www.elsevier.com/locate/yjtbi

Piecewise virus-immune dynamic model with HIV-1 RNA-guided therapy

Biao Tang, Yanni Xiao, Robert A. Cheke, Ning Wang

PII: S0022-5193(15)00177-0
DOI: http://dx.doi.org/10.1016/j.jtbi.2015.03.040
Reference: YJTBI8156

To appear in: *Journal of Theoretical Biology*

Received date: 14 October 2014
Revised date: 3 February 2015
Accepted date: 24 March 2015

Cite this article as: Biao Tang, Yanni Xiao, Robert A. Cheke, Ning Wang, Piecewise virus-immune dynamic model with HIV-1 RNA-guided therapy, *Journal of Theoretical Biology,* http://dx.doi.org/10.1016/j.jtbi.2015.03.040

Piecewise virus-immune dynamic model with HIV-1 RNA-guided therapy

Biao Tang[†] Yanni Xiao[†1] Robert A. Cheke[§], Ning Wang[‡]

† School of Mathematics and Statistics, Xi'an Jiaotong University
Xi'an, 710049, P.R. China

§ Natural Resources Institute, University of Greenwich at Medway
Chatham Maritime, Chatham, Kent, ME4 4TB, UK

‡ National Center for AIDS/STD Prevention and Control,
Chinese Center for Disease Control and Prevention, 27 Nanwei Rd,
Beijing 100050, PRChina

[1]Corresponding author. E-mail: yxiao@mail.xjtu.edu.cn
Tel: 0086(29)82663156

Abstract

Clinical studies have used CD4 T cell counts to evaluate the safety or risk of plasma HIV-1 RNA-guided structured treatment interruptions (STIs), aimed at maintaining CD4 T cell counts above a safe level and plasma HIV-1 RNA below a certain level. However, quantifying and evaluating the impact of STIs on the control of HIV replication and on activation of the immune response remains challenging. Here we extend the virus-immune dynamic system by including a piecewise smooth function to describe the elimination of HIV viral loads and the activation of effector cells under plasma HIV-1 RNA-guided therapy, in order to quantitatively explore the STI strategies. We theoretically investigate the global dynamics of the proposed Filippov system. Our main results indicate that HIV viral loads could either go to infinity or be maintained below a certain level or stabilize at a previously given level, depending on the threshold level and initial HIV virus loads and effector cell counts. This suggests that proper combinations of threshold and initial HIV virus loads and effector cell counts, based on threshold policy, can successfully preclude exceptionally high growth of HIV virus and, in particular, maximize the controllable region.

Keywords Structured treatment interruptions; Filippov system; sliding mode; pseudo-equilibrium

1 Introduction

Highly active antiretroviral therapy (HAART) has been shown to significantly improve survival and reduce morbidity in HIV patients (Palella et al., 1998; Mocroft et al., 1998). However, long-term HAART continues to be associated with many problems such as adherence difficulties and the evolution of drug resistance (Zhang et al., 1999; Carr et al., 1999; Harrington and Carpenter, 2000; Johnson et al., 2004). Structured therapy interruptions (STIs) have been suggested as being capable of achieving sustained specific immunity for early therapy in HIV infection. As an alternative strategy, STI is a good choice for some chronically infected individuals who may need to take drugs throughout their lives, and it is beneficial for the patients' immune reconstruction during the period when they are not taking the drugs (Maggiolo et al., 2009). Therefore, drug therapies targeted at boosting a virus specific immune response have attracted more and more attention.

Recently, several clinical studies have been aimed at comparing STI strategies with continuous antiretroviral therapy, but conflicting results have been reported (Maggiolo et al., 2009; Ruiz et al., 2007; EL-Sadr et al., 2006; Anaworanich et al., 2006; Guerrero et al., 2005; Hadjiandreou et al., 2009; Lori et al., 2000; Maggiolo et al., 2004). In particular, Ruiz et al. (Ruiz et al., 2007) designed an experiment to evaluate the safety of CD4 cell counts and plasma HIV-1 RNA-guided structured treatment interruptions (STIs) aiming to maintain CD4 T cell counts higher than 350 cells/μl and plasma HIV-1 RNA less than 100,000 copies/μl. They concluded that STIs were not as safe as continuing therapy. Although many mathematical models have been formulated to model continuous therapy(Kuznetsov et al., 1994; Blower et al., 2000; Rong et al., 2007; Tian and Liu, 2014), few attempts have been made to model structured treatment interruptions. In 2012, the authors (Tang et al., 2012) proposed a piecewise system to describe the CD4 cell count-guided STIs, to quantitatively explore STI strategies and to investigate their dynamic behaviors, which explained some controversial conclusions from different clinical studies. To the authors' best knowledge, no mathematical model has yet been proposed to model a plasma HIV-1 RNA-guided structured treatment strategy. An additional challenge remains regarding examination of whether the virus-guided structured treatment can successfully maintain plasma HIV-1 RNA below a certain level or not, and to determine under what conditions patients are suitable for structured treatment interruptions. Quantifying these issues through a mathematical modeling framework is the main objective of this study.

The purpose of the study is to propose a mathematical model to describe plasma HIV-1 RNA-guided structured treatment, and examine the efficacy of this treatment for maintaining plasma HIV-1 RNA below a certain level. The paper is organized as follows. In the next section, we propose our model, provide the definitions for our Filippov system of the virus-immune system and describe the main dynamics of two subsystems. Then the sliding domain and the sliding dynamics are discussed in section 3. In section 4, we investigate the global dynamics of the proposed system. Finally, the biological meaning and the concluding remarks are discussed in section 5.

2 Model equations and preliminaries

The virus dynamic system was formulated to investigate the interaction between the virus and the effector cells (Pugliese and Gandolfi, 2008; Boer and Perelson, 1998). The model equations without considering density dependent inhibition of the virus are as follows

$$\begin{cases} \dot{x} &=& rx - \beta xy, \\ \dot{y} &=& \frac{\rho xy}{1+\omega x} - \mu xy - \delta y, \end{cases} \tag{1}$$

where x and y represent the HIV viral loads and the density of effector cells, respectively and r is the growth rate of HIV virus which incorporates both multiplication and death of HIV virus, δ is the death rate of the effector cells, β denotes the rate of binding of the effector cells to the HIV viruses. As shown in (Abrahms and Brahmi , 1988; Callewaert et al., 1988; Komarova et al., 2003; Shu et al., 2014), the effector cells seem to have a limited ability to repeatedly kill target cells during the interaction of the effector cells and target cells, which shows the inactivation of effector cells. Here, let μ represent the rate of inactivation of the effector cells. Note that that when the virus load is low, the level of immune response is simply proportional to both the viral load, x, and the density of effector cells, y. However, effector cell multiplication due to immune response has a maximum value as HIV viral load gets sufficiently large. Therefore, it is reasonable to suggest the nonlinear form $\rho xy/(1+\omega x)$ to model this (Shu et al., 2014).

Based on the above virus dynamic system (1), we model the plasma HIV-1 RNA-guided therapy in order to maintain the amount of virus below a certain level and to activate the immune system. To this end, whenever the virus load exceeds a critical level (or threshold level T_c), antiretroviral drugs are applied to inhibit growth of the virus, and simultaneously interleukin (IL)-2 treatment is used to activate the immune response (e.g., promote maturation and cytotoxicity of CD4 cells (effector cells))(Marchetti et al., 2005; Napolitano, 2003). Hence the HIV virus dynamic system with HIV-1 RNA-guided therapy can be described following piecewise model

$$\left\{\begin{array}{lcl} \dot{x} & = & rx - \beta xy - \varepsilon_1 \Psi x, \\ \dot{y} & = & \frac{\rho xy}{1+\omega x} - \mu xy - \delta y + \varepsilon_2 \Psi y \end{array}\right. \tag{2}$$

with

$$\Psi = \begin{cases} 0, & \text{if } \; H(x) = x - T_c < 0, \\ 1, & \text{if } \; H(x) = x - T_c > 0, \end{cases} \tag{3}$$

and parameter ε_1 represents the rate of elimination of HIV virus due to antiretroviral therapy and ε_2 denotes the growth rate of the effector cells due to interleukin (IL)-2 treatment.

System (2) with (3), a particular form of a Filippov system, can also be theoretically investigated by using a general dynamical method but this requires complicated and elaborate mathematical techniques (see details in (Kuznetsov et al., 2003; Bernardo et al,, 2008; Padmanabhan and Singh, 1995)). The following definitions on all types of equilibria of non-smooth system (2) with (3) are necessary throughout the rest of this paper.

Let $R^2_+ = \{X = (x,y)|x \geq 0, y \geq 0\}$, $S_1 = \{X \in R^2_+|H(X) < 0\}$, and $S_2 = \{X \in R^2_+|\ H(X) > 0\}$ with $H(X)$ as a smooth scale function. Moreover, the discontinuity boundary Σ separating the two regions is described as

$$\Sigma = \{X \in R^2_+|\ H(X) = 0\}.$$

It is easy to see that $R^2_+ = S_1 \bigcup \Sigma \bigcup S_2$. Consider the following generic planar Filippov system

$$\dot{X} = \begin{cases} F_{S_1}(X), & X \in S_1, \\ F_{S_2}(X), & X \in S_2, \end{cases} \tag{4}$$

and denote

$$\sigma(X) = \langle H_X(X), F_{S_1}(X)\rangle\langle H_X(X), F_{S_2}(X)\rangle,$$

where $\langle \cdot, \cdot \rangle$ is the standard scalar product and $H_X(X)$ represents the gradient of $H(X)$ which is non-vanishing on Σ. Then the sliding mode domain is defined as

$$\Sigma_S = \{X \in \Sigma|\sigma(X) \leq 0\}.$$

In what follows we will use the notation $F_{s_i} \cdot H(X) = \langle H_X(X), F_{S_i}(X)\rangle$ for $i = 1, 2$.

Definition 1. *A point X^* is called a regular equilibrium of system (4) if $F_{S_1}(X^*) = 0$, $H(X^*) < 0$ or $F_{S_2}(X^*) = 0$, $H(X^*) > 0$; A point X^* is called a virtual equilibrium of system (4) if $F_{S_1}(X^*) = 0$, $H(X^*) > 0$ or $F_{S_2}(X^*) = 0$, $H(X^*) < 0$.*

Definition 2. *A point X^* is called a pseudo-equilibrium if it is an equilibrium of the sliding mode of system (4), i.e. $\lambda F_{S_1}(X^*) + (1-\lambda)F_{S_2}(X^*) = 0, H(X^*) = 0$ with $0 < \lambda < 1$ and*

$$\lambda = \frac{\langle H_X(X^*), F_{S_2}(X^*)\rangle}{\langle H_X(X^*), F_{S_2}(X^*) - F_{S_1}(X^*)\rangle}.$$

A point X^* is called a boundary equilibrium of system (4) if $F_{S_1}(X^*) = 0$,

$H(X^*) = 0$ or $F_{S_2}(X^*) = 0,\ H(X^*) = 0$.

Definition 3. *A point X^* is called a $\Sigma-$contact point of system (4) if $X^* \in \Sigma_S$ and $[F_{s_1} \cdot H(X^*)][F_{s_2} \cdot H(X^*)] = 0$. A $\Sigma-$contact point X^* is called a $\Sigma-$fold point of F_{S_1} if $F_{S_1} \cdot H(X^*) = 0$ but $F^2_{s_1} \cdot H(X^*) \neq 0$. Moreover, X^* is called a visible (invisible) $\Sigma-$fold point of F_{S_1} if $F_{S_1} \cdot H(X^*) = 0$ but $F^2_{S_1} \cdot H(X^*) > 0$ ($F^2_{s_1} \cdot H(X^*) < 0$). We call X^* a $\Sigma-$fold point of the system (4) if it is a $\Sigma-$fold point either of F_{S_1} or of F_{S_2}.*

3 Dynamics of two subsystems

For convenience, we call the Filippov system (2) with (3) defined in the region S_1 as subsystem S_1, and the system defined in the region S_2 as subsystem S_2. Moreover, we assume that $\rho - \mu - \delta\omega > 2\sqrt{\mu\delta\omega}$, $\delta > \varepsilon_2$ and $r > \varepsilon_1$ hold true throughout this work, which guarantee that subsystem $S_1(S_2)$ exists two positive equilibria, denoted by $E_{11} = (x_{11}, y_{11})$ and $E_{12} = (x_{12}, y_{12})(E_{21} = (x_{21}, y_{21})$ and $E_{22} = (x_{22}, y_{22}))$, respectively. Here for $i = 1, 2$ we have

$$x_{1i} = \frac{\rho - \mu - \delta\omega \mp \sqrt{(\rho - \mu - \delta\omega)^2 - 4\mu\delta\omega}}{(2\mu}, y_{1i} = \frac{r}{\beta}$$

and

$$x_{2i} = \frac{\rho - \mu - (\delta - \varepsilon_2)\omega \mp \sqrt{(\rho - \mu - (\delta - \varepsilon_2)\omega)^2 - 4\mu(\delta - \varepsilon_2)\omega}}{2\mu}, y_{2i} = \frac{r - \varepsilon_1}{\beta}.$$

Thus, we have the following conclusions on the existence and stability of the equilibria of the two subsystems.

Proposition 1. *For the subsystem S_1 (S_2) there exists a trivial equilibrium $E_{10} = (0, 0)$ ($E_{20} = (0, 0)$) which is a saddle point; The subsystem S_1 (S_2) has two positive equilibria E_{11} (E_{21}) which is a center, and $E_{12}(E_{22})$ which is a saddle point. Also, there exists a homoclinic orbit with respect to E_{12} (E_{22}), denoted as*

$\Gamma^1_{S_1}$ ($\Gamma^1_{S_2}$).

The topological structure of the orbits of the both subsystems is shown in Fig.1. From which we can see that there is an intersection point of the homoclinic orbit $\Gamma^1_{S_1}$ ($\Gamma^1_{S_2}$) with the line $y = r/\beta$ ($y = (r - \varepsilon_1)/\beta$), which is denoted by $E_{13} = (x_{13}, r/\beta)$ ($E_{23} = (x_{23}, (r - \varepsilon_1)/\beta)$).

Lemma 1. *The horizontal components of four positive equilibria of the two subsystems satisfy* $x_{21} < x_{11} < x_{12} < x_{22}$.

Proof. Consider the function

$$f(z) = \rho - \mu - \omega z - \sqrt{(\rho - \mu - \omega z)^2 - 4\mu\omega z}.$$

By simple calculations we have $f'(z) > 0$ if and only if $\rho - z > -1$. Thus, the function $f(z)$ is strictly monotonically increasing when $z < \rho + 1$. It follows from the existence conditions of the positive equilibria of the two subsystems that $\rho - (\delta - \varepsilon_2) > \rho - \delta > \mu\beta > -1$. Therefore, $x_{11} > x_{21}$ always holds true. Further, we can verify that $x_{12} < x_{22}$ is always true whenever they exist. This completes the proof.

According to the definitions above, we have that if $T_c < x_{21}$, then both the equilibria E_{21} and E_{22} are regular equilibria while E_{11} and E_{12} are virtual equilibria. As T_c increases and exceed x_{21}, then the equilibrium E_{21} becomes a virtual equilibrium. If T_c continuously increases and crosses x_{11}, equilibrium E_{11} becomes a regular equilibrium while the equilibrium E_{12} becomes a regular equilibrium too when $T_c > x_{12}$. Furthermore, if $T_c > x_{22}$ holds true, the equilibrium E_{22} is a virtual equilibrium. Therefore, if we let the parameter T_c vary and fix all other parameters we have five different types of the regular/virtual equilibria of system (2) with (3) which are shown in Table 1.

If we consider the subsystem S_1 in the phase space, then y can be seen as a function of x with the following differential equation

$$\frac{dy}{dx} = \frac{y}{x} \frac{\frac{\rho x}{1+\omega x} - \mu x - \delta}{r - \beta y},$$

and integrating above equation from (x_1, y_1) to (x, y), one yields

$$\int_{x_1}^{x} \left(\frac{\rho}{1+\omega x} - \mu - \frac{\delta}{x} \right) = \int_{y_1}^{y} \left(\frac{r}{y} - \beta \right).$$

That is, the first integral $H_1(x, y)$ of subsystem S_1 is as follows

$$H_1(x, y) = -\frac{\rho}{\omega} \ln(1 + \omega x) + \delta \ln(x) + \mu x + r \ln(y) - \beta y = h_1, \tag{5}$$

where $h_1 = H_1(x_1, y_1)$ is a constant. Similarly, the subsystem S_2 also has the following first integral

$$H_2(x, y) = -\frac{\rho}{\omega} \ln(1 + \omega x) + (\delta - \varepsilon_2) \ln(x) + \mu x + (r - \varepsilon_1) \ln(y) - \beta y = h_2 \tag{6}$$

with constant $h_2 = H_2(x_2, y_2)$.

Thus, according to the definition of the Lambert W function (Appendix A) and solving $H_1(x, y) = h_1$ with respect to y, one yields two roots

$$y_l^{S_1} = -\tfrac{r}{\beta} W \left[0, -\tfrac{\beta}{r} \exp\left(\tfrac{\rho \ln(1+\omega x) - \delta\omega \ln(x) - \mu\omega x + h_1\omega}{r\omega}\right)\right] \tag{7}$$

and

$$y_u^{S_1} = -\tfrac{r}{\beta} W \left[-1, -\tfrac{\beta}{r} \exp\left(\tfrac{\rho \ln(1+\omega x) - \delta\omega \ln(x) - \mu\omega x + h_1\omega}{r\omega}\right)\right]. \tag{8}$$

Similarly, solving $H_2(x, y) = h_2$ with respect to y, one has

$$y_l^{S_2} = -\tfrac{r-\varepsilon_1}{\beta} W \left[0, -\tfrac{\beta}{r-\varepsilon_1} \exp\left(\tfrac{\rho \ln(1+\omega x) - (\delta-\varepsilon_2)\omega \ln(x) - \mu\omega x + h_2\omega}{(r-\varepsilon_1)\omega}\right)\right] \tag{9}$$

and

$$y_u^{S_2} = -\tfrac{r-\varepsilon_1}{\beta} W \left[-1, -\tfrac{\beta}{r-\varepsilon_1} \exp\left(\tfrac{\rho \ln(1+\omega x) - (\delta-\varepsilon_2)\omega \ln(x) - \mu\omega x + h_2\omega}{(r-\varepsilon_1)\omega}\right)\right]. \tag{10}$$

In order to show that $y_l^{S_i}$ and $y_u^{S_i}$ $(i = 1, 2)$ are well defined, the domains of the Lambert W function and its properties are used, which have been addressed in detail in Appendix A.

4 Basic properties of Filippov system (2)

Based on the definitions and discussions in section 2, the interior of the sliding domain can be defined as

$$int\Sigma_s = \{X \in \Sigma| \ \ \sigma(X) < 0\}$$

and according to the definition of $\sigma(X)$ we have

$$\sigma(X) = (rx - \beta xy)(rx - \beta xy - \varepsilon_1 x).$$

Solving the inequality $\sigma(X) < 0$, one yields $(r - \varepsilon_1)/\beta < y < r/\beta$.

Therefore, the sliding mode domain of Filippov system (2) with (3) can be defined as

$$\Sigma_S = \left\{(x, y) \in R_+^2| \ \ x = T_c, \ \frac{r - \varepsilon_1}{\beta} \leq y \leq \frac{r}{\beta}\right\}.$$

Denote $A = (T_c, r/\beta)$, $B = (T_c, (r - \varepsilon_1)/\beta)$, which are the two end-points of sliding segment Σ_S. By simple calculation we have $F_{s_1} \cdot H(A) = 0$, $F_{s_1}^2 \cdot H(A) > 0$, $F_{s_1} \cdot H(B) = 0$ and $F_{s_1}^2 \cdot H(B) > 0$. Therefore A (B) is a $\Sigma-$fold point of subsystem S_1 (S_2) which is visible.

Next, we employ Utkin's equivalent control method introduced in (Utikin et al., 2009) to obtain the sliding dynamics in the region Σ_S. It follows from $H = 0$ and the first equation of system (2) that

$$\frac{dH}{dt} = \frac{dx}{dt} = rx - \beta xy - \Psi\varepsilon_1 x = 0. \tag{11}$$

Solving equation (11) with respect to Ψ yields

$$\Psi = \frac{r - \beta y}{\varepsilon_1}.$$

Substituting Ψ into the second equation of system (2) gives

$$\frac{dy}{dt} = y\left(-\frac{\varepsilon_2\beta}{\varepsilon_1}y + \frac{\varepsilon_2 r}{\varepsilon_1} + \frac{\rho T_c}{1 + \omega T_c} - \mu T_c - \delta\right).$$

Therefore, the vector field of Filippov model (2) defined on the sliding domain can be described as follows:

$$\dot{Z}(t) = F_s(X), \quad X \in int\Sigma_s,$$

where $F_s(X) = (P_s(X), Q_s(X))$ with
$Q_s(X) = y(-\varepsilon_2\beta y/\varepsilon_1 + \varepsilon_2 r/\varepsilon_1 + \rho T_c/(1+\omega T_c) - \mu T_c - \delta)$ and $P_s(X) = 0$.
Therefore, the sliding mode dynamics are described by $dy/dt = Q_s(X)$. There exist two roots of $Q_s = 0$ given as follows:

$$y_0 = 0,\ y_c = \frac{r}{\beta} + \frac{\varepsilon_1}{\beta\varepsilon_2}\left(\frac{\rho T_c}{1+\omega T_c} - \mu T_c - \delta\right).$$

Theorem 1 *If $x_{21} \le T_c \le x_{11}$ or $x_{12} \le T_c \le x_{22}$ holds true, then there exists one and only one pseudo-equilibrium $E_c = (T_c, y_c)$ of Filippov system (2) with (3), which is stable on the sliding domain Σ_S. Further, if $T_c = x_{21}$ (x_{11}, x_{12}, x_{22}) holds true, then the positive equilibrium E_{21} (E_{11}, E_{12}, E_{22}), the boundary point B (A, A, B) and the pseudo-equilibrium E_c will coincide into together.*

Proof. Define the function

$$g_1(x) = \frac{\rho x}{1+\omega x} - \mu x - \delta.$$

According to Proposition 1 if $\rho - \mu - \delta\omega > 2\sqrt{\mu\delta\omega}$, then there would be two positive roots of the equation $g(x) = 0$, which are x_{11} and x_{12}. Simple analysis shows that if $x_{11} < x < x_{12}$, then $g_1(x) > 0$; If $x < x_{11}$ or $x > x_{12}$, then $g_1(x) < 0$. This indicates that if $T_c < x_{11}$ or $T_c > x_{12}$ then $y_c < r/\beta$; If $x_{11} < T_c < x_{12}$ then $y_c > r/\beta$.

Rearranging y_c yields

$$y_c = \frac{r-\varepsilon_1}{\beta} + \frac{\varepsilon_1}{\beta\varepsilon_2}\left(\frac{\rho T_c}{1+\omega T_c} - \mu T_c - \delta + \varepsilon_2\right).$$

Similarly, we can define the function

$$g_2(x) = \frac{\rho x}{1+\omega x} - \mu x - \delta + \varepsilon_2.$$

Again from Proposition 1 if $x_{21} < x < x_{22}$ then $g_2(x) > 0$; If $x_{21} < x$ or $x > x_{22}$ then $g_2(x) < 0$. This implies that if $x_{21} < T_c < x_{22}$ then $y_c > (r - \varepsilon_1)/\beta$; If $x_{21} < T_c$ or $T_c > x_{22}$ then $y_c < (r - \varepsilon_1)/\beta$.

Based on the above discussions, if $x_{21} < T_c < x_{11}$ or $x_{12} < T_c < x_{22}$, then we have $(r - \varepsilon_1)/\beta < y_c < r/\beta$. That is, when $x_{21} < T_c < x_{11}$ or $x_{12} < T_c < x_{22}$, then $E_c = (T_c, y_c)$ is a pseudo-equilibrium of system (2) with (3).

Moreover, it is easy to have

$$\left.\frac{dQ_s}{dy}\right|_{(T_c,y_c)} = -\frac{\varepsilon_2\beta}{\varepsilon_1}T_c < 0,$$

which shows that the pseudo-equilibrium E_c is locally stable on the sliding domain Σ_S whenever it exists.

Therefore, if we choose $T_c = x_{21}$, then the boundary point B will coincide with the equilibrium E_{21} according to the definition of the sliding domain. Moreover, when $T_c = x_{21}$, then $g_2(x) = 0$ (i.e. $y_c = (r - \varepsilon_1)/\beta$) holds true. Therefore, the boundary point B of the sliding domain will also coincide with the pseudo-equilibrium E_c when $T_c = x_{21}$. Thus, the three points including the boundary point B, the positive equilibrium E_{21} and the pseudo-equilibrium E_c coincide into one point as $T_c = x_{21}$. A similar thing happens for $T_c = x_{11}$, $T_c = x_{12}$ and $T_c = x_{22}$. This completes the proof.

5 Global analysis of Filippov system (2)

In this section we discuss the global dynamics of Filippov system (2) with (3). It is interesting to note that here an important curve, which consists of some orbits of system (2) and/or of some segments of orbits of system (2), can be defined to identify the different dynamic behaviours. In order to define this key curve denoted by Υ and examine the global dynamics of Filippov system (2), we consider three different cases: (a) $T_c < x_{23}$; (b) $T_c > x_{22}$; and (c) $x_{23} < T_c < x_{22}$.

Case (a): $T_c < x_{23}$. For this case there must be an orbit Γ^4 of subsystem S_2 tangent to $x = T_c$ at point B shown in Fig.2. Let Γ^4_u and Γ^4_l represent the upper and lower branches of the orbit Γ^4, respectively. According to the topological structure of the subsystems, we have that there must be an orbit of subsystem S_1 initiating from B, denoted as Γ^5, and it intersects with line $x = T_c$ at another point $E_4 = (T_c, y_4)$. It follows from the first integral of subsystem S_1 and equation (8) that y_4 can be calculated as

$$y_4 = -\frac{r}{\beta} W\left[-1, -\frac{\beta}{r}\exp\left(\frac{\rho\ln(1+\omega T_c) - \delta\omega\ln(T_c) - \mu\omega T_c + h_{11}\omega)}{r\omega}\right)\right] \qquad (12)$$

with $h_{11} = H_1(T_c, r/\beta)$.

Similarly, there should exist an orbit of the subsystem S_2 passing through the point E_4, and we denote it as Γ^6. Therefore, the curve Υ can be defined as $\Gamma^6 \cup \Gamma^5 \cup \Gamma^4_u$ in this case. Define the region inside the curve Υ as D_Υ, the region inside the orbit Γ^4 as D_{Γ^4} and the region inside the homoclinic orbit $\Gamma^1_{S_i}$ $(i = 1, 2)$ as $D_{\Gamma^1_{S_i}}$ $(i = 1, 2)$.

Moreover, the orbits initiating from D_{Γ^4} can not reach the line $x = T_c$, and hence are free from switching. Therefore, the equilibrium E_{21} is a regular equilibrium which is locally stable within the region $D_{\Gamma^1_{S_2}}$. The orbits of subsystem S_2 starting from the region $D_{\Gamma^4} \setminus D_{\Gamma^1_{S_2}}$ will tend to $(\infty, 0)$, shown in Fig.2. Then, we consider the orbits inside the curve Υ (i.e. in D_Υ). All orbits starting from the region D_Υ either directly reach the segment $\overline{BA}$ or enter into the region S_1 by crossing the segment $\overline{AE_4}$, then follows the dynamics of subsystem S_1, and finally reaches the segment $\overline{BA}$. Furthermore, any trajectory initiating from the segment $\overline{BA}$ slides down and approaches point B due to $dy/dt = Q_S < 0$. Therefore all the orbits initiating from the region D_Υ will approach the point B and finally tend to $(\infty, 0)$ along Γ^4_l.

It follows from Fig.2 that any orbit starting from the region above the curve Υ initially reaches the switching line on $\{(T_c, y) : y > y_4\}$, enters the region S_1 and follows the dynamics of subsystem S_1, then crosses the switching line again on $\{(T_c, y) : 0 < y < (r - \varepsilon_1)/\beta\}$ and enters S_2 finally tending to $(\infty, 0)$ along the dynamics of subsystem S_2. Based on the above discussion, we have the following conclusion.

Theorem 2 *If $T_c < x_{23}$ holds true, then the equilibrium E_{21} is a center and locally stable in $D_{\Gamma^1_{S_2}}$. All other orbits initiating from $R^2_+ \setminus D_{\Gamma^1_{S_2}}$ will tend to $(\infty, 0)$. The global attractor of the Filippov system (2) is $D_{\Gamma^1_{S_2}} \cup \{(\infty, 0)\}$.*

Case (b): $T_c > x_{22}$. It is similar to case (a) and so there must be an orbit of subsystem S_2 tangent to the line $x = T_c$ at point B shown in Fig.3, which we also denoted as Γ^4. The definition of the curve Υ is also the same to case (a). In such a case, equilibrium E_{11} is a regular equilibrium which is a center and locally stable within the region $D_{\Gamma^1_{S_1}}$. Any orbit in the region D_{Γ^4} is free from switching and tends to $(\infty, 0)$ along subsystem S_2. Similarly, any orbit initiating from the region $D_\Upsilon \setminus D_{\Gamma^1_{S_1}}$ will first approach point B and then tend to $(\infty, 0)$ along the orbit Γ^4_l as shown in Fig.3. So when $T_c > x_{22}$, the global dynamics of system (2) can be concluded as following results.

Theorem 3 *If $T_c > x_{22}$ holds true, then the equilibrium E_{11} is a center and locally stable in $D_{\Gamma^1_{S_1}}$. All other orbits starting from $R^2_+ \setminus D_{\Gamma^1_{S_1}}$ will tend to $(\infty, 0)$. And the global attractor of the switching system (2) is $D_{\Gamma^1_{S_1}} \cup \{(\infty, 0)\}$.*

Case (C): $x_{23} < T_c < x_{22}$. In this scenario, there would be two intersection points between the homoclinic orbit $\Gamma^1_{S_2}$ and line $x = T_c$ denoted by $E_5 = (T_c, y_5)$ and $E_6 = (T_c, y_6)$ respectively, shown in Figs.4-8. It follows from the first integral of subsystem S_2 and equations (9) and (10) that y_5 and y_6 can be calculated respectively as

$$y_5 = -\frac{r-\varepsilon_1}{\beta} W\left[0, -\frac{\beta}{r-\varepsilon_1}\exp\left(\frac{\rho\ln(1+\omega T_c) - (\delta-\varepsilon_2)\omega\ln(T_c) - \mu\omega T_c + h_{21}\omega)}{(r-\varepsilon_1)\omega}\right)\right] \tag{13}$$

and

$$y_6 = -\frac{r-\varepsilon_1}{\beta} W\left[-1, -\frac{\beta}{r-\varepsilon_1}\exp\left(\frac{\rho\ln(1+\omega T_c) - (\delta-\varepsilon_2)\omega\ln(T_c) - \mu\omega T_c + h_{21}\omega)}{(r-\varepsilon_1)\omega}\right)\right] \tag{14}$$

with $h_{21} = H_2(x_{22}, y_{22})$.

According to the topological structure of subsystem S_1, there must be an orbit Γ^7 of subsystem S_1 initiating from the point E_5, and it intersects with line $x = T_c$ at another point $E_7 = (T_c, y_7)$. And we can conclude that $y_7 > y_6$ holds true by using the Lemma 2 (see appendix B). It follows from the first integral of subsystem S_1 and the equation (8) that y_7 can be solved as:

$$y_7 = -\frac{r}{\beta} W \left[-1, -\frac{\beta}{r} \exp\left(\frac{\rho ln(1+\omega T_c) - \delta \omega ln(T_c) - \mu \omega T_c + h_{12}\omega)}{r\omega} \right) \right] \tag{15}$$

with $h_{12} = H_1(T_c, y_5)$.

Similarly, there must exist an orbit of subsystem S_2 passing through the point E_7, which is denoted by Γ^8. Then the curve Υ for this scenario can be defined as $\Gamma^8 \cup \widehat{E_7 E_5}|_{S_1} \cup \widehat{E_5 E_{22}}|_{S_2} \cup \Gamma^2_{S_2}$. In the following we specify four subcases in terms of relationships among T_c, x_{21}, x_{11} and x_{12}.

Subcase $(C1)$: Suppose $x_{23} < T_c < x_{21}$ holds true. Then there exists a closed orbit ζ_1 of subsystem S_2 which is tangent to line $x = T_c$ at the point B shown in Fig.4. As we have discussed in section 3, B is a boundary point of the sliding domain which is also a visible $\Sigma-$fold point. Therefore the closed orbit ζ_1 is a touching cycle of the Filippov system (2) (see (Kuznetsov et al., 2003)). Define the region inside the cycle ζ_1 as D_{ζ_1}. The equilibrium E_{21} is a regular equilibrium which is a center and locally stable in D_{ζ_1}. Then we will show that all the orbits in $D_\Upsilon \setminus D_{\zeta_1}$ tend towards the touching cycle ζ_1. To verify this conclusion, we need to consider two different situations:

When $y_6 < r/\beta$, any orbit initiating from $D_\Upsilon \setminus D_{\zeta_1}$ either directly reaches the segment $\overline{BA}$ or enters into the region S_1 by crossing the segment $\overline{AE_7}$ as shown in Fig.4(a). Note that the orbit of subsystem S_1 initiating from $\overline{AE_7}$ either directly reaches the segment $\overline{BA}$ or enters into the region S_2 by crossing the segment $\overline{E_5B}$, follows the dynamics of subsystem S_2, and finally reaches the segment $\overline{BA}$. Therefore, all the orbits in the region $D_\Upsilon \setminus D_{\zeta_1}$ will first reach the segment $\overline{BA}$. Furthermore, any trajectory initiating from the segment $\overline{BA}$ will slide down to point B, and then remain at the touching cycle ζ_1, due to $dy/dt = Q_S < 0$. This verified the conclusion under this situation.

When $y_6 > r/\beta$, similarly, any orbit starting from the region $D_\Upsilon \setminus D_{\zeta_1}$ (see Fig.4(b)) will (i) directly reach the segment $\overline{BA}$, or (ii) enter into the region S_1 by crossing the segment $\overline{AE_7}$, and follow the dynamics of system S_1 then approach the segment $\overline{BA}$ or enter to the region S_2 by crossing the segment $\overline{E_5B}$, and follow the dynamics of system S_2 then reaches to the segment $\overline{BA}$ or enter the region S_1 again by crossing the segment $\overline{AE_7}$ and then it follows Lemma 2 (see Appendix B) that we can deduce that it will finally reach the segment $\overline{BA}$. Moreover it is similar to the former case that any trajectory initiating from the segment $\overline{BA}$ will slide down and reach the touching cycle ζ_1. This verified the conclusion for this situation.

Next, we consider where the orbits initiating from the region $R_+^2 \setminus D_\Upsilon$ go. Definitely, any orbit initiating from the region between $\Gamma_{S_2}^2$ and $\Gamma_{S_2}^3$ is free from switching, follows the dynamics of system S_2 and finally tends to $(\infty, 0)$. Any orbit starting from the region above the curve Υ firstly reaches the switching line on $\{(T_c, y)|y > y_7\}$, enters the region S_1 and follows the dynamics of system S_1, then crosses the switching line again on $\{(T_c, y)|0 < y < y_5\}$ and enters the region S_2 again, finally tending to $(\infty, 0)$. Then we conclude that trajectories initiating from different region will approach the different states. It is more interesting to show the various simulations in Fig.5(a-b) in which all the parameter values are fixed as in Fig.4(a). It follows from Fig.5(a) that the viral load fluctuates at a certain level while (b) demonstrates that the viral load goes to infinity. Then the global dynamics of system (2) when $x_{23} < T_c < x_{21}$ can be concluded as follows:

Theorem 4 *If $x_{23} < T_c < x_{21}$ holds true, then system (2) has a touching cycle ζ_1. The equilibrium E_{21} is a regular equilibrium which is a center and locally stable in D_{ζ_1}. The orbits initiating from the region $D_\Upsilon \setminus D_{\zeta_1}$ will tend towards the touching cycle ζ_1, and the other orbits starting from $R_+^2 \setminus D_\Upsilon$ finally tend towards $(\infty, 0)$. The global attractor of the Filippov system (2) is $D_{\zeta_1} \cup \{(\infty, 0)\}$.*

Subcase $(C2)$: Suppose $x_{21} < T_c < x_{11}$ holds true, it follows from theorem 1 that there exists a pseudo-equilibrium E_c which is locally stable on the sliding domain. Fig.6 shows that the orbits starting from D_Υ initially reach the switching segment $\overline{BA}$, and then slide down or up to the pseudo-equilibrium E_c. Simultaneously, we have that all the other orbits will tend to $(\infty, 0)$. Further, Fig.5(c) illustrates that the viral load is successfully controlled and stabilizes at a level of T_c and (d) shows that the viral load goes to infinity. Then we have the following conclusion.

Theorem 5 *If $x_{21} < T_c < x_{11}$ holds true, then there exists a pseudo-equilibrium E_c of system (2) which is locally asymptotically stable in D_Υ. Any orbit initiating from $R_+^2 \setminus D_\Upsilon$ tends to $(\infty, 0)$. The global attractor of the Filippov system (2) is $\{E_c, (\infty, 0)\}$.*

Subcase $(C3)$: Suppose $x_{11} < T_c < x_{12}$ holds true. Then there exists a closed orbit ζ_2 of subsystem S_1 which is tangent to line $x = T_c$ at the point A. Based on the discussion in section 3, the point A is a boundary point while it is also a visible $\Sigma-$fold point of subsystem S_1. Therefore, the closed orbit ζ_2 is a touching cycle of system (2) (see (Kuznetsov et al., 2003)). Define the region bounded by the touching cycle ζ_2 as D_{ζ_2}. At this time, any orbit initiating from the segment $\overline{BA}$ will slide up and reach the point A. The global dynamics of the Filippov system (2) are similar to those of theorem 4. Here we conclude as follows:

Theorem 6 *If $x_{11} < T_c < x_{12}$ holds true, then there also exists a touching circle ζ_2. The equilibrium E_{11} is a regular equilibrium and is locally stable in D_{ζ_2}. Any orbit initiating from the region $D_\Upsilon \setminus D_{\zeta_2}$ finally tends to the touching cycle ζ_2, all the other orbits starting from $R_+^2 \setminus D_\Upsilon$ will tend to $(\infty, 0)$. The global attractor of the Filippov system (2) is $D_{\zeta_2} \cup \{(\infty, 0)\}$.*

Subcase $(C4)$: Suppose $x_{12} < T_c < x_{22}$. Then the equilibrium E_{11} is a regular equilibrium, and there also exists a pseudo-equilibrium E_c which is stable on the sliding domain according to theorem 1. The global dynamics of the Filippov system (2) are similar to the case when $x_{11} < T_c < x_{12}$, and we then omit the proof. So we have the following conclusion.

Theorem 7 *If $x_{12} < T_c < x_{22}$ holds true, then the equilibrium E_{11} is a regular equilibrium which is also a center and locally stable in $D_{\Gamma^1_{S_1}}$. There also exists a pseudo-equilibrium E_c with any orbit starting from the region $D_\Upsilon \setminus D_{\Gamma^1_{S_1}}$ finally tending to it. All the other orbits starting from $R_+^2 \setminus D_\Upsilon$ will tend to $(\infty, 0)$. The global attractor of the Filippov system (2) is $D_{\Gamma^1_{S_1}} \cup \{E_c, (\infty, 0)\}$.*

In summary, we have examined the global dynamics of the Filippov system (2). It has been shown that for relatively low or large level of threshold (i.e. $T_c < x_{23}$ or $T_c > x_{22}$) the Filippov system (2) behaves either like the controlled subsystem S_2 or free subsystem S_1. It indicates that the region $D_{\Gamma^1_{S_1}}$ (or $D_{\Gamma^1_{S_2}}$) bounded by homoclinic orbit $\Gamma^1_{S_1}$ (or $\Gamma^1_{S_2}$) is the only invariant set from which HIV virus load remains bounded. While for intermediate levels of threshold (i.e. $x_{23} < T_c < x_{22}$), a new phenomenon was observed for this virus-guided therapy. In particular, we obtained a much bigger region D_Υ bounded by the critical curve Υ from which HIV virus load can be maintained below a certain level, and hence we name the region D_Υ as the controllable region. It is interesting to examine how the region D_Υ change as the threshold T_c, elimination rate ε_1 and growth rate of the effector cells ε_2 vary. Since the region D_Υ increases with increasing x, we choose a certain constant and sufficiently large value of x, say $x = T_\Upsilon$, such that the region is closed and can be evaluated.

Without lose of generality, we assume that $T_\Upsilon > x_{22}$ always holds true. Then the line $x = T_\Upsilon$ divides the region D_Υ into two subregions and we denote the left subregion as D^L_Υ. Simultaneously, there exists an intersection point of the curve Υ to the line $y = r/\beta$, denoted as $E_\Upsilon = (x_\Upsilon, r/\beta)$ (shown in Fig.4-8) and x_Υ satisfy the following equation

$$-\frac{\rho}{\omega}\ln(1+\omega x_\Upsilon) + \delta \ln x_\Upsilon + \mu x_\Upsilon + r\ln(\frac{r}{\beta}) - r - h_{12} = 0. \tag{16}$$

It follows from the first integral of the subsystems and the definition of the Lambert W function that we can calculate the area of region $S_{D^L_\Upsilon}$ as follows

$$\begin{aligned} S_{D^L_\Upsilon} = & \int_{x_\Upsilon}^{T_c} \Big(\tfrac{r}{\beta} \Big(W \Big[0, -\tfrac{\beta}{r} \exp\Big(\tfrac{\rho \ln(1+\omega x) - \delta\omega \ln(x) - \mu\omega x + h_{12}\omega}{r\omega} \Big)\Big] \\ & - W \Big[-1, -\tfrac{\beta}{r-\varepsilon_1} \exp\Big(\tfrac{\rho \ln(1+\omega x) - \delta\omega \ln(x) - \mu\omega x + h_{12}\omega}{r-\omega} \Big)\Big]\Big)\Big) dx + \\ & \int_{T_c}^{T_\Upsilon} \Big(\tfrac{r-\varepsilon_1}{\beta} \Big(W \Big[0, -\tfrac{\beta}{r-\varepsilon_1} \exp\Big(\tfrac{\rho \ln(1+\omega x) - (\delta-\varepsilon_2)\omega \ln(x) - \mu\omega x + h_{21}\omega}{(r-\varepsilon_1)\omega} \Big)\Big] \\ & - W \Big[-1, -\tfrac{\beta}{r-\varepsilon_1} \exp\Big(\tfrac{\rho \ln(1+\omega x) - (\delta-\varepsilon_2)\omega \ln(x) - \mu\omega x + h_{22}\omega}{(r-\varepsilon_1)\omega} \Big)\Big]\Big)\Big) dx \end{aligned} \tag{17}$$

where

$$h_{12} = H_1(T_c, y_5), \quad h_{21} = H_2(x_{22}, y_{22}), \quad h_{22} = H_2(T_c, y_7). \tag{18}$$

Due to highly nonlinear properties of $S_{D^L_\Upsilon}$, we numerically investigate the variation in $S_{D^L_\Upsilon}$ with parameters $T_c, \varepsilon_1, \varepsilon_2$. Fig.9(A-B) shows that for a given ε_i $(i = 1, 2)$, $S_{D^L_\Upsilon}$ initially increases and then turns to decline as the threshold T_c increases. This means that there exists an optimal threshold such that the area of the controllable region D^L_Υ maximizes and hence in such a scenario HIV virus is maximally controlled. It follows from Fig.9(C-D) that for a given threshold T_c, $S_{D^L_\Upsilon}$ becomes large as ε_1 or ε_2 increases.

Remark: Based on the above discussions, it is interesting to observe that several bifurcations occur if we let bifurcation parameter T_c increase and keep all other parameters fixed. As T_c increases and exceeds x_{23}, a touching cycle appears. When T_c reaches x_{21}, the touching cycle disappears, and the pseudo-equilibrium appears and coincides with the boundary equilibrium B of the sliding domain. Then system (2) with (3) undergoes a sliding grazing bifurcation and the boundary center bifurcation at $T_c = x_{21}$. As T_c increases and exceeds x_{21}, the pseudo-equilibrium E_c coincides with the boundary point A for $T_c = x_{11}$ at which the boundary center bifurcation occurs. As T_c continuously increases, a touching cycle appears again and it will disappear at $T_c = x_{12}$. When T_c exceeds x_{12} the pseudo-equilibrium E_c appears, and the pseudo-equilibrium will coincide with the boundary saddle point A at $T_c = x_{12}$. Finally, when $T_c = x_{22}$, the pseudo-equilibrium coincides with the boundary saddle point B and the system (2) undergoes a boundary saddle bifurcation.

6 Biological implications and discussion

Although the strategies of STIs of antiretroviral therapies have been proposed for clinical management of HIV infected patients, clinical studies on STIs failed to achieve a consistent conclusion for this strategy. Many researchers suggested that in order to evaluate the benefits and risks of STIs, long-term studies are necessary and the choice of threshold may be pivotal for successful STIs (Maggiolo et al., 2009; Danel et al., 2006; DART Trial Team , 2008; Hirschel and Flanigan, 2009).

In this study we have proposed and analyzed a viral dynamic model with a piecewise control function concerning a threshold policy for an HIV management strategy. The proposed model extends the classic model by including a piecewise elimination rate of HIV virus and growth rate of effector cells to represent therapy strategies (antiretroviral drugs and interleukin (IL)-2 treatment) being triggered once the HIV virus load exceeds a threshold level.

We examined the sliding domain and the sliding dynamics of system (2), and then the global dynamics of system (2) is discussed by considering several different cases. Note that the pseudo-equilibrium E_c is feasible and is locally asymptotically stable for $x_{21} < T_c < x_{11}$ or $x_{12} < T_c < x_{22}$. In particular, when $T_c < x_{23}$ (or $T_c > x_{22}$), the region $D_{\Gamma^1_{S_2}}$ ($D_{\Gamma^1_{S_1}}$) bounded by the homoclinic orbit $\Gamma^1_{S_2}$ ($\Gamma^1_{S_1}$) is an invariant set, all other orbits initiating from $R^2_+ \setminus D_{\Gamma^1_{S_2}}$ ($R^2_+ \setminus D_{\Gamma^1_{S_1}}$) approach $(\infty, 0)$. When the threshold satisfies $x_{23} < T_c < x_{22}$, the critical curve Υ consisting of several critical orbits was defined, by which the global dynamics of system (2) can easily be obtained. It has been shown that the orbits starting from D_Υ either (i) approach the pseudo-equilibrium E_c (Fig.6) or (ii) approach or remain in the invariant set $D_{\zeta_i}(i = 1, 2)$ (shown in Fig.4, 7), or (iii) approach the pseudo-equilibrium E_c or remain in invariant set $D_{\Gamma^1_{S_1}}$ (Fig.8), depending on the threshold and initial data. In such a scenario, other orbits starting from $R^2_+ \setminus D_\Upsilon$ also approach $(\infty, 0)$. It is worth mentioning that choosing an appropriate threshold level for making the decision to trigger the intervention and for its suspension is crucial (Canchemez et al., 2009; Day et al., 2006; Wang and Xiao, 2013; Tang and Liang, 2013; Xiao et al., 2012, 2015).

It is important to emphasize that this policy led to interesting biological interpretations which can help us to develop an optimal treatment strategy. For a relatively low level of threshold T_c (e.g. $T_c < x_{23}$), then orbits of the system (2) may finally either remain in the invariant set $D_{\Gamma^1_{S_2}}$ or approach $(\infty, 0)$. This indicates that any patient whose initial viral loads and effector cells lie in the region $D_{\Gamma^1_{S_2}}$ could successfully maintain their viral loads less than a certain level under such a treatment regime. Whereas other patients, whose initial viral loads and effector cells lie in the region $R_+^2 \setminus D_{\Gamma^1_{S_2}}$, may fail to control the increase in the viral loads. For a relatively high level of threshold T_c (e.g. $T_c > x_{22}$), it follows from Fig.3 that the dynamics of the switching system (2) are the same as those for subsystem S_1. This means that therapy is actually not triggered when the threshold is relatively large.

For an intermediate threshold (e.g. $x_{23} < T_c < x_{22}$), any orbit initiating from D_Υ remains bounded, which implies that the HIV viral loads can be controlled. It also implies that any patient with initial HIV virus and effector cell populations in the region D_Υ can maintain his/her HIV virus population less than a low level by carrying out HIV virus-guided therapy with a suitable threshold. Whereas, patients with initial HIV virus and effector cell populations outside the region D_Υ can not prevent their HIV virus loads from increasing to infinity. Therefore, region D_Υ can be thought of as a controllable region. In such a scenario, for a previous given threshold T_c (therapy regime is fixed), different patients may have very different treatment outcomes. For a given intensity of therapy (ε_1 and ε_2 fixed) there is an optimal threshold such that the area of region D_Υ maximizes (as shown in Fig.9(A-B)). This indicates that for a patient with an initial HIV virus load may or may not maintain the growth of HIV virus loads, depending on the threshold level. Therefore, an individualized therapy is suggested, which indicates that the optimal choice of a treatment strategy for a given patient should depend on HIV virus and effector cell populations at outset and the proposed threshold level.

When therapy is implemented continuously, system (2) is actually subsystem S_2 and only those orbits starting from the invariant set $D_{\Gamma^1_{S_2}}$ can be controllable. However, by using proper HIV virus-guided therapy strategy (i.e. for $x_{23} < T_c < x_{22}$) the controllable region can be greatly enlarged. Moreover, it is worth mentioning that when $x_{21} < T_c < x_{11}$ the pseudo-equilibrium E_c is an unique attractor within the region D_Υ, in which the HIV virus stabilizes at the previously given value T_c. This suggests that proper combinations of threshold and initial HIV virus loads and effector cell counts based on a threshold policy can either preclude the uninhibited growth of HIV virus or lead to the HIV virus decreasing to a previously chosen level.

The work presented here is an approach to the dynamics of HIV management when plasma HIV-1 RNA-guided therapy is initiated. Our main results indicate that HIV viral loads could be maintained either below a certain level or stabilize at a previously given level or go to infinity (corresponding to the effector cells vanishing), depending on the threshold level and the initial HIV virus load and effector cell counts. This would explain why some clinical studies support the implementation of STIs while others do not, mainly due to various threshold levels or recruited patients with differing initial HIV virus loads and effector cell counts. Therefore, the findings suggest that it is essential to carefully choose the thresholds of plasma HIV-1 RNA copies and individualize the STIs for each patient based on their initial plasma HIV-1 RNA copies and effector cell counts.

Acknowledgements

The authors are supported by the national Megaproject of Science Research no. 2012ZX10001-001, by the National Natural Science Foundation of China(NSFC, 1171268(YX)), by the International Development Research Center, Ottawa, Canada(104519-010).

References

Abrahms, S. I., Brahmi, Z., 1988. Mechanism of K562-induced human natural killer cell inactivation using highly enriched effector cells isolated via a new single-step sheep erythrocyte rossette assay. Ann. Inst. Pasteur. Immunol. 139, 361-381.

Anaworanich, J., Nuesch, R., Le Braz, M., et al., 2006. CD4 guided scheduled treatment interruption compared to continuous therapy: results of the Staccato Trial. Lancet 368, 459-465.

Bernardo, M.D., Budd, C.J., Champneys, A.R., Kowalczyk, P., Nordmark, A.B., Tost, G.O., Pllrolnen, P.T., 2008. Bifurcations in nonsmooth dynamical system, SIAM Review 50(4), 629-701.

Blower, S.M., Gershengorn, H.B., Grant, R.M., 2007. A tale of two futures: HIV and antiritroviral therapy in San Francisco. Science 287, 650-654.

Callewaert, D.M., Meyers, P., Hiernaux, J., Radcliff, G., 1988. Kinetics of cellular cytotoxicity mediated by cloned cytotoxic T lymphocytes. Immunobiol. 178, 203-214.

Cauchemez, S., et al., 2009. Household transmission of 2009 pandemic influenza A (H1N1) virus in the United States. N. Engl. J. Med. 361, 2619-2627.

Carr, A., Samaras, K., Thorisdottir, A., Kaufmann, G.R., Chisholm, D.J., Cooper, D.A., 1999. Diagnosis, prediction, and natural course of HIV-1 protease-inhibitor-associated lipodystrophy, hyperlipidaemia, and diabetes mellitus: a cohort study. Lancet 353, 2093-2099.

Corless, R.M., Connet, G.H., Hare, D.E.G., Jeffrey, D.J., Knuth, D.E., 1996. On the Lambert W function. Adv. Comput. Math. 5, 329-359.

Danel, C., Moh, R., Minga, A., et al., 2006. CD4-guided structured antiretroviral treatment interruption stategy in HIV-infected adults in west Africa (Trivacan ANRS 1269 trial): a randomised trial. Lancet 367, 1981-1989.

DART Trial Team, 2008. Fixed duration interruptions are inferior to continuous treatment in African adults starting therapy with CD4 cell counts<200 cells/μl. AIDS 22, 237-247.

Day, T., Park, A., Madras, N., Gumel, A., Wu, J.H., 2006. When is quarantine a useful control strategy for emerging infectious diseases? Am. J. Epidemiol. 163, 479-485.

De Boer, R.J., Perelson, A.S., 1998. Target cell limited and immune control models of HIV infection: a comparison. J. Theor. Biol. 190, 201-214.

EL-Sadr, W.M., Lundgren, J.D., Neaton, J.D., et al., 2006. $CD4^+$ count-guided interruption of antiretroviral treatment. The strategies for Management of antiretroviral therapy study grou.. N. Engl. J. Med. 355, 2283-2296.

Guerrero Fernandez, M.L., Rivas, P., Molina, M., Garcia, R., De Gorgolas, M., 2005. Long-term follow-up of asymptomatic HIV-infected patients who discontinued antiretroviral therapy. Clin. Infect. Dis. 41, 390-394.

Hadjiandreou, M.M., Conejeros, R., Ian Wilson, D., 2009. Long-term HIV dynamics subject to continous therapy and structured treatment interruption. Chem. Eng. Sci. 64, 1600-1617.

Harrington, M., Carpenter, C., 2000. Hit HIV-1 hard, but only when necessary. Lancet 355, 2147-2152.

Hirschel, B., Flanigan, T., 2009. Is it smart to continue to study treatment interruptions? AIDS 23, 757-759.

Johnson, V.A., Brun-Vezinet, F., Clotet, B., Conway, B., et al., 2004. Update of the drug resistance mutations in HIV-1: 2004. Topics HIV Med 12, 119-124.

Komarova, N.L., Barnes, E., Klenerman, P., Wodarz, D., 2003. Boosting immunity by antiviral drug therapy: a simple relationship among timing, efficacy, and success. Proc. Natl. Acad. Sci. USA 100, 1855-1860.

Kuznetsov, V.A., Makalkin, I.A., Taylor, M.A., Petelson, A.S., 1994. Nonlinear dynamics of immunogenic tumors: parameter estimation and global bifurcation analysis, Bull. Math. Biol. 56(2), 295-321.

Kuznetsov, Yu. A., Rinaldi, S., Grafnani, A., 2003. One parameter bifurcations in planar Filippov systems. Inter. J. Bif. Chaos. 13, 2157-2188.

Lori, F., Maserati, R., Foli, A., et al., 2000. Structured treatment interruptions to control HIV-1 infection. Lancet 355, 287-288.

Maggiolo, F., Airoldia, M., Callegaro, A., et al., 2009. CD4 cell-guided scheduled treatment interruptions in HIV-infected patients with sustained immunologic response to HAART. AIDS 23, 799-807.

Maggiolo, F., Ripamonti, D., Grellegaro, A., et al., 2004. Effect of prolonged discontinuation of successful antiretroviral therapy on CD4 cells, a controlled, prospective trial. AIDS 18, 439-446.

Marchetti, G., Franzetti, F., Gori, A., 2005. Partial immune reconstitution following highly active antiretroviral therapy: can adjuvant interleukin-2 fill the gap? J. Antimicrob. Chemother. 55, 401-409.

Mocroft, A., Vella, S., Benfield, T.L., Chiesi, A., Miller, V., Gargalianos, P., et al., 1998. Changing patterns of mortality across Europe in patients infected with HIV-1. Lancet 352, 1725-1730.

Napolitano, L.A., 2003. Approaches to immune reconstitution in HIV infection. Topics in HIV Medicine: a Publication of the International AIDS Society, USA 11(5), 160-163.

Padmanabhan, C., Singh, R., 1995. Dynamics of a piecewise nonlinear system subject to dual harmonic excitation using parametric continuation. J. Sound. Vib. 184(5), 767-799.

Palella, F.J., Delaney, K.M., Moorman, A.C., Loveless, M.O., Fuhrer, J., Satten, G.A., et al., 1998. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. New. Engl. J. Med. 338, 853-860.

Pugliese, A., Gandolfi, A., 2008. A simple model of pathogen-immune dynamics including specific and non-specific immunity. Math. Biosci. 214, 73-80.

Rong, L.B., Feng, Z.L., Perelson, A.S., 2007. Emergence of HIV-1 Drug resistance during antiretroviral treatment. Bull. Math. Biol. 69, 2027-2060.

Ruiz, L., Paredes, R., Gómez, G., Romeu, J., et al., 2007. Antiretroviral therapy interruption guided by CD4 cell counts and plasma HIV-1 RNA levels in chronically HIV-1-infected patients. AIDS 21, 169-178.

Shu, H.Y., Wang, L., Watmough, J., 2014. Sustained and transient oscillations and chaos induced by delayed antiviral immune response in an immunosuppressive infection model. J. Math. Biol. 68, 477-503.

Tang, S.Y., Liang, J.H., 2013. Global qualitative analysis of a non-smooth Gause predatorCprey model with a refuge. Nonlin. Anal. TMA 76, 165-180.

Tang, S.Y., Xiao, Y.N., Wang, N., Wu, H.L., 2012. Piecewise HIV virus dynamic model with $CD4^+$ T cell count-guided therapy: I. J. Theor. Biol. 308, 123-134.

Tian, Y., Liu, X., 2014. Global dynamics of a virus dynamical model with general incidence rate and cure rate, Nonl. Anal. RWA 16, 17-26.

Utikin, V.L., Guldner, J., Shi, J.X., 2009. Sliding mode control in Electro-Mechanical system, second ed. Taylor & Fraacis Group, 2009.

Wang, A.L., Xiao, Y.N., 2013. Sliding bifurcation and global dynamics of a Filippov epidemic model with vaccination. Inter. J. Bif. Chaos. 23(8), Doi: 10.1142/S0218127413501447.

Xiao, Y.N., Xu, X.X., Tang, S.Y., 2012. Sliding Mode Control of Outbreaks of Emerging Infectious Diseases. Bull. Math. Biol. 74, 2403-2422.

Xiao, Y., Tang, S., Wu, J., 2015. Media impact switching surface during an infectious disease outbreak. Sci. Rep. 5, 7838; DOI:10.1038/srep07838.

Zhang, L., Ramratnam, B., Tenner-Racz, K., et al., 1999. Quantifying residual HIV-1 replication in patients receiving combination antiretroviral therapy, New. Engl. J. Med. 340, 1605-1613.

Appendix A

(A_1) **The property of the Lambert W function**

The Lambert W function (Corless et al., 1996) is defined to be a multivalued inverse of the function $z \mapsto ze^z$ satisfying

$$\mathrm{LambertW}(z)\exp(\mathrm{LambertW}(z)) = z.$$

And we denote it as W for simplicity. Note that the function $z\exp(z)$ has the positive derivative $(z+1)\exp(z)$ when $z > -1$. Define the inverse function of $z\exp(z)$ restricted on the interval $[-1, +\infty)$ to be $W(0, z)$. Similarly, we define the inverse function of $z\exp(z)$ restricted on the interval $(-\infty, -1]$ to be $W(-1, z)$. The branch $W(0, z)$ is defined on the interval $[-e^{-1}, +\infty)$ and it is monotonically increasing with respect to z. And the branch $W(-1, z)$ is defined on the interval $[-e^{-1}, 0)$ and it is a monotonically decreasing function with respect to z.

(A_2) **The definition domain of $y_l^{S_1}$ and $y_u^{S_1}$**

Let's consider the following equation

$$e^{-1} = -\frac{\beta}{r}\exp\left(\frac{\rho\ln(1+\omega x) - \delta\omega\ln(x) - \mu\omega x + h_1\omega}{r\omega}\right),$$

and rearranging it gives

$$\rho\ln(1+\omega x) - \delta\omega\ln(x) = \mu\omega x - h_1\omega + r\omega\ln\left(\frac{re^{-1}}{\beta}\right).$$

Denote

$$G_1(x) = \rho\ln(1+\omega x) - \delta\omega\ln(x), G_2(x) = \mu\omega x - h_1\omega + r\omega\ln\left(\frac{re^{-1}}{\beta}\right),$$

then by simple calculations we have

$$G_1'(x) = \frac{\rho\omega}{1+\omega x} - \frac{\delta\omega}{x}, \quad G_1''(x) = -\frac{\rho\omega^2}{(1+\omega x)^2} + \frac{\delta\omega}{x^2}.$$

Solving $G_1'(x) = 0$ with respect to x, we get an extreme point, denoted by $x_G = \delta/(\rho - \delta\omega)$, and $x_G > 0$ holds true due to $\rho - \mu - \delta\omega > 0$. Further, solving $G_1''(x) = 0$ yields two inflexion points, denoted by x_I^1 and x_I^2, where

$$x_I^1 = \frac{\delta\omega + \sqrt{\rho\delta\omega}}{\omega(\rho - \delta\omega)}, \quad x_I^2 = \frac{\delta\omega - \sqrt{\rho\delta\omega}}{\omega(\rho - \delta\omega)}$$

with $x_I^1 < x_G < x_I^2$.

Moreover, it is easy to see that $\lim_{x\to 0^+} G_1(x) = +\infty$, and solving $G_1'(x) = G_2'(x)$ with respect to x, yields two roots, which are exactly the first components of the two interior equilibria E_{11} and E_{12}. Let

$$l_1 = H_1(x_{11}, y_{11}), \quad l_2 = H_1(x_{12}, y_{12}),$$

then the family of closed orbits of subsystem S_1 is

$$\Gamma_h = \{(x, y)|H_1(x, y) = h, \ l_2 < h < l_1\}.$$

Furthermore, Γ_h converts to the equilibrium E_{11} as $h \to l_1$, and Γ_h becomes the homoclinic cycle as $h \to l_2$.

Therefore, the two curves are tangent at $x = x_{11}$ or x_{12}. If we choose h as a bifurcation parameter, then the domains of two branches of $y_l^{S_1}$ and $y_u^{S_1}$ can be determined as follows:

1. If $l_2 < h < l_1$, then there exist three intersect points between the two functions G_1 and G_2, denoted by $x_{\min}$, x_{mid} and $x_{\max}$. In this case, the two branches of $y_l^{S_1}$ and $y_u^{S_1}$ are well defined for all $x \in [x_{\min}, x_{\text{mid}}] \cup [x_{\max}, +\infty)$ with $y_l^{S_1} < r/\beta < y_u^{S_1}$.

2. If $h \leq l_2$ or $h \geq l_1$, then there is one intersect point between the two functions G_1 and G_2, denoted by $x_{\min}$. In this situation, we have that the two branches of $y_l^{S_1}$ and $y_u^{S_1}$ are well defined for all $x \in [x_{\min}, \infty)$ with $y_l^{S_1} < r/\beta < y_u^{S_1}$.

Similar results for $y_l^{S_2}$ and $y_u^{S_2}$ can be obtained by using the same methods as above.

Appendix B

Lemma 2: If the solution trajectory initiating from the point $P^{S_2} = (T_c, y_P^{S_2})$ on the segment $\{(T_c, y) : y_5 < y < (r - \varepsilon_1)/\beta\}$ first reaches the switching line $x = T_c$ at $P_1 = (T_c, y_{P_1})$ on the segment $\overline{AE}_7$ along the system S_2, and enters the region S_1 by crossing the switching line $x = T_c$, and then approaches the switching line $x = T_c$ again at the point $P_2 = (T_c, y_P^{S_1})$ on the segment $\{(T_c, y) : y_5 < y < (r - \varepsilon_1)/\beta\}$ along the system S_1, then we have $y_p^{S_2} < y_P^{S_1}$.

Proof. It follows from the first integral of the two subsystems and equations (8) and (10), we have that

$$\begin{aligned} y_{P_1} &= -\tfrac{r-\varepsilon_1}{\beta} W\left[-1, -\tfrac{\beta}{r-\varepsilon_1}\exp\left(\tfrac{\rho\ln(1+\omega T_c)-(\delta-\varepsilon_2)\omega\ln(T_c)-\mu\omega T_c+h_{P_1}^{S_2}\omega)}{(r-\varepsilon_1)\omega}\right)\right] \\ &= -\tfrac{r}{\beta} W\left[-1, -\tfrac{\beta}{r}\exp\left(\tfrac{\rho\ln(1+\omega T_c)-\delta\omega\ln(T_c)-\mu\omega T_c+h_{P_1}^{S_1}\omega)}{r\omega}\right)\right], \end{aligned} \tag{B.1}$$

which indicates that

$$\begin{aligned} &W\left[-1, -\tfrac{\beta}{r-\varepsilon_1}\exp\left(\tfrac{\rho\ln(1+\omega T_c)-(\delta-\varepsilon_2)\omega\ln(T_c)-\mu\omega T_c+h_{P_1}^{S_2}\omega)}{(r-\varepsilon_1)\omega}\right)\right] < \\ &W\left[-1, -\tfrac{\beta}{r}\exp\left(\tfrac{\rho\ln(1+\omega T_c)-\delta\omega\ln(T_c)-\mu\omega T_c+h_{P_1}^{S_1}\omega)}{r\omega}\right)\right]. \end{aligned} \tag{B.2}$$

Then, according to the property of the Lambert W function, we have that

$$\begin{aligned} &-W\left[0, -\tfrac{\beta}{r-\varepsilon_1}\exp\left(\tfrac{\rho\ln(1+\omega T_c)-(\delta-\varepsilon_2)\omega\ln(T_c)-\mu\omega T_c+h_{P_1}^{S_2}\omega)}{(r-\varepsilon_1)\omega}\right)\right] < \\ &-W\left[0, -\tfrac{\beta}{r}\exp\left(\tfrac{\rho\ln(1+\omega T_c)-\delta\omega\ln(T_c)-\mu\omega T_c+h_{P_1}^{S_1}\omega)}{r\omega}\right)\right]. \end{aligned} \tag{B.3}$$

That is, we have

$$\begin{aligned} &-\tfrac{r-\varepsilon_1}{\beta} W\left[0, -\tfrac{\beta}{r-\varepsilon_1}\exp\left(\tfrac{\rho\ln(1+\omega T_c)-(\delta-\varepsilon_2)\omega\ln(T_c)-\mu\omega T_c+h_{P_1}^{S_2}\omega)}{(r-\varepsilon_1)\omega}\right)\right] < \\ &-\tfrac{r}{\beta} W\left[0, -\tfrac{r}{\beta}\exp\left(\tfrac{\rho\ln(1+\omega T_c)-\delta\omega\ln(T_c)-\mu\omega T_c+h_{P_1}^{S_1}\omega)}{r\omega}\right)\right]. \end{aligned} \tag{B.4}$$

Then it follows from equations (7) and (9) that there is $y_P^{S_2} < y_P^{S_1}$. This completes the proof.

Table 1: The different types of all possible equilibria of system (2)

Values of T_c	E_{11}	E_{12}	E_{21}	E_{22}
$x_{21} > T_c$	VE	VE	RE	RE
$x_{21} < T_c < x_{11}$	VE	VE	VE	RE
$x_{11} < T_c < x_{12}$	RE	VE	VE	RE
$x_{12} < T_c < x_{22}$	RE	RE	VE	RE
$x_{22} < T_c$	RE	RE	VE	VE

Note that 'RE' denotes regular equilibrium and 'VE' represents virtual equilibrium

Figure legend

Figure 1:

The illustration of topological structure of the orbits of the subsystems. We denote the homoclinic orbit of the equilibrium E_{i2} as $\Gamma^1_{S_i}$ $(i = 1, 2)$. And we denote the stable codimension-1 manifolds and the unstable codimension-1 manifolds with respect to E_{i2} as $\Gamma^2_{S_i}$ and $\Gamma^3_{S_i}$ $(i = 1, 2)$, respectively. Here the curves are plotted using subsystem S_1 and the parameter values as $r = 1.8, \beta = 0.6, \omega = 0.55, \rho = 0.8, \mu = 0.23, \delta = 0.3$.

Figure 2:

The topological structure of the Filippov system (2) when $T_c < x_{23}$. All the parameter values are fixed as $r = 2.6, \beta = 1, \rho = 0.5, \omega = 0.1, \mu = 0.23, \delta = 0.5, \varepsilon_1 = 0.8, \varepsilon_2 = 0.1, T_c = 0.7$.

Figure 3:

The topological structure of the Filippov system (2) when $T_c > x_{22}$. Here $T_c = 10.5$ and other parameters are fixed as those in Fig.2.

Figure 4:

The topological structure of the Filippov system (2) when $x_{23} < T_c < x_{21}$ with (a) showing $y_6 < r/\beta$ and (b) showing $y_6 > r/\beta$. Here parameter $T_c = 1.2$, $r = 2.6$ in (a) and $r = 2$ in (b) and other parameters are fixed as those in Fig.2.

Figure 5:

Solutions of the Filippov system (2) when $x_{23} < T_c < x_{21}$ in subplot (a-b) and $x_{21} < T_c < x_{11}$ in subplot (c-d). Here $T_c = 1.2$ in (a) and (b) with the initial conditions of $(9, 3.2)$ and $(9, 5)$ respectively, and $T_c = 2.5$ in (c) and (d) with the initial conditions of $(9, 3.2)$ and $(9, 5)$ respectively. All the other parameters are fixed as those in Fig.2.

Figure 6:

The topological structure of the Filippov system (2) when $x_{21} < T_c < x_{11}$. Here $T_c = 2.5$ and other parameters are fixed as those in Fig.2.

Figure 7:

The topological structure of the Filippov system (2) when $x_{11} < T_c < x_{12}$. Here $T_c = 5.5$ and other parameters are fixed as those in Fig.2.

Figure 8:

The topological structure of the Filippov system (2) when $x_{12} < T_c < x_{22}$. Here $T_c = 6.5$ and other parameters are fixed as those in Fig.2.

Figure 9:

(A) The curves of $S_{D^L_\Upsilon}$ as T_c increases with ε_2 are fixed as 0.1; (B) The curves of $S_{D^L_\Upsilon}$ as T_c increases with ε_1 are fixed as 0.8; (C) The curves of $S_{D^L_\Upsilon}$ as ε_1 increases where $\varepsilon_2 = 0.1$; (D) The curves of $S_{D^L_\Upsilon}$ as ε_2 increases where $\varepsilon_1 = 0.8$. All other parameters are fixed as $r = 2.6, \beta = 1, \omega = 0.1, \rho = 0.5, \mu = 0.23, \delta = 0.5, T_\Upsilon = 10$.