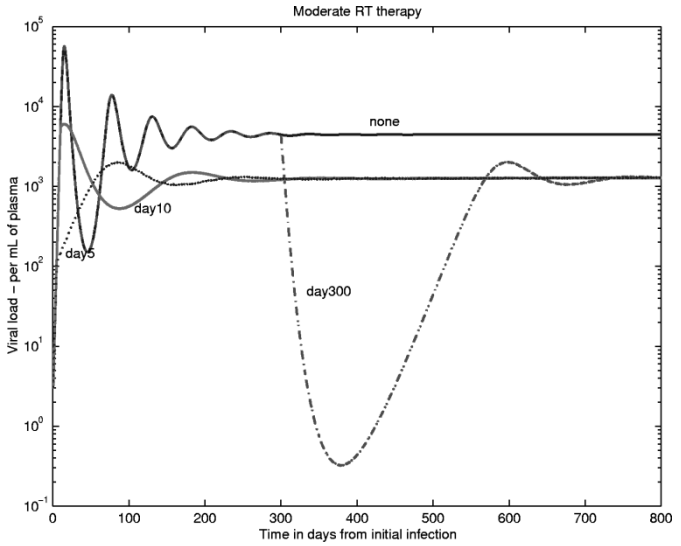

Fig. 5. Viral load response to moderate therapy when initiated at days 5, 10, and 300.

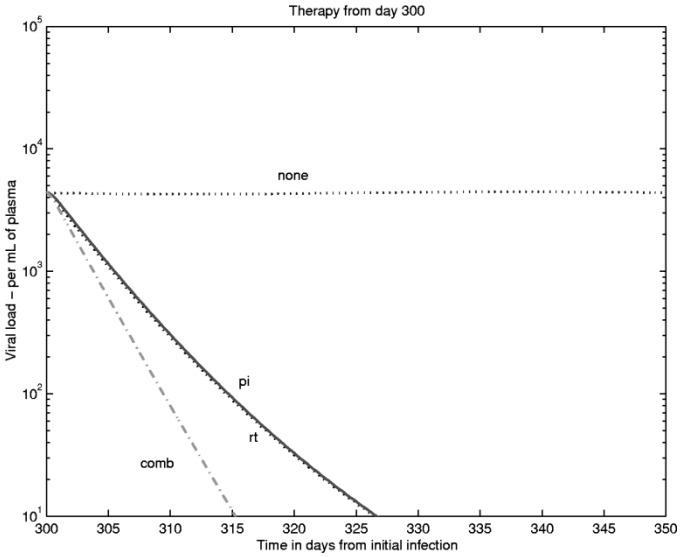

Fig. 6. Comparison between PI and RT with same efficacy (mod) and their combination.

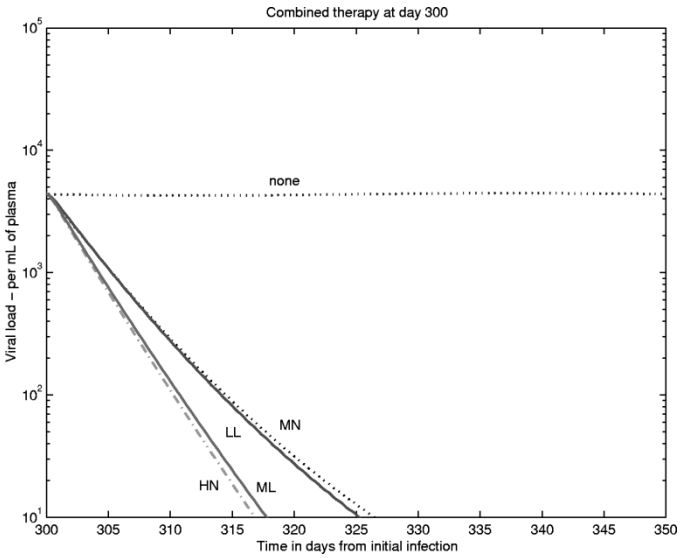

Fig. 7. Effect of combining varying efficacies of RT and PI., low = 40%, mod = 62.5%, high = 80% effectiveness.

(high) therapy fluctuates between 0.15 and 0.25 with 80% average effectiveness.

Fig. 4, showing the virus response when RT therapy is initiated from day 300, is used as the reference response graph. Fig. 5 shows a comparison between initiating moderate RT therapy at different days as the HIV infection progresses. The viral load continuously increases for the duration of the therapy when therapy is initiated at day 5. This is because the viral load at day 5 is lower than the treatment steady-state. The viral load at day 10 is almost the same as that at day 300. However, controllability at day 300 is lower and, therefore, the viral load takes longer to settle at the new steady-state value when therapy is initiated at day 300.

Fig. 6 shows a comparison between using moderate doses of an RT inhibitor(s) alone, PI(s) alone, and a combination of the two moderate inhibitors from both classes. Combined therapy, as expected, is superior to mono-class therapy. Combination therapy using two low doses can be as good as using a single-class moderate therapy as shown in Fig. 7 where the first letter in the graph labels refers to the RT inhibitor and the second to the PI. N, L, M, and H mean none, low, moderate, and high, respectively. For example, ML means combined therapy using a moderate RT inhibitor with a low PI. Using mismatched doses has some interesting outcomes. Using a combined low- and moderate-dose therapy can be as good as using a single-class high-dose therapy. Also, using a combined low- and moderate-dose therapy can be as good as using a single high-dose therapy. Low-efficacy drugs can, therefore, be used to augment the performance of other drugs with higher efficacies.

V. Conclusion

The following conclusions can be drawn from this study.

1) Even though any viral load for all stages of the disease, apart from when the associated T cell count is zero, is theoretically controllable, some stages are more controllable than others.

2) The very early acute infection stage and the late advanced stages are the most difficult to control.
3) The acute infection stage, when the viral load is very high is the easiest stage to control.
4) Combination therapy, as expected, is superior to monoclass therapy.
5) Using two weak therapies from each class can be as good as using moderate therapy from a single class. Furthermore, using weak and moderate therapies from different classes can be as good as using strong therapy from a single class.
6) From a viral load controllability point of view, therapy is best initiated when the viral load is easier to control because this implies the use of lower drug doses and consequently bearable sides effects.
7) This study seems to indicate that when therapy is initiated at the appropriate time, the use of highly potent HAART may not be necessary. However, caution is needed due to resistance issues.

REFERENCES

[1] J. Alvarez-Ramirez, M. Meraz, and J. X. Velasco-Hermandez, "Feedback control of the chemotherapy of HIV," *Int. J. Bifurcation Chaos*, vol. 10, pp. 2207–2219, 2000.

[2] S. L. Becker and R. M. W. Hoetelmas. Exploiting Pharmacokinetics to Optimize Antiretroviral Therapy, Medscape's CME Activity. [Online]. Available: http://www.medscape.com/viewprogram/703index

[3] S. Bonhoeffer and M. A. Nowak, "Pre-existence and emergence of drug resistance in HIV-1," *Proc. Roy. Soc. London B*, vol. 264, pp. 631–637, 1997.

[4] M. E. Brandt and G. Chen, "Feedback control of a biodynamical model of HIV-1," *IEEE Trans. Biomed. Eng.*, vol. 48, pp. 754–759, July 2001.

[5] S. Butler, D. Kirschner, and S. Lenhart, "Optimal control of chemotherapy affecting the infectivity of HIV," in *Advances in Mathematical Populations Dynamics: Molecules, Cells, Man*, Singapore: World Scientific, 1997, pp. 104–120.

[6] T. W. Chun, D. Engel, M. M. Berrey, T. Shea, L. Corey, and A. S. Fauci, "Early establishment of a pool of latently infected, resting $\mathrm{CD4}^+$ T cells during primary HIV-1 infection," *Proc. Nat. Acad Sci USA*, vol. 95, pp. 8869–8873, 1998.

[7] "MRC/ANRS randomized double-blind controlled trial of immediate and deferred AZT in symptom free HIV infection," *Lancet*, vol. 343, pp. 871–881, 1994.

[8] D. C. Douek, R. D. McFarland, and P. H. Keiser, "Changes in thymic function with the age and during treatment of HIV infection," *Nature*, vol. 396, pp. 690–695, 1998.

[9] J. A. M. F. de Souza, M. A. L. Caetano, and T. Yoneyama, "Optimal control theory applied to the anti-viral treatment of AIDS," presented at the *IEEE Conf. Decision and Control*, Sydney, Australia, 2000.

[10] N. M. Ferguson, C. Fraser, and R. M. Anderson, "Viral dynamics and anti-viral pharmacodynamics: Rethinking *in vitro* measures of drug potency," *Trends Pharm. Sci.*, vol. 22, pp. 97–100, 2001.

[11] D. Finzi, M. Hermankova, T. Pierson, L. M. Carruth, C. Buck, R. E. Chaisson, T. C. Quinn, K. Chadwick, J. Margolick, R. Brookmeyer, J. Gallant, M. Markowitz, D. D. Ho, D. D. Richman, and R. F. Siliciano, "Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy," *Science*, vol. 278, pp. 1295–1300, 1997.

[12] R. M. W. Hoetelmans *et al.*, "The effect of plasma drug concentrations on HIV-1 clearanace rate during quadruple drug therapy," *J. AIDS*, vol. 12, pp. 111–115, 1998.

[13] R. S. Hogg, B. Yip, E. Wood, K. Chan, K. J. P. Craib, M. V. O'Shaughnessy, and J. S. G. Montaner, "Diminished effectiveness of antiretroviral therapy among patients initiating therapy with $\mathrm{CD4}^+$ cell counts below $200/\mathrm{mm}^3$," presented at the *8th Conf. Retroviruses and Opportunistic Infections*, Chicago, IL, Feb. 2001.

[14] T. Hraba, J. Dolezal, and S. Celikovsky, "Model based analysis of $\mathrm{CD4}^+$ lymphocyte dynamics in HIV infected individuals," *Immunology*, vol. 181, pp. 108–118, 1990.

[15] M. Jeffrey, X. Xia, and I. K. Craig, "Controllability analysis of the chemotherapy of HIV/AIDS," in *Proc. 15th IFAC Triennial World Congress Automatic Control*, vol. Q, July 2002, pp. 127–132.

[16] A. S. Joshi *et al.*, "Population pharmacokinetics of efavirenzin phase II studies and relationship with efficacy," in *Prog. and Abstr. 30th ICAAC*, San Francisco, CA, Sept. 1999.

[17] J. Kaplan, D. Hanson, J. Karon, D. Cohn, M. Thompson, S. Buskin, P. Fleming, and M. Dworkin, "Late initiation of antiretroviral therapy (at $\mathrm{CD4}^+$ lymphocyte count $<200\ \mathrm{cells/mL}$) is associated with increased risk of death," presented at the *8th Conf. Retroviruses and Opportunistic Infections*, Chicago, IL, Feb. 2001.

[18] D. Kirschner, "Using mathematics to understand HIV immune dynamics," *AMS Notices*, pp. 191–202, 1996.

[19] D. Kirschner, S. Lenhart, and S. Serbin, "Optimal control of the chemotherapy of HIV," *J. Math. Biol.*, vol. 35, pp. 775–792, 1997.

[20] D. Kirschner and G. F. Webb, "A model for treatment strategy in the chemotherapy of AIDS," *J. Math. Biol .*, vol. 58, pp. 367–390, 1996.

[21] D. Kirschner, G. F. Webb, and M. Cloyd, "Model of HIV-1 disease progression based on virus-induced lymph node homing and homing-induced apoptosis of $\mathrm{CD4}^+$ lymphocytes," *J. AIDS*, vol. 24, pp. 352–362, 2000.

[22] G. X. McLeod and S. M. Hammer, "Zidovudine: 5 years later," *Ann. Int. Med.*, vol. 117, pp. 487–501, 1992.

[23] M. A. Nowak and R. M. May, *Virus Dynamics: Mathematical Principles of Immunology and Virology*. New York: Oxford Univ. Press, 2000.

[24] M. Opravil, B. Ledergerber, H. Furrer, S. Gallant, B. Hirschel, F. Meienberg, T. Wagels, E. Bernasconi, M. Rickenbach, and R. Weberand, "Clinical benefit of early initiation of HAART in patients with asymptomatic HIV infection and $\mathrm{CD4}^+$ counts $>350/\mathrm{mm}^3$," presented at the *8th Conf. Retroviruses and Opportunistic Infections*, Chicago, IL, Feb. 2001.

[25] A. S. Perelson, D. Kirschner, and R. D. Boer, "Dynamics of HIV infection of $\mathrm{CD4}^+$ T cells," *Math. Biosci.*, vol. 114, pp. 81–125, 1993.

[26] A. S. Perelson and P. W. Nelson, "Mathematical analysis of HIV-1 dynamics in vivo," *SIAM Rev.*, vol. 41, pp. 3–44, 1999.

[27] A. S. Perelson *et al.*, "HIV-1 dynamics in vivo: Virion clearance rate, infected cell life-span and viral generation time," *Science (Online Abstr.)*, vol. 271, pp. 1582–1586, 1996.

[28] B. Ramratnam *et al.*, "Rapid production and clearance of HIV-1 and hepatitis C virus assessed by large volume apheresis," *Lancet*, vol. 354, pp. 1782–1785, 1999.

[29] S. Skogestad and I. Postlethwaite, *Multivariable Feedback Control: Analysis and Design*, Chichestre, U.K.: Wiley, 1996.

[30] T. R. Sterling, R. E. Chaisson, J. G. Bartlett, and R. D. Moore, "$\mathrm{CD4}^+$ lymphocyte level is better than HIV-1 plasma viral load in determining when to initiate HAART," presented at the *8th Conf. Retroviruses and Opportunistic Infections*, Chicago, IL, Feb. 2001.

[31] W. Y. Tan and H. Wu, "Stochastic modeling of the dynamics of $\mathrm{CD4}^+$ T cells by HIV and some Monte Carlo studies," *Math. Biosci.*, vol. 147, pp. 173–205, 1998.

[32] (2003) Guidelines for the Use of Antiretroviral Agents in HIV-Infected Adults and Adolescents, HIV/AIDS Treatment Information Service. USPHS. [Online]. Available: http://www.aidsinfo.nih.gov

[33] L. M. Wein, S. A. Zenios, and M. A. Nowak, "Dynamic multidrug therapies for HIV: A control theoretic approach," *J. Theor. Biol.*, vol. 185, pp. 15–29, 1997.

[34] J. K. Wong, M. Hezareh, H. F. Gunthard, D. V. Havlir, C. C. Ignacio, C. A. Spina, and D. D. Richman, "Recovery of replication-competent HIV despite prolonged suppression of plasma viremia," *Science*, vol. 278, pp. 1291–1295, 1997.

[35] X. Xia, "Estimation of HIV/AIDS parameters," in *Proc. 15th IFAC Triennial World Congress*, vol. Q, July 2002.

Annah M. Jeffrey received the B.Eng. degree in electrical and electronic engineering from Ryerson Polytechnical University, the B.A. in mathematics from York University, North York, ON, Canada, and the M.Sc. degree in systems engineering from the University of Wales, Cardiff. Currently, she is a Lecturer at the University of Botswana and pursuing the Ph.D. degree in electronic engineering at the University of Pretoria, South Africa.

Her research interests include intelligent sensor design, pH control of waste material, and control applications in biomedical systems.

Xiaohua Xia (M'98–SM'98) received the Ph.D. degree in engineering from the Beijing University of Aeronautics and Astronautics, Beijing, China.

He is presently a Professor at the University of Pretoria, Pretoria, South Africa, and is leading the research on control system applications to HIV/AIDS. His other current research interests include control theory, intelligent transportation systems and telecommunication.

Dr. Xia currently serves as the IEEE South Africa Control Chapter chair.

Ian K. Craig (M'86–SM'97) received the B.Eng. degree in electronic engineering from the University of Pretoria (UP), Pretoria, South Africa, the S.M. degree from the Massachusetts Institute of Technology, Cambridge, and the Ph.D. and M.B.A. degrees from the University of the Witwatersrand, Johannesburg, South Africa.

He has been Professor and Section Head: Control Systems in the Department of Electrical, Electronic and Computer Engineering at UP since 1995. Previously he was Group Leader in the Measurement Control Division of Mintek, where he was involved in the design and implementation of advanced controllers for the mineral processing industry. His research interests include industrial process control as applied to metals and mineral processing, the economic evaluation of control systems, and lately the control of biomedical systems.

Prof. Craig is a past member of the IFAC Technical Board and currently serves on the IFAC Council. He is Deputy Editor-in-Chief of *Control Engineering Practice* and Editor for Control of the *Transactions of the South African Institute of Electrical Engineers*. He is a registered Professional Engineer in South Africa, and a Council Member and Fellow of the South African Institute of Electrical Engineers.