Chapter 4

Model Analysis

MODEL analysis will start with the latently infected cell model (3.1)-(3.4), with and without therapy. As pointed out before, the simpler model provides analytical insights more readily, which will then be extended to the co-circulating target cell extended model (3.38)-(3.43).

The models (3.1)-(3.4) and (3.38)-(3.43) are inherently nonlinear. Often in this chapter, approximate solutions will be derived by applying linear analytical tools to the linearized models. Linear analytical tools enable easier derivation of an analytic solution. Approximate as the solution may be, it does present the result in a simpler, more intuitive way, at least within a certain locality, given the diverse backgrounds of the intended audience. The solutions that will be obtained, or the predictions that will be made from the linear analysis will be verified by applying them to the nonlinear model. Instances where these linear analytical tools are applied will be pointed out.

As stated before, the concepts presented in this and the following chapter can be applied to any other model.

4.1   Steady State Analysis

Steady state analysis [170] for the HIV/AIDS models will be conducted in order to determine the dependence of the treatment viral load steady state on drug efficacy. The variation of the treatment steady state when therapy is initiated at various stages of the infection will also be examined. In particular, this section focuses on the attainment of durable suppression of the viral load. The intention is to determine when, as the HIV infection progresses, initiating therapy is most likely to attain durable suppression of the viral load. Analytical solutions for the expected steady state viral load suppression will be derived. If steady state is related to the infection stage, then this study can help decide when best to initiate therapy.

The latently infected cell model (3.1) - (3.4) in section 3.1 has two steady states or equilibrium points. The first is the pre-infection steady state and the other is the infected steady state $[\mathrm{T}_{ss}, \mathrm{T}_{lss}, \mathrm{T}_{ass}, \mathrm{V}_{ss}]^T$, which is given by (4.1)-(4.4). The pre-infection steady state is unstable (once infected), while the infected steady state is stable. There is therefore, only one infected stable steady state and its global asymptotic stability (attractiveness) has been proven or discussed (without T_l) by [59], as well as [84].

$$T_{ss} \;=\; \frac{c\delta_a}{r_T\beta_T q} \tag{4.1}$$

$$T_{lss} \;=\; \frac{q_l c\delta_a}{r_T q(\delta_l + k)} V_{ss} \tag{4.2}$$

$$T_{ass} \;=\; \frac{c}{r_T} V_{ss} \tag{4.3}$$

$$V_{ss} \;=\; \frac{r_T s_T q}{c\delta_a} + \frac{p - d_T}{\beta_T} - \frac{pc\delta_a}{r_T\beta_T^2 q T_m} \tag{4.4}$$

where $q = q_a + q_l \frac{k}{\delta_l + k}$. T_{ss}, T_{lss}, T_{ass}, and V_{ss} are expressions for the uninfected CD4$^+$ T cell, latently infected CD4$^+$ T cell, actively infected CD4$^+$ T cell and viral load steady states, respectively. Equations (4.2) and (4.3) are expressed in that manner to show the proportional relationship between the virus and infected cell steady states.

The infected steady state is determined by both the host cell and virus parameters. Variations in steady state set points from one individual to the other, and ultimately variations in the response to therapy, can therefore be attributed to inter-individual variations in model parameters as has been previously observed by [41, 71, 171, 172].

A point worth noting is that the CD4$^+$ T cell steady state value (4.1) does not depend on CD4$^+$ T cell specific parameters s_T, p, T_m and d_T.

4.1.1   Analysis with Replication Cycle Based HAART

Replication cycle based HAART in this thesis, entails the concomitant use of multiple drugs from the reverse transcriptase and/or protease inhibitor classes. Model (4.5)-(4.9) incorporates the effect of HAART for the latently infected cell model (3.1)-(3.4) that was presented in section 3.1.

$$\frac{dT}{dt} = s_T + pT(1 - T/T_{max}) - d_T T - \beta_T T V_i \tag{4.5}$$

$$\frac{dT_l}{dt} = (1 - \eta_{rt})q_l \beta_T T V_i - k T_l - \delta_l T_l \tag{4.6}$$

$$\frac{dT_a}{dt} = (1 - \eta_{rt})q_a \beta_T T V_i + k T_l - \delta_a T_a \tag{4.7}$$

$$\frac{dV_i}{dt} = (1 - \eta_{pi})r_T T_a - c V_i \tag{4.8}$$

$$\frac{dV_n}{dt} = \eta_{pi} r_T T_a - c V_n \tag{4.9}$$

The treatment steady states for the uninfected CD4$^+$ T cells, infectious virus, non infectious virus and total viral load are given by:

$$T_{ss}(\eta) = \frac{c\delta_a}{(1 - \eta_{co})q\beta_T r_T} \tag{4.10}$$

$$V_{iss}(\eta) = \frac{(1 - \eta_{co})r_T s_T q}{c\delta_a} + \frac{p - d_T}{\beta_T} - \frac{pc\delta_a}{(1 - \eta_{co})r_T \beta_T^2 q T_m} \tag{4.11}$$

$$V_{nss}(\eta) = \frac{\eta_{pi}}{1 - \eta_{pi}} V_{iss} \tag{4.12}$$

$$V_{tss}(\eta) = V_{iss} + V_{nss} = \frac{1}{1 - \eta_{pi}} V_{iss} \tag{4.13}$$

where $\eta_{co} = \eta_{rt} + \eta_{pi} - \eta_{rt}\eta_{pi}$ is the combined efficacy when both reverse transcriptase and protease inhibitors are used in the regimen. Note also that $0 \leq \eta_{co} \leq 1$.

It is clear from the above equations, why and how HAART reduces the viral load and increases CD4$^+$ T cell counts. Therapy with HAART therefore, moves the states from one point to another, and the new on treatment steady states are determined by the combined drugs efficacy as illustrated in figure 4.1.

An important point to note from equations (4.10)-(4.13), is that the treatment steady states are independent of when therapy is initiated. Initiating therapy at a stage where 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. However, initiating therapy at a stage where 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, given the same fixed drug efficacy.
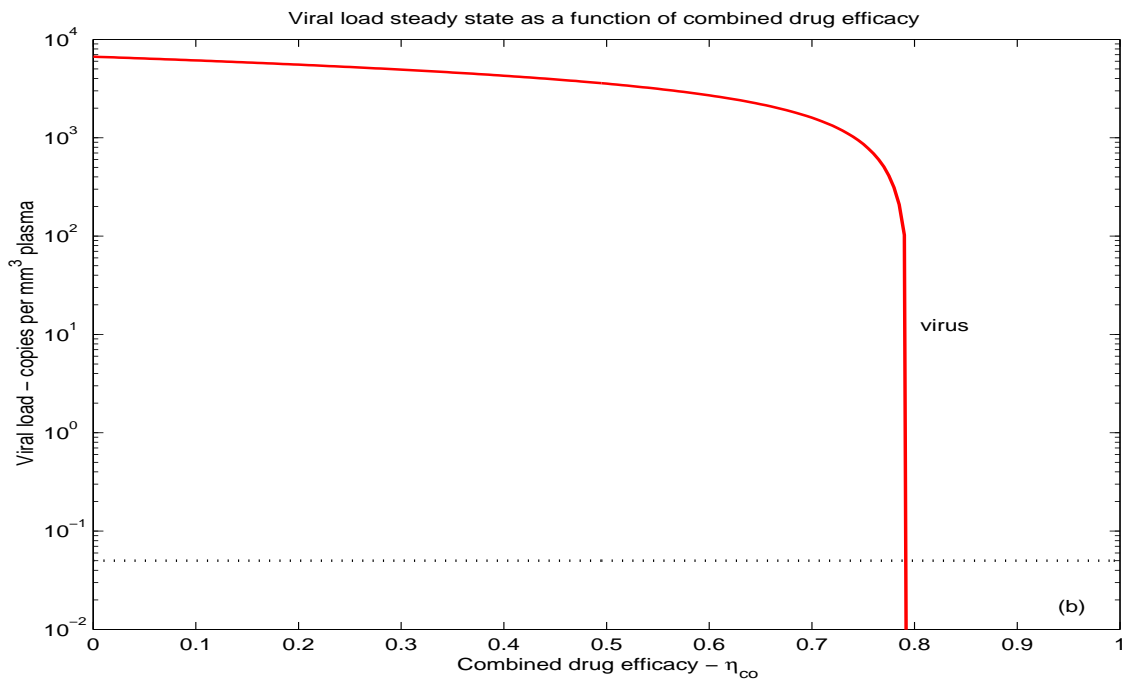
Equation (4.11) can be solved to get

$$\eta_{zss} = 1 + \frac{c\delta_a}{2\beta_T r_T s_T q}[(p - d_T) - \sqrt{(p - d_T)^2 + \frac{4s_T p}{T_m}}] \tag{4.14}$$

as the minimum combined drug efficacy that is required to obtain an on treatment steady state viral load of zero.
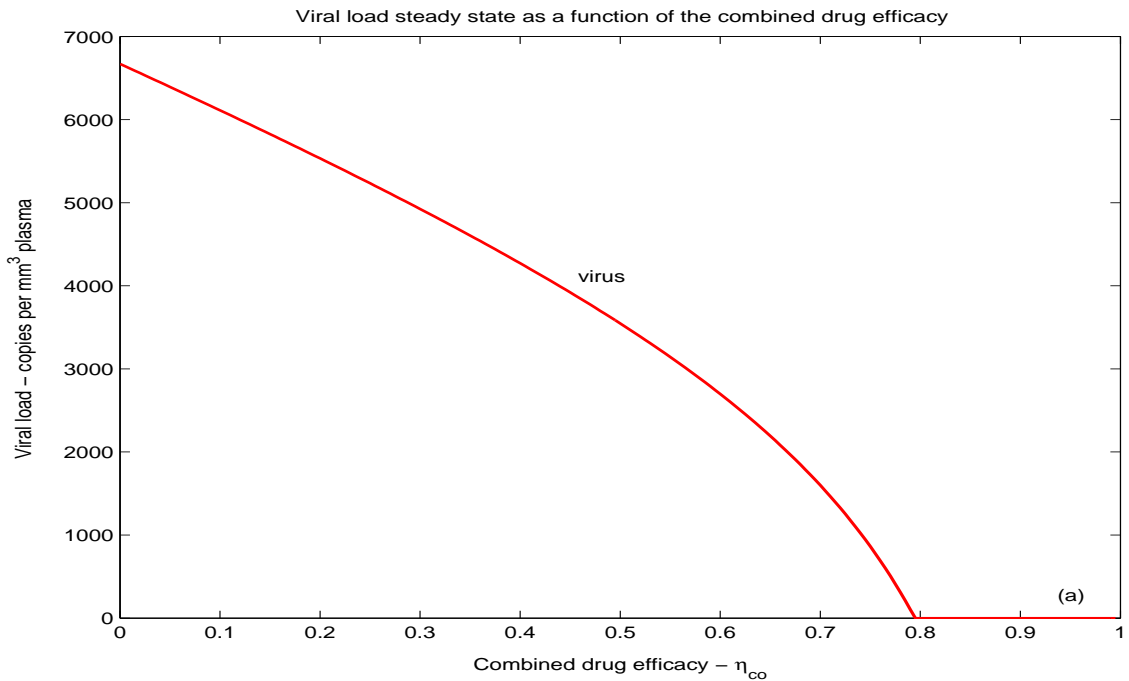
(a)

(b)

Figure 4.1: Viral load steady state for the latently infected cell model as determined by the combined HAART efficacy. (a) Linear plot. (b) Logarithmic plot. Parameters are in Table A.1.

Equation (4.11) can also be solved to get

$$\eta_{sup} \;\; = \;\; 1 + \frac{c\delta_a}{2\beta_T r_T s_T q}[(p - d_T - 50\beta_T) - \sqrt{(p - d_T - 50\beta_T)^2 + \frac{4s_T p}{T_m}}] \quad (4.15)$$

as the minimum combined drug efficacy that is required to obtain a treatment steady state viral load that is below detection (below 50 copies mL^{-1}), as depicted in figure 4.1.

It should be noted that even though attaining a zero viral load steady state corresponds to zero infected cell steady states, that does not necessarily imply perfect inhibition of virus replication, as for most parameter combinations, $\eta_{zss} < 1$. However, increasing the combined drug efficacy to a value above η_{zss} will have no further long term suppression of the virus that is produced from the CD4$^+$ T cells. This therefore means that any circulating or detectable plasma viremia is from other sources such as macrophages, follicular dendritic cells, resting memory T cells and other sources that are known to harbour pro-viral DNA [39, 40, 113, 117]. Higher drug doses are used though, because there is differential drug penetration into some compartments [34, 38, 173].

The combined drug efficacy η_{zss} (4.14) for a zero steady state viral load, as well as η_{sup} (4.15) for a below detection steady state viral load are parameter dependent. They will therefore vary from one individual to another. η_{sup} (4.15) can result in a combined drug efficacy as low as 60% to higher than 95% for some parameter combinations. This explains why some individuals experience virologic failure on therapy that is highly effective on others. From a vaccination point of view, η_{zss} (4.14) can be interpreted as the minimum drug efficacy that is required to prevent the initial virus inoculation from successfully replicating. This, however, is so if one assumes that virus replication starts in the CD4$^+$ T cells before spreading to other compartments.

For drug efficacies in the vicinity of η_{sup} (4.15), the viral load steady state is very sensitive to small changes in drug efficacy, as it has been previously observed by [71]. However, [71] have since demonstrated that this viral load steady state sensitivity to small changes in drug efficacy, can be reduced or eliminated when the extended model (3.38)-(3.43) is used instead. This reduced sensitivity becomes more apparent when differential drug penetration into target cells is modelled. This realization has been used to further explain why, contrary to as suggested by equation (4.14), virus eradication is not possible with the use of HAART and replication competent virus can still be recovered from individuals on potent HAART, as it has been reported in [28, 31, 33]. In any case, this variation in steady state response occurs at very low viral loads. So the use of very high drug doses above η_{sup} or close to η_{zss}, bearing in mind that virus eradication does not seem attainable with HAART, may from a clinical and toxicity perspective, not

be necessary. The drug efficacy cutoff η_{sup} as given by equation (4.15) is therefore, a reasonable upper limit when defining an operating therapeutic range for an individual, for each class of inhibitor that is used in the combined HAART regimen.

Equation (4.10) implies that it is possible for the uninfected CD4$^+$ T cells to rebound to pre-infection values under therapy. However, this is not true for most infected individuals as many attain virologic success with inadequate cell rebounds. Clinical studies suggest that HIV damages the immune system and the extent of T cell rebound therefore, depends on the extent to which the immune system is repairable, and not necessarily on the ability of the drugs to suppress virus replication. There is therefore, a need to find ways of quantifying the health or damage to the immune system, so that it can be factored into the equations describing the CD4$^+$ T cell dynamics. From inspecting equation (4.1), one could expect that β_T and infected CD4$^+$ T cell specific parameters δ_a and q would correlate to the health or damage to the immune system.

4.1.2   Analysis with Immune Based Therapies

Latently Infected Cell Model

Immune based CD4$^+$ T cell specific therapies, as discussed before, could entail:

- Reducing CD4$^+$ T cell source rate s, hence referred to as source rate limiting therapy. η_{sl} is the efficacy (inhibitory effect) of the source suppressing drugs used and $0 \leq \eta_{sl} < 1$.

- Reducing CD4$^+$ T cell proliferation rate p, hence referred to as proliferation suppressive therapy. η_{ps} is the efficacy of the proliferation suppressing drugs used and $0 \leq \eta_{ps} < 1$.

- Accelerating infected CD4$^+$ T cell death rates δ_l and δ_a, hence referred to as infected cell death accelerating therapy. η_{da} is the percentage rate at which CD4$^+$ T cell death rate is increased, or the death acceleration factor and $\eta_{da} \geq 0$.

- Accelerating target CD4$^+$ T cell death rate d_T, hence referred to as apoptosis inducing therapy. η_{ap} is the percentage rate at which apoptosis of target CD4$^+$ T is induced and $\eta_{ap} \geq 0$.

Immune based therapies in current practice, generally entail the independent use of the immunosuppressive drug as hyroxyurea (cell proliferation and maturation suppressor) and the immune stimulant IL-2 (infected cell death accelerator). Model (4.16)-(4.19) incorporates the effect of immune therapies for the latently infected cell model (3.1)-(3.4) that was presented in section 3.1.

$$\frac{dT}{dt} = (1 - \eta_{sl})s_T + (1 - \eta_{ps})pT(1 - T/T_{max}) - (1 + \eta_{ap})d_TT - \beta_TTV_i \quad (4.16)$$

$$\frac{dT_l}{dt} = q_l\beta_TTV_i - kT_l - (1 + \eta_{da})\delta_lT_l \quad (4.17)$$

$$\frac{dT_a}{dt} = q_a\beta_TTV_i + kT_l - (1 + \eta_{da})\delta_aT_a \quad (4.18)$$

$$\frac{dV_i}{dt} = r_TT_a - cV_i \quad (4.19)$$

When immunosuppressive therapy is used, then the uninfected CD4$^+$ T cell and viral load steady states will be:

$$T_{ss}(\eta) = \frac{c\delta_a}{r_T\beta_Tq} \quad (4.20)$$

$$V_{ss}(\eta) = \frac{(1 - \eta_{sl})r_Ts_Tq}{c\delta_a} - \frac{d_T}{\beta_T} + \frac{(1 - \eta_{ps})p}{\beta_T}(1 - \frac{c\delta_a}{r_T\beta_TqT_m}) \quad (4.21)$$

Similarly, the steady states with immune stimulants and apoptosis inducing therapy will be:

$$T_{ss}(\eta) = \frac{(1 + \eta_{da})c\delta_a}{r_T\beta_Tq_{da}} \quad (4.22)$$

$$V_{ss}(\eta) = \frac{r_Ts_Tq_{da}}{(1 + \eta_{ap})c\delta_a} - \frac{(1 + \eta_{ap})d_T}{\beta_T} + \frac{p}{\beta_T}(1 - \frac{(1 + \eta_{da})c\delta_a}{r_T\beta_Tq_{da}T_m}) \quad (4.23)$$

where, $q_{da} = q_a + q_l(\frac{k}{k+(1+\eta_{da})\delta_l})$.

As with replication cycle based therapies, the steady states are parameter dependent, and are also independent of when, during the infection progression, therapy is initiated. As noted before, T_{ss} (4.20,4.22), unlike V_{ss} (4.21,4.23), does not depend on parameters s_T, p and d_T. Immunosuppressive therapies that reduce s_T and p, or therapies that accelerate d_T will reduce the viral load set point V_{ss}, but will have no effect on the CD4$^+$ T cell steady state T_{ss}, as illustrated by figure 4.3. One would expect however, that the introduction of these therapies will transiently perturb the CD4$^+$ T cell dynamics, but the cell count will settle at the steady state T_{ss}, as depicted by figure 4.2.

Increasing the infected cell death rates δ_l and δ_a will on the other hand, increase the CD4$^+$ T cell steady state as well as reduce the viral load set point. So the above considered CD4$^+$ T cell specific therapies will reduce the viral load, but unlike replication cycle based HAART, some will not increase the CD4$^+$ T cell count in the long term.

Figures 4.2 and 4.3 show that, from an end point efficacy perspective, accelerating infected CD4$^+$ T cell death rates, is better than proliferation suppressive therapy at
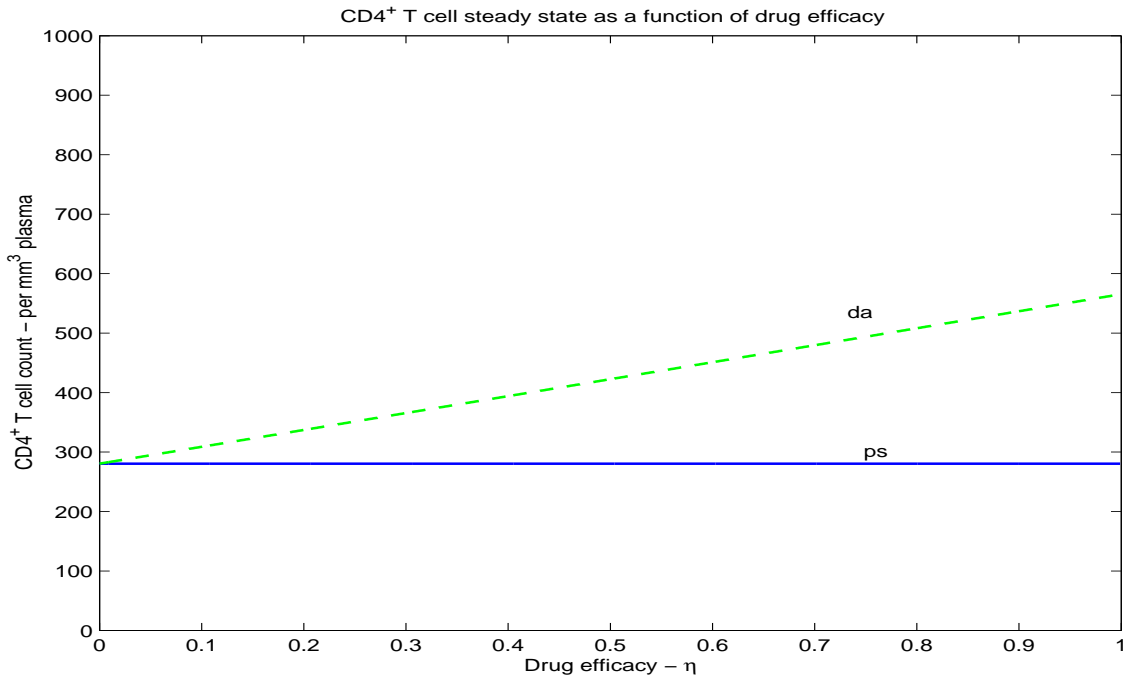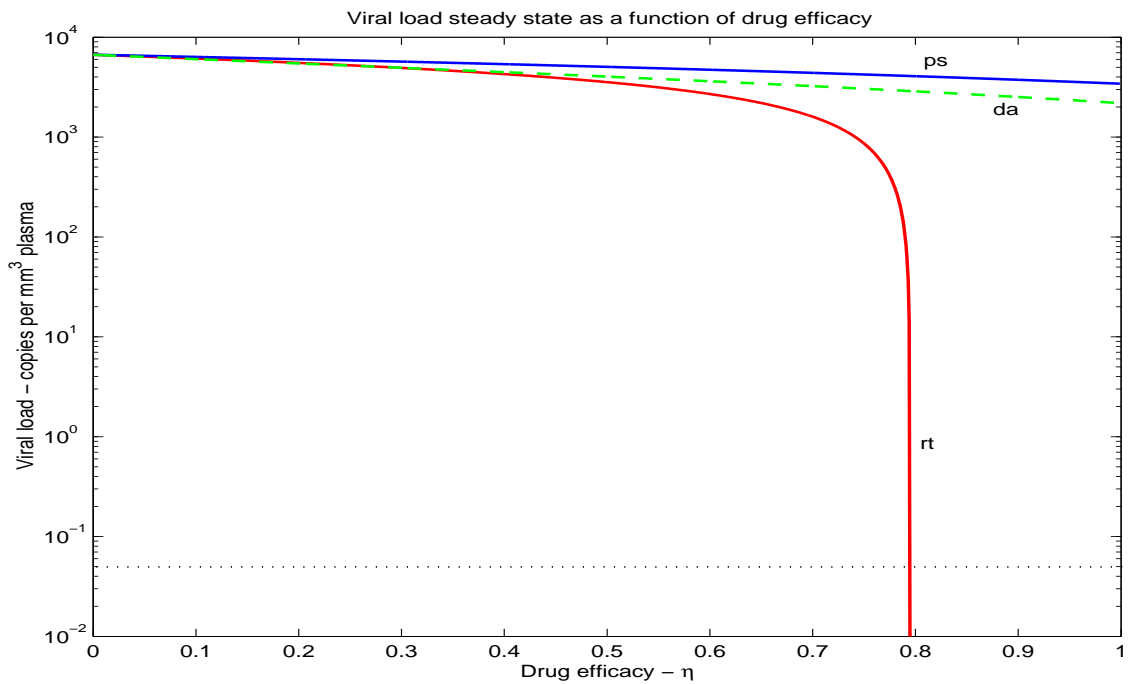
Figure 4.2: CD4$^+$ T cell and viral load steady states as determined by the immune therapy drug efficacy. ps: proliferation suppressing therapy, da: infected cell death accelerating therapy, rt: reverse transcriptase. (a) CD4$^+$ T cells. (b) Viral load. Parameters are in Table A.1.
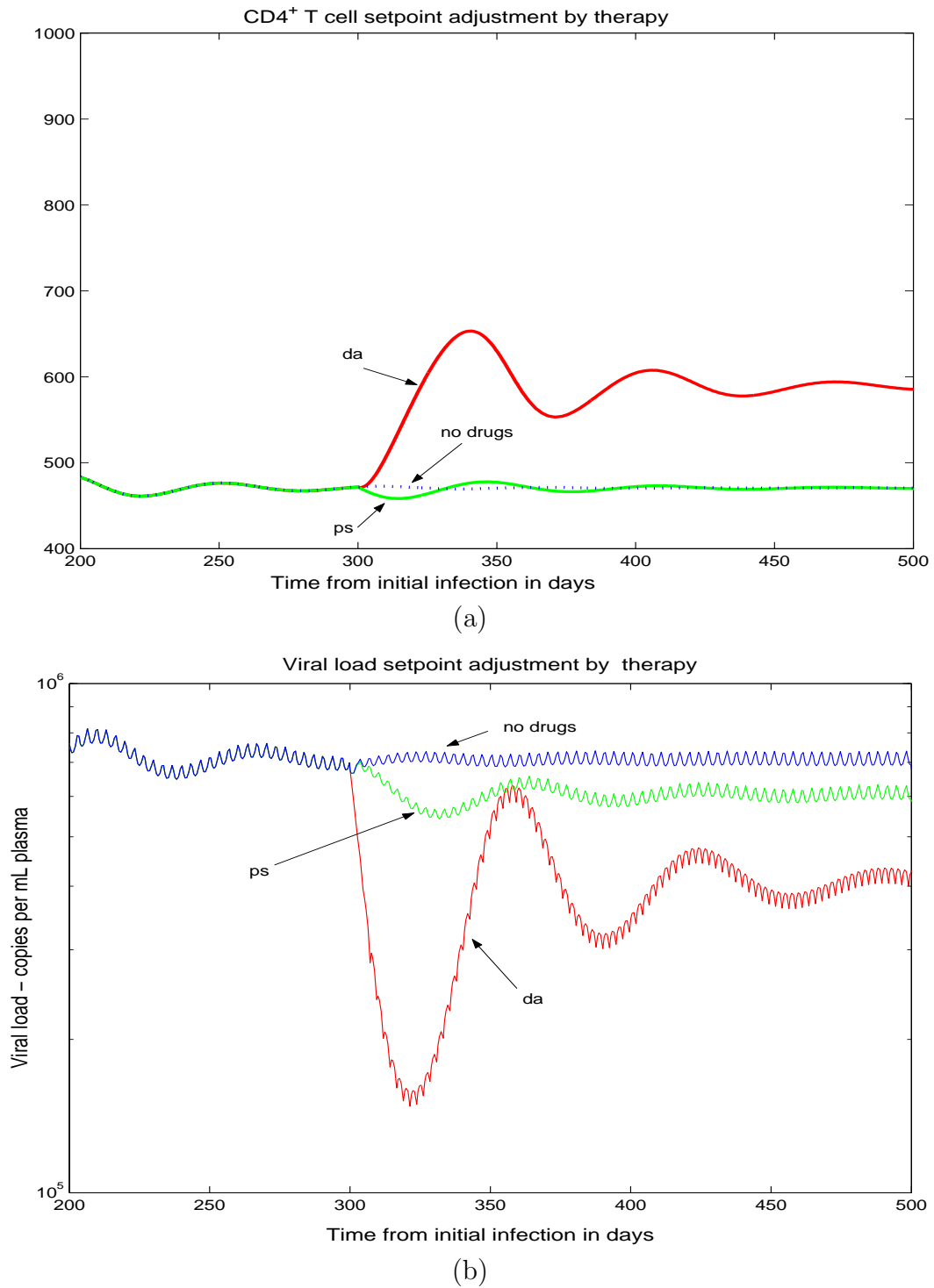
(a)

(b)

Figure 4.3: Pre-treatment set point and set point adjustment by immune therapies. ps: proliferation suppressing therapy, da: infected cell death accelerating therapy. $\eta_{ps} = \eta_{da} = 0.25$ (a) CD4$^+$ T cell count. (b) Viral load. Parameters are in Table A.2.

reducing the viral load, given the same drug efficacy. Furthermore, cell death accelerating therapy has an added advantage of being able to also increase the $CD4^+$ T cell count. Care should be taken though, not to confuse a lower drug efficacy to automatically imply a lower pill or drug intake.

There is currently no clinical evidence that shows how immunosuppressive and cell death rate accelerating therapies affect macrophages. It can be intuitively assumed though, that these drugs would have some effect on macrophage cells as they do on $CD4^+$ T cells. However, the effect of immune based therapies on macrophage cells will not be considered in this thesis.

4.1.3   Combining HAART with Immune Based Therapies

There are numerous studies that have been undertaken on the concomitant use of some immune based therapies with HAART. In particular, there are some studies that have investigated the possible inclusion of the immunosuppressive drugs in HAART regimens. The intention is to improve the efficacy of HAART. This section discusses some of the results that were obtained from such clinical trials.

The non-dependence of the $CD4^+$ T cell steady state on the $CD4^+$ T cell specific therapy efficacy (η_{sl} and η_{ps}) explains why some of the clinical trial outcomes have reported increased drug toxicity with no added efficacy, when the immunosuppressive drug hydroxyurea is added to HAART regimens [122, 123, 124]. As pointed out before in section 4.1.2, proliferation suppressive therapy has marginal reduction of the viral load, and is not capable of modifying the $CD4^+$ T cell steady state count. This analysis suggests that the addition of proliferation suppressive or source rate limiting therapy to replication cycle based HAART regimens, is an ill conceived idea. An option that could be considered is adding infected $CD4^+$ T cell death accelerators to the HAART regimen.

4.1.4   Conclusions

The following conclusions can be drawn from the analysis that was carried out in this section.

1. The end result of HIV therapy is to move the pre-treatment viral load to a treatment steady state. This treatment steady state is independent of when therapy is initiated, but is dependent on the individual's viral and host cell parameters, as well as the combined drug efficacy.

2. Initiating therapy when the viral load is below this treatment steady state will result in an increasing viral load, which will be perceived as failure to control the

viral load. When therapy is initiated with a viral load that is higher than this treatment steady state though, there will be an observable reduction in the viral load.

3. This dependence of the treatment steady state on model parameters and drug efficacy, also gives an indication whether or not durable viral load suppression to below detectable levels is attainable for a particular individual.

4. The best way to ensure durable suppression of the viral load to below 50 copies therefore, is to select a drug dosage that has a treatment steady state of less than 50 copies per mL of plasma.

5. The dependence of the steady states on the individuals model parameters explains why there is inter-individual variations in viral load and CD4$^+$ T cell set-points, as well as why some individuals have virologic failure on therapy that is highly effective on others.

6. For replication cycle based therapy with the latently infected cell model, the viral load steady state is sensitive to small changes in drug efficacy when one is operating around the detection cutoff (50 copies mL^{-1} steady state efficacy. However, [71] has since shown that this sensitivity is eliminated for the extended model.

7. Immunosuppressive therapy has no effect on the CD4$^+$ T cell treatment steady state. This suggests that its concomitant use with replication cycle based HAART is a bad idea. This is further supported by some clinical trial outcomes that show increased toxicity with no added efficacy with their use in HAART regimens. The better option is adding infected CD4$^+$ T cell death accelerators to the HAART regimen. This would particularly be useful in cases where the individual has marginal cell gains with HAART.

8. The use of target cell death accelerating therapies at first sight may appear counter intuitive because current focus of therapy is to increase CD4$^+$ T cell counts. However, these therapies will not 'harm' the immune system in so far as the CD4$^+$ T cell counts are concerned.

9. Generally, immune based therapies, from an end point efficacy perspective, are not capable of suppressing the viral load as effectively as replication cycle based therapies.

When using replication cycle based drugs (RTI and PI), the viral load steady state is determined by the combined efficacy for these classes of antiretroviral agents. This implies that using two low efficacy drugs from each class can be as good as using moderate

efficacy therapy from a single class. Furthermore, using low efficacy and moderate efficacy from different classes can be as good as using a high efficacy therapy from a single class.

The analysis suggests that it is possible for the uninfected CD4$^+$ T cells to rebound to pre-infection values under therapy. However, this is not true for many infected individuals who attain maximal viral load suppression with inadequate CD4$^+$ T cell gains (virologic success with immunologic failure). There is therefore, a need to find ways of quantifying organ health or damage to the immune system, so that it can be factored into the equations describing the CD4$^+$ T cell dynamics.

4.2   Transient Response Analysis

The steady state analysis presented in section 4.1 has shown that the long term end effect of therapy is to move the viral load and target cell counts from their pre-treatment values to their treatment steady states. However, the steady state analysis does not explain the dynamics underlying this transition. Transient response analysis [170] for the HIV/AIDS models therefore needs to be conducted. Maximal suppression of the viral load to below detectable levels using HAART is attainable. However, the durability of such suppression has proved in many cases, to be elusive as the virus often rebounds (blips) after periods of effective suppression [22, 23, 24, 25]. Issues that are of concern therefore, are the suppression of the viral load to below detectable levels and the ability of the drugs to maintain such suppression once attained.

This section focuses on the attainment of maximal and durable suppression of the viral load. Transient response analysis of an HIV/AIDS model will be used to determine the extent to which the viral load is suppressible at different stages of the progression of the HIV/AIDS infection. The intention is to determine when, as the HIV infection progresses, initiating therapy is most likely to attain maximal and durable suppression of the viral load, as well as the conditions that are conducive to the minimization of viral load blips. If response is related to the infection stage, then this study can also help decide when best to initiate therapy. Simulations are used to show the effect on the viral load of initiating therapy at different stages of the HIV infection.

4.2.1   Analysis with the Latently Infected Cell Model

An approximate analysis by linearizing the nonlinear equations in (4.5)-(4.9) with the combined use of reverse transcriptase and protease inhibitors can be obtained. The Jacobian when evaluated at an operating point $[T^o, T_l^o, T_a^o, V_i^o, V_n^o]^T$ along the state trajectory

---

from initial infection is given by:

$$
\mathbf{A_L} = \begin{bmatrix}
\kappa_1 & 0 & 0 & -\beta_T T^o & 0 \\
q_l \beta_T V_i^o & -(\delta_l + k) & 0 & q_l \beta_T T^o & 0 \\
q_a \beta_T V_i^o & k & -\delta_a & q_a \beta_T T^o & 0 \\
0 & 0 & r_T & -c & 0 \\
0 & 0 & 0 & 0 & -c
\end{bmatrix}
$$

where $\kappa_1 = p(1 - 2T^o/T_m) - d_T - \beta_T V_i^o$. The assumption is that the individual is treatment naïve, that is, has no prior exposure to therapy.

It is clear from the entries in matrix $\mathbf{A_L}$, that the eigenvalues are functions of the viral load V^o and the CD4$^+$ T cell count T^o, at the instance when therapy is initiated. Since these measurements vary with the stage of the HIV infection, then the eigenvalues can also be expected to vary with the stage of the HIV infection, that is, with when therapy is initiated.

The first eigenvalue is $\lambda_1 = -c$, while the remaining four are the solutions to

$$
\begin{aligned}
(\lambda + c)(\lambda + \delta_a)(\lambda + \delta_l + k)(\lambda - \kappa_1) & \\
-r_T \beta_T T^o(\lambda - \kappa_1 - \beta_T V_i)(kq_l + q_a(\lambda + \delta_l + k)) &= 0
\end{aligned} \tag{4.24}
$$

The eigenvalues are either all real or have a complex pair depending on the stage as the infection progresses. There is therefore, a bifurcation of λ_4 and λ_5 as they are distinct at some infection stages and are a conjugate pair at other infection stages, as depicted in figure 4.4. For the infection stage where λ_4 and λ_5 are a conjugate pair, the viral load transient response will have the general form

$$
\tilde{v}_c(t) = A_1 e^{\lambda_1 t} + A_2 e^{\lambda_2 t} + A_3 e^{\lambda_3 t} + A_4 e^{\sigma t} \cos(\omega t + \phi) + v_{ss} \tag{4.25}
$$

where λ_1, λ_2 and λ_3 are the real eigenvalues, while ω is the frequency of oscillation and σ is the transient decay rate for the complex pair. This implies that when therapy is initiated, the response will oscillate about the steady state before settling. The magnitude of the oscillation and its decay rate will be parameter, drug efficacy and infection stage (timing) dependent. Figure 4.5 illustrates the variation of ω and σ with the combined HAART (RTI and PI) efficacy when therapy is initiated at the asymptomatic stage.
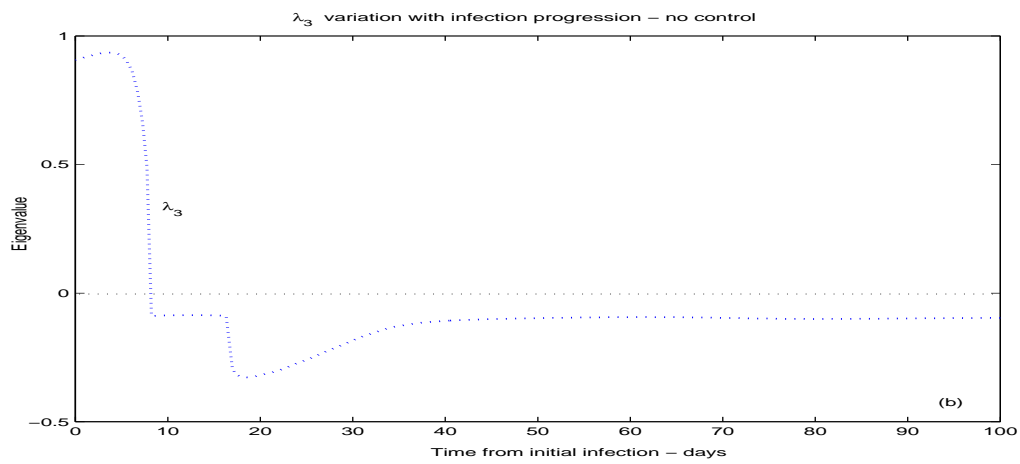
For the infection stages where λ_4 and λ_5 are distinct, the viral load transient response will have the general form

$$
\tilde{v}_r(t) = A_1 e^{\lambda_1 t} + A_2 e^{\lambda_2 t} + A_3 e^{\lambda_3 t} + A_4 e^{\lambda_4 t} + A_5 e^{\lambda_5 t} + v_{ss} \tag{4.26}
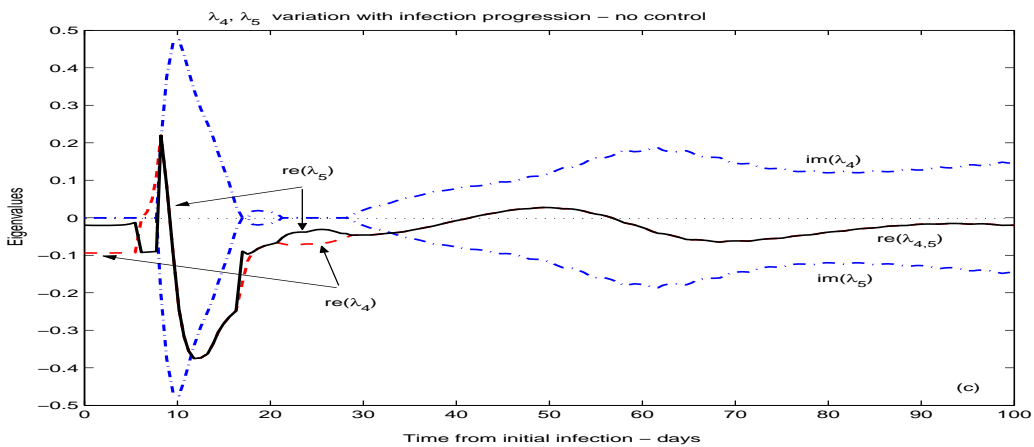$$

In this case, one can expect the transition to the treatment steady state to be smooth.
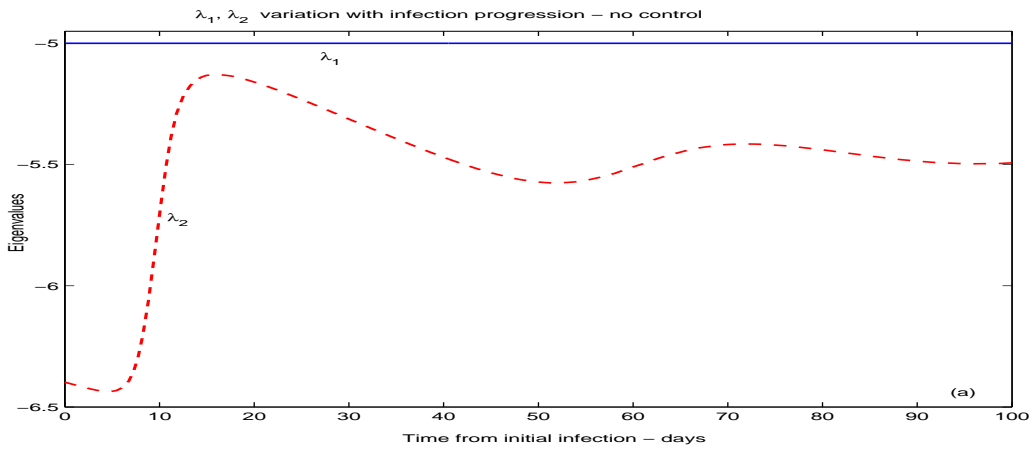
(a)

(b)

(c)

Figure 4.4: Eigenvalue variation with infection progression for the latently infected cell model. $X(0) = [1000, 0, 0, 1]^T$. Parameters are in Table A.1.
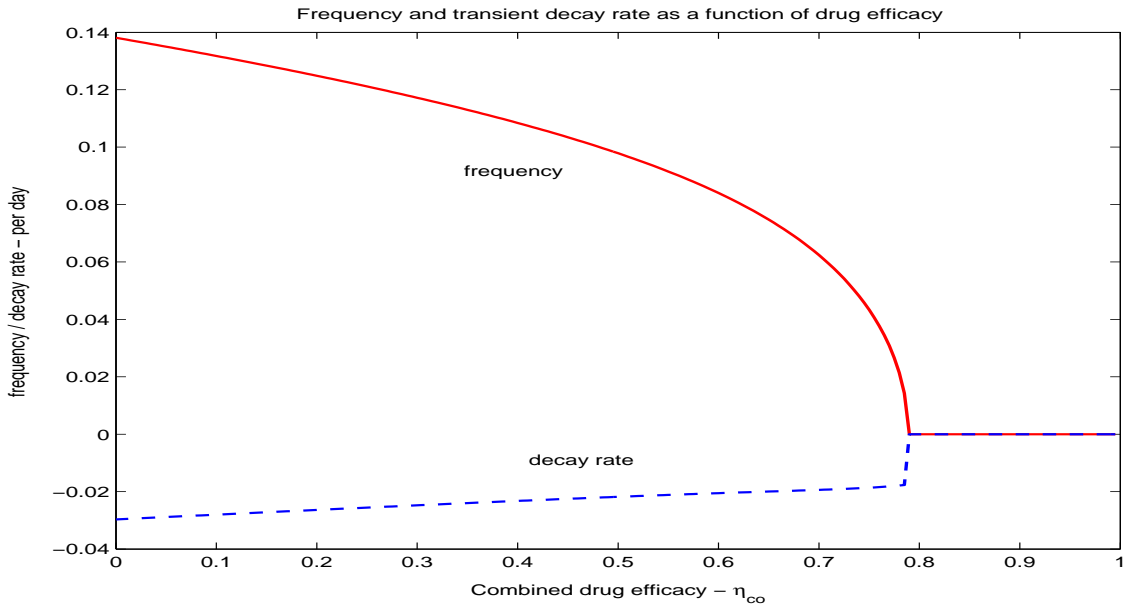
Figure 4.5: Complex eigenvalue variation with combined HAART (RTI and PI) efficacy when therapy is initiated at the asymptomatic stage. Parameters are in Table A.1.

The absolute values for λ_1 and λ_2 are relatively large, hence their respective transients will die out quickly ($\tau \leq 0.2$ days). The variation in response is therefore mainly due to the variation in λ_3 and the complex eigenvalue parameters ω and σ. This variation in ω and σ also implies that initiating therapy at some infection stages can be expected to result in a more oscillatory transition to the steady state than at other stages. For instances where all the eigenvalues are real, then the transition to the treatment steady state should be smoother. This analysis is consistent with the often observed 'viral load blips' under HAART, where a blip is defined as a transient rebound of plasma viremia after suppression has been attained.

The solutions (4.25) and (4.26) are approximations of the true response. However, the intention was more to understand the nature or form of the response that one can expect, than it was to obtain an expression for the true response. Now that the type of response one can expect is determined by this approximate linearization, it will then be verified using the nonlinear model.

Figure 4.6 shows how the viral load responds to therapy when it is initiated at various stages of the infection, using the same drug efficacy. In particular, the figure shows how viral load suppression depends on when therapy is initiated. It can be seen that a fixed drug dosage can be suppressive at one stage of the infection, but fail when therapy is initiated too early.
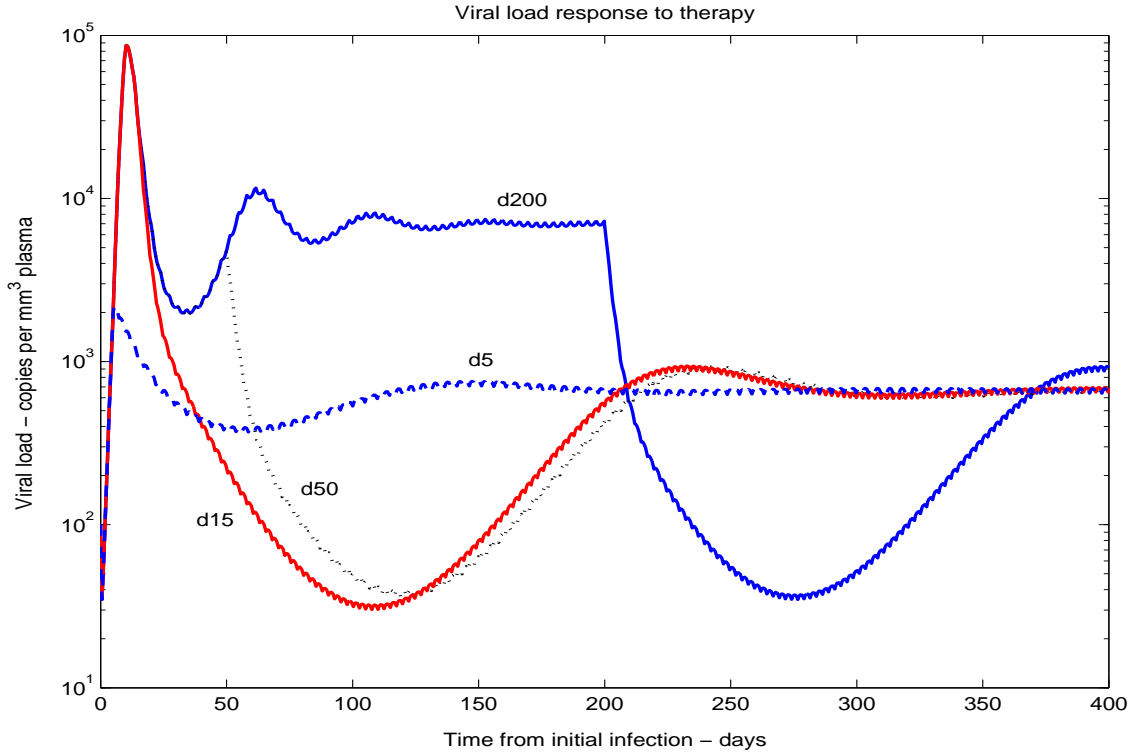
Figure 4.6: Response to therapy when initiated at various stages of the HIV infection. d200:day 200; d5:day 5; d50:day 50; d15:day 15. Drug efficacy $\eta_{rt} = 0.7$ Parameters are in Table A.1.

On the other hand, figure 4.7 shows how the viral load responds to therapy when it is initiated at the asymptomatic stage of the infection. In this case, the figure shows how viral load suppression depends on drug efficacy.

4.2.2   Analysis with the Extended Model

The Jacobians for the extended model when evaluated at an operating point $[T^o, T_l^o, T_a^o, M^o, M^{*o}, V_i^o, V_n^o]^T$ along the state trajectory from initial infection is given by:

$$
\mathbf{A_E} =
\begin{bmatrix}
\kappa_1 & 0 & 0 & 0 & 0 & -\beta_T T_{ss} & 0 \\
q_l \beta_T V^o & -(k+\delta_l) & 0 & 0 & 0 & q_l \beta_T T^o & 0 \\
q_a \beta_T V^o & k & -\delta_a & 0 & 0 & q_a \beta_T T^o & 0 \\
0 & 0 & 0 & -(d_M + \beta_M V^o) & 0 & -\beta_M M^o & 0 \\
0 & 0 & 0 & q_M \beta_M V^o & -\mu & q_M \beta_M M^o & 0 \\
0 & 0 & r_T & 0 & r_M & -c & 0 \\
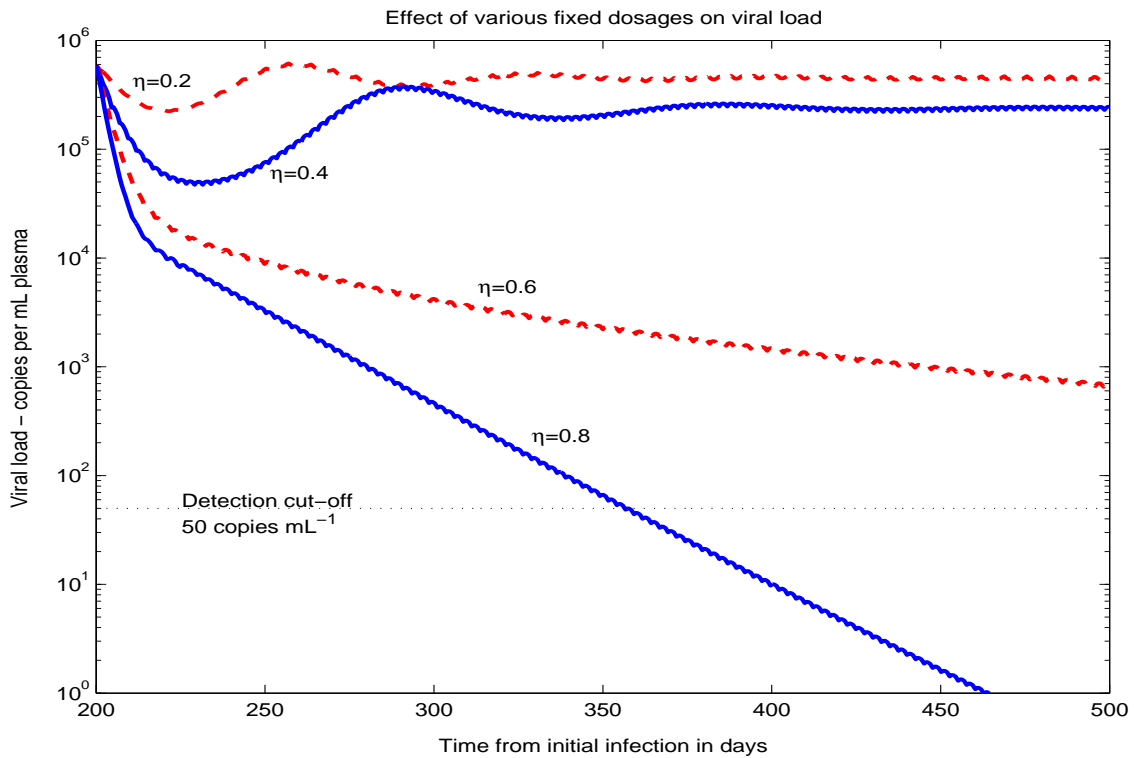0 & 0 & 0 & 0 & 0 & 0 & -c
\end{bmatrix}
$$

Figure 4.7: Response to therapy when initiated at the asymptomatic stage of the infection using a varying drug efficacy. Parameters are in Table A.2.

The entries in matrix $\mathbf{A_E}$ show that the eigenvalues will be functions of the viral load V^o, the uninfected macrophage cell count M^o, and the uninfected CD4$^+$ T cell count T^o, at the instance therapy is initiated. And similarly, the eigenvalues, and hence the response to therapy, can also be expected to vary depending on when therapy is initiated.

4.2.3   On Attaining Maximal and Durable Suppression of the Viral Load

Viral load suppression is considered to be maximal when the viral load reaches below levels of detection by the currently available assays. Given a fixed drug efficacy, the initial viral load and CD4$^+$ T cell count, the minimum viral load and the duration of viral load suppression to below 50 copies per mL of plasma, if attainable, can be determined from either (4.25) or (4.26). Even though this minimum value below 50 copies per mL that the viral load can be reduced to is currently not clinically quantifiable as it can not be readily measured. However, an estimate of the minimum value that the viral load can be reduced to is useful in determining whether viral load suppression to below detectable levels can be attained, for the given drug dosage and instance when therapy is initiated.

Eigenvalue variation gives an indication of how the duration of viral load suppression is expected to vary as the infection progresses. Higher values of ω indicate that if viral load suppression is attained, then it will be short lived. Similarly, higher values of σ indicate that the transition to the steady state will be more rapid. Very early initiation of therapy can therefore be expected to result in a shorter viral load suppression period and a rapid transition to the steady state, which is indicated by the relatively higher values of ω and σ. Late therapy during the asymptomatic stage will most likely result in a prolonged viral load suppression period. Increasing the drug dose results in a longer period of viral load suppression, which is indicated by the decreasing values of ω and σ.

Given the foregoing, maximal suppression of the viral load is most likely to be attained when therapy is initiated during the asymptomatic, very early and late acute infection stages of the infection, or when high drug doses are used. It is therefore possible for a drug dosage to be suppressive at one stage of the infection, but fail when therapy is initiated at a different stage. For a given instance when therapy is initiated, then the degree and duration of viral load suppression is drug efficacy dependent.

Durable Suppression

The best way to ensure a more durable suppression of the viral load will be to select a drug dosage that has a treatment steady state of at most 50 copies per mL of plasma. Alternatively, durable viral load suppression can be attained when therapy is initiated at a time when the associated complex eigenvalues have a smaller frequency component. If any viral load suppression to below the treatment steady state is attained, then it will be long lived as the virus slowly rebounds before settling to the treatment steady state value. It is reasonable to assume that the lower to below detectable levels the viral load is suppressed to, the longer the duration of suppression will be. Maximal suppression therefore, generally implies durable suppression. Durable suppression is therefore also most likely to be attained when therapy is initiated during the asymptomatic, very early and late acute infection stages if the infection, or when a drug dosage that can attain a viral load treatment steady state below 50 copies per mL of plasma is used.

Viral Load Rebounds Under Therapy

Since a viral load blip is a transient rebound of the viral load to above a specified value after maximal suppression has been attained, this means that viral load blips are more likely to occur when drug doses with steady states that are higher than the set value are used (usually over 400 or 500 copies per mL of plasma). For a fixed drug dosage and

constant frequency ω, blips are most likely to occur when therapy is initiated at a stage where the associated transient decay rate σ is low. Rebounds can be expected to be more pronounced when transient viral load suppression is attained with lower drug doses, as is the case when therapy is initiated during the asymptomatic stage. Initiating therapy very early during the acute infection stage could minimize viral load blips, but result in a shorter duration of suppression. The difference in peak viral load rebounds for the different infection stages, except for the very early acute infection stage, is not really that significant, which seems to suggest that viral load rebounds are more drug efficacy and parameter variation dependent, than they are timing dependent. The sensitivity of the steady state to slight changes in efficacy could also be the cause of transient rebounds. These slight efficacy variations could be due to daily variations in drug metabolism, for example.

4.2.4   Conclusions

The following conclusions can be drawn from the analysis presented in this section:

1. The transition from the pre-treatment steady state to the treatment steady state is oscillatory. The response is parameter and drug efficacy dependent, but unlike the treatment steady state, it is also dependent on when therapy is initiated.

2. If a drug is capable of driving the viral load to a particular steady state value, then the duration of the viral load suppression to below this steady state value can be maximized by choosing the right time to initiate therapy.

3. Initiating therapy during the mid acute infection stage when the viral load is high, results in a faster transition to the treatment steady state. This however, implies a shorter viral load suppression period.

4. Initiating therapy during the asymptomatic stage of the infection however, will result in a more durable suppression of the viral load. Furthermore, maximal suppression implies a more durable suppression of the viral load.

5. Viral load blips will occur whenever the drug efficacy is such that the viral load treatment steady state is higher than the figure that is set as indicative of a blip. This value in most cases is set at 500 copies per mL of plasma.

6. A point to note is that, from the steady state analysis, there is not much difference in the required drug efficacy to attain steady state viral loads of 50 or 500 copies per mL of plasma. Therefore, it is possible that blips could be caused by slight variations in efficacy due to daily variations in drug pharmacokinetics, or other such extenuating factors.

7. The stage of the infection where viral load blips are most likely to occur is not very clear, but it is clear that viral load rebounds under therapy do not necessarily imply virologic failure. It could be that the drug dosage is not capable of continued viral suppression.

8. However, if the individual's parameters are known, it is then possible to estimate the duration of viral load suppression, as well as anticipate the magnitude and timing of the viral load blip.

9. This study therefore, puts emphasis on the need to estimate parameters and individualize antiretroviral therapy.

4.3   Interruption of Highly Active Antiretroviral Therapy

The viral load will rebound and CD4$^+$ T cell counts will decline during HAART interruption. This is considered as undesirable, especially when the CD4$^+$ T cell count did not rebound adequately while one was still on HAART. There is therefore, need to investigate options, therapeutic or otherwise, that will slow down the viral load rebound and/or CD4$^+$ T cell decline during HAART interruption.

It has been explained in section 4.1.3 why adding immunosuppressive drugs to HAART regimens that are already capable of maximally suppressing the viral load has marginal benefits. In any case, these drugs are on their own, only capable of slightly reducing the viral load set point. This section deals specifically with investigating the possible use of CD4$^+$ T cell specific immune based therapies as adjuvant to HAART interruptions. The intention is to slow down viral load rebounds when HAART is interrupted.

4.3.1   Anti-CD4 Therapy as Adjuvant to HAART Interruption

Viral load rebounds occur because HAART interruptions "induce sudden antigenic activation with high peaks of viral load, up to a set-point of viral load or higher, which infects new populations of activated CD4 T cells .." [120]. More precisely, viral load rebounds occur because (refer to section 2.4):

1. Viral load suppression with HAART does not necessarily imply a reconstitution of HIV specific immune responses [127].

2. There is over stimulation of the immune system during infection [92].

3. The availability of new target cells due to cell gains incurred during HAART [174].

Strategies that reduce cell activation or limit the population of available target cells can therefore be expected to slow down the rebound rate. These immune based CD4$^+$ T cell specific therapies are referred to as anti-CD4 therapy, since they are perceived to counteract the increase of CD4$^+$ T cells. These therapies are considered to have no anti-viral activities because they do not directly interfere with the viruses replication cycle. Instead, they manipulate the population dynamics of the target CD4$^+$ T cells. As explained before, anti-CD4 therapy could include:

1. Reducing CD4$^+$ T cell proliferation rate p, referred to as proliferation suppressive therapy.
2. Reducing CD4$^+$ T cell source rate s_T, referred to as source limiting therapy.
3. Accelerating infected CD4$^+$ T cell death rates δ_l and δ_a, referred to as cell death accelerating therapy.
4. Accelerating target CD4$^+$ T cell death rate d_T, referred to as apoptosis inducing therapy.

These therapies are referred to as Anti-CD4 therapy, as they are perceived to be suppressing the expansion of the CD4$^+$ T cell pools.

4.3.2   HAART Interruption with the Latently Infected Cell Model

When HAART is on, the steady states for the viral load and uninfected CD4$^+$ T cell count for the latently infected cell model were given by equations (4.1) and (??) in section 4.1.1, and are repeated here for ease of reference.

$$T_{ss} \quad = \quad \frac{c\delta_a}{(1-\eta_{co})q\beta_T r_T} \tag{4.27}$$

$$V_{iss} \quad = \quad \frac{(1-\eta_{co})qr_T s_T}{c\delta_a} - \frac{d_T}{\beta_T} + \frac{p}{\beta_T}(1 - \frac{c\delta_a}{(1-\eta_{co})q\beta_T r_T T_m}) \tag{4.28}$$

where $q = q_a + q_l(\frac{k}{k+\delta_l})$. When HAART is interrupted, the system will move to another steady state. When proliferation suppressive therapy is used during this period, then the steady states will be :

$$T_{ss} \quad = \quad \frac{c\delta_a}{q\beta_T r_T} \tag{4.29}$$

$$V_{iss} \quad = \quad \frac{qr_T s_T}{c\delta_a} - \frac{d_T}{\beta_T} + \frac{(1-\eta_{ps})p}{\beta_T}(1 - \frac{c\delta_a}{q\beta_T r_T T_m}) \tag{4.30}$$

Similarly, the steady states when cell death rate accelerating therapy is used will be:

$$T_{ss} \quad = \quad \frac{(1+\eta_{da})c\delta_a}{q_{da}\beta_T r_T} \tag{4.31}$$

$$V_{iss} \quad = \quad \frac{q_{da}r_T s_T}{(1+\eta_{da})c\delta_a} - \frac{(1+\eta_{ap})d_T}{\beta_T} + \frac{p}{\beta_T}(1 - \frac{(1+\eta_{da})c\delta_a}{q_{da}\beta_T r_T T_m}) \tag{4.32}$$

where, $q_{da} = q_a + q_l(\frac{k}{k+(1+\eta_{da})\delta_l})$.
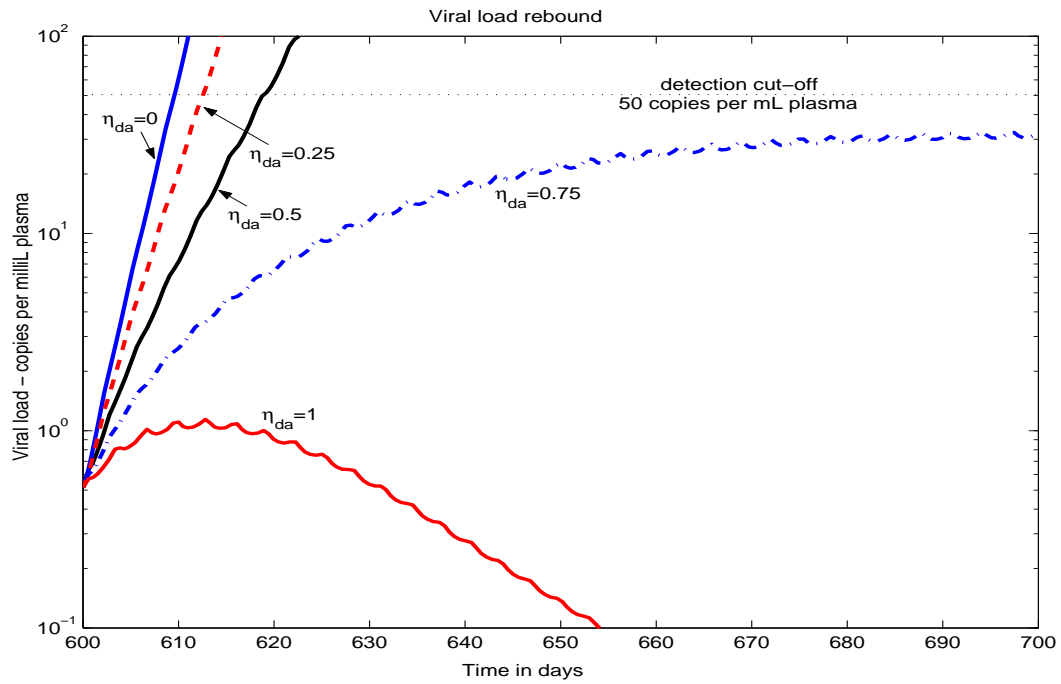
The use of anti-CD4 therapies at first sight may appear counter intuitive because current focus of therapy is to increase $CD4^+$ T cell counts. However, anti-CD4 therapies will not 'harm' the immune system. This is made apparent by examining equations (4.29) and (4.31). These equations show that when proliferation suppressive therapy is used during HAART interruption, then the $CD4^+$ T cell count will decline to the pre-HAART steady state value. The $CD4^+$ T cell count will also decline when cell death accelerators are used during HAART interruption. However, unlike with proliferation suppressors, this value will be higher than the pre-HAART steady state, and the percentage increase will be equal to the drug efficacy η_{da}.

On a similar note, it is clear from equations (4.30) and (4.32) how the use of anti-CD4 therapies during HAART interruption can prevent the viral load from rebounding to pre-HAART values. Their ability to reduce the viral load set point could imply that their use during HAART interruptions could reduce the peak viral load. This in turn could imply that the rate of viral load rebound is reduced, and consequently this would prolong the duration of HAART interruption. It is also clear that cell death accelerators, from an end point efficacy perspective, would have better viral load control than proliferation suppressors.

Figure 4.8(a) shows how one could expect viral load rebound to be slowed down when cell death accelerating anti-CD4 therapy is used during HAART interruption. For this case, viral suppression can be maintained for almost 10 days without therapy. Also, significantly longer HAART interruption periods can be attained by the use of anti-CD4 therapy. One however, needs to weigh the advantages of increasing OFF HAART periods against the sacrifice of having no drug free days.

Figure 4.8(b) shows that using anti-CD4 therapy results in a rapid initial decline in $CD4^+$ T cell counts. This is followed by a slower decline as $CD4^+$ T cell counts will eventually settle at a value that is higher than when no drugs are used during HAART interruption. When considering the sacrifice of having no drug free days then, one should also consider the benefits of long term $CD4^+$ T cell gains.

Analysis of the virus and target cell dynamics during HAART interruption with the extended model should illustrate the same concept: that antiCD4 therapy could be used to slow down viral load rebounds that occur when HAART is interrupted.
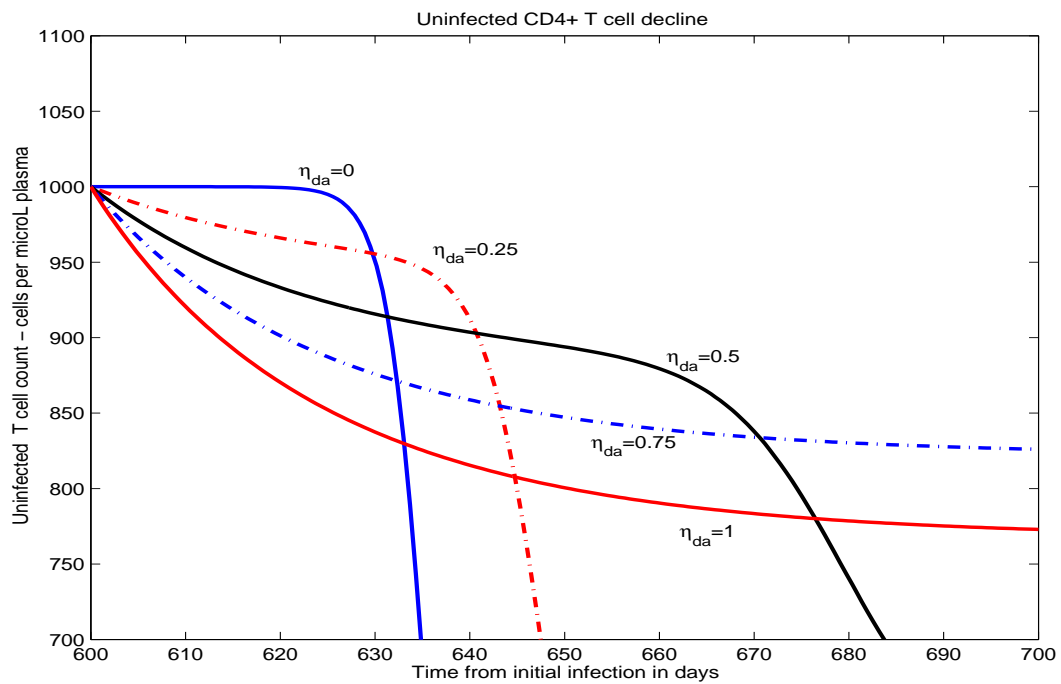
Figure 4.8: Viral load rebound and CD4$^+$ T cell decline when anti-CD4 therapy is used during HAART interruption. $\eta_{da} = \eta_{ap}$: cell death acceleration factors. HAART was previously ON for 300 days with $\eta_{rt} = 0.7$ (a) Viral load (b) CD4$^+$ T cells. Parameters are in Table A.2.

4.3.3   Conclusions

It is not possible for some HIV infected individuals to attain long term viral load control when HAART is interrupted. A possible way to prolong OFF HAART periods for these individuals is to use anti-CD4 immune based therapies when HAART is interrupted. The use of anti-CD4 therapies at first sight may appear counter intuitive because current focus of therapy is to increase CD4$^+$ T cell counts. However, anti-CD4 therapy has a potential to prolong the duration of HAART interruptions without excessive reduction in CD4$^+$ T cell counts. Even though anti-CD4 therapy initially accelerates CD4$^+$ T cell depletion, it does not necessarily, in the long term harm the immune system in so far as CD4$^+$ T cell counts are concerned.

What remains then, is to determine the efficacy of the anti-CD4 therapy required for pre-determined ON/OFF HAART periods. Conversely, if the efficacy is known, then the ON/OFF HAART period can be determined. The optimal ON/OFF HAART periods also need to be determined. For the protocol considered in this section, it was assumed that the immune based therapy is to be used for the entire duration that HAART is interrupted. The possibility of whether this therapy's use during OFF HAART periods can be limited to include schedules with drug holidays, where the individual is either ON HAART, or ON Anti-CD4 therapy, or completely OFF all drugs, needs to be explored.

4.4   Controllability Analysis

The viral load is considered to be controllable if the control law in use can reduce it by 90% in 8 weeks from the time treatment is initiated and continue to suppress it to below 50 copies per mL of plasma in less than 6 months [1]. There are numerous reported cases where the viral load of individuals has been suppressed to below detectable levels with the use of HAART. So, from a medical/clinical perspective, it seems that the viral load of many HIV infected persons is controllable by the available antiretroviral drugs.

Controllability analysis for the basic 3D HIV/AIDS model has been previously addressed by [43, 107]. Analysis will therefore be carried out for the latently infected cell model Σ_L (4.33) and the co-circulating target cell extended model Σ_E (4.34) with both HAART and immune based therapies.

$$\Sigma_L = \begin{cases} \frac{dT}{dt} &= s_T + (1-\eta_{ps})pT(1-\frac{T}{T_m}) - d_TT - \beta_TTV_i \\ \frac{dT_l}{dt} &= (1-\eta_{rt})q_l\beta_TTV_i - (1+\eta_{da})\delta_lT_l - kT_l \\ \frac{dT_a}{dt} &= (1-\eta_{rt})q_a\beta_TTV_i - (1+\eta_{da})\delta_aT_a + kT_l \\ \frac{dV_i}{dt} &= (1-\eta_{pi})r_TT_a - cV_i \end{cases} \tag{4.33}$$

$$\Sigma_{Eco} = \begin{cases} \frac{dT}{dt} & = & s_T + (1 - \eta_{ps})pT(1 - \frac{T}{T_m}) - d_T T - \beta_T TV_i \\ \frac{dT_l}{dt} & = & (1 - \eta_{rt})q_l\beta_T TV_i - (1 + \eta_{da})\delta_l T_l - kT_l \\ \frac{dT_a}{dt} & = & (1 - \eta_{rt})q_a\beta_T TV_i - (1 + \eta_{da})\delta_a T_a + kT_l \\ \frac{dM}{dt} & = & s_M - d_M M - \beta_M MV_i \\ \frac{dM^*}{dt} & = & (1 - \alpha_{rt}\eta_{rt})q_M\beta_M MV_i - \mu M^* \\ \frac{dV_i}{dt} & = & (1 - \eta_{pi})r_T T_a + (1 - \alpha_{pi}\eta_{pi})r_M M^* - cV_i \end{cases} \tag{4.34}$$

4.4.1   Controllability

Controllability, by definition, "is a property of a system by which an input can be found that takes every state variable from a desired initial state to a desired final state in finite time" [170].

There are ways to determine if a system is controllable. The use of the controllability matrix is one such way. The controllability matrix for a linear system $\dot{x} = \mathbf{A}x + \mathbf{B}u$ is given by $\mathbf{M_c} = [\mathbf{B} \ \mathbf{AB} \ \mathbf{A}^2\mathbf{B} \ ... \ \mathbf{A}^{n-1}\mathbf{B}]$, where n is the system's order. Then the system is said to be controllable if rank $\mathbf{M_c} = n$. That is, if $\mathbf{M_c}$ is not singular. In this section therefore, a linear system's analytical tool will be applied to the linearized models.

For analysis with the latently infected cell model, the linearized model matrix $\mathbf{A_L}$ at an operating point $[\bar{T} \ \bar{T}_l \ \bar{T}_a \ \bar{V}_i]^T$ along the state trajectory from initial infection is:

$$\mathbf{A_L} = \begin{bmatrix} \kappa_1 & 0 & 0 & -\beta_T\bar{T} \\ q_l\beta_T\bar{V}_i & -(k+\delta_l) & 0 & q_l\beta_T\bar{T} \\ q_a\beta_T\bar{V}_i & k & -\delta_a & q_a\beta_T\bar{T} \\ 0 & 0 & r_T & -c \end{bmatrix}$$

where, $\kappa_1 = p(1 - 2\frac{\bar{T}}{T_m}) - (d_T + \beta_T\bar{V}_i)$. $\bar{T}$, and $\bar{V}_i$ are the uninfected CD4$^+$ T cell and viral load measurements at the respective point in the HIV infection progression.

Similarly for analysis with the extended model, the linearized model matrix $\mathbf{A_E}$ at an operating point $[\bar{T} \ \bar{T}_l \ \bar{T}_a \ \bar{M} \ \bar{M}^* \ \bar{V}_i]^T$ along the state trajectory from initial infection is:

$$\mathbf{A_E} = \begin{bmatrix} \kappa_1 & 0 & 0 & 0 & 0 & -\beta_T\bar{T} \\ q_l\beta_T\bar{V}_i & -(k+\delta_l) & 0 & 0 & 0 & q_l\beta_T\bar{T} \\ q_a\beta_T\bar{V}_i & k & -\delta_a & 0 & 0 & q_a\beta_T\bar{T} \\ 0 & 0 & 0 & -(d_M+\beta_M\bar{V}_i) & 0 & -\beta_M\bar{M} \\ 0 & 0 & 0 & \beta_M\bar{V}_i & -\mu & \beta_M\bar{M} \\ 0 & 0 & r_T & 0 & r_M & -c \end{bmatrix}$$

where, $\bar{M}$ is the uninfected macrophage measurement at the respective point in the HIV infection progression.

The assumption is that the individual is treatment naïve. The input matrices $\mathbf{B_L}$ for the latently infected cell model and $\mathbf{B_E}$ for the extended model, will depend on the type or combination of therapy that is under consideration.

4.4.2 Analysis with Replication Cycle Based HAART

Under mono class therapy using Reverse transcriptase inhibitors exclusively (PI sparing regimens), the input matrices for the latently infected cell model $\mathbf{B_{Lrt}}$ and the extended model $\mathbf{B_{Ert}}$ are respectively given by:

$$\mathbf{B_{Lrt}} = \beta_T \bar{V}_i \bar{T} \begin{bmatrix} 0 \\ -q_l \\ -q_a \\ 0 \end{bmatrix}$$

$$\mathbf{B_{Ert}} = \bar{V}_i \begin{bmatrix} 0 \\ -q_l \beta_T \bar{T} \\ -q_a \beta_T \bar{T} \\ 0 \\ -\alpha_{rt} \beta_M \bar{M} \\ 0 \end{bmatrix}$$

The controllability matrix $\mathbf{M_{Lrt}} = [\mathbf{B_{Lrt}} \quad \mathbf{A_L B_{Lrt}} \quad \mathbf{A_L}^2 \mathbf{B_{Lrt}} \quad \mathbf{A_L}^3 \mathbf{B_{Lrt}}]$ for the latently infected cell model is given by:

$$\mathbf{M_{Lrt}} = \beta_T \bar{T} \bar{V}_i \begin{bmatrix} 0 & 0 & q_a \beta_T \bar{T} r_T & \beta_T \bar{T} r_T (q_a \kappa_1 - \kappa_4) \\ -q_l & (k+\delta_l)q_l & -q_l \kappa_2 & q_l \beta_T \bar{T} r_T (q_a \beta_T \bar{V}_i q_a + \kappa_4) + (k+\delta_l)q_l \kappa_2 \\ -q_a & -kq_l + \delta_a q_a & \kappa_3 & q_a \beta_T \bar{T} r_T (q_a \beta_T \bar{V}_i q_a + \kappa_4) - kq_l \kappa_2 - \delta_a \kappa_3 \\ 0 & -r_T q_a & r_T \kappa_4 & r_T(\kappa_3 - c\kappa_4) \end{bmatrix}$$

where

$\kappa_2 = (k+\delta_l)^2 + q_a \beta_T \bar{T} r_T$

$\kappa_3 = kq_l(k+\delta_l+\delta_a) - q_a \delta_a^2 - q_a^2 \beta_T \bar{T} r_T$

$\kappa_4 = -kq_l + \delta_a q_a + q_a c.$

The matrix $\mathbf{M_{Lrt}}$ is not of full rank only when the viral load is zero or when the uninfected CD4$^+$ T cell count is zero. A zero viral load is invalid because the individual is assumed to be actively infected. When the CD4$^+$ 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 uninfected CD4$^+$ T cell count or the viral load is zero, are therefore controllable.

The controllability matrix $\mathbf{M_{Ert}} = [\mathbf{B_{Ert}} \quad \mathbf{A_E B_{Ert}} \quad \mathbf{A_E}^2 \mathbf{B_{Ert}} \quad \cdots \quad \mathbf{A_E}^5 \mathbf{B_{Ert}}]$ for the extended model can be determined in a similar manner. The matrix $\mathbf{M_{Ert}}$ is also not of full rank only when the viral load is zero or when both the uninfected CD4$^+$ T cell and uninfected macrophage counts are zero. Again, a zero viral load is invalid and when both the uninfected CD4$^+$ T cell and macrophage counts are zero, then the immune system is completely damaged and there is no point trying to control the virus. All other states apart from when the uninfected cell counts or the viral load are zero, are therefore controllable.

Under mono class therapy using protease inhibitors exclusively, the input matrices are respectively given by

$$\mathbf{B_{Lpi}} = r_T \bar{T}_a \begin{bmatrix} 0 \\ 0 \\ 0 \\ -1 \end{bmatrix}$$

$$\mathbf{B_{Epi}} = (r_T \bar{T}_a + \alpha_{pi} r_M \bar{M}^*) \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -1 \end{bmatrix}$$

where, $\bar{T}_a$ and $\bar{M}^*$ are the actively infected CD4$^+$ T cell and macrophage measurements at the respective point in the HIV infection progression.

The controllability matrix $\mathbf{M_{Lpi}} = [\mathbf{B_{Lpi}} \quad \mathbf{A_L B_{Lpi}} \quad \mathbf{A_L}^2 \mathbf{B_{Lpi}} \quad \mathbf{A_L}^3 \mathbf{B_{Lpi}}]$ for the latently infected cell model is given by:

$$\mathbf{M_{Lpi}} = r_T \bar{T}_a \begin{bmatrix} 0 & \beta_T \bar{T} & \beta_T \bar{T}(\kappa_1 - c) \\ 0 & -q_l \beta_T \bar{T} & q_l \beta_T \bar{T} \kappa_6 \\ 0 & -q_a \beta_T \bar{T} & \beta_T \bar{T}(q_a \beta_T \bar{V}_i + \kappa_4) \\ -1 & c & -\kappa_5 \end{bmatrix}$$

$$\begin{matrix} \beta_T \bar{T}(\kappa_1^2 - c\kappa_1 + \kappa_5) \\ q_l \beta_T \bar{T}(\beta_T \bar{V}_i(\kappa_1 - c) - \kappa_6(k + \delta_l) - \kappa_5) \\ \beta_T \bar{T}(q_a(\beta_T \bar{V}_i(\kappa_1 - c - \delta_a) - \kappa_5) + \kappa_6 k q_l - \delta_a \kappa_4) \\ \beta_T \bar{T} r_T(q_a \beta_T \bar{V}_i + \kappa_4) - c\kappa_5 \end{matrix} \Bigg]$$

where

$$\kappa_5 = q_a\beta_T\bar{T}r_T + c^2$$
$$\kappa_6 = \beta_T\bar{V}_i + k + \delta_l + c$$

Matrix $\mathbf{M_{Lpi}}$ is not of full rank when the actively infected CD4$^+$ T cell is zero or when the uninfected CD4$^+$ T cell count is zero. When the actively infected CD4$^+$ T cell count is zero, then no virus particles can be produced. All other states apart from when the actively infected or uninfected CD4$^+$ T cell counts are zero, are controllable.

The controllability matrix $\mathbf{M_{Epi}} = [\mathbf{B_{Epi}} \quad \mathbf{A_E}\mathbf{B_{Epi}} \quad \mathbf{A_E}^2\mathbf{B_{Epi}} \cdots \mathbf{A_E}^5\mathbf{B_{Epi}}]$ for the extended model can be determined in a similar manner. The matrix $\mathbf{M_{Epi}}$ is also not of full rank when both the actively infected CD4$^+$ T cell and infected macrophages counts are zero or when both the uninfected CD4$^+$ T cell and uninfected macrophage cell counts are zero. Again, when the infected macrophage cell and actively infected CD4$^+$ T cell counts are both zero, then no virus particles can be produced. All other states are controllable.

Under combined HAART using both reverse transcriptase and protease inhibitors, the input matrices are respectively given by

$$\mathbf{B_{Lco}} = \begin{bmatrix} 0 & 0 \\ -q_l\beta_T\bar{V}_i\bar{T} & 0 \\ -q_a\beta_T\bar{V}_i\bar{T} & 0 \\ 0 & -r_T\bar{T}_a \end{bmatrix}$$

$$\mathbf{B_{Eco}} = \begin{bmatrix} 0 & 0 \\ -q_l\beta_T\bar{V}_i\bar{T} & 0 \\ -q_a\beta_T\bar{V}_i\bar{T} & 0 \\ 0 & 0 \\ -\alpha_{rt}\beta_M\bar{V}_i\bar{M} & 0 \\ 0 & -(r_T\bar{T}_a + \alpha_{pi}r_M\bar{M}^*) \end{bmatrix}$$

The expression for the controllability matrix $\mathbf{M_{Lco}} = [\mathbf{B_{Lco}} \quad \mathbf{A_L}\mathbf{B_{Lco}} \quad \mathbf{A_L}^2\mathbf{B_{Lco}} \quad \mathbf{A_L}^3\mathbf{B_{Lco}}]$ for the latently infected cell model, as well as the expression for the controllability matrix $\mathbf{M_{Eco}} = [\mathbf{B_{Eco}} \quad \mathbf{A_E}\mathbf{B_{Eco}} \quad \mathbf{A_E}^2\mathbf{B_{Eco}} \ldots \mathbf{A_E}^5\mathbf{B_{Eco}}]$ for the extended model can likewise be determined.

4.4.3   Analysis with Immune Based Therapies

Without going into much detail, the input matrix $\mathbf{B_{Lps}}$ and controllability matrix $\mathbf{M_{Lps}}$ for the latently infected cell model, when proliferation suppressors are exclusively used, are respectively given by:

$$\mathbf{B_{Lps}} = \phi(\bar{T}) \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\mathbf{M_{Lps}} = \phi(\bar{T}) \begin{bmatrix} -1 & -\kappa_1 & -\kappa_1^2 & -\kappa_1^3 + q_a\beta_T^2\bar{T}\bar{V}_i r_T \\ 0 & -q_l\beta_T\bar{V}_i & q_l\beta_T\bar{V}_i(k+\delta_l-\kappa_1) & -q_l\beta_T\bar{V}_i(\kappa_1^2+\kappa_1(k+\delta_l)+\kappa_2) \\ 0 & -q_a\beta_T\bar{V}_i & \beta_T\bar{V}_i(q_a(\delta_a-\kappa_1)-q_l k) & \beta_T\bar{V}_i(q_a\kappa_1^2+\kappa_1(q_a\delta_a-kq_l)-\kappa_3) \\ 0 & 0 & -q_a\beta_T\bar{V}_i r_T & \beta_T\bar{V}_i r_T(\kappa_4-q_a\kappa_1) \end{bmatrix}$$

where
$\phi(\bar{T}) = p\bar{T}(1 - \frac{\bar{T}}{T_m})$

The matrix $\mathbf{M_{Lps}}$ is not of full rank only when the viral load is zero or when the uninfected CD4$^+$ T cell count is zero or when the uninfected CD4$^+$ T cell count has reached the proliferation shut down cell count. In the latter case, then no proliferation can take place at any rate, suppressed or otherwise. All other states, except when $V_i = 0$, or when $T = T_m$ or when $T = 0$ are controllable.

Similarly, the input matrix $\mathbf{B_{Lda}}$ and controllability matrix $\mathbf{M_{Lda}}$ for the latently infected cell model, when cell death accelerators are used, are respectively given by:

$$\mathbf{B_{Lda}} = \begin{bmatrix} 0 \\ -\delta_l\bar{T}_l \\ -\delta_a\bar{T}_a \\ 0 \end{bmatrix}$$

$$\mathbf{M_{Lda}} = \begin{bmatrix} 0 & 0 & \beta_T\bar{T}r_T\delta_a\bar{T}_a & \beta_T\bar{T}(\kappa_1 r_T\delta_a\bar{T}_a - \kappa_9) \\ -\delta_l\bar{T}_l & \delta_l\bar{T}_l(k+\delta_l) & \kappa_7 & q_l\beta_T\bar{T}(\beta_T\bar{V}r_T\delta_a\bar{T}_a+\kappa_9) - \kappa_7(k+\delta_l) \\ -\delta_a\bar{T}_a & -k\delta_l\bar{T}_l+\delta_a^2\bar{T}_a & \kappa_8 & q_a\beta_T\bar{T}(\beta_T\bar{V}r_T\delta_a\bar{T}_a+\kappa_9) + k\kappa_7 - \delta_a\kappa_8 \\ 0 & -r_T\delta_a\bar{T}_a & \kappa_9 & r_T\kappa_8 - c\kappa_9 \end{bmatrix}$$

where
$\kappa_7 = -\delta_l\bar{T}_l(k+\delta_l)^2 - q_l\beta_T\bar{T}r_T\delta_a\bar{T}_a$
$\kappa_8 = k\delta_l\bar{T}_l(k+\delta_l) + \delta_a(k\delta_l\bar{T}_l - \delta_a^2\bar{T}_a) - q_a\beta_T\bar{T}r_T\delta_a\bar{T}_a$

$$\kappa_9 = r_T(\delta_a^2 \bar{T}_a - k\delta_l \bar{T}_l) + cr_T\delta_a\bar{T}_a)$$

The controllability matrices $\mathbf{M_{Eps}}$ and $\mathbf{M_{Eda}}$ for the extended model, or any other desired matrix, can be determined and analyzed in a similar manner.

4.4.4   Singular Value Decomposition

Minimum singular value (MSV) decomposition [175] will be used as a measure of controllability. Since controllability varies along the state trajectory, this means that minimum singular value as a measure of controllability is an instantaneous concept. When minimum singular value decomposition is applied to the controllability matrices, an estimate measure of how controllable the system is at a particular time during the progression of the infection can be obtained, or the extent to which the various stages of the infection progression are controllable, can be determined. Singular value decomposition as a measure of controllability is a valid method to use here because the same variables are used to measure controllability and compared at different times as the infection progresses. It is usually desirable to have a relatively large singular value, as that implies an easier to control infection stage. The reasoning is that, the more controllable the viral load at a particular instance when therapy is initiated, the faster the transition to the treatment steady state, the lower the total drug intake during the transition period.

These 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 infection progresses. Comparison will be made between when therapy is initiated in the acute infection and asymptomatic stages of the infection. Comparisons will also be made between the different types and combinations of therapy.

The graphs labelled *pi*, *rt* and *co* in figure 4.9(b) show how the minimum singular value varies with time for the latently infected cell model. These graphs were generated using matrices $\mathbf{M_{Lrt}}$, $\mathbf{M_{Lpi}}$ and $\mathbf{M_{Lco}}$, respectively. Similarly, figure 4.10 shows how the minimum singular value varies with time for the extended model. Graphs were generated using matrices $\mathbf{M_{Ert}}$, $\mathbf{M_{Epi}}$ and $\mathbf{M_{Eco}}$, respectively.

The controllability profiles for the reverse transcriptase and protease inhibitors are similar in that where one controls most effectively, the other one does also. One can not however, draw any conclusions on which type of drug is more capable of suppressing the virus. Singular value decomposition as a controllability measure is not valid in this case.

Some stages of the HIV infection are more (or less) controllable than others. This is so because, for a particular drug, the minimum singular value varies with the instance
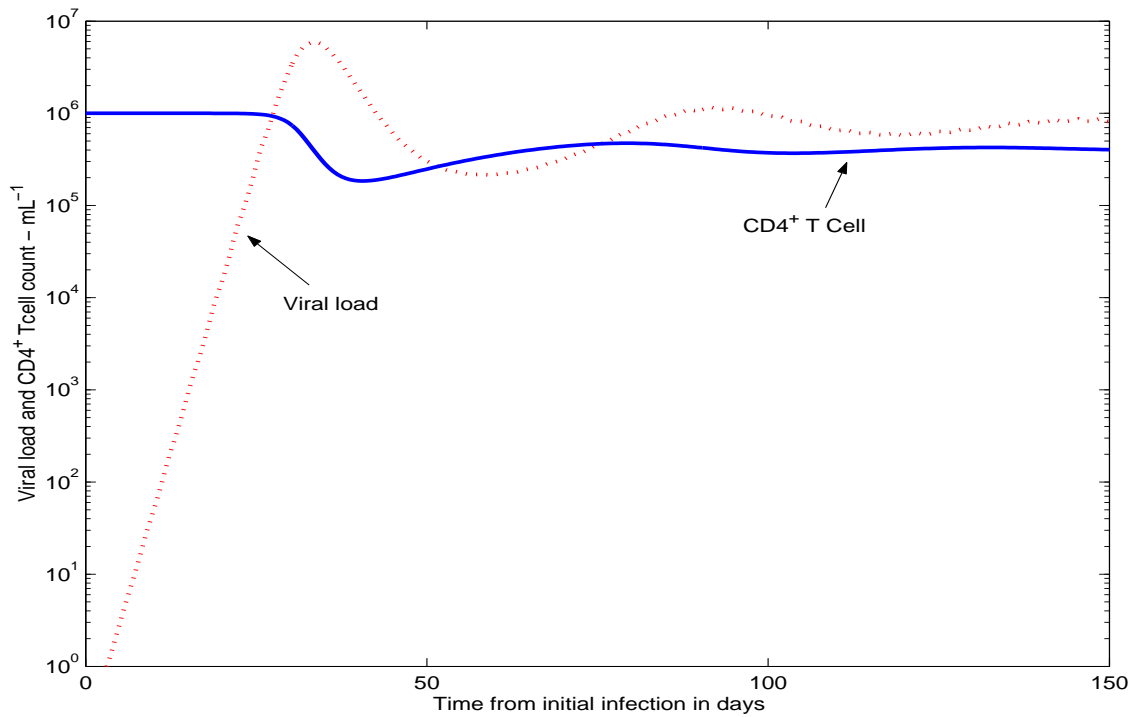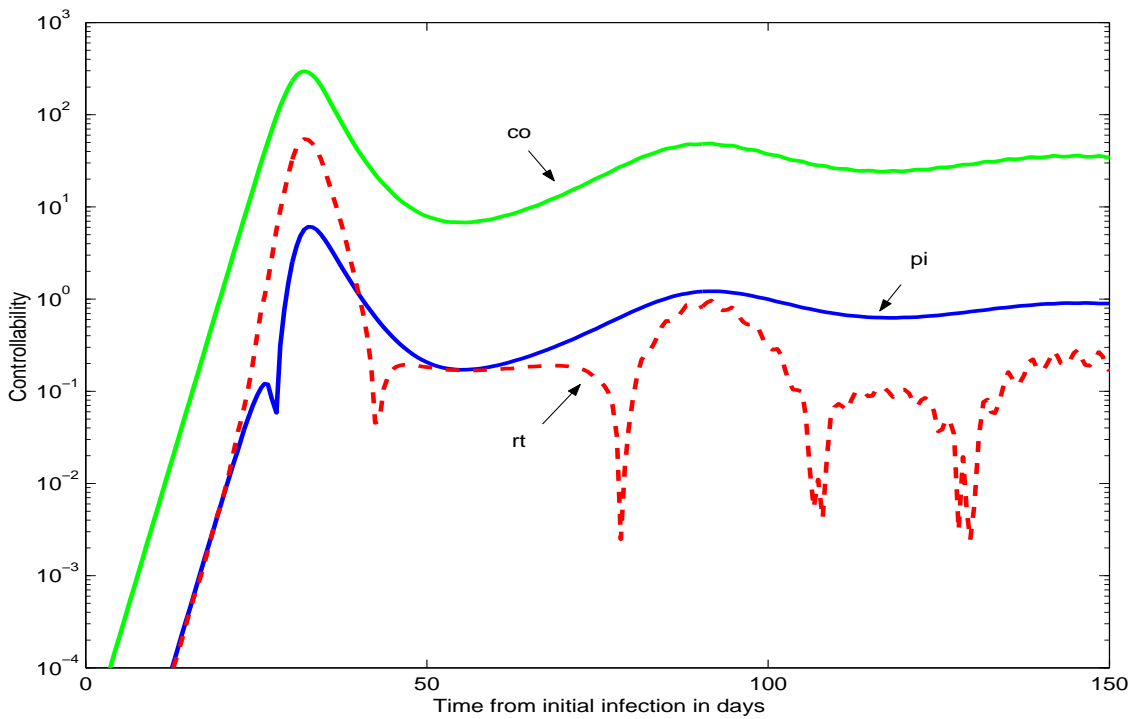
Figure 4.9: (a) Viral load and CD4$^+$ T cells. (b) Controllability to the asymptomatic stage for the latently infected cell model with replication cycle based therapies. rt: reverse transcriptase inhibitors, pi: protease inhibitors, co: combined rt and pi. Parameters are in Table A.4.
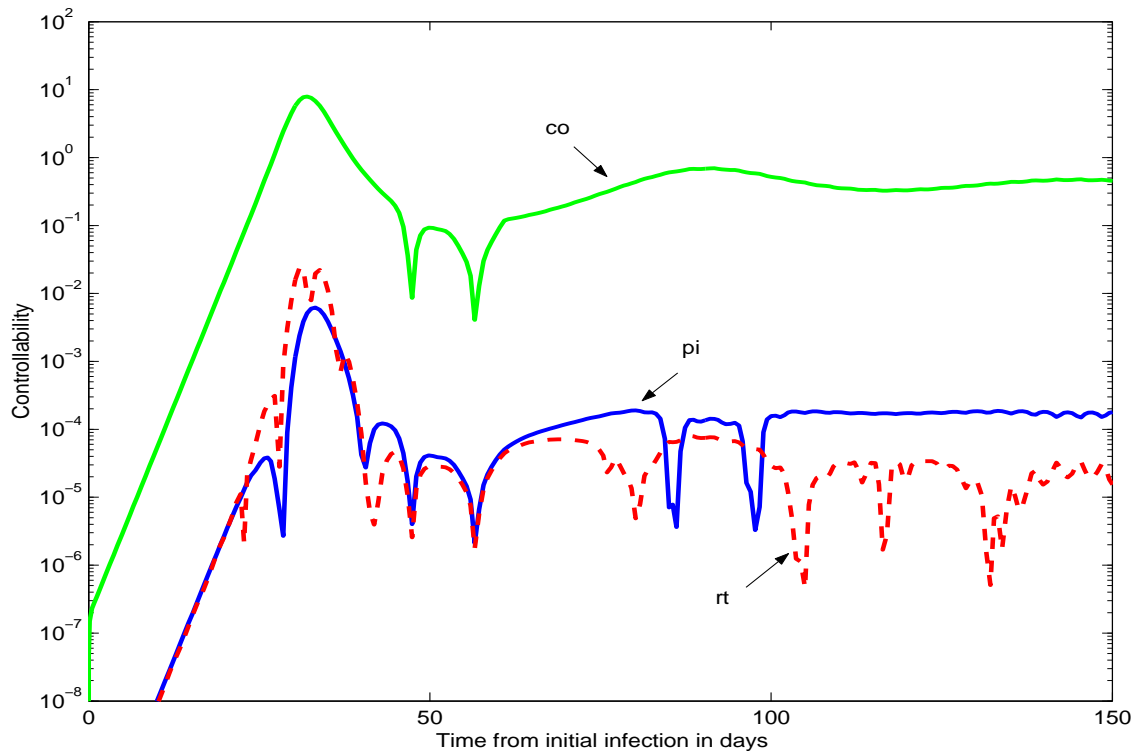
Figure 4.10: Controllability to the asymptomatic stage for the extended model with replication cycle based therapies. rt: reverse transcriptase inhibitors, pi: protease inhibitors, co: combined rt and pi. Parameters are in Table A.4.

during the infection, when therapy is initiated. 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. Initiating therapy therefore, when the viral load is easier to control implies the use of lower drug doses and consequently more bearable side effects.

Figures 4.9(b) and 4.10 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 from figure 4.9 that up to the asymptomatic stage, controllability and viral load are correlated, whereas there is no obvious correlation between the CD4$^+$ T cell or macrophage cell count with controllability. So it seems that from a viral load controllability perspective, the measured viral load at the initiation of therapy is a better prognostic indicator of virologic success, when compared to the CD4$^+$ T cell or macrophage cell counts.

For controllability with the use of immune based therapies, graphs 'ps' and 'da' in
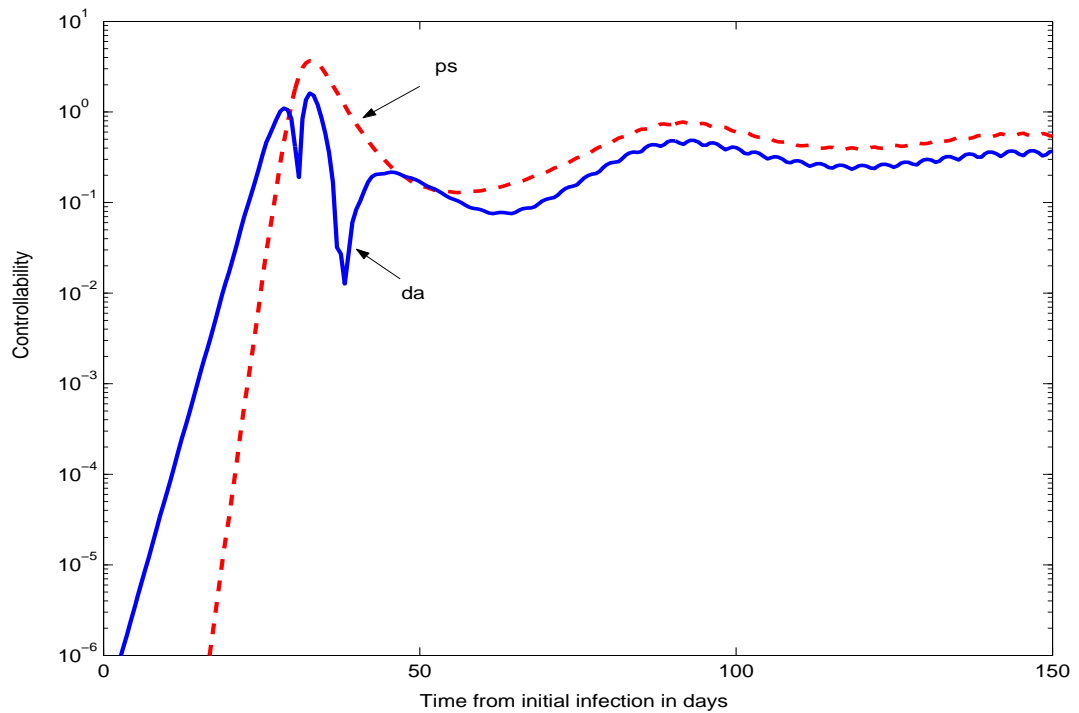
Figure 4.11: Controllability to the asymptomatic stage for the latently infected cell model with immune based therapies. ps: proliferation suppressors, da: death accelerators. Parameters are in Table A.4.

figure 4.11 were generated using matrices $\mathbf{M_{Lps}}$ and $\mathbf{M_{Lda}}$, respectively. When figure 4.11 is compared with figure 4.9, the results show that viral load controllability characteristics are independent of the type of immune based therapy that is used, as it was also the case with replication cycle based therapy.

4.4.5   Controllability to the Advanced Stage

Model parameters are thought to vary with time as has been shown in [84, 167]. For this section, in an attempt to account for the slow decline in CD4$^+$ T cells counts during the asymptomatic stage, and for the rapid increase of the viral load at the advanced stage of the infection, parameters r_T and β_T are assumed to increase linearly with time as given by equation (4.35) as

$$ y(t) \;\; = \;\; \begin{cases} y_o, & 0 \le t \le t_b \\ y_o(1 + m_y(t - t_b)), & t > t_b \end{cases} \tag{4.35} $$

where the variable y represents the concerned parameter (r_T or β_T), y_o is the original value, and m_y is the rate at which the parameter changes. All parameters are assumed

to be constant during the acute infection stage. The parameters change at a slower rate during the asymptomatic stage, which is then followed by a higher rate during the advanced stage. These assumptions, though not clinically validated, do give virus and $CD4^+$ T cell profiles that comply with clinical observations.

For illustrative purposes, clinical latency is attained after 200 days and the immune system is taken to break down 400 days from initial infection. Figure 4.12(a) illustrates the relatively slower $CD4^+$ T cell decline during the asymptomatic stage, and figure 4.12(b) shows the minimum singular value plots to the advanced stage.

It can be seen that in all cases, the advanced stage is not as controllable as the asymptomatic stage and that controllability now correlates with the $CD4^+$ T cells. There is a period of time from just when the virus rebounds and the $CD4^+$ T cells decline, when the viral load controllability slightly increases.

Virus dynamics are oscillatory, so the response when therapy is initiated will initially either overshoot or undershoot before settling to the new steady state. Initiating therapy when the viral load is easier to control implies that the viral load will rebound and settle to the treatment steady state faster than when therapy is initiated where the viral load is more difficult to control. If any viral load suppression to below the treatment steady state is attained, then it will be short lived as the virus quickly rebounds and settles. This will call for early changes or an increase in dosage in order to re-suppress the virus.

On the other hand, initiating therapy when the viral load is more difficult to control implies that the viral load will slowly rebound and settle to the treatment steady state. Any viral load suppression attained will therefore, be relatively durable as the virus slowly rebounds and settles.

The reciprocal of the minimum singular value will give an indication of the viral load settling time and consequently, an indication of the duration of viral load suppression. In terms of the objectives of therapy, therapy is best initiated at a time when treatment will effectively suppress the viral load for as long as possible before it rebounds. If a drug is capable of driving the viral load to a particular steady state value, then the durability of the viral load suppression to below this steady state value can be maximized by choosing the right time to initiate therapy. This would be at a time when the viral load is higher than this steady state value and more difficult to control.

The viral load controllability analysis that has been carried out in this section complements the transient response analysis that was presented in the previous section 4.2. This controllability analysis further explains why differing responses to therapy, as previously illustrated in figure 4.6 (section 4.2.2), can be attained depending on when, during

(a)

(b)
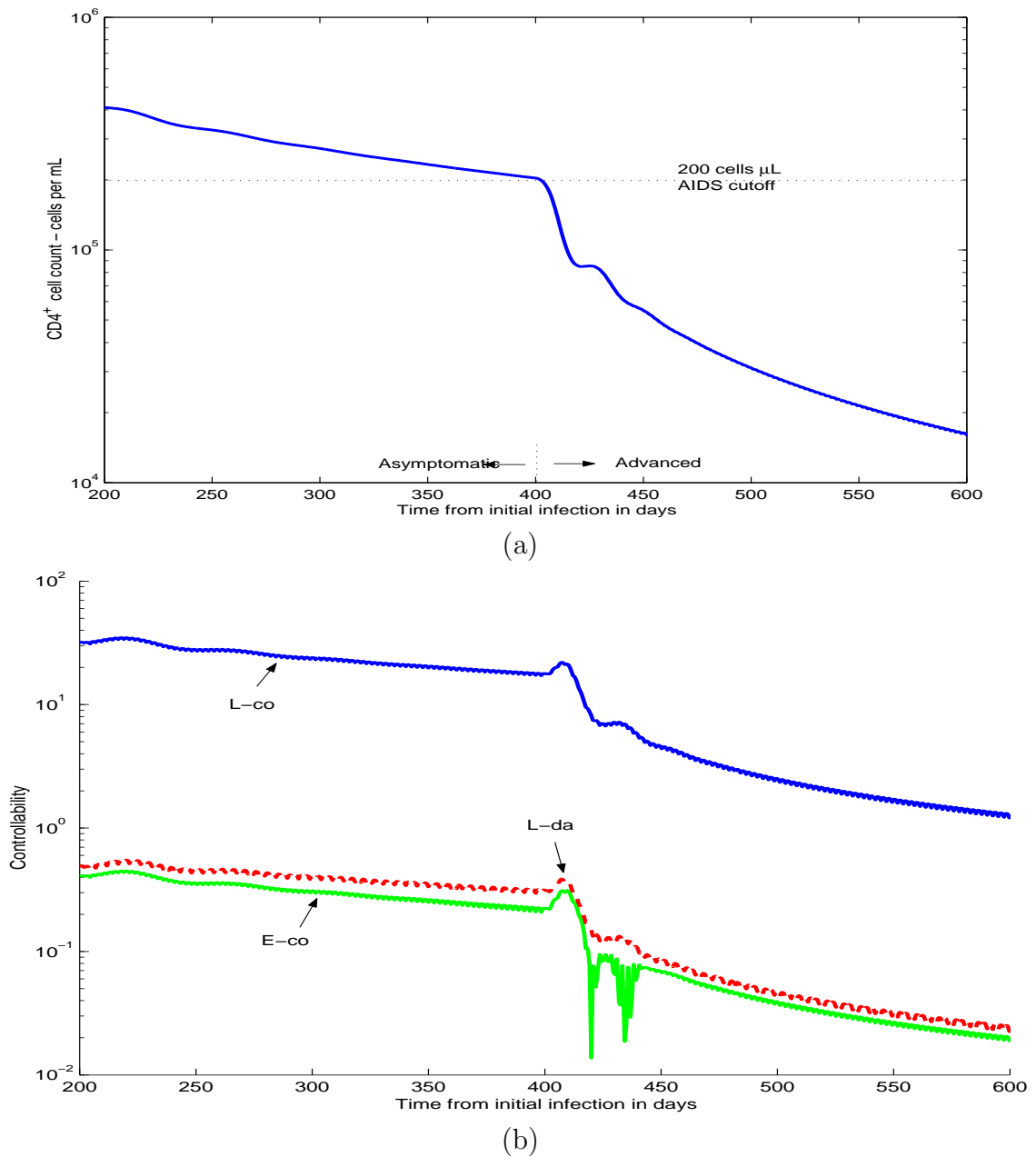
Figure 4.12: Controllability to the advanced stage with a progressive asymptomatic stage. Parameters r_T and β_T slowly changing at rates $m_{r1} = m_{\beta1} = 0.005$, then rapidly at rates $m_{r2} = m_{\beta2} = 0.1$. (a) CD4$^+$ T cell decline. (b) Controllability. L-co: Latently infected cell model with combined rt and pi; L-da: Latently infected cell model with cell death accelerators; E-co: Extended model with combined rt and pi. Parameters are in Table A.4.

the HIV infection, therapy is initiated.

4.4.6   Conclusions

The following conclusions can be drawn from this study.

1. It is apparent that the controllability profiles for the different therapies are similar, or that viral load controllability characteristics are independent of the type of therapy that is used.

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

3. The early acute infection stage and the advanced stages are the most difficult stages to control.

4. The mid-acute infection stage, when the viral load is very high is the easiest stage to control.

5. There is a strong correlation between the viral load and controllability up to the asymptomatic stage, while controllability strongly correlates with the CD4$^+$ T cell count during the advanced stage of the infection.

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.

4.5   Identifiability Analysis

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 carried out 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 before hand. 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 costs less, could then be selected. All the foregoing will be investigated in this chapter using a well established nonlinear system identifiability theory.

Identifiability is a basic system property of whether all parameters can be calculated from the measured output. The identifiability concept applied here was presented and analyzed by [63]. 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.

There are other identifiability analysis tools for nonlinear systems that are available in literature besides the one that is adopted in this thesis. However, the approach as presented by [63] is more practically applicable because it explicitly addresses the issue of the minimum number of measurements that would be required for the eventual determination of the model parameters.

According to [63], if the system is algebraically identifiable, then this enables one to construct the parameters from solving algebraic equations depending only on the information of the input and output. Also, if the system is geometrically identifiable, then this enables one to construct the parameters from solving algebraic equations depending on the information of the input, the output and the initial conditions of the system.

4.5.1   The Need for Parameter Estimates

The preceding model analysis of sections 4.1 to 4.4, as well as the eventual design of individualized dosage schedules will all depend on model parameters. 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 [166]. 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, [50, 84, 85, 176]. Not much effort however, has been put into simultaneously estimating all of these parameters [63]. The available 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 [98, 171]. Furthermore, parameters are thought to vary from one stage to the next as the infection progresses [80, 84]. There is a possibility though, that what seems to be changes in model parameters with time could just be the effect of un-modelled dynamics or external disturbances. Obtaining an individual's parameters at different stages of the viral infection could therefore, settle this issue.

HIV drug pharmacokinetics, pharmacodynamics and adverse reactions are genetically predisposed [41, 126]. Furthermore, the response time to therapy, for example, the time to effectively suppress the viral load are parameter dependent [60]. 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 antiretroviral therapy [177, 178]. There will be a need then, for test measurements to be done over a short period of time.

All variables in the models presented in this and chapter 3 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.

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 ($< 1\%$) of the total CD4$^+$ T cell count [26, 161]. 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. The same is assumed to apply to macrophage measurements. This thesis has therefore, opted to take the uninfected CD4$^+$ T cell count as the measured output instead of the total CD4$^+$ T cell count. The same applies to macrophage measurements. The actual parameter extraction is outside the scope of this thesis.

4.5.2   Identifiability Properties of the Latently Infected Cell Model

The identifiability analysis of the latently infected cell model Σ_{Lnp}(4.36) has already been carried out by [63]. A summary of their findings will be presented in this section for comparative purposes.

Table 4.1: Identifiability of system Σ_{Lnp} with some known parameters.

Known parameters	Identifiability of remaining parameters
q_l, q_a	identifiable
q_l, δ_l	identifiable
q_l, k	identifiable
q_l, δ_a	not identifiable
q_l, r_T	identifiable
q_l, c	not identifiable
q_a, δ_l	identifiable
q_a, k	identifiable
q_a, δ_a	not identifiable
q_a, r_T	identifiable
q_a, c	not identifiable
δ_l, k	not identifiable
δ_l, δ_a	not identifiable
δ_l, r_T	identifiable
δ_l, c	not identifiable
k, δ_a	not identifiable
k, r_T	identifiable
k, c	not identifiable
δ_a, r_T	not identifiable
δ_a, c	not identifiable

[63]

$$\Sigma_{Lnp} = \begin{cases} \frac{dT}{dt} &=& s_T - d_T T - \beta_T TV \\ \frac{dT_l}{dt} &=& q_l \beta_T TV - kT_l - \delta_l T_l \\ \frac{dT_a}{dt} &=& q_a \beta_T TV + kT_l - \delta_a T_a \\ \frac{dV}{dt} &=& r_T T_a - cV \end{cases} \qquad (4.36)$$

If the viral load (V) and uninfected CD4$^+$ T cell count (T) are taken as the measured system outputs, system Σ_{Lnp} has been shown not to be algebraically identifiable. Besides the identification of s_T, d_T and β_T, five of the seven remaining parameters may be computed in terms of the measurements if two of the remaining parameters are known. A summary on the identifiability of system Σ_{Lnp} when various combinations of two model parameters are known is presented in Table 4.1.

However, Σ_{Lnp}(4.36) has been shown to be geometrically identifiable. Therefore, all the original parameters of Σ_{Lnp} 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. And as [63] put it, one needs to

Table 4.2: Minimum number of measurements for the latently infected cell model Σ_{Lnp}.

Measured	Property	T	T_l	T_a	V	Total
T, V	Geometric	6	1	1	8	16
[63]						

have a "comprehensive test" done before measurements are taken.

The required minimum number of measurements for a complete first determination of all ten parameters are as summarized in Table 4.2.

4.5.3   Identifiability Properties of the Extended Model

An analysis of the extended model Σ_{Enp} in (4.37) will be presented for when different combinations of model variables are the measured outputs.

$$\Sigma_{Enp} = \begin{cases} \frac{dT}{dt} & = & s_T - d_T T - \beta_T TV \\ \frac{dT_l}{dt} & = & q_l \beta_T TV - kT_l - \delta_l T_l \\ \frac{dT_a}{dt} & = & q_a \beta_T TV + kT_l - \delta_a T_a \\ \frac{dM}{dt} & = & s_M - d_M M - \beta_M MV \\ \frac{dM^*}{dt} & = & q_M \beta_M MV - \mu M^* \\ \frac{dV}{dt} & = & r_T T_a + r_M M^* - cV \end{cases} \tag{4.37}$$

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. That is,

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

For output y_1, compute

$$\dot{y}_1 = s_T - d_T y_1 - \beta_T y_1 y_2. \tag{4.38}$$

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

$$\ddot{y}_1 = -d_T \dot{y}_1 - \beta_T (y_1 y_2)^{(1)}, \tag{4.39}$$

$$y_1^{(3)} = -d_T \ddot{y}_1 - \beta_T (y_1 y_2)^{(2)}. \tag{4.40}$$

where $(y)^{(n)}$ is the nth derivative of y.

Now there are three equations (4.38), (4.39) and (4.40) with three unknown parameters.

Parameters 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

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

The three equations (4.38), (4.39) and (4.40) can be solved to get a unique solution for s_T, d_T and β_T. 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 twelve parameters $(\delta_l, \delta_a, q_l, q_a, k, r_T, s_M, d_M, \beta_M, r_M, \mu, c)$, compute

$$\dot{y}_2 = r_T T_a + r_M M^* - c y_2, \tag{4.41}$$
$$\vdots$$
$$\begin{aligned} y_2^{(5)} =\ & r_T q_a \beta_T (y_1 y_2)^{(3)} + r_T \beta_T (\theta_2 + \mu q_a)(y_1 y_2)^{(2)} \\ & + r_T \beta_T \mu \theta_2 (y_1 y_2)^{(1)} - (\theta_1 + c + \delta_a) y_2^{(4)} - (\theta_4 + (k + \delta_l)\theta_3) y_2^{(3)} \\ & + (\theta_5 + s_M - (k + \delta_l)\theta_4)\ddot{y}_2 + (k + \delta_l)\theta_5 \dot{y}_2 \\ & + s_M(\Sigma + (\psi_1 - \beta_M y_2)\ddot{y}_2) + \frac{\dot{\Sigma}}{\Sigma} - (d_M - \beta_M y_2)\Lambda, \end{aligned} \tag{4.42}$$

where,

$$\begin{aligned} \theta_1 &= k + \delta_l + \mu, \\ \theta_2 &= k q_l + (k + \delta_l) q_a, \\ \theta_3 &= \mu + c + \delta_a, \\ \theta_4 &= \mu c + \delta_a(\mu + c), \\ \theta_5 &= r_M \beta_M s M - \mu c \delta_a, \\ \psi_1 &= \delta_a - d_M - \beta_M y_2, \\ \Sigma &= \ddot{y}_2 + (\psi_1 - \beta_M y_2)\dot{y}_2, \\ \Lambda &= y_2^{(4)} - r_T q_a \beta_T (y_1 y_2)^{(2)} - r_T \beta_T (\theta_2 + \mu q_a)(y_1 y_2)^{(1)} \\ &\quad - r_T \beta_T \mu \theta_2 y_1 y_2 + (k + \delta_l + \theta_3) y_2^{(3)} + (\theta_4 + \theta_3(k + \delta_l))\ddot{y}_2 \\ &\quad - (\theta_5 + s_M - \theta_4(k + \delta_l))\dot{y}_2 - \theta_5(k + \delta_l) y_2 - \psi_1 s_M y_2. \end{aligned}$$

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.

Identifiability with viral load and uninfected CD4^+ T cell count and macrophage measurements

The identifiability property of system Σ_{Enp} (4.37) can be improved by also measuring the uninfected macrophages. Then, taking the outputs as

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

and compute

$$\dot{y}_1 \quad = \quad s_T - d_T y_1 - \beta_T x_1. \tag{4.43}$$

Then, s_T, d_T and β_T can therefore, as illustrated before, 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

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

For output y_3, compute

$$\dot{y}_3 \quad = \quad s_M - d_M y_3 - \beta_T x_2. \tag{4.44}$$

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

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

For the remaining nine parameters $(\delta_l, \delta_a, q_l, q_a, k, r_T, r_M, \mu, c)$, define

$$
\begin{aligned}
\theta_1 &= k + \delta_l + \mu, \\
\theta_2 &= kq_l + (k + \delta_l)q_a, \\
\theta_3 &= \mu + c + \delta_a, \\
\theta_4 &= \mu c + \delta_a(\mu + c), \\
\theta_5 &= \mu c(\mu + c)(k + \delta_l), \\
\theta_6 &= c\delta_a(c + \delta_a)(k + \delta_l), \\
\theta_7 &= \delta_a\mu(\delta_a + \mu)(k + \delta_l).
\end{aligned}
$$

Then compute

$$
\dot{y}_2 = r_T T_a + r_M M^* - cy_2, \tag{4.45}
$$

$$
\begin{aligned}
\ddot{y}_2 = {}& q_a r_T \beta_T x_1 + r_M \beta_M x_2 + r_T k T_l \\
& - (\mu + c)\dot{y}_2 - \mu c y_2 - r_T(\delta_a - \mu)T_a,
\end{aligned} \tag{4.46}
$$

$$
\begin{aligned}
y_2^{(3)} = {}& q_a r_T \beta_T \dot{x}_1 + r_M \beta_M \dot{x}_2 + r_T \beta_T (kq_l + \mu q_a)x_1 \\
& + r_M \beta_M \delta_a x_2 - \theta_3 \ddot{y}_2 - \theta_4 \dot{y}_2 - \delta_a \mu c y_2 \\
& - r_T k(k + \delta_l - \mu)T_l,
\end{aligned} \tag{4.47}
$$

$$
\begin{aligned}
y_2^{(4)} = {}& q_a r_T \beta_T \ddot{x}_1 + r_M \beta_M \ddot{x}_2 + r_T \beta_T(\theta_2 + \mu q_a)\dot{x}_1 \\
& + r_M \beta_M (k + \delta_l + \delta_a)\dot{x}_2 + r_T \beta_T \mu \theta_2 x_1 \\
& + r_M \beta_M \delta_a(k + \delta_l)x_2 - (\theta_3 + k + \delta_l)y_2^{(3)} \\
& - (\theta_4 + \theta_3(k + \delta_l))\ddot{y}_2 - (\delta_a \mu c + \theta_4(k + \delta_l))\dot{y}_2 \\
& - (k + \delta_l)\delta_a \mu c y_2.
\end{aligned} \tag{4.48}
$$

The remaining nine parameters can therefore be computed from any persistently exciting trajectory $y(t)$ such that rank $\partial(y_2^{(4)}, \ldots, y_2^{(12)})/\ \partial(\delta_l, \delta_a, q_l, q_a, k, r_T, r_M, \mu, c) = 9$.
That is, if

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

where,

$$
\begin{aligned}
\psi_1 &= r_T\beta_T q_a(\dot{x}_1 + \mu x_1) + r_M\beta_M(\dot{x}_2 + \delta_a x_2) \\
&\quad -y_2^{(3)} - \theta_3\ddot{y}_2 - \theta_4\dot{y}_2 - \delta_a\mu c y_2, \\
\psi_2 &= r_M\beta_M(\dot{x}_2 + (k+\delta_l)x_2) - y_2^{(3)} \\
&\quad -(\theta_4 + c)\ddot{y}_2 - \theta_5\dot{y}_2 - (k+\delta_l)\mu c y_2, \\
\psi_3 &= r_T\beta_T k(\dot{x}_1 + \mu x_1), \\
\psi_4 &= r_T\beta_T(\ddot{x}_1 + \theta_1\dot{x}_1 + \mu(k+\delta_l)x_1), \\
\psi_5 &= r_T\beta_T(q_l + q_a)(\dot{x}_1 + \mu x_1) + r_M\beta_M(\dot{x}_2 + \delta_a x_2) \\
&\quad -y_2^{(3)} - \theta_3\ddot{y}_2 - \theta_4\dot{y}_2 - \delta_a\mu c y_2, \\
\psi_6 &= \beta_T(q_a\ddot{x}_1 + (\theta_3 + \mu q_a)\dot{x}_1 + \mu\theta_3 x_1), \\
\psi_7 &= \beta_M(\ddot{x}_2 + \delta_a(k+\delta_l)x_2), \\
\psi_8 &= r_T\beta_T(q_a\dot{x}_1 + \theta_3 x_1) - y_2^{(3)} \\
&\quad -(k+\delta_l+\mu+\delta_a)\ddot{y}_2 - \theta_6\dot{y}_2 - (k+\delta_l)\delta_a c y_2, \\
\psi_9 &= -y_2^{(3)} - (\delta_a + \theta_1)\ddot{y}_2 - \theta_7\dot{y}_2 - (k+\delta_l)\delta_a\mu y_2.
\end{aligned}
$$

System Σ_E (4.37) can be shown to be algebraically identifiable, and all fifteen parameters can be computed for measurements of the uninfected CD4$^+$ T cells, the viral load and uninfected macrophages, if

$$
r_T\beta_T\beta_M k q_l q_a (k - q_l)(k + \delta_l - \delta_a) \neq 0,
$$

$$
\delta_a \neq c \quad \text{and} \quad c \neq \mu.
$$

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.

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

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

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_M, d_M and β_M is as presented earlier in this

section 4.5.3. For identifiability of the remaining twelve parameters, define

$$\triangle_1 = \frac{\dot{y}_2}{y_2} - d_M - \beta_M y_2,$$
$$\triangle_2 = \ddot{y}_2 - r_T \dot{y}_4 + (\mu + c)\dot{y}_2 - r_T \mu y_4 + \mu c y_2,$$

and compute

$$\dot{y}_2 = r_T y_4 + r_M M^* - c y_2, \tag{4.49}$$

$$\ddot{y}_2 = r_T \dot{y}_4 + r_T \mu y_4 - (\mu + c)\dot{y}_2 - \mu c y_2 + r_M \beta_M y_2 M, \tag{4.50}$$

$$y_2^{(3)} = r_T \ddot{y}_4 - (\mu + c + d_M)\ddot{y}_2 + r_T(\mu + d_M)\dot{y}_4$$
$$-(\mu c + d_M(\mu + c))\dot{y}_2 + r_T d_M \mu y_4$$
$$-(d_M \mu c - r_M \beta_M s_M)y_2 + (\frac{\dot{y}_2}{y_2} - \beta_M y_2)\triangle_2. \tag{4.51}$$

Then the seven parameters $(r_T, r_M, \mu, c, s_M, d_M, \beta_M)$ in (4.51) can be computed from any persistently exciting trajectory $y(t)$ such that
rank $\partial(y_2^{(3)}, \ldots, y_2^{(9)})/\partial(r_T, r_M, \mu, c, s_M, d_M, \beta_M) = 7$. That is, if

$$\text{rank} \begin{bmatrix} v_1 & s_M v_6 & v_2 & v_3 & r_M v_6 & v_4 & v_5 \\ \dot{v}_1 & s_M \dot{v}_6 & \dot{v}_2 & \dot{v}_3 & r_M \dot{v}_6 & \dot{v}_4 & \dot{v}_5 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ v_1^{(6)} & s_M v_6^{(6)} & v_2^{(6)} & v_3^{(6)} & r_M v_6^{(6)} & v_4^{(6)} & v_5^{(6)} \end{bmatrix} = 7,$$

where,

$$v_1 = \ddot{y}_4 + (\mu\triangle_1)\dot{y}_4 - \mu\triangle_1 y_4,$$
$$v_2 = -\ddot{y}_2 + r_T \dot{y}_4 - (c - \triangle_1)\dot{y}_2 - r_T\triangle_1 y_4 + \triangle_1 c y_2,$$
$$v_3 = -\ddot{y}_2 - (\mu - \triangle_1)\dot{y}_2 + \triangle_1 \mu y_2,$$
$$v_4 = -\ddot{y}_2 + r_T \dot{y}_4 - (\mu + c)\dot{y}_2 + r_T \mu y_4 - \mu c y_2,$$
$$v_5 = (r_M s_M + \triangle_2)y_2,$$
$$v_6 = \beta_M y_2.$$

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 r_M in order to determine the other six parameters.

For the still remaining parameters $(\delta_l, \delta_a, q_l, q_a, k)$, compute

$$\dot{y}_4 = q_a \beta_T x_1 + kT_l - \delta_a y_4, \tag{4.52}$$

$$\ddot{y}_4 = q_a \beta_T \dot{x}_1 + \beta_T(kq_l + (k + \delta_l)q_a)x_1$$
$$-(k + \delta_l + \delta_a)\dot{y}_4 - \delta_a(k + \delta_l)y_4. \tag{4.53}$$

The five parameters can be computed from any persistently exciting trajectory $y(t)$ such that

rank $\partial(\ddot{y}_4, \ldots, y_4^{(6)})/\partial(\delta_l, \delta_a, q_l, q_a, k) = 5$.

That is, if

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

where,

$$\begin{aligned} \phi_1 &= -\dot{y}_4 + \beta_T q_l x_1 - \delta_a y_4, \\ \phi_1 &= -\dot{y}_4 - (k + \delta_l) y_4, \\ \phi_3 &= \beta_T(\dot{x}_1 + (k + \delta_l) x_1. \end{aligned}$$

However, the above matrix has rank < 5, and therefore not all five remaining parameters can be estimated. One needs prior knowledge of either δ_l, q_l 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 (4.50) and (4.52)

$$\begin{aligned} \dddot{y}_2 &= r_T \dot{y}_4 + r_T \mu y_4 \\ & \quad -(\mu + c)\ddot{y}_2 - \mu c y_2 + r_M \beta_M y_2 M, \\ \dot{y}_4 &= q_a \beta_T x_1 + k T_l - \delta_a y_4. \end{aligned}$$

to generate a 7th equation for the y_2 dynamics and a 5th 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_l 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 be measured is necessary. Table 4.3 summarizes the results for when various model variables are the measured outputs.

Table 4.3: Minimum number of measurements for the extended model Σ_{Enp}.

Measured	Property	T	T_l	T_a	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_a, V	Geometric	5	1	6	1	1	10	24

4.5.4   When to Take Measurements

If either one of the measured outputs y_1, y_2 or y_3 is constant, the higher order derivatives 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 Figure 4.13, 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.

4.5.5   Identifiability With the Use of Antiretroviral Agents

It has again been observed that in the short period following the initiation of therapy, the CD4$^+$ T cell count does not change much [90, 179]. Therefore, during this short period, a complete determination of all the parameters will not be possible. 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. It would be better then, for parameter estimation purposes, if reverse transcriptase inhibitors were exclusively used when measurements are taken.

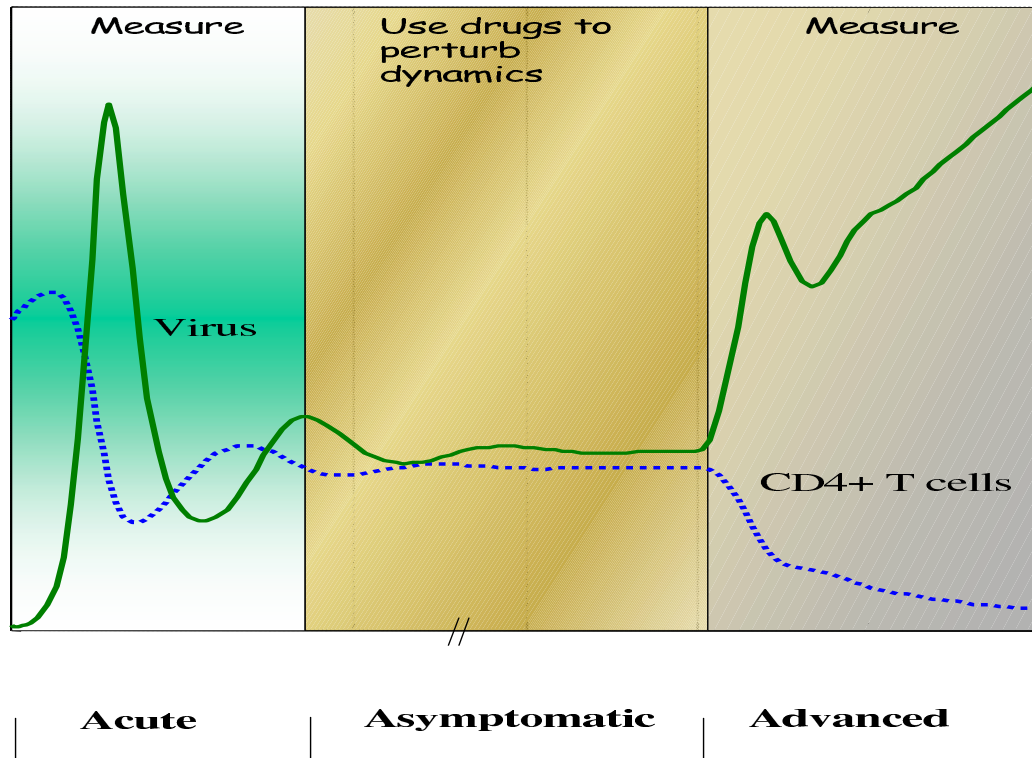A summary of the model parameters that are affected by therapy has been presented in Table 3.3 in section 3.10.

Figure 4.13: When to take measurements.

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 models Σ_{Lnp} (4.36) and Σ_{Enp} (4.37), with the use of replication cycle based reverse transcriptase and protease inhibitors, the identifiable parameters will be the composite parameters. That is, $\eta_{rt}q_l$, $\eta_{rt}q_a$, $\eta_{pi}r_T$, $\alpha_{rt}\eta_{rt}q_M$ and $\alpha_{pi}\eta_{pi}r_M$ would be obtainable, instead of the original parameters q_l, q_a, r_T, q_M and r_M. This means that the drug control effect or efficacy can not be separated from the parameter it affects.
The same holds when immune based therapies are used.

It would be better when using replication cycle based therapies, to use reverse transcriptase inhibitors exclusively. Protease inhibitors increase the proportion of non-infectious virus particles in the measured viral.

4.5.6   Conclusions

In this section, 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 the conditions under which such measurements can be taken.

The extended system Σ_{Enp} given by equation (4.37) is identifiable with proper measurements. The identifiability property of the system has been analyzed when various system variable combinations are taken as the measured outputs.

- System Σ_{Enp} is geometrically identifiable when the viral load and the uninfected CD4$^+$ T cells are the measured system outputs. That is, it is possible to obtain all 16 model parameters from viral load and uninfected CD4$^+$ T cell count measurements if the initial conditions for both the latently and actively infected CD4$^+$ T cells, and both the uninfected and infected macrophages are known.

- System Σ_{Enp} is algebraically identifiable when the viral load, the uninfected CD4$^+$ T cells and the uninfected macrophages are the measured system outputs. It is therefore possible to obtain all model parameters from the measured system outputs.

- System Σ_{Enp} is geometrically identifiable when the viral load, the uninfected CD4$^+$ T cells and actively infected CD4$^+$ T cells are the measured system outputs. It is possible to obtain all the model parameters from the measured system outputs if the initial conditions for both the latently infected CD4$^+$ T cells, and both the uninfected and infected macrophages are known.

This shows that careful consideration of what should be to measured is necessary.

If antiretroviral drugs are used when measurements are taken, then composite parameters $\eta_{rt}q_l$, $\eta_{rt}q_a$, $\eta_{pi}r_T$, $\alpha_{rt}\eta_{rt}q_M$ and $\alpha_{pi}\eta_{pi}r_M$ would be obtainable, instead of the original parameters q_l, q_a, r_T, q_M and r_M.

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

4.6   Model Reduction

The HIV/AIDS models analyzed in this thesis have been shown to be identifiable with proper measurements (refer to section 4.5), and the conditions that apply have been outlined. However, there are too many parameters to be identified, and in some cases, too many measurements would be required to obtain these parameters for some medical settings. The model reduction exercise was carried out to reduce the total number of parameters and lessen the eventual cost of the parameter estimation. However, the resulting or composite parameter set must have clinical significance or meaning.

4.6.1   Residualization of the Latently Infected Cell Model

Reduction of the latently infected cell model Σ_{Lnp}(4.36) without proliferation as it was presented and analyzed in section 4.5 will be carried out. The model is re-presented here for ease of reference as:

$$\frac{dT}{dt} = s_T - d_T T - \beta_T TV \tag{4.54}$$

$$\frac{dT_l}{dt} = q_l \beta_T TV - kT_l - \delta_l T_l \tag{4.55}$$

$$\frac{dT_a}{dt} = q_a \beta_T TV + kT_l - \delta_a T_a \tag{4.56}$$

$$\frac{dV}{dt} = r_T T_a - cV \tag{4.57}$$

Equations (4.54) - (4.57) are a minimal realization of the system. The Jacobians at some operating point $\bar{\mathbf{x}} = [T^o, T_l^o, T_a^o, V_i^o]^T$ are given by

$$\mathbf{A} = \begin{bmatrix} -(d_T + \beta_T V_i^o) & 0 & 0 & -\beta_T T^o \\ q_l \beta_T V_i^o & -(\delta_l + k) & 0 & q_l \beta_T T^o \\ q_a \beta_T V_i^o & k & -\delta_a & q_a \beta_T T^o \\ 0 & 0 & r_T & -c \end{bmatrix}$$

The eigenvalues of matrix $\mathbf{A}$ are the solutions to

$$(\lambda + c)(\lambda + \delta_a)(\lambda + k + \delta_l)(\lambda + d_T + \beta_T V_i^o) - r_T \beta_T T^o(\lambda + d_T)(q_l k - q_a(\lambda + k + \delta_l)) = 0$$

When suppressive therapy is initiated at the asymptomatic stage, then for a typical parameter estimate set in table A.3,

$$0 < r_T \beta_T T^o q_l k < 0.005$$

The eigenvalues can therefore be approximated very well by

$$\bar{\lambda}_1 = -c; \; \bar{\lambda}_2 = -\delta_a; \; \bar{\lambda}_3 = -(k + \delta_l); \; \bar{\lambda}_4 = -d_T$$

From all the available parameter estimates,

$$\bar{\lambda}_1 < \bar{\lambda}_2 < \bar{\lambda}_3 < \bar{\lambda}_4$$

The dynamics of T_l and T are much slower than those of V and T_a, hence T_l and T can be residualized.

Note that residualizing the slower transients T_l and T is equivalent to assuming that their values remain constant for some time after therapy is initiated, as has been observed in practice. A residualization can be obtained by setting $\frac{dT_l}{dt}$ and $\frac{dT}{dt}$ to zero to obtain

$$T = \frac{s_T}{d_T + \beta_T V} \tag{4.58}$$

$$T_l = \frac{q_l \beta_T s_T V}{(d_T + \beta_T V)(k + \delta_l)} \tag{4.59}$$

A second order approximation of the full order system is therefore given by

$$\frac{dV}{dt} = r_T T_a - cV \tag{4.60}$$

$$\frac{dT_a}{dt} = \frac{\beta_T s_T}{(d_T + \beta_T V)}(q_l \frac{k}{k + \delta_l} + q_a)V - \delta_a T_a \tag{4.61}$$

and has non trivial steady states given by

$$V_{ss} = \frac{s_T r_T}{c \delta_a}(q_l \frac{k}{k + \delta_l} + q_a) - \frac{d_T}{\beta_T} \tag{4.62}$$

$$T_{ass} = \frac{c}{r_T} V_{ss} \tag{4.63}$$

The nonlinear term in Equation (4.61) can be expanded to give

$$\frac{dT_a}{dt} = \frac{\beta_T s_T}{d_T}(q_l \frac{k}{k + \delta_l} + q_a)V(1 - \frac{\beta_T}{d_T}V + \cdots) - \delta_a T_a \tag{4.64}$$

and when the higher order terms are ignored, the reduced order system can be linearized as

$$\frac{dV}{dt} = r_T T_a - cV \tag{4.65}$$

$$\frac{dT_a}{dt} = \frac{\beta_T s_T}{d_T}(q_l \frac{k}{k + \delta_l} + q_a)V - \delta_a T_a \tag{4.66}$$

This linearized system's characteristic equation is given by

$$\lambda^2 + (c + \delta_a)\lambda + c\delta_a(1 - R_0) = 0 \tag{4.67}$$

where

$$R_0 = \frac{\beta_T s_T r_T}{c \delta_a d_T}(q_l \frac{k}{k + \delta_l} + q_a) \tag{4.68}$$

is the basic reproductive number [85] and defined as the number of secondary infections resulting from one infected T cell. At steady state, then on average, only one secondary infection results per infected CD4$^+$ T cell. That is, $R_{0(ss)} = 1$.

4.6.2   Response Time Estimation with Reduced Model

Medically, the viral load is considered to be controllable if the antiretroviral drugs in use can reduce the viral load by 90% or by $1\log_{10}$ scale, in 8 weeks from when treatment is initiated and continue to suppress it to below 50 copies per milliliter of plasma in less than 6 months [1]. The time taken from when therapy is initiated to reduce the viral load by 90% is here referred to as the response time, t_{res} and the time taken to suppress the viral load to below 50 copies is the suppression time, t_{sup}. Similarly, the time from when suppressive therapy is terminated to when the viral load rebounds to more than 50 copies is referred to as the rebound time, t_{reb}. This section only estimates the response and suppression times t_{res} and t_{sup} when therapy is initiated at the asymptomatic stage of the infection.

Equations (4.65) and (4.66) under combined antiretroviral therapy are expressed as

$$\frac{dV}{dt} = u_{pi}r_T T_a - cV \tag{4.69}$$

$$\frac{dT_a}{dt} = \frac{u_{rt}\beta_T s}{d}(q_l\frac{k}{k+\delta_l} + q_a)V - \delta_a T_a \tag{4.70}$$

where $u_{pi} = (1 - \eta_{pi})$ is the control effort for the protease inhibitors used, while $u_{rt} = (1 - \eta_{rt})$ is for the reverse transcriptase. For a reduction in viral load therefore, as is the case with effective therapy, the clearance rate of the virus must exceed its replication rate [85]. This means that

$$c > \frac{u\beta_T s_T r_T}{\delta_a d_T}(q_l\frac{k}{k+\delta_l} + q_a)$$

where $u = u_{rt}u_{pi}$ is the combined control effect of the drug combination used and $0 \leq u \leq 1$. The system is therefore stable and its eigenvalues are

$$\lambda_{1,2} = \frac{-c - \delta_a \pm \sqrt{(c+\delta_a)^2 - 4c\delta_a(1 - R_{0(u)})}}{2}$$

where

$$R_{0(u)} = uR_{0(ss)} = u$$

is the number of secondary infections resulting from an infected CD4$^+$ T cell under therapy.

The solution for the viral load $V(t)$ for equations (4.69) and (4.70) has the form

$$V(t) = A_1 e^{\lambda_1 t} + A_2 e^{\lambda_2 t} \tag{4.71}$$

where

$$A_1 = \frac{\lambda_2}{\lambda_2 - \lambda_1}\bar{V} \tag{4.72}$$

$$A_2 = \frac{\lambda_1}{\lambda_1 - \lambda_2}\bar{V} \tag{4.73}$$

---

and $\bar{V}$ could be the steady state or viral load measurement before therapy.

A first estimate for the response time to therapy as well as the time to suppress the viral load can be obtained by solving equation (4.71). In order to do that, one only needs to know the death rate of the actively infected CD4$^+$ T cells δ_a, the clearance rate of the virus c, as well as the reproductive number $R_{0(u)}$. It is interesting to note that these three parameter estimates are attainable. For how to estimate δ_a, c and $R_{0(u)}$ from viral load measurements under therapy, refer to [84, 85, 90]. Methods for obtaining parameter estimates by using control theory techniques are presented in [62, 64, 66].

The estimation error due to residualization is bounded [175]. If one assumes that the absolute error in time estimation is constant for high enough drug efficacies, including $u = 0$, (even though perfect inhibition of virus replication is not practically attainable) then, this could be utilized to derive an expression for the difference in time estimates due to residualization. For $u = 0$, the equations for the residualization difference model are given by

$$\frac{dV_r}{dt} = r_T T_{ar} - cV_r \tag{4.74}$$

$$\frac{dT_{ar}}{dt} = -\delta_a T_{ar} + kT_{lr} \tag{4.75}$$

$$\frac{dT_{lr}}{dt} = -kT_{lr} - \delta_l T_{lr} \tag{4.76}$$

These equations are linear and the solution for the residualization difference in viral load $V_r(t)$ has the form

$$V_r(t) = B_1 e^{\gamma_1 t} + B_2 e^{\gamma_2 t} + B_3 e^{\gamma_3 t} \tag{4.77}$$

where

$$\gamma_1 = -\delta_a; \quad \gamma_2 = -(k + \delta_l); \quad \gamma_3 = -c$$

and

$$B_1 = \frac{\gamma_1}{\gamma_2 - \gamma_1} \frac{\gamma_3}{\gamma_3 - \gamma_1} \bar{V} \tag{4.78}$$

$$B_2 = \frac{\gamma_1}{\gamma_1 - \gamma_2} \frac{\gamma_3}{\gamma_3 - \gamma_2} \bar{V} \tag{4.79}$$

$$B_3 = \frac{\gamma_1}{\gamma_1 - \gamma_3} \frac{\gamma_3}{\gamma_2 - \gamma_3} \bar{V} \tag{4.80}$$

Time estimates can be obtained by adding Equations (4.71) and (4.77) and solving for t_{res} and t_{sup}. This however, would increase the number of required parameters to include a not readily attainable estimate for the combined rate $k + \delta_l$, at which latently infected
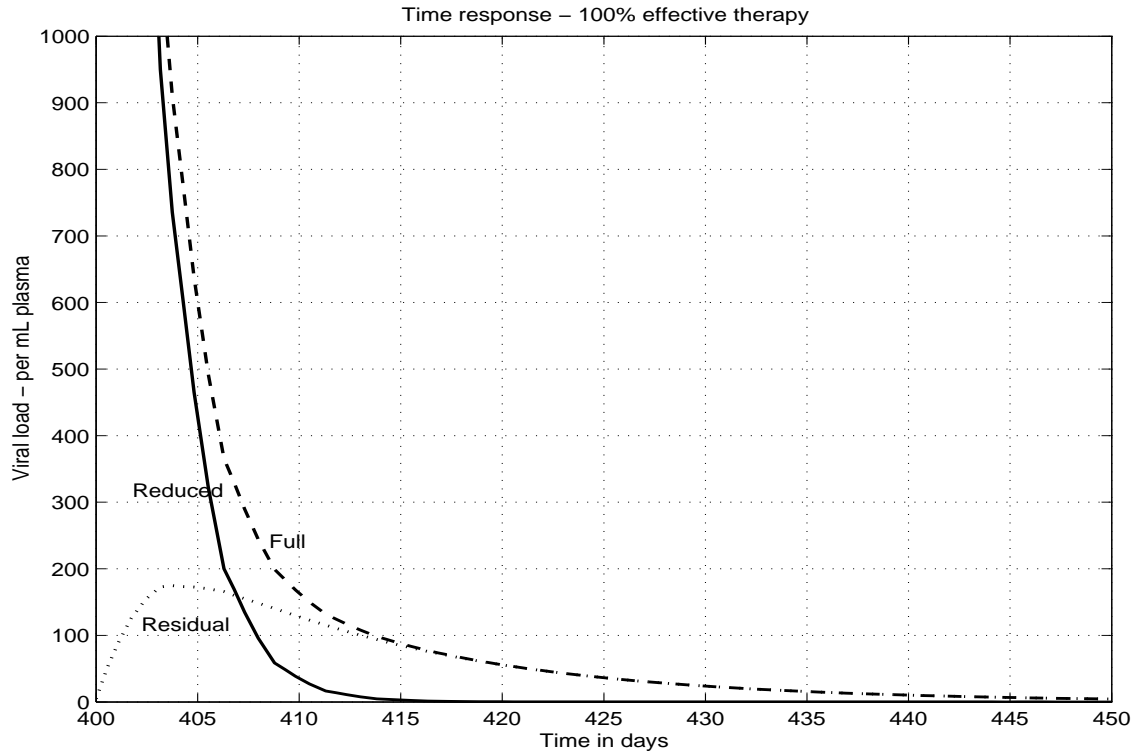
Figure 4.14: Viral load time response for $u = 0$. Full (4D nonlinear), reduced (2D linear) and difference(residual) models. Viral load at start of therapy is 5000 copies per mL plasma. Parameters are in Table A.3.

CD4$^+$ T cells are cleared from plasma. Figure 4.14 shows the viral load time response for when perfect inhibition is assumed for the full, reduced and difference models.

An alternate approach is to consider that simulations show that the differences in response and suppression time estimates due to residualization can be approximated as

$$dif_{res} = t_{res}/R_0 \qquad (4.81)$$

$$dif_{sup} = t_{sup}(1 + 1/R_0) \qquad (4.82)$$

where R_0 can be obtained. This means that residualization differences in time estimates can be obtained if the basic reproductive number is known, by adding the respective difference to the appropriate solution of equation (4.71).

Figures 4.15, 4.16 and 4.17 show the response and suppression time graphs for the full and the reduced order systems when therapy is on and varying drug efficacies are used. Table 4.4 summarizes the results for the response time estimates, with the associated absolute errors in brackets. Table 4.5 does likewise for the suppression time estimates.
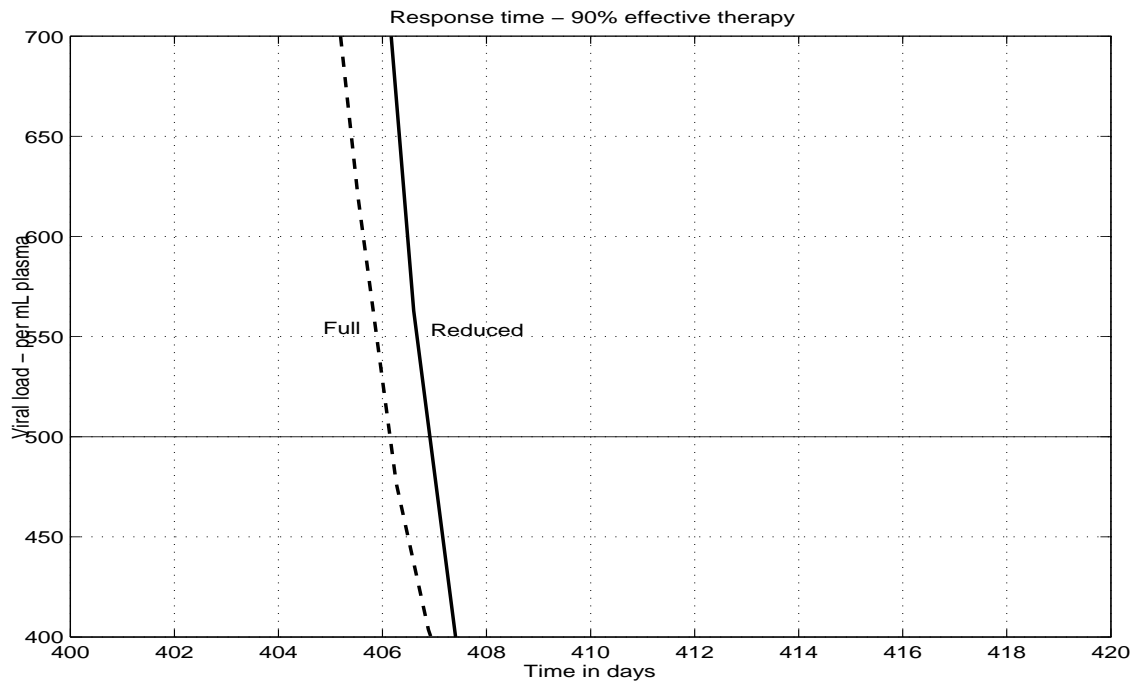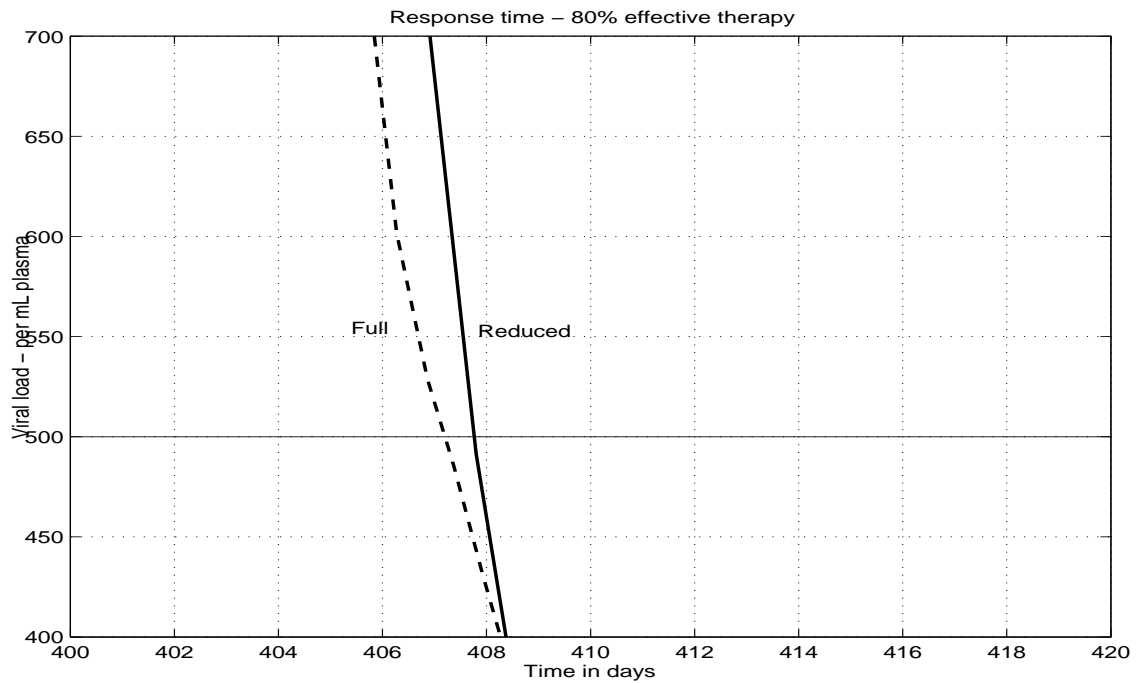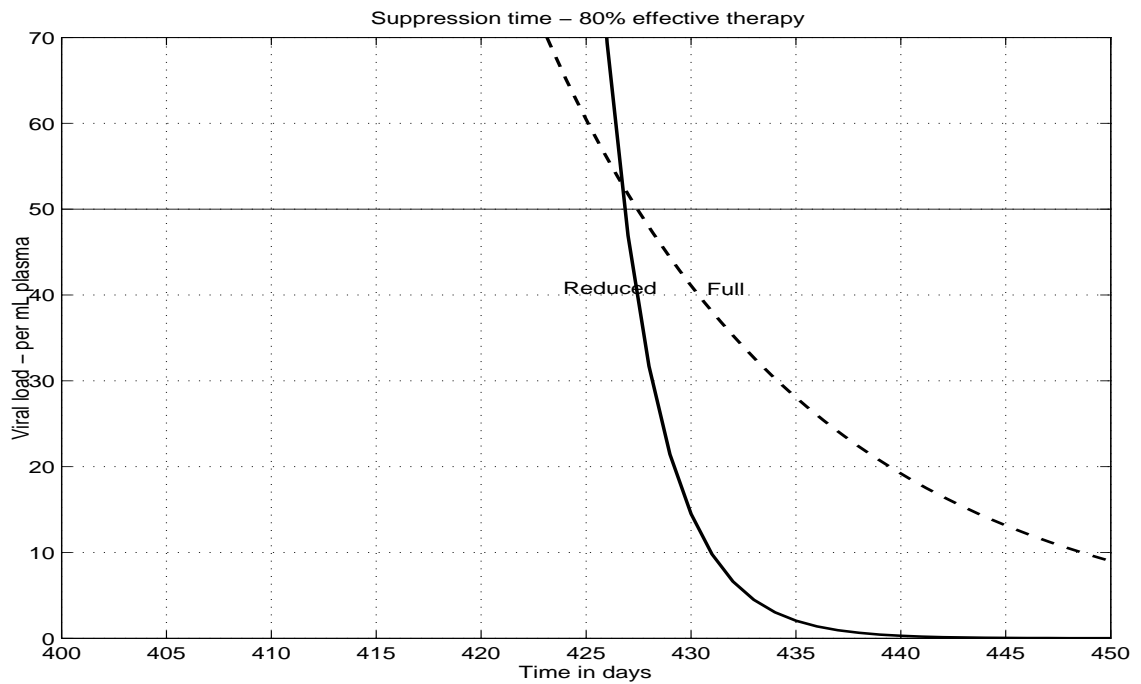
(a)
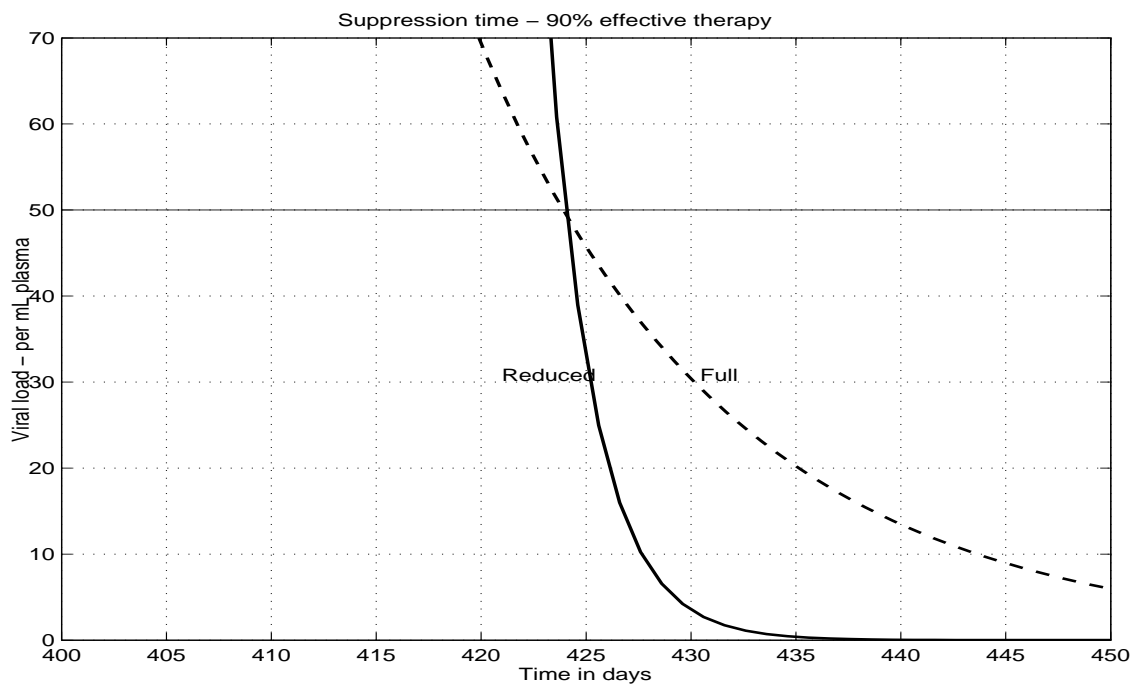
(b)

Figure 4.15: Response and suppressions times for full (4D nonlinear) and reduced (2D linear) order systems: chemotherapy effectiveness: 90% ($u = 0.1$). Viral load at start of therapy (day 400) is 5000 copies per mL plasma. Parameters are in Table A.3.

(a)

(b)

Figure 4.16: Response and suppression times for full (4D nonlinear) and reduced (2D linear) order systems: chemotherapy effectiveness: 80% ($u = 0.2$). Viral load at start of therapy (day 400) is 5000 copies per mL plasma. Parameters are in Table A.3.
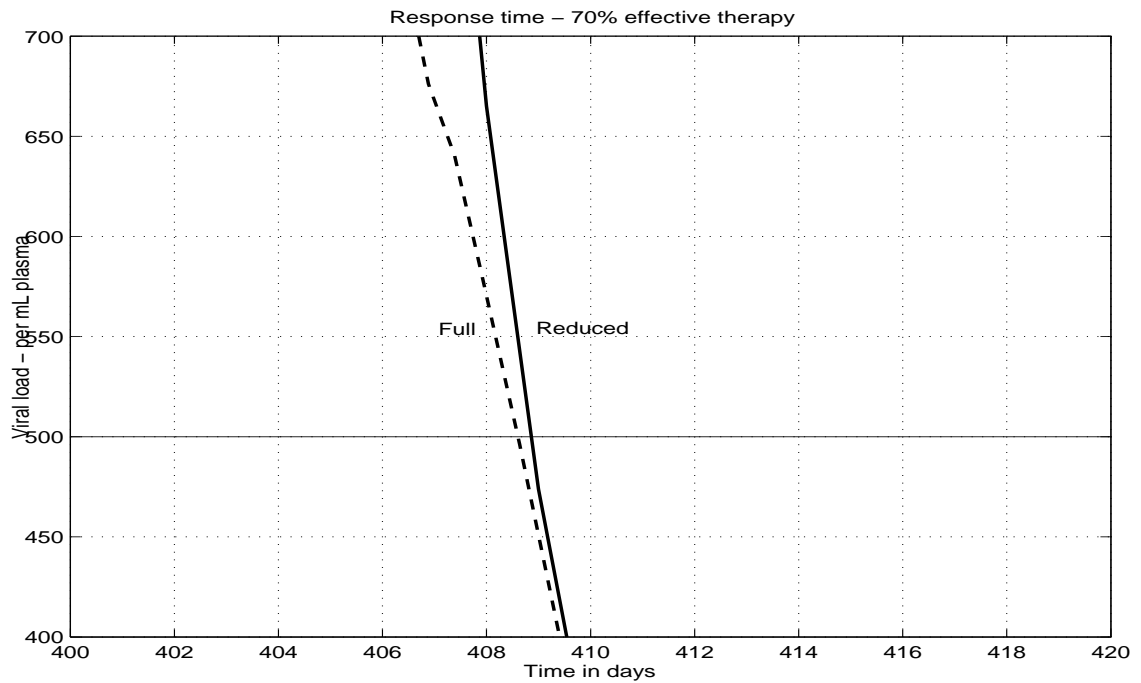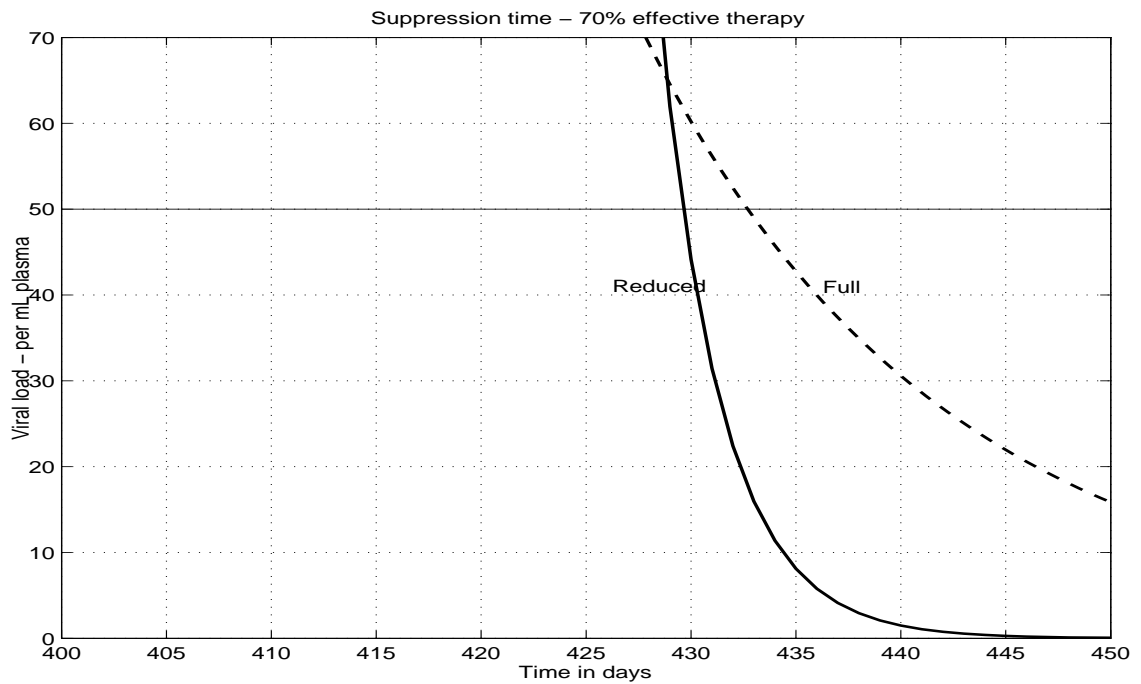
(a)

(b)

Figure 4.17: Response and suppression times for full (4D nonlinear) and reduced (2D linear) order systems: chemotherapy effectiveness: 70% ($u = 0.3$). Viral load at start of therapy (day 400) is 5000 copies per mL plasma. Parameters are in Table A.3.

Table 4.4: Response times

Control u(%effect)	Full system days	Reduced days(error)
0.05 (95%)	6	7 (1)
0.10 (90%)	6	7 (1)
0.15 (85%)	7	8 (1)
0.20 (80%)	7	8 (1)
0.25 (75%)	8	9 (1)
0.30 (70%)	9	9 (0)

Table 4.5: Suppression times

Control u(%effect)	Full system days	Reduced days(error)
0.05 (95%)	23	23 (0)
0.10 (90%)	24	24 (0)
0.15 (85%)	26	25 (-1)
0.20 (80%)	28	26 (-2)
0.25 (75%)	30	28 (-2)
0.30 (70%)	33	29 (-4)

4.6.3   Conclusions

The following conclusions can be drawn from this study.

1. Estimates for the response time to therapy and the time to suppress the viral load, can be determined from values for the death rate of the actively infected CD4$^+$ T cells δ_a, the clearance rate of the virus c, the drug efficacy u and either, the basic reproductive number R_0, or the combined rate $k + \delta_l$ at which latently infected CD4$^+$ T cells are cleared from plasma.

2. The time estimates are parameter dependent and will therefore vary from one individual to the other.

3. The estimated response times, unlike the suppression times, do not exhibit any significant variation with drug efficacy.

4. The initial response is heavily influenced by the virus clearance rate constant c, which is much larger than the actively infected CD4$^+$ T cell clearance rate constant δ_a.

5. The estimated suppression times are shorter than the actual suppression times. When designing a therapy based on the estimates, the drug dosage is always

conservative. In some of the cases of HIV patients, this is desirable.

6. The error in suppression time estimation increases with decreasing drug efficacy. This is because the perfect inhibition assumption used to determine the difference in estimation due to residualization, does not hold for low drug efficacies.

7. This approach enables the determination of the drug efficacy in order to obtain predetermined response and suppression times.

8. This approach can be incorporated into an interruptible control strategy for the viral load.

9. Viral response and suppression time estimates can aid clinicians in scheduling therapy and viral load measurements.

4.7   Chapter Summary

Linear analytical tools were used in some sections of this chapter to derive approximate solutions. The approximate solutions that were obtained, or the predictions that were made from the linear analysis were verified by applying them to the nonlinear model. Since the response was usually as expected or predicted, one can conclude that the linear solution gives a very good approximation.

The following is a summary of the antiretroviral issues that have been addressed by the model analysis that was carried out in this chapter.

4.7.1   Persistent Virus Replication under HAART

The continuous virus replication that has been observed under potent HAART is because the currently used antiretroviral agents are not capable of completely inhibiting virus replication in all the virus producing compartments. However, if the reported virus replication occurs elsewhere and not in the CD4$^+$ T cells, then this suggests that the drugs' efficacy at least equals that required for a zero viral load steady state, as given by equation (4.14).

4.7.2   Variable Response to Therapy

Results from the steady state analysis in section 4.1 indicate that antiretroviral therapy moves the viral load from the pre-treatment value to a treatment steady state. This steady state is drug efficacy and parameter dependent, and will therefore, vary from one individual to the other, given the same drug efficacy. However, the treatment steady state is independent of when therapy in initiated during the course of the infection.

The dependence of the treatment steady states on the model parameters and drug efficacy suggests that if the said model parameters and drug efficacy are known, then it is possible to assess whether or not the individuals viral load can be suppressed to below detectable levels, as well as to determine the individual's operating therapeutic range. This steady state analysis can conversely, be used to determine the drug efficacy required in order to attain and maintain a certain degree of viral load suppression.

The analysis therefore, provides some insight on the issue of variability in response between individuals on the same regimen. The study confirms previous observations that, for a fixed drug efficacy, this variation in response to therapy between individuals is due to inter-individual variations in parameters.

Results from the transient response analysis in section 4.2 show that the transition from the pre-treatment viral load to the treatment steady state is generally oscillatory. The magnitude and frequency of these oscillations are parameter, drug efficacy and timing dependent. The controllability analysis that was presented in section 4.4 shows that from a viral load controllability perspective, some stages of the HIV infection are more controllable than others. The transition from the pre-treatment viral load to the treatment steady state will therefore depend on when therapy is initiated. The transient response and controllability analysis therefore, collectively suggest that a particular individual can have variable response to antiretroviral therapy depending on when therapy is initiated.

4.7.3   Transient Viral Load Rebounds or Virologic Failure?

As discussed before, for a fixed drug efficacy, the transition from the pre-treatment viral load to the treatment steady state is generally oscillatory. The viral load will therefore transiently oscillate about the treatment steady state before settling. When therapy is initiated therefore, the viral load will initially be suppressed to a value that is below the treatment steady state, then rebound to a value above this treatment steady state, before settling. Viral load rebounds are therefore, from a treatment steady state perspective, all transient.

As the viral load oscillates about the steady state value, some oscillations will be larger than others and the viral load will take longer to settle in some cases, depending on the individuals parameters, drug efficacy and when therapy is initiated. However, given that the treatment steady state is independent of when therapy is initiated, it can be concluded that what a viral load rebound is indicative of, does not depend on when therapy is initiated.

This generally means that for an individual, if the drug efficacy is such that the treatment steady state will be at or above a value that is considered indicative of virologic failure, then the viral 'blip' will, in the clinical context, also be indicative of virologic failure. On the other hand, if the drug efficacy is such that the treatment steady state is below detectable levels, then the viral blip will be, in the clinical context again, transient.

4.7.4   Indicators of Virologic and Immunologic Success

Given the preceding analysis, how can one tell in advance, what benefits an individual initiating therapy can expect to obtain, and who will fail to attain viral load suppression? On an individual basis, the attainment of maximal suppression of the viral load, as well as the duration of such suppression once attained, depend on:

1. The HIV infection stage at which therapy is initiated. This stage in turn is defined by the viral load cell count measurements at that time.
2. The virus and target cell parameters.
3. The combined efficacy of the drugs used in the regimen.

However, the maximum cell count to which the cell counts can rebound to depend on the drug efficacy.

This analysis therefore, seems to confirm the findings of some clinical studies that suggests that the virologic and immunologic conditions (viral load and T cell counts) at the start of therapy, determine the outcome of therapy. On the other hand, durable suppression depends on the treatment steady state that the given drug combination efficacy can attain. This also depends on the model parameters and is independent of when therapy is initiated. Does this then disqualify the viral load and cell counts as the indicators of virologic success?

It appears that viral load and T cell counts at the start of therapy are more indicators of virologic response than they are of virologic success. True virologic success, for a particular individual, is determined by the end point drug efficacy. And this lends support to findings that have suggested that virologic failure can be attributed to the regimen used.

It has been argued before that the basic reproductive ratio R_o, is one of the prognostic indicators of virologic success. Given that R_o is exclusively determined by the individual's model parameters, which are in turn co-determinants of virologic success, then the argument is justified.

4.7.5   When Best to Initiate Antiretroviral Therapy?: Clarifying the Confusion

Results from all the preceding sections in this chapter will be analyzed together in order to determine the best time to initiate therapy during the course of the HIV infection. This will be compared with the recommendations from the guidelines, as well as outcomes of some clinical trials. Timing the initiation of therapy can be looked at from different perspectives.

A maximal viral load suppression perspective:
For a particular regimen, the maximality of viral load suppression to below the treatment steady state is highest when therapy is initiated during the early-acute, late-acute and asymptomatic stages of the infection.

A durable viral load suppression perspective:
For a truly durable suppression of the viral load, the drug efficacy should be high enough to attain a treatment steady state that is below levels of detection. The treatment steady state does not depend on when therapy is initiated.

A viral load point to point controllability perspective:
The degree of viral load controllability is highest during the mid-acute infection stage, as the viral load will reach the treatment steady state in a shorter period of time. In this case, one can know if there is need to adjust dosage sooner if viral load suppression is not attained. Also, initiating therapy at this stage will minimize viral load blips.

An Immunologic success perspective:
It appears target cells will rebound irrespective of when therapy is initiated. The maximum to which they can rebound to depends on the model parameters and drug efficacy. No conclusion on timing the initiation of therapy can be drawn from this perspective.

A Virologic success perspective: True virologic success is attained if the drug efficacy is such that the treatment steady state is below levels of detection.

A clinical trial perspective:
From the trial outcomes with STI for autoimmunization, the best time to initiate therapy is during the early acute infection stage.

Guidelines recommendation:

Not very clear. 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. But the guidelines are more inclined towards initiating therapy during the mid-asymptomatic stage.

Compromise: When best to initiate?

The compromise is objective driven and some objectives are in conflict. There is need therefore, to prioritize one's objectives in order to find an optimal solution.

4.7.6   The Possibility of Individualizing Antiretroviral Therapy

The bottom line is that, therapy should be individualized and objectives need to be prioritized. Then the benefits of therapy can be maximized by appropriate drug dosing to attain the desired effect as well as selecting the right time to initiate therapy. Determining the appropriate dose however, places an emphasis on the need to obtain the individual's parameter estimates.

Since the identifiability and model reduction studies have indicated the possibility of attaining the full parameters estimates from system measurements, and the reducibility of this parameter set if need be, then this affirms that it is possible to individualize antiretroviral therapy. If parameters are obtained, then one can determine the individuals therapeutic range. However, if antiretroviral drugs are used when measurements are taken, then it may not be possible to separate the drug efficacy from the model parameter that it affects.