

Optimal and MPC Switching Strategies for Mitigating Viral Mutation and Escape^{*}

Esteban A. Hernandez-Vargas * Richard H. Middleton * Patrizio Colaneri **

 * Hamilton Institute, National University of Ireland, Maynooth, Co. Kildare, Ireland (e-mail: Richard.Middleton@nuim.ie, abelardo_81@hotmail.com)
 ** DEI, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy (e-mail: colaneri@elet.polimi.it)

Abstract: This paper is motivated by the study of mutation in HIV infection. Combination antiretroviral therapy slows the clinical progression of HIV infection, however drug resistance due to viral mutation is a challenging problem. Some studies have speculated that alternating between drug regimens on a fixed schedule might forestall therapeutic failure. To further analyze this speculation, we consider a model of 64 viral strains with 3 drug combinations to analyze drug regimens to maximise the delay till viral escape. A model predictive control scheme is proposed for determining near optimal switching drug schedules. This technique is compared with an optimal control approach and with the strategy commonly used in clinical practice.

Keywords: Biological Systems, HIV Mutation, MPC

1. INTRODUCTION

HIV can only replicate inside cells. The process typically begins when a virus particle contacts an immune system cell that carries on its surface a special protein called CD4. Different drugs have been proposed to affect specific parts of the HIV life cycle, see Fig.1. Highly active antiretroviral therapy (HAART) is a combination therapy used to reduced viral replication and to delay the progression of the infection. However, HAART is not always successful, many patients experience virologic failure, defined as the inability to sustain suppression of HIV RNA levels to less than 50 copies/ml. Temporary suppression followed by increased viral load is referred to as viral rebound (Department of Health [2009]). This had lead to the conclusion that switching therapeutic options will be required lifelong in order to prevent HIV disease progression.

Antiretroviral drug sequencing provides a strategy to deal with virologic failure and anticipates that therapy will fail in a proportion of patients due to resistant mutations. The primary objectives of therapy sequencing are the avoidance of accumulation of mutations and control of multi-drugresistant viruses Martinez et al. [2008]. Using mathematical modeling, in Amato et al. [1998] it was hypothesized that alternating HAART regimens, even while plasma HIV RNA levels were lower than 50 copies/ml, would further reduce the likelihood of the emergence of resistance. This concept has preliminary support from a clinical trial Mar-





tinez et al. [2003], the SWATCH (SWitching Antiviral Therapy Combination against HIV) study.

The design of drug sequencing has been a point of discussion. Ryan et al. [2007] proposed a pattern of structured treatment interruptions preceding the introduction of the new regimen can significantly decrease the risk of resistance emerging to new regimens. However, structured treatment interruption were prohibited by the Department of Health [2009].

Motivated by switching regimens, Middleton et al. [2010] proposed a 4 variant, 2 drug combination model, which can be seen as a switched system. Using optimal control

 $^{^{\}star}$ This work was supported by Science Foundation of Ireland 07/PI/I1838.

strategies was shown the importance of alternating drug regimens.

However the solution to optimal control problems frequently results in a two point boundary condition problem, which can not be solved using regular integration techniques. Numerical algorithms have been proposed to determine optimal trajectories. Iterative solutions based on Pontryagin's maximum principle have been proposed, for example Kirk [1998], but without any guarantee of convergence. On the other hand, dynamic programming may be useful for problems of reasonable dimension Kirk [1998] and Middleton et al. [2010a].

Model Predictive Control (MPC) appears to be suitable for a suboptimal application to the biomedical area, due to its robustness to disturbances, model uncertainties and the capability of handling constraints. Using MPC techniques to plan treatment applications for HIV is not a new idea Zurakowski et al. [2006], Landi et al. [2010]. However, the models used in these previous approaches do not accurately reflect the interaction between different genotypes and drug treatments, and consequently do not predict the possibility of the appearance of highly resistant genotypes.

The paper is organized as follows. The notation used throughout the paper is introduced in Section 2. A virus mutation treatment model and the cost function to be minimized are presented in Section 3. The optimal control problem and the algorithm to compute optimal trajectories are reviewed in Section 4. A MPC algorithm for the computation of antiretroviral drug sequencing is proposed in Section 5. Simulations results to the mitigation of viral escape is introduced in Section 6. The paper is finalized in Section 7.

2. NOTATION

Throughout this paper, \mathbb{R} denotes the field of real numbers, \mathbb{R}^n stands for the vector space of all *n*-tuples of real numbers, $\mathbb{R}^{n \times n}$ is the space of $n \times n$ matrices with real entries, and \mathbb{N} denotes the set of natural numbers. Matrices or vectors are said to be positive (non-negative) if all their entries are positive (non-negative). We write A' for the transpose of A, and exp(A) for the matrix exponential of A.

3. VIRUS MUTATION TREATMENT MODEL

We consider a model with n different viral genotypes, with viral populations, $V_i : i = 1, ...n$; and N different possible drug therapies that can be administered, represented by $\sigma(t) \in \{1, ..., N\}$, where σ is permitted to change with time, t. In the model equations below, T_H represents the uninfected CD4+T cell population, T_i^* are the infected cells and L_i represents the latently infected cells. The equations for the HIV treatment model are given below:

$$\dot{T}_{i}^{*} = \psi k_{1,\sigma}^{i} T_{H} V_{i} + a_{1} L_{i} - d_{1} T_{i}^{*} + \sum_{1}^{n} \mu m_{i,j} V_{j} T_{H} \quad (1)$$

$$\dot{L}_i = (1 - \psi)k_{1,\sigma}^i T_H V_i - a_1 L_i - d_2 L_i \quad (2)$$

$$\dot{V}_i = k_{2,\sigma}^i T_i^* - k_3 T V_i - d_3 V_i \quad (3)$$

HIV can infect a number of different cells; activated CD4+T cell, resting CD4+T cell, quiescent CD4+T cell, macrophages and dentritic cells. For simplicity, we consider only CD4+T cells as viral hosts. The infectivity rate can be expressed as $k_{1,\sigma}^i$, this parameter depends of the genotype and the therapy that is being used. Once CD4+T cells are infected, a proportion of cells, ψ passes into the infected cells population, whereas a proportion $(1 - \psi)$ passes into the latently-infected cell population. These latently-infected cells might be activated later and start reproducing virus, this activation is represented through the a_1 term. Viral proliferation is achieved in infected activated CD4+T cells, this is represent by $k_{2,\sigma}$, which depends on the fitness of the genotype and the therapy. The mutation rate is represented by μ , the death or decay rates are represented by d_1, d_2, d_3 for T^*, L, V respectively. $m_{i,j} \in \{0,1\}$ represents the genetic connections between genotypes.

Under normal treatment circumstances both simulations and clinical data suggest that healthy CD4+T counts are approximately constant Middleton et al. [2010]. This assumption allows us to simplify the dynamics to being essentially linear. Then the system (1-3) can be rewritten as follows;

$$\dot{x} = \begin{bmatrix} \Lambda_{1,\sigma} & 0 & \dots & 0 \\ 0 & \Lambda_{2,\sigma} & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \Lambda_{n,\sigma} \end{bmatrix} x + \mu T_H M x$$
(4)

where $M = m_{i,j}$, $\Lambda_{i,\sigma}$ is given by

$$\Lambda_{i,\sigma} = \begin{bmatrix} -d_1 & a_1 & \psi k_{1,\sigma}^i T_H \\ 0 & -(a_1 + d_2) & (1 - \psi) k_{1,\sigma}^i T_H \\ k_{2,\sigma}^i & 0 & -(k_3 T_H + d_3) \end{bmatrix}$$
$$x' = [T_1^*, L_1, V_1, \dots, T_n^*, L_n, V_n].$$

3.1 A 64 variant, 3 drug combination model

We propose a model with 64 genetic variants, that is n = 64, and 3 possible drug therapies, N = 3. The viral variants (also called 'genotypes') are organized in a threedimensional lattice as is shown in Fig. 2. This lattice is based on the simplifying assumption that multiple independent mutations are required to achieve resistance to all therapies.

The wild type genotype (g_1) would be the most prolific variant in the absence of any drugs, however, it is also the variant that all drug combinations have been designed to combat, and therefore is susceptible to all therapies. On the other hand, after several mutations the highly resistant genotype (g_{64}) is a strain with low proliferation rate,

and



Fig. 2. Mutation tree for a 64 genotype and 3 drug combination model

but that is resistant to all drug therapies. For simulation purposes, we shall use parameters shown in Table 1.

Table 1. Parameters Values

Parameter	Value	Value taken from:
k_1	3.8×10^{-3}	Conejeros et al. [2007]
k_2	6.45×10^{-1}	Fitted
k_3	7.79×10^{-6}	Conejeros et al. [2007]
d_1	0.4	Conejeros et al. [2007]
d_2	0.005	Conejeros et al. [2007]
d_3	2.4	Conejeros et al. [2007]
a	3×10^{-4}	Conejeros et al. [2007]
ψ	0.97	Conejeros et al. [2007]

In this model therapies are composed of reverse transcriptase inhibitors and protease inhibitors, which can affect two specific parts of the HIV cycle. This is represented by

$$k_{1,\sigma}^{i} = k_{1} f_{i} \eta_{\sigma,i} \tag{5}$$

$$k_{2,\sigma}^{i} = k_{2} f_{i} \theta_{\sigma,i} \tag{6}$$

where $\eta_{\sigma,i}$ represents the infection efficiency for genotype i under treatment σ , and $\theta_{\sigma,i}$ expresses the production efficiency for the genotype i under treatment. We assume that in the absence of treatment mutation reduces the fitness of the genotype, thus we use linear decreasing factors for f_i , which represents the fitness of the genotype i. The drug efficiencies can be seen in Fig. 2, where the arrows indicate the efficiency of the drug. For instance, the genotypes $g_1, g_5, ..., g_{61}$ are all on one face of the lattice, and are fully susceptible to therapy 1. The opposite face, $g_4, g_8, ..., g_{64}$ describes all genotypes highly resistant to therapy 1.

3.2 Cost Function Motivation

We consider a cost function of the form

$$J := c'x(T) \tag{7}$$

where c = [0, 0, 1, ..., 0, 0, 1], and T is an appropriate final time. This cost was proposed by Middleton et al. [2010],

and penalises the total viral load. For biological reasons, if the total viral load is small enough during a finite time of treatment, then there is a significant probability that the total viral load becomes zero. Munoz et al. [1997] noted that the stronger the suppression of viral replication the less likely a significant resistance will emerge.

4. OPTIMAL CONTROL PROBLEM REVIEW

The model for the treatment of viral mutation given in (4) is described in continuous time. In practice, measurements can only reasonably be made at certain periodic intervals. For simplicity, we consider a regular treatment interval τ , during which treatment is fixed. If we use $k \in \mathbb{N}$ to denote the number of intervals since t = 0, then

$$x(k+1) = A_{\sigma(k)}x(k) \tag{8}$$

where $x(k) = x(k\tau)$ is the sampled state, $A_{\sigma} := exp((diag(\Lambda_{n,\sigma}) + \mu M)\tau), \sigma(k)$ is the switching sequence $\sigma(k) \in \{1, 2, ..., N\}$, and $x(0) = x_0$ is the initial condition. Clearly, $\sigma(k)$ constrains $A_{\sigma(k)}$ to jump among the N vertices of the matrix polytope $A_1, ..., A_N$.

The cost function (7) is to be minimized over all admissible switching sequences. The optimal switching signal, the corresponding trajectory and the optimal cost function will be denoted as $\sigma^{o}(k)$, $x^{o}(k)$ and $J(x_0, x^0, \sigma^0)$ respectively. Letting $u = \sigma(k)$, and using the Hamilton-Jacobi-Bellman equation for the discrete case, we have;

$$V(x,k) = \min_{u \in U} \{ V(k+1, x(k+1)) \}$$
(9)

where denoting the costate vector by p(k), the general solution for this system is

$$V(x(k),k) = p(k)'x(k)$$
(10)

Using equations (7), (8), (9) and (10), we obtain the following system

$$x^{o}(k+1) = A_{\sigma^{o}(k)}x^{o}(k)$$

$$p^{o}(k) = A'_{\sigma^{o}(k)}p^{o}(k+1)$$
(11)
$$\sigma^{o}(k) = \operatorname{argmin}\{p^{o}(k+1)'A_{s}x^{o}(k)\}$$

with boundary conditions $x(0) = x_0$ and p(T) = c. **Remark 1.** Notice that equations (11) are inherently nonlinear. The system must be integrated forward whereas the co-state equation must be integrated backward, both according to the coupling condition given by the switching rule. (11) is a two point boundary value problem, and can not be solved using regular integration techniques.



4.1 Algorithm for Computing Optimal Trajectories

Because of the discrete nature of the input, a possible numerical solution for the optimal control problem stated in (11) is to make an exhaustive examination of every possible switching rule. Given the cost (7) and the initial i_T, i_T

5

condition x(0) the optimal control problem can be written as

$$\min_{i_1,\dots,i_1 \in \{1,\dots,N\}} c' A_{i_T} A_{i_{T-1}} \dots A_{i_1} x(0)$$
(12)

Let us recursively define the sequence of matrices

$$\Omega_T = c$$

$$\Omega_{k-1} = [A'_1 \Omega_k \ A'_2 \Omega_k \ \dots \ A'_N \Omega_k]$$

Then we have that $V(x,0) = \min_i \Omega'_{0,i} x$, where $\Omega_{0,i}$ is the *i*th column of Ω_0 and in general

$$V(x,k) = \min \Omega'_{T-k,i} x(k) \tag{13}$$

At each step of the evolution the feedback strategy can be computed as

$$u(x(k)) = \operatorname*{argmin}_{i} \Omega'_{T-k,i} x(k) \tag{14}$$

namely selecting the smallest component of the vector $\Omega'_{T-k}x(k)$. The implementation of this strategy that we call "brute force" requires storing the columns of $\Omega'_{T-k}x(k)$ whose number would be $1, N, N^2, N^3, \ldots, N^T$. In order to reduce computational and storage demands we can combine the algorithm above with a dual version Middleton et al. [2010a].

5. MODEL PREDICTIVE CONTROL

Model predictive control problem can be formulated as solving on-line a finite horizon open-loop optimal control problem subject to system dynamics and constraints involving states and controls Allgower et al. [2002]. The basic idea of MPC is based on measurements obtained at time t, the controller predicts the future dynamic behavior of the system over a prediction horizon T_p and computes an open-loop optimal control problem with control horizon T_c , to predict the future input for the system, see Fig.3.





Due to disturbances, measurement noise and model-plant mismatch, the true system behavior is different from the predicted one. In order to incorporate a feedback mechanism, the first step of the optimal control sequence is implemented. When the next measurement becomes available, at time $t + \tau$, the whole procedure -prediction and optimization- is repeated to find a new input function with the control and prediction horizons moving forward.

5.1 Mathematical Formulation of MPC

From the biological nature of HIV infection, the system (8) is unstable and in fact not stabilizable. This is because of the existence of a highly resistant genotype that is not affected by any treatment. Therefore, once the highly resistant mutant has "emerged" the population will "blow up" after a period of time. The objective of MPC is to suppress the total viral load, (7) for as long as possible. In order to distinguish the real system and the system model used to predict the future for the controller, we denote the internal variables in the controller by a bar $(\bar{x}, \bar{\sigma})$. We can formulate the following model predictive control problem;

Problem 1 Find

$$\min J(x(t), \bar{\sigma}; T_c, T_p),$$

with

$$J(x(t),\bar{\sigma};T_p,T_c) := c'x(t+T_p)$$

$$\bar{x}(k+h+1/k) = A_{\bar{\sigma}(k+h/k)}\bar{x}(k+h/k)$$
$$\bar{\sigma}(k+h/k) \in D$$
$$\bar{x}(k+h/k) \in X$$

For numerical solution, we use the next algorithm;

MPC Algorithm

- (1) Compute the open-loop optimal control for a receding horizon time T_p
- (2) Apply only the first input of the optimal command sequence to the system
- (3) The remaining optimal inputs are disregarded
- (4) Collect the new measurement from the system
- (5) Continue with point (1) until the final time is reached

In this algorithm, T_p has to be chosen in advance. It is important to mention that the shorter the horizon, the less costly the solution of the on-line optimization problem. The method to solve the open-loop optimal problem using (14) has an exponential growth. Then, it is desirable use short horizons MPC schemes for computational reasons. In general it is not true that a repeated minimization over a finite horizon objective in a receding horizon manner leads to an optimal solution for the the infinite horizon problem Bitmead et al. [1990]. In fact, both solutions differ significantly if a short horizon is chosen.

5.2 MPC Treatment Scheduling

MPC algorithms typically achieve superior performance with respect to other control strategies when manipulated and controlled variables have constraints to meet. A thorough overview of the MPC history can be found in Morari et al. [1999]. However, for this particular application the system is not stabilizable. The MPC objective is to delay as far as possible the escape of the system. In mathematical terms, the problem could be thought as a minimization of the worst eigenvalue of A_{σ} under the switching σ . However, this would only be expected to be optimal over an infinite time horizon. In general the finite set of possible control values causes problems for many control design techniques, but for this problem helps MPC by making the optimization easier to solve.

6. SIMULATION STUDIES

We constrain the control to take the values of 1, 2 or 3, this represent the possible treatments. It is important to mention that patient can not be under two or more treatments at the same moment. To be consistent with application in a clinical setting, we simulate measurements and new clinical decisions being made every 3 months. Thus the decision time τ is fixed to 90 days.

6.1 Switch on Virologic Failure

For comparison purposes, we introduce a "switched on virologic failure" strategy. This is based on the guidelines for the use of antiretrovirals agents in HIV-1 infected adults presented by the Department of Health [2009]. Virologic failure means virologic rebound after complete suppression. There is no clear consensus on the optimal time to change therapy for virologic failure. The most accepted approach allows detectable viremia up to an arbitrary level (e.g., 1000-5000 copies/ml). For simulation purpose, we will consider as a virologic failure when viral load is > 1000 copies/ml. This is because, the ongoing viral replication in the presence of antiretroviral drugs promotes the selection of drug resistance mutations and may limit future treatment options.

Using the virologic failure Treatment, see in Fig.4, there are three changes in therapy in a period of 4 years. The first therapy keeps the viral concentration below 1000 copies/ml for 1 year, however resistant genotypes appear. Then the second treatment is introduced, the viral population is decreased for a while, but because there is a high concentration of infected cells already, virologic failure occurs in a shorter period of time. It is then necessary to introduce the third treatment. However, we can observe how the viral load starts to escape after 4 years in this model. An important fact is that latently infected cells remain almost constant, this fits with clinical and theoretical studies which agree that these cells played a very important role for late stage HIV infection.



Fig. 4. Simulation of switch virologic failure treatment strategy

6.2 Implementation of MPC Algorithms

Using 10 months as prediction horizon for the MPC, we present in Fig. 5 the closed-loop response of T^* , L, V and the switched drug therapy over a period of 14 years. The MPC algorithm accomplishes the main goal of containing the viral load as long as possible.



Fig. 5. Closed loop treatment using MPC

In fact, for this example the virologic failure would occur after 12 years, this means that we can extend the virologic failure for 8 years more compared to the current clinical assessment. In addition the total population of infected CD4+T cells is decreased for a period of 6 years. These results are consistent with the preliminary clinical SWATCH trial Martinez et al. [2003]. This trial concluded that proactive alternation of antiretroviral regimens might extend the long-term effectiveness of treatment options without adversely affecting patients' adherence or quality of life. We may notice in Fig. 5 how the therapy switching is irregular for the first 4 years, but it remains quite regular for the remaining years. This means that perhaps using an oscillating treatment could be a good strategy to mitigate the viral replication when it is not possible to measure the complete HIV state.

In fact, oscillating strategy is very similar to the trial SWATCH proposed by Martinez et al. [2003]. We notice in Table 2 that the oscillating drug regimen gives very close results to the MPC and optimal control. Based on simulation results and clinical trials we could infer that this approach would minimize emergence of drug-resistant virus better than frequent monitoring for viral rebound. For comparison purpose, we test one step ahead MPC strategy, which gives very low computational time and has good results with respect to the other strategies.

Table 2. Simulation results for a 5 years period treatment

Strategy	Viral Load	Computational Time
Virologic Failure	5.14e28	1.2 sec
MPC	2.31e-9	4.4 sec
Optimal Control	2.20e-9	32 min
Oscillating	6.22e-9	1.5 sec
One step ahead	6.74e-8	1.9 sec

6.3 Computational Resources

From these results, we conclude that alternating between drugs regimens is important to suppress HIV RNA levels maximally and prevent further selection of resistant mutations. However, to find the optimal trajectories can require high computational resources. For example, using a "brute force" approach which analyzes all possible combinations for the rapy 1, 2 and 3 with decision time $\tau = t_d$ for a period of T, there are $3^{\frac{T}{t_d}}$ possible combinations. Considering a period treatment of 5 years and decision time of 3 months, we can compare in Table 2 the computational time for different strategies. For instance, to compute the open-loop optimal switching trajectory requires 32 min simulation time, while on the other hand the MPC just require 4.4 sec; the difference on the viral load between these techniques is very small. Then we can infer that relaxing the condition of optimality, MPC may reduce dramatically simulation resources.

7. CONCLUSIONS

In this paper, a 64 variant with 3 drug combination model is presented to analyze different drug regimens to mitigate HIV escape. Simulation results show the potential of proactive switching and alternation of antiretroviral regimens with drugs to extend the overall long-term effectiveness of treatment options. These results are consistent with preliminary clinical trials. At the expense of some conservatism in the cost function, MPC is a very appropriate framework for this problem, because it provides close results to the open-loop optimal control problem, extends the virologic failure time and reduces computational resources.

REFERENCES

- Panel of Antiretroviral Guidelines for Adults and Adolescents. Guidelines for the use of Antiretroviral agents in HIV-1 infected adults and adolescents. *Department of Health and Human Services*, December 1, 2009
- J. Martinez-Cajas and M.A. Wainberg. Antiretroviral Therapy: Optimal Sequencing of Therapy to Avoid Resistance, Drugs, 139:81-89, 2003
- R.M. D'Amato, R.T. D'Aquila and L.M. Wein. Management of antiretroviral therapy for HIV infection: modelling when to change therapy. *Antivir Ther.* 3:147-158, 1998
- J. Martinez-Picado, E. Negredo and et al. Alternation of Antiretroviral Drug Regimens for HIV infection. Annals of Internal Medicine, 68(1):43-72, 2003
- R. Zurakowski and D. Wodarz. Treatment interruptions to decrease risk of resistance emerging during therapy switching in HIV treatment. *CDC*, New Orleans, USA, 2007
- E.A. Hernandez-Vargas, R.H. Middleton, P. Colaneri, and F. Blanchini. Discrete-time control for switched positive systems with application to mitigating viral escape. *International Journal of Robust and Nonlinear Control*, 2010
- D.E. Kirk. *Optimal Control theory*, Dover Publications, New York, 1998.
- E.A. Hernandez-Vargas, R.H. Middleton, P. Colaneri, and F. Blanchini, Dyanmic Optimization Algorithm to Mitigate HIV escape. *CDC*, Atlanta, USA, 2010
- R. Zurakowski and A.R. Teel. A Model Predictive Control based scheduling method for HIV therapy. *Journal of Theoretical Biology*, 238:368-382, 2006
- G. Pannocchia, M. Laurino and A. Landi. A Model Predictive Control Strategy Toward Optimal Structured Treatment Interruptions in Anti-HIV Therapy. *IEEE Transactions on Biomedical Engineering*, 57(5):1040-1050, 2010
- J. Mellors, A. Munoz, J. Giorgi, et al. Plasma viral load and CD4+ lymphocytes as prognostic markers of HIV-1 infection. Ann Intern Med, 126(12):946-954, 1997
- R. Findeisen and F. Allgower. An introduction to Nonlinear Model Predictive Control. 21st Benelux Meeting on Systems and Control, Veldhoven, 2002
- R. R. Bitmead, M. Gevers and V. Wertz. Adaptive Optimal Control. *The thinking Man's GPC*, Prentice Hall, New york, 1990
- M. Morari and J.H. Lee. Model Predictive Control: past, present and future. Computers & Chemical Engineering, 23:667-682, 1999
- M. Hadjiandreou, R. Conejeros and V. Vassiliadis. Towards a Long-Term Model Construction for the Dynamic Simulation of HIV Infection. *Mathematical Bio*science and Engineering, 4:489-504, 2007.