Analysis of a Cell-Cycle Specific Model for Cancer Chemotherapy

Urszula Ledzewicz
Department of Mathematics and Statistics,
Southern Illinois University at Edwardsville,
Edwardsville, Illinois, 62026-1653
e-mail: uledzew@siue.edu

Heinz Schättler
Department of Systems Science and Mathematics,
Washington University,
St. Louis, Missouri, 63130-4899
e-mail: hms@cec.wustl.edu

Abstract

A class of mathematical models for cancer chemotherapy which has been described in the literature takes the form of an optimal control problem with dynamics given by a bilinear system. In this paper we analyze a three-dimensional model in which the cell-cycle is broken into three compartments. The cytostatic agent used as control to kill the cancer cells is active in a compartment which combines the second growth phase and mitosis where cell-division occurs. A blocking agent is used as a second control to slow down the transit of cells during synthesis, but does not kill cells. The cumulative effect of the killing

agent is used to model the negative effect of the treatment on healthy cells. It is shown that singular controls are not optimal. This eliminates treatments where during some time only a portion of the full drug dose is administered. Consequently only treatments which alternate between a full and no dose, i.e. so-called bang-bang controls, can be optimal for this model. Both necessary and sufficient conditions for optimality of treatment schedules of this type are given.

1 Introduction

Models for cancer chemotherapy have a long history in mathematical biology (see, for instance, [6, 9, 20, 22, 28, 29]). While these approaches based on quantifying dynamics and objective have their critics in the medical community (e.g. [35, 37]), in recent years there has been renewed interest due to better models, but also due to a refinement of the techniques which can be used to analyze the problems. In addition to various new models for the chemotherapy of cancer [3, 7, 19], the modeling of HIV has become an important and very active area [11, 12, 13]. For example, Kirschner, Lenhart and Serbin [12] consider an optimal control problem for chemotherapy of HIV which accounts for latently infected and actively infected $CD4^+$-cells and Fister and Panetta [7] analyze a mathematical model for cancer chemotherapy which takes into account bone-marrow destruction.

Most mathematical models for optimal control of cancer chemotherapy are based on cell-cycle kinetics and treat this cell cycle as the object of control [9, 31, 32]. We briefly describe the model and some of the biology behind it. The *state* variable is given by the *number of cancer cells* and the *controls* are the *drug dosages*. The main active ingredient is a cytostatic agent which kills cancer cells and healthy cells alike. The goal is to maximize

the number of cancer cells which the agent kills while keeping the toxicity to the normal tissues acceptable. The dynamics of the cell cycle within limits indeed allows for such a feature by having the agents be active at different stages. The cell cycle is modeled in the form of *compartments* which describe the different cell phases or, in simplified models, combine phases of the cell cycle into clusters. Each cell passes through a sequence of phases from cell birth to cell division. Starting point is a growth phase G_1, after which the cell enters a phase S where DNA synthesis occurs. Then a second growth phase G_2 takes place in which the cell prepares for mitosis or phase M. Here cell division occurs. Each of the two daughter cells can either reenter phase G_1 or may simply lie dormant for some time in a separate phase G_0 until reentering G_1, thus starting the entire process all over again. In the two-compartment model phases G_0, G_1 and S are combined into the first compartment while G_2 and M define the second compartment; in the three-compartment model the growth phases G_0 and G_1 are separated from the synthesis phase S. In both models the killing agent is applied in the G_2/M phase when the cell is most vulnerable. While this is the only control considered in the two-compartment model, in the three-compartment model in addition a blocking agent is introduced which slows down the development of cells in the synthesis phase S [32]. An example of such a drug is, for instance, Cyclophosphamide, which affects DNA replication, while one of the most commonly used killing agents in chemotherapy is Taxol [7].

So far analytical approaches to problems of optimal control for cancer chemotherapy of the two- and three-compartment models developed by Kimmel and Swierniak [10, 30, 31, 32] were limited to applications of the Pontryagin Maximum Principle [25]. Beyond that controls and trajectories resulting from the Maximum Principle were only analyzed numerically

[4, 5]. This analysis centered around both bang-bang and singular controls. While bang-bang controls correspond to treatment protocols which alternate maximum doses of chemotherapy with rest periods when no drug is administered, singular controls correspond to applying varying doses at less than their maximum. Although bang-bang controls, which are widely used as protocols in medical treatments, seemed to be the more natural choice as candidates for optimality, and it was even observed numerically that singular protocols actually give the worst performance [32], the question of optimality of singular controls was left open [33, 34]. In recent papers we have analyzed the structure of optimal controls for the 2-compartment model of cancer chemotherapy. Using high-order necessary conditions for optimality we showed that singular arcs are not optimal [15] and then developed sharp necessary and sufficient conditions for local optimality of bang-bang controls [18]. In this paper we consider the three-compartment model also due to Kimmel and Swierniak [31, 32]. Compared with the simpler two-compartment model, this problem is a multi-input control system since in addition to the killing agent also a blocking agent is considered. We again exclude the optimality of singular controls using high-order necessary conditions for optimality. Consequently an analysis of bang-bang controls, the natural candidates for optimality, is given. Using results from [24] we develop simple and easily verifiable conditions which allow to determine the local optimality of these controls. Summarizing, our results show that protocols which would only use a partial drug dose at some time are not optimal and for those protocols which alternate full drug doses with rest periods in between we can determine which schedules provide at least a local minimum and which ones are not optimal at all.

2 The 3-Compartment model

We give a brief review of the model based on [32]. The first compartment corresponds to the dormant phase G_0 and the first growth phase G_1, the second to the synthesis phase S, and the second growth phase G_2 and mitosis M are combined into the third compartment. Let $N_i(t)$, $i = 1, 2, 3$, denote the number of cancer cells in the i-th compartment at time t. The transit times of cells through phases of the cell cycle vary, particularly in malignant cells. In the simplest models an exponential distribution is used to model the transit times and the expected number of cells exiting the i-th compartment is given by $a_i N_i(t)$, where a_i is the parameter of the exponential distribution related to the inverse of the transit time. Assuming for the moment that no external stimuli are present, the inflow of the second (respectively third) compartment equals the outflow of the first (respectively second) and we therefore have for $i = 2, 3$, that

$$\dot{N}_i(t) = -a_i N_i(t) + a_{i-1} N_{i-1}(t). \tag{2.1}$$

Cell division is represented by a factor 2 in the equation which describes the flow from the third into the first compartment:

$$\dot{N}_1(t) = -a_1 N_1(t) + 2a_3 N_3(t). \tag{2.2}$$

Hence, the unperturbed dynamics of the cell cycle, or the number of cells in a particular compartment, can be represented by a system of linear ordinary differential equations if there are no external stimuli present.

Drug treatment influences the cell cycle in many possible ways and in this model only the two most fundamental aspects are considered, *cell-killing* and *cell-arrest*. The implications of cell-killing are obvious; cell arrest tries

to slow down the growth of malignant cells in the sense that it prevents cells from reaching the phase where cell division occurs. In this model cell arrest in phase S is considered and cells are released when a second G_2/M-specific cytotoxic agent is at a maximum destroying potential. This also will allow better protection for the normal cells since they will be less dispersed, thus traveling faster through the G_2/M phases. Specifically, a cytostatic blocking agent v is applied to slow the transit times of cancer cells during the synthesis phase S. As a result the flow of cancer cells from the second into the third compartment is reduced by a factor $v(t)$ to $v(t)a_2N_2(t)$, $0 < v_{\min} \leq v(t) \leq 1$, of its original flow. The control $v(t) = 1$ corresponds to no drug being applied while a reduction to $v_{\min}$ occurs with a full dose. Cells are released when a second G_2/M-specific cytotoxic agent u is at a maximum destroying potential. This makes sense from a biological standpoint for a couple reasons. First, in mitosis M the cell becomes very thin and porous; hence, the cell is more vulnerable to an attack while there will be a minimal effect on the normal cells since due to cell-arrest they will be less dispersed, thus traveling faster through the G_2/M phases. Second, chemotherapy during mitosis will prevent the creation of daughter cells. The control u represents the dose of the drug administered with the value $u = 0$ corresponding to no treatment and $u = 1$ corresponding to a maximum dose. It is assumed that the dose stands in a direct relation to the fraction of cells which are being killed in the G_2/M phase. Therefore only the fraction $(1 - u)a_3N_3$ of cells reenters phase G_1 and undergoes cell division. All cells

a_3N_3 leave compartment G_2/M. Thus the mathematical model becomes

$$\dot{N}_1 = -a_1N_1 + 2(1-u)a_3N_3, \tag{2.3}$$

$$\dot{N}_2 = -va_2N_2 + a_1N_1, \tag{2.4}$$

$$\dot{N}_3 = -a_3N_3 + va_2N_2, \tag{2.5}$$

with initial conditions $N_i(0) = N_{i0}$, $i = 1, 2, 3$, all positive.

The performance index or objective to be minimized is chosen as

$$J = r_1N_1(T) + r_2N_2(T) + r_3N_3(T) + \int_0^T u(t)dt \to \min. \tag{2.6}$$

The coefficients r_i are weights and the penalty term $r_1N_1(T) + r_2N_2(T) + r_3N_3(T)$ represents a weighted average of the total number of cancer cells at the end of an assumed fixed therapy interval $[0, T]$. The integrand u represents the fraction of ineffective cell divisions and thus corresponds to the number of cancer cells killed. Since the drug kills healthy cells alike, this is also used to model the negative effect of the drug on the normal tissues or its toxicity. Thus the Lagrangian models the cumulative negative effects of the treatment. The blocking agent v does not kill cells and therefore is not included in the objective.

3 The structure of the mathematical model

The dynamics of the three compartment model is described by a *bilinear system* [21]. If we set $N = (N_1, N_2, N_3)^T$, $r = (r_1, r_2, r_3)$ and consider a multi-dimensional control variable (u, v), then the general form of the dynamics is

$$\dot{N}(t) = (A + uB_1 + vB_2)N(t), \quad N(0) = N_0, \tag{3.7}$$

where A, B_1 and B_2 are fixed (3×3)-matrices given by

$$A = \begin{pmatrix} -a_1 & 0 & 2a_3 \\ a_1 & 0 & 0 \\ 0 & 0 & -a_3 \end{pmatrix}, \tag{3.8}$$

and

$$B_1 = \begin{pmatrix} 0 & 0 & -2a_3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; B_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -a_2 & 0 \\ 0 & a_2 & 0 \end{pmatrix}. \tag{3.9}$$

The objective is to minimize

$$J(u) = rN(T) + \int_0^T u(t)dt \tag{3.10}$$

subject to the dynamics (3.7) and initial condition $N(0) = (N_{10}, N_{20}, N_{30})$ over all Lebesgue measurable functions (u, v) which take values in $[0, 1] \times [v_{\min}, 1]$.

Note that since admissible controls (u, v) take values in a compact set in $\mathbb{R}^2$, it follows that the norm of matrix $A + uB_1 + vB_2$ is bounded over the interval $[0, T]$ and therefore the right-hand side of the differential equation (3.7) is linearly bounded. It follows from well-known results about ordinary differential equations that for any admissible control the corresponding trajectory (i.e. solution to the dynamics) exists on all of $[0, T]$.

Obviously only states $N(t)$ for which each component is positive are meaningful. However, it is not necessary to add this condition as extra state-space constraint since it is automatically satisfied because of structural properties of the model. Let $\mathbb{P} = \mathbb{R}^3_+ = \{N \in \mathbb{R}^3 : N_i > 0 \; for \; i = 1, 2, 3\}$ denote the first quadrant. It is easy to see [16] that, regardless of the control u which is applied, all coordinates of $N(t)$ remain positive for all times $t \geq t_0$

if each coordinate of $N(t_0)$ is positive. Hence *the physically meaningful part of the state space is positively invariant* for the control system.

The first-order necessary conditions for optimality are given by the Pontryagin Maximum Principle [25]: If (u_*, v_*) is an optimal control, then there exists an absolutely continuous function λ, which we write as row-vector, $\lambda : [0,T] \to (\mathbb{R}^3)^*$, satisfying the adjoint equation with transversality condition,

$$\dot{\lambda} = -\lambda(A + u_* B_1 + v_* B_2), \qquad \lambda(T) = r, \tag{3.11}$$

i.e. componentwise

$$\dot{\lambda}_1 = a_1(\lambda_1 - \lambda_2), \qquad \lambda_1(T) = r_1, \tag{3.12}$$

$$\dot{\lambda}_2 = a_2 v_*(\lambda_2 - \lambda_3), \qquad \lambda_2(T) = r_2, \tag{3.13}$$

$$\dot{\lambda}_3 = a_3(\lambda_3 - 2(1 - u_*)\lambda_1), \qquad \lambda_3(T) = r_3, \tag{3.14}$$

such that the following condition is satisfied: the optimal controls (u_*, v_*) minimize the Hamiltonian

$$H = u + \lambda(A + uB_1 + vB_2)N \tag{3.15}$$

over the control set along $(\lambda(t), N_*(t))$. The same structure of the equations which implies positive invariance of $\mathbb{P}$ for the flow of trajectories also gives negative invariance of the first quadrant in the dual space, $\mathbb{P}^* = \{\lambda \in \left(\mathbb{R}^3\right)^* : \lambda_i > 0 \text{ for } i = 1,2,3\}$, for the adjoint flow (3.11), i.e., if $\lambda(T) \in \mathbb{P}^*$, then $\lambda(t) \in \mathbb{P}^*$ for all times $t \leq T$. Hence, since $N_0 \in \mathbb{P}$ and $r \in \mathbb{P}^*$, we have that

Proposition 3.1 *All states N_i and costates λ_i are positive over $[0,T]$.* $\square$

Optimal controls (u_*, v_*) must satisfy the minimum condition of the Maximum Principle. Since the Hamiltonian is linear in the controls, this is

equivalent to the following two conditions

$$(1 + \lambda(t)B_1N(t))u_*(t) = \min_{0\le u\le 1}(1 + \lambda(t)B_1N(t))u. \tag{3.16}$$

$$\lambda(t)B_2N(t)v_*(t) = \min_{v_{\min}\le v\le 1}\lambda(t)B_2N(t)v \tag{3.17}$$

Thus, if we define the so-called *switching functions* Φ_1 and Φ_2 by

$$\Phi_1(t) = 1 + \langle\lambda(t), B_1N(t)\rangle = 1 - 2a_3\lambda_1N_3 \tag{3.18}$$

and

$$\Phi_2(t) = \langle\lambda(t), B_2N(t)\rangle = a_2(\lambda_3 - \lambda_2)N_2, \tag{3.19}$$

then the optimal controls are given as

$$u_*(t) = \begin{cases} 0 & \text{if } \Phi_1(t) > 0 \\ 1 & \text{if } \Phi_1(t) < 0 \end{cases} \tag{3.20}$$

and

$$v_*(t) = \begin{cases} v_{\min} & \text{if } \Phi_2(t) > 0 \\ 1 & \text{if } \Phi_2(t) < 0 \end{cases}. \tag{3.21}$$

A priori the controls are not determined by the minimum condition at times where $\Phi_i(t) = 0$. However, if $\Phi_i(t)$ vanishes on an open interval, also all its derivatives must vanish and this may determine the control. Controls of this kind are called *singular* while we refer to piecewise constant controls as *bang-bang* controls. Optimal controls then need to be synthesized from these and possibly other more complicated candidates. The structure of the optimal controls is determined by the switching functions and their derivatives. For instance, if $\Phi_i(t) = 0$, but $\dot{\Phi}_i(t) \neq 0$, then the i-th control has a switch at time t. To analyze the structure of the optimal controls we therefore need

to analyze the switching function and its derivatives. The following lemma, which allows to calculate first and higher order derivatives of the switching function simply by calculating commutators of matrices, is verified by an elementary direct calculation.

Lemma 3.1 *Suppose M is a constant matrix and let $\Psi(t) = \lambda(t)MN(t)$, where N is a solution to the system equation (3.7) for control u and λ is a solution to the corresponding adjoint equation. Then*

$$\dot{\Psi}(t) = \lambda(t)[A + uB_1 + vB_2, M]N(t), \tag{3.22}$$

where $[A, M]$ denotes the commutator of the matrices A and M defined as $[A, M] = MA - AM$.

Note that we have chosen the order in the commutator so that it is consistent with the Lie derivative of the linear vector fields $f(N) = AN$ and $g(N) = MN$. For,

$$[f, g] = Dg(N)f(N) - Df(N)g(N) = MAN - AMN = [A, M]N. \tag{3.23}$$

4 Singular Controls

Although singular controls exist for the 3-compartment model, using high-order necessary conditions for optimality we now show that they are not optimal.

Proposition 4.1 *The control v_* cannot be singular on any open interval I.*

Proof: Suppose v is singular on an open interval I, i.e. $\Phi_2(t) \equiv 0$ on I. Since $N_2(t)$ is positive, it then follows that $\Psi_2(t) = \lambda_3(t) - \lambda_2(t)$ vanishes

identically on I. Thus also

$$\dot{\lambda}_2(t) = a_2 v(t)(\lambda_2(t) - \lambda_3(t)) \equiv 0 \tag{4.24}$$

and hence $\lambda_3(t)$ and $\lambda_2(t)$ are actually constant and equal on I. Furthermore

$$0 \equiv \dot{\lambda}_3(t) = a_3[\lambda_3(t) - 2(1 - u(t))\lambda_1(t)] \tag{4.25}$$

and thus

$$\lambda_3(t) = 2(1 - u(t))\lambda_1(t). \tag{4.26}$$

Since λ_3 is positive by Proposition 3.1, we have $u(t) < 1$. If $u \equiv 0$ on any subinterval $J \subset I$, then $\lambda_1(t)$ also must be constant on J and thus

$$0 \equiv \dot{\lambda}_1(t) = a_1[\lambda_1(t) - \lambda_2(t)] \tag{4.27}$$

implying $\lambda_1 = \lambda_2 = \lambda_3 = const$ on J. But this contradicts (4.26). Hence u_* also must be singular on I.

In the case when both u_* and v_* are singular on I, it is a high-order necessary condition for optimality, the so-called *Goh-condition* [14, Cor. 6.3], that

$$\frac{\partial}{\partial v}\frac{d}{dt}\frac{\partial H}{\partial u}(\lambda(t), N(t), u_*(t), v_*(t)) \equiv 0, \tag{4.28}$$

i.e. this quantity vanishes identically along the optimal trajectory. For our system

$$\frac{\partial H}{\partial u}(\lambda(t), N(t), u_*(t), v_*(t)) = \lambda(t) B_1 N(t). \tag{4.29}$$

Using Lemma 3.1 we have

$$\frac{d}{dt}\frac{\partial H}{\partial u}(\lambda(t), N(t), u_*(t), v_*(t)) = \lambda(t)[A, B_1]N(t) + v(t)\lambda(t)[B_2, B_1]N(t) \tag{4.30}$$

and thus

$$\frac{\partial}{\partial v}\frac{d}{dt}\frac{\partial H}{\partial u}(\lambda(t), N(t), u_*(t), v_*(t)) = \lambda(t)[B_2, B_1]N(t). \tag{4.31}$$

A direct calculation gives

$$[B_2, B_1] = -2a_2a_3 \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{4.32}$$

and thus

$$\lambda(t)[B_2, B_1]N(t) = -2a_2a_3\lambda_1(t)N_2(t) < 0 \tag{4.33}$$

violating the Goh condition. Hence v cannot be singular on any interval I. □

Proposition 4.2 *The control u_* cannot be singular on any open interval I.*

Proof: Suppose u_* is singular on an open interval I, i.e. $\Phi_1(t) \equiv 0$ on I. It follows from the above argument that v_* cannot be singular on any subinterval and thus, without loss of generality, we may assume that v_* is constant on I and given by either $v_{\min}$ or 1. Let $f(N) = (A+v_*B_2)N$ and let $g(N) = B_1N$ denote the drift and control vector fields for the corresponding single-input systems with u_* as control. Since Φ_1 vanishes identically on I, all its time derivatives must vanish as well. Using Lemma 3.1 it follows that the first two derivatives are given by

$$\dot{\Phi}_1(t) = \langle \lambda(t), [f + u_*g, g](N)(t) \rangle = \langle \lambda(t), [f, g](N)(t) \rangle , \tag{4.34}$$

$$\ddot{\Phi}_1(t) = \langle \lambda(t), [f + u_*g, [f, g]](N)(t) \rangle . \tag{4.35}$$

In our case

$$[g,[f,g]](N) = [B_1,[A + v_* B_2, B_1]]N \tag{4.36}$$

and a direct calculation shows that both $[B_1,[A,B_1]]$ and $[B_1,[B_2,B_1]]$ are zero. Thus $[g,[f,g]]$ vanishes identically. Hence, and using the notation $adf(g) = [f,g]$ for the Lie bracket, we have

$$\ddot{\Phi}_1(t) = \langle \lambda(t), [f,[f,g]](N)(t) \rangle = \left\langle \lambda(t), ad^2 f(g)(N)(t) \right\rangle . \tag{4.37}$$

Differentiating once more gives

$$\Phi_1^{(3)}(t) = \langle \lambda(t), [f + u_* g, [f,[f,g]]](N)(t) \rangle \tag{4.38}$$

But it follows from the Jacobi identity for the Lie bracket of vector fields that

$$[g, ad^2 f(g)] = -[f,[g,[f,g]]] = -[f,0] = 0 \tag{4.39}$$

and thus

$$\Phi_1^{(3)}(t) = \left\langle \lambda(t), ad^3 f(g)(N)(t) \right\rangle \tag{4.40}$$

Differentiating one more time yields

$$\begin{aligned} \Phi_1^{(4)}(t) &= \left\langle \lambda(t), [f + u_* g, ad^3 f(g)](N) \right\rangle \\ &= \left\langle \lambda(t), ad^4 f(g)(N) \right\rangle + u_* \left\langle \lambda(t), [g, ad^3 f(g)](N) \right\rangle \end{aligned} \tag{4.41}$$

If we set $\bar{A} = A + v_* B_2$, then an explicit calculation verifies that

$$[B_1, ad^3 \bar{A}(B)] = 12 a_1 a_2 a_3^2 v_* \begin{pmatrix} 0 & -a_1 & a_1 - a_3 \\ 0 & 0 & -a_1 \\ 0 & 0 & 0 \end{pmatrix} \tag{4.42}$$

$$= 6 a_1 a_2 a_3 v_* \left\{ [\bar{A}, B_1] + 2 a_3 \begin{pmatrix} 0 & -a_1 - a_2 v_* & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right\}. \tag{4.43}$$

But along the singular we have

$$\lambda(t)[\bar{A}, B_1] N(t) = \dot{\Phi}_1(t) \equiv 0 \tag{4.44}$$

and thus

$$\big\langle \lambda(t), [g, ad^3 f(g)](N)(t) \big\rangle = -12 a_1 a_2 a_3^2 v_* (a_1 + a_2 v_*) \lambda_1(t) N_2(t) < 0.$$

Thus equation (4.41) can be solved to determine the explicit form of the singular control u_*. It thus follows that the singular control u_* is of order 2. However, the generalized Legendre-Clebsch condition [14, Thm 5.2] states that it is a necessary condition for optimality of the singular arc that

$$\big\langle \lambda(t), [g, ad^3 f(g)](N)(t) \big\rangle \geq 0 \tag{4.45}$$

and this condition is violated. (The sign reverses in our application compared with [14, Thm 5.2] since we minimize the Hamiltonian while it is maximized in [14, Thm 5.2].) Hence a singular control u_* cannot be optimal in any interval I. $\square$

These results imply that, once a maximum acceptable level of drug dose is agreed upon (which is the meaning of $u = 1$), then it will not be optimal to ever administer lower doses. Of course, what this maximum level is, depends on the patient's overall health and is a matter for the doctor to determine.

5 Bang-bang Controls

Consequently, and although more complicated structures cannot be excluded, bang-bang trajectories become the prime candidates for optimality. However, apriori it is not possible to restrict the number of switchings. Yet some restrictions on the structure of switchings exist. For, let $\Psi_2 = \lambda_3 - \lambda_2$ be the reduced switching function for the control v. Using the adjoint equation it follows that

$$\dot{\Psi}_2 = a_2 v \Psi_2 + a_3 \left(\lambda_3 - 2(1-u)\lambda_1\right). \tag{5.46}$$

Along $u = 1$ the right-hand side is positive and thus Ψ_2 is strictly increasing. But then we have the following property of optimal controls:

Lemma 5.1 *If $u_*(t) \equiv 1$ on an interval I, then v_* is bang-bang on I with at most one switching from $v = 1$ to $v = v_{\min}$.* $\square$

Recall that $v = 1$ stands for no reduction in the outflow of cells from the second compartment, i.e. the blocking agent is not being used. Thus Lemma 5.1 states that it will never be optimal to withdraw the blocking agent v if both the killing and blocking agents are administered at some given time. The killing agent should be withdrawn first before the blocking agent can be terminated. However, and we will give an example verifying this below, optimal protocols may initiate the blocking agent in addition to an already active killing agent.

We now develop conditions under which a bang-bang control, i.e. a protocol which alternates between full drug doses and restperiods, indeed gives a relative minimum over a set of controls whose trajectories lie in some neighborhood of the reference trajectory in the time-state space. We will do

this by constructing a parametrized family of extremals in the sense defined in [24] by integrating the dynamics and the adjoint equation backward from the terminal time T with the terminal condition $N(T) = p \in \mathbb{P}$ as free parameter. (Once this is being done, however, positivity of the trajectories is no longer guaranteed and needs to be enforced.) For the moment, and denoting the corresponding minimizing controls by $u = u(t,p)$ and $v = v(t,p)$, $0 \leq t \leq T$, the dynamics for the system and adjoint equation are given by

$$\dot{N}(t,p) = (A + uB_1 + vB_2)N(t,p), \tag{5.47}$$

$$\dot{\lambda}(t,p) = -\lambda(t,p)(A + uB_1 + vB_2), \tag{5.48}$$

with terminal values

$$N(T,p) = p \quad \text{and} \quad \lambda(T,p) = r. \tag{5.49}$$

For bang-bang controls these are linear equations with constant matrices between the switching times, hence easily solved. The crucial step is to determine the controls $u(t,p)$ and $v(t,p)$, i.e. the correct switching sequence depending on the terminal value of the trajectory. This will be done through the minimum condition of the Maximum Principle. Then the issue becomes whether the flow map σ of the trajectories defined as

$$\begin{aligned} \sigma :&[0,T] \times \mathbb{P} \to [0,T] \times \mathbb{P}, \\ &(t,p) \mapsto \sigma(t,p) = (t, N(t,p)) \end{aligned} \tag{5.50}$$

defines a field, i.e. whether the trajectories cover the state-space injectively. It is shown in [23, 24] that this will be the case if the restrictions of the flow away from the switching surfaces are diffeomorphisms and if proper transversality conditions which guarantee that the combined flows are locally injective are satisfied at the switching surfaces. Under these conditions

a differentiable solution to the Hamilton-Jacobi-Bellman equation can be constructed and these controls are optimal [2]. In [18] this construction has been carried out in detail for the 2-compartment model. Here we generalize this construction to the 3-compartment model, but we refer to [18] for some of the details of the construction.

Let $(N_*, (u_*, v_*))$ be a reference extremal pair where the controls (u_*, v_*) are bang-bang with switchings at times t_k, $k = 1, \dots, m$, $0 < t_m < \cdots < t_1 < t_0 = T$, and N_* is the corresponding trajectory. Henceforth we always **assume** that

(A1) the controls u_* and v_* do not have simultaneous switchings.

We then associate with (u_*, v_*) a map $\iota : \{1, \dots, m\} \to \{1, 2\}$ which assigns to every switching time t_k the index $\iota(k)$ corresponding to the control which has a switch at time t_k, i.e. $\iota(k) = 1$ if the switch is in the control u and $\iota(k) = 2$ if the switch is in the control v. We also assume that all switchings are strict in the sense that the derivative of the corresponding switching function at t_i does not vanish. Thus, if λ_* denotes the corresponding adjoint variable, then we also assume that

(A2) for $k = 1, \dots, m$, and $\iota = \iota(k)$ we have that

$$\dot{\Phi}_\iota(t_k) = \lambda_*(t_k)[A + u_* B_1 + v_* B_2, B_\iota] N_*(t_k) \neq 0. \tag{5.51}$$

Notice that, since $[B_\iota, B_\iota] = 0$, the derivative of the switching function Φ_1 for the control u involves the control v, but not u itself and the derivative of the switching function Φ_2 for v involves only u. Thus Φ_1 will not be differentiable at switchings in v and Φ_2 will not be differentiable at switchings in u. However, since we assume that there are no simultaneous switchings,

Φ_1 and Φ_2 are differentiable at switchings in u and v respectively . Hence Φ_ι is differentiable at t_k.

Regarding a medical interpretation, assumption (A2) simply enforces that at the indicated times indeed administration of the blocking or killing agent has to be initiated or terminated. Condition (A1) can easily be enforced by giving the drugs in sequence. Mathematically, conditions (A1) and (A2) will be satisfied generically (i.e. except in a few special cases) and they imply the most regular structure for the field around the reference trajectory. Specifically, call a triple $\Gamma = (N_*, (u_*, v_*), \lambda_*)$ for which (A1) and (A2) are satisfied a *strictly bang-bang extremal lift without simultaneous switchings*. Let $p_* = N(T)$ and for p in a neighborhood W of p_* integrate the equations (5.47) and (5.48) backward while chosing the controls $u = u(t,p)$ and $v = v(t,p)$ to maintain the minimum condition of the Maximum Principle. Thus $u(t,p_*)$ and $v(t,p_*)$ are given by the reference controls u_* and v_* and $N(t,p_*)$ and $\lambda(t,p_*)$ are the reference trajectory and corresponding multiplier. The following lemma asserts that a strictly bang-bang extremal lift without simultaneous switchings can be embedded into a family of strictly bang-bang extremal lifts without simultaneous switchings which have switchings in the same order by varying the terminal condition of the reference trajectory.

Lemma 5.2 *Let $\Gamma = (N(\cdot,p_*), (u(\cdot,p_*), v(\cdot,p_*)), \lambda(\cdot,p_*))$ be a strictly bang-bang extremal lift without simultaneous switchings and denote the switching times of the controls by t_k, $k = 1,\dots,m$, $0 < t_m < \cdots < t_1 < t_0 = T$. Then there exists a neighborhood W of p_* and continuously differentiable functions τ_k defined on W, $k = 1,\dots,m$, such that for $p \in W$ the controls $u(\cdot,p)$ and $v(\cdot,p)$ are bang-bang with switchings in the same order as the reference*

control at the times $0 < \tau_m(p) < \cdots < \tau_1(p) < T$ *and the corresponding triples* $\Gamma_p = (N(\cdot,p),(u(\cdot,p),v(\cdot,p)),\lambda(\cdot,p))$ *for* $p \in W$ *are strictly bang-bang extremal lifts without simultaneous switchings.*

The proof of Lemma 5.2 is rather straightforward and we refer the reader to [24, Lemma 5.1] for the details. However, we briefly discuss the significance of assumptions (A1) and (A2): It follows from condition (A2) and the implicit function theorem that for every switching time t_k there exists a neighborhood W_k of p_* on which the equation $\Phi_\iota(t,p) = 0$ can locally be solved for t as a differentiable function of p, $t = \tau_k(p)$, with $t_k = \tau_k(p_*)$. Hence there exist well-defined switching surfaces

$$\mathcal{S}_k = \{(t,N) : t = \tau_k(p), N = N(\tau_k(p),p), p \in W_k\} \tag{5.52}$$

near $(t_k, N_*(t_k))$ in the combined (t,N)-space which are 3-dimensional imbedded submanifolds. Even in the case of simultaneous switchings the existence of these surfaces can easily be guaranteed by making analogous assumptions on the derivatives from the right for the switching functions. Condition (A1) is equivalent to the geometric property that these switching surfaces $\mathcal{S}_k$ do not intersect if a sufficiently small neighborhood W of p_* is chosen. This implies the simplest (and locally the most typical) form of a local synthesis for the field of extremals around the reference bang-bang extremal lift Γ as described in Lemma 5.2. This structure is no longer valid if two switching surfaces $\mathcal{S}_{k_1}$ and $\mathcal{S}_{k_2}$ intersect in $(t_k, N_*(t_k))$. In such a case it matters *how* these surfaces intersect and the geometric properties of their intersection need to be analyzed in detail. These, however, will depend on the specific values of the parameters and states and costates at the switching. Therefore a general analysis of these situations is not feasible, but instead a case-by-case analysis becomes necessary. For a specific trajectory the local structure

of the field of extremals around a reference bang-bang extremal lift Γ which has simultaneous switchings in the controls can still be determined under suitable assumptions on higher order derivatives of the switching functions involved. Then the construction of a regular synthesis [1] can be pursued. Our results, making assumption (A1), present the most typical behavior which can be analyzed in general. While the theorems and optimality conditions formulated in our paper are for this scenario only and do not apply to the case of simultaneous switchings in the control, the construction itself does. But it then requires modifications of the proofs and the calculations to fit the specific cases.

While the states and costates remain continuous at the switching surfaces, their partial derivatives become discontinuous because of the different controls. For later use we need to update the p-partials at a switching surface. Let $t = \tau(p)$ parametrize a switching surface. Without loss of generality assume the control u switches and denote the constant controls for $t > \tau(p)$ and $t < \tau(p)$ by u_+ and u_-, respectively. Also let $\Delta u = u_+ - u_-$ denote the jump in the control. For $p \in W$, let $N^+(t,p)$ and $N^-(t,p)$ denote the solutions to the system equation with controls u_+ and u_-, respectively, and initial condition $N^\pm(t,p) = N(t,p)$ for $t = \tau(p)$. The control v has the value of the reference control v_* which is constant in a sufficiently small interval around $t_* = \tau(p_*)$. Hence the trajectories of the system are given by $N^+(t,p)$ for $t \geq \tau(p)$ and by $N^-(t,p)$ for $t \leq \tau(p)$. In particular, N^+ and N^- are differentiable functions which agree on the set $S = \{(t,p) : t = \tau(p)\}$. It is an elementary argument (for example, see [24, Lemma 2.3]) to see that for every component $i = 1, 2, 3,$ of N there exists a continuous real-valued

function $\kappa_i(t,p)$ defined near S such that for $(t,p) \in S$ we have

$$grad N_i^-(t,p) = grad N_i^+(t,p) + \kappa_i(t,p)(1, -\nabla\tau(p)), \tag{5.53}$$

where $grad$ denotes the gradient in both t and p and $\nabla\tau$ denotes the gradient of the function τ of p, both written as row vectors. Thus we have

$$\dot{N}_i^-(t,p) = \dot{N}_i^+(t,p) + \kappa_i(t,p) \tag{5.54}$$

and for $j = 1,2,3$,

$$\frac{\partial N_i^-}{\partial p_j}(t,p) = \frac{\partial N_i^+}{\partial p_j}(t,p) - \kappa_i(t,p)\frac{\partial \tau}{\partial p_j}(p). \tag{5.55}$$

Using the differential equation for $\dot{N}$, and reverting back to vector notation writing $\kappa = (\kappa_1, \kappa_2, \kappa_3)^T$, it follows that

$$\begin{aligned} \kappa(t,p) &= (A + u_- B_1 + vB_2)\, N(t,p) - (A + u_+ B_1 + vB_2)\, N(t,p) \\ &= -\Delta u B_1 N(t,p), \quad (t,p) \in S. \end{aligned} \tag{5.56}$$

Hence we get for the matrices of the partial derivatives with respect to p that

$$\frac{\partial N^-}{\partial p}(\tau(p),p) = \frac{\partial N^+}{\partial p}(\tau(p),p) + \Delta u B_1 N(\tau(p),p)\nabla\tau(p). \tag{5.57}$$

An analogous formula holds for the multipliers $\lambda^{\pm}(t,p)$ defined correspondingly. Note that $\frac{\partial N^-}{\partial p}(\tau(p),p)$ is a rank 1 correction of $\frac{\partial N^+}{\partial p}(\tau(p),p)$. For our construction we need that the matrix $\frac{\partial N}{\partial p}(t,p)$ remains invertible. The partial derivative $\frac{\partial N}{\partial p}$ is the solution of the variational equation for (5.47). Since this equation is linear and since the controls are constant, the variational equation is identical and we have for $t \le \tau(p)$ that

$$\frac{\partial N}{\partial p}(t,p) = \exp\left((A + u_- B_1 + vB_2)(t - \tau(p))\right) \cdot \frac{\partial N^-}{\partial p}(\tau(p),p). \tag{5.58}$$

Hence $\frac{\partial N}{\partial p}(t,p)$ will be invertible if $\frac{\partial N^-}{\partial p}(\tau(p),p)$ is. Assuming inductively that $\frac{\partial N^+}{\partial p}(\tau(p),p)$ is nonsingular, it is a well-known result from Linear Algebra [8] that $\frac{\partial N^-}{\partial p}(\tau(p),p)$ is nonsingular if and only if

$$1+\nabla\tau(p)\left(\frac{\partial N^+}{\partial p}(\tau(p),p)\right)^{-1}\Delta uB_1N(\tau(p),p)\neq 0. \tag{5.59}$$

In fact, the inverse is given by

$$\left(\frac{\partial N^-}{\partial p}(\tau(p),p)\right)^{-1}=\left(\frac{\partial N^+}{\partial p}(\tau(p),p)\right)^{-1} \cdot\left(Id-\frac{\Delta uB_1N(\tau(p),p)\nabla\tau(p)\left(\frac{\partial N^+}{\partial p}(\tau(p),p)\right)^{-1}}{1+\nabla\tau(p)\left(\frac{\partial N^+}{\partial p}(\tau(p),p)\right)^{-1}\Delta uB_1N(\tau(p),p)}\right). \tag{5.60}$$

The following result has been proven for bang-bang trajectories for general systems in [24]:

Theorem 5.1 *[24] Let $\Gamma=(N_*,(u_*,v_*),\lambda_*)$ be a strictly bang-bang extremal lift without simultaneous switchings and denote the switching times of the controls by t_k, $k=1,\dots,m$, $0<t_m<\cdots<t_1<t_0=T$. Let $p_*=N_*(T)$ and let $\tau_k=\tau_k(p)$, $k=1,\dots,m$, $0<\tau_m(p)<\cdots<\tau_1(p)<T$ be parametrizations of the switching times defined on some neighborhood of p_*. For the k-th switching let $\iota(k)$ be the indicator of the control that switches and denote the jump in the control by*

$$\Delta_\iota=\begin{cases} u(t_k+)-u(t_k-) & \text{if } \iota(k)=1\\ v(t_k+)-v(t_k-) & \text{if } \iota(k)=2\end{cases}.$$

If for every switching $k=1,\dots,m$, we have that

$$1+\nabla\tau_k(p_*)\left(\frac{\partial N^+}{\partial p}(t_k,p_*)\right)^{-1}\Delta_\iota B_\iota N_*(t_k)>0, \tag{5.61}$$

then (u_*, v_*) is a relative minimum. More precisely, there exists a neighborhood W of p_* such that the flow σ:$[0,T] \times W, (t,p) \mapsto (t, N(t,p))$, defines a field of extremals and (u_*, v_*) is optimal relative to any other control whose corresponding trajectory lies in the image $R = \sigma([0,T] \times W)$.

Idea of the Proof: Since $\frac{\partial N}{\partial p}(T,p) = Id$, the identity matrix, the matrix $\frac{\partial N}{\partial p}(t,p)$ is invertible for $\tau_1(p) \leq t \leq T$. It follows inductively from (5.59) that the matrix $\frac{\partial N}{\partial p}(t,p)$ is nonsingular on all intervals $[\tau_{i+1}(p), \tau_i(p)]$, $i = 1, \ldots, m-1$, and on $[0, \tau_m(p)]$. The specific sign in the transversality condition (5.61) guarantees that the combined flow transversally crosses the switching surfaces. This allows to combine the separate bang-bang flows into one injective flow. Furthermore, using (5.61) it can be shown that the corresponding value-function remains continuously differentiable at the switching surface. Optimality of (u_*, v_*) relative to R then follows from standard Hamilton-Jacobi theory. The details of the argument can be found in [24]. $\square$

For this result to be of practical use we need an efficient way to calculate the quantities

$$\nabla \tau_k(p_*) \left(\frac{\partial N^+}{\partial p}(t_k, p_*) \right)^{-1}. \tag{5.62}$$

Such a procedure exists and it requires neither the computation of partial derivatives with respect to p nor matrix inversion. Indeed, since only one of the controls switches at every switching surface, with a minor modification for the overall dynamics, these calculations are identical to those for the 2-compartment model [18] and we only summarize the resulting formulas.

Proposition 5.1 *Let ι denote the index of the control which switches at $t_k = \tau_k(p_*)$. Then for p in a sufficiently small neighborhood W of p_* we*

have

$$\nabla\tau_k(p)\left(\frac{\partial N^+}{\partial p}(\tau_k(p),p)\right)^{-1} = -\frac{1}{\dot\Phi_\iota(\tau_k(p),p)}\Big(\lambda(\tau_k(p),p)B_\iota + N^T(\tau_k(p),p)B_\iota^T S^+(\tau_k(p),p)\Big) \tag{5.63}$$

where $S^+(\tau_k(p),p)$ *(respectively* $S^-(\tau_k(p),p)$*) denotes the limit of the function*

$$S(t,p) = \frac{\partial\lambda^T}{\partial p}(t,p)\left(\frac{\partial N}{\partial p}(t,p)\right)^{-1}. \tag{5.64}$$

from the right (respectively left) at $t=\tau_k(p)$*. The matrix* S *can be calculated inductively on* $[\tau_k(p),\tau_{k-1}(p)]$ *as the solution to the Lyapunov equation*

$$\begin{aligned}\dot S(t,p) + S(t,p)(A+u(t,p)B_1+v(t,p)B_2) \\ + (A+u(t,p)B_1+v(t,p)B_2)^T S(t,p) \equiv 0\end{aligned} \tag{5.65}$$

with terminal conditon $S^-(\tau_{k-1}(p),p)$*, i.e.*

$$\begin{aligned}S(t,p) = \exp\left((A+u(t,p)B_1+v(t,p)B_2)^T(\tau_{k-1}(p)-t)\right) \\ \cdot S^-(\tau_{k-1}(p),p)\cdot\exp\left((A+u(t,p)B_1+v(t,p)B_2)(\tau_{k-1}(p)-t)\right).\end{aligned} \tag{5.66}$$

In particular, no partial derivatives of the state or costate with respect to the parameter p need to be calculated nor do we need to calculate the inverse. Everything is subsumed in the calculation of the matrix $S(t,p)$. However, since S is defined in terms of the p-partials, it will no longer be continuous at switching times and we need to use the correct update formulas. As above, let $S^+(t,p)$ and $S^-(t,p)$ denote the matrices S when constructed with $N^\pm$ and $\lambda^\pm$ respectively. Then using (5.60) the following formulas can be verified:

Proposition 5.2 *Let ι denote the index of the control which switches at $t_k = \tau_k(p_*)$ and define*

$$G_k(p) = \frac{\Delta_\iota}{\dot{\Phi}_\iota(\tau_k(p), p)} \left(\lambda(\tau_k(p), p)B_\iota + N^T(\tau_k(p), p)B_\iota^T S^+(\tau_k(p), p)\right). \tag{5.67}$$

Then we have for p in a sufficiently small neighborhood W of p_* that

$$\begin{aligned} S^-(\tau_k(p), p) = &\left(B_\iota^T \lambda^T(\tau_k(p), p)G_k(p) + S^+(\tau_k(p), p)\right) \\ &\cdot \left(Id + \frac{B_\iota N(\tau_k(p), p)G_k(p)}{1 - G_k(p)B_\iota N(\tau_k(p), p)}\right). \end{aligned} \tag{5.68}$$

The formulas developed so far are valid in general for strictly bang-bang extremal lifts without simultaneous switchings for arbitrary bilinear systems. A number of simplifications occur for the 3-compartment model for cancer chemotherapy based on the special strucure of the matrices and we summarize them in the next Theorem:

Theorem 5.2 *Let $\Gamma = (N_*, (u_*, v_*), \lambda_*)$ be a strictly bang-bang extremal lift without simultaneous switchings and let $\Phi_1^*(t) = 1 + \lambda_*(t)B_1N_*(t)$ and $\Phi_2^*(t) = \lambda_*(t)B_2N_*(t)$ be the switching functions associated with the controls u_* and v_* respectively. Denote the switching times of the controls by t_k, $k = 1, \ldots, m$, $0 < t_m < \cdots < t_1 < t_0 = T$ and let (u_k, v_k) denote the constant values of the controls on the interval (t_k, t_{k-1}). For the k-th switching let $\iota = \iota(k)$ be the indicator of the control that switches and denote the absolute jump in the control by*

$$\theta_\iota = \begin{cases} 1 & \text{if} \quad \iota(k) = 1 \\ 1 - v_{\min} & \text{if} \quad \iota(k) = 2 \end{cases}.$$

Set $S_0^- = 0$ *and for* $k = 1, \dots, m$, *define*

$$S_k^+ = \exp\left((A + u_k B_1 + v_k B_2)^T (t_{k-1} - t_k)\right) S_{k-1}^- \exp\left((A + u_k B_1 + v_k B_2)(t_{k-1} - t_k)\right), \tag{5.69}$$

$$G_k = -\frac{\theta_\iota}{\left|\dot{\Phi}_\iota^*(t_k)\right|}\left(\lambda_*(t_k) B_\iota + N_*^T(t_k) B_\iota^T S_k^+\right), \tag{5.70}$$

$$S_k^- = \left(B_\iota^T \lambda_*^T(t_k) G_k + S_k^+\right)\left(Id + \frac{B_\iota N_*(t_k) G_k}{1 - G_k B_\iota N_*(t_k)}\right). \tag{5.71}$$

If for $k = 2, \dots, m$, *we have that*

$$\left|\dot{\Phi}_\iota^*(t_k)\right| + \theta_\iota N_*^T(t_k) B_\iota^T S_k^+ B_\iota N_*(t_k) > 0, \tag{5.72}$$

then all the matrices S_k^-, $k = 2, \dots, m$, *are well-defined and* (u_*, v_*) *is a relative minimum for the 3-compartment model. More precisely, there exists a neighborhood* W *of* $N_*(T)$ *such that the flow* σ *restricted to* $[0, T] \times W$ *defines a field of strictly bang-bang extremals without simultaneous switchings and* (u_*, v_*) *is optimal relative to any other control whose trajectory lies in the image* $R = \sigma([0, T] \times W)$.

Proof: The theorem follows from Theorem 5.1 using special properties of the model. Set $p_* = N_*(T)$ and let $\tau_k = \tau_k(p)$, $k = 1, \dots, m$, $0 < \tau_m(p) < \cdots < \tau_1(p) < T$ be parametrizations of the switching times defined on some neighborhood of p_*. In the terminolgy of Theorem 5.1 we then have $S^-(T, p_*) = S_0^-$ and for $k = 1, \dots, m$, $S_k^\pm = S^\pm(\tau_k(p_*), p_*)$. We need to verify that condition (5.61) is satisfied at every switching, i.e., for $k = 1, \dots, m$, we have

$$1 + \nabla\tau_k(p_*)\left(\frac{\partial N^+}{\partial p}(t_k, p_*)\right)^{-1} \Delta_\iota B_\iota N_*(t_k)$$
$$= 1 - \frac{\Delta_\iota}{\dot{\Phi}_\iota^*(t_k)}\left(\lambda_*(t_k) B_\iota + N_*^T(t_k) B_\iota^T S_k^+\right) B_\iota N_*(t_k) > 0. \tag{5.73}$$

By the minimum condition of the Maximum Principle we always have that

$$\Delta_\iota = -sgn\left(\dot{\Phi}^*_\iota(t_k)\right)\theta_\iota \tag{5.74}$$

(for example, if $u_k = 1$ and $u_{k+1} = 0$, then the switching function Φ_1 decreases at t_k) and therefore

$$-\frac{\Delta_\iota}{\dot{\Phi}^*_\iota(t_k)} = \frac{\theta_\iota}{\left|\dot{\Phi}^*_\iota(t_k)\right|}.$$

Furthermore, for our model $B_1^2 = 0$ and therefore these terms drop out trivially. Also the corresponding terms for B_2 drop out, but for a different reason. A direct calculation verifies that $B_2^2 = -a_2 B_2$ and therefore we have at any switching in the control v that

$$\lambda_*(t_k)B_2 \cdot B_2 N_*(t_k) = -a_2\lambda_*(t_k)B_2 N_*(t_k) = -a_2\Phi^*_2(t_k) = 0 \tag{5.75}$$

since t_k is a switching time. Summarizing, we therefore have in general that

$$1 + \nabla\tau_k(p_*)\left(\frac{\partial N^+}{\partial p}(t_k, p_*)\right)^{-1}\Delta_\iota B_\iota N_*(t_k) =$$
$$1 + \frac{\theta_\iota}{\left|\dot{\Phi}^*_\iota(t_k)\right|}N_*^T(t_k)B_\iota^T S_k^+ B_\iota N_*(t_k)$$

and thus conditions (5.61) and (5.72) are equivalent. In particular, these conditions inductively guarantee that the matrices S_k^-, $k = 1, \ldots, m$, are well-defined. Finally, at the terminal time we have $N(T,p) = p$ and $\lambda(T,p) = r$ giving $S^-(T,p) \equiv 0$. In particular, it follows that $S^+(\tau_1(p),p) \equiv 0$ and therefore the first transversality (5.72) will always be satisfied for $k = 1$. □

Corollary 5.1 *Every strictly bang-bang control with at most one switching is a relative minimum.* □

Remark: The conditions (5.72) ensure that the switching surfaces are crossed in the same direction by the two flows involved in the switching. They form a true requirement which in general need not be satisfied. As a case in point notice that (since $S_1^+ = 0$)

$$S_1^- = -\frac{\theta_\iota B_\iota^T \lambda_*^T(t_1)\lambda_*(t_1)B_\iota}{\left|\dot{\Phi}_\iota^*(t_1)\right|} \tag{5.76}$$

is negative semi-definite (and so is then S_2^+) so that the second term in (5.72) is non-positive for $k = 2$. In general, due to the continued addition of rank 1 matrices at the switching times, the matrices S_k^+ will have no definiteness properties.

If the transversality condition (5.72) is violated with a negative value, then indeed local optimality of the bang-bang trajectory ceases at this switching.

Theorem 5.3 *With the notation of Theorem 5.2 assume that the transversality condition*

$$\left|\dot{\Phi}_\iota^*(t_k)\right| + \theta_\iota N_*^T(t_k) B_\iota^T S_k^+ B_\iota N_*(t_k) > 0 \tag{5.77}$$

is satisfied for $k = 2, \ldots, \ell - 1$, but that

$$\left|\dot{\Phi}_\iota^*(t_\ell)\right| + \theta_\iota N_*^T(t_\ell) B_\iota^T S_\ell^+ B_\iota N_*(t_\ell) < 0. \tag{5.78}$$

Then there exists a neighborhood W of $p_* = N_*(T)$ such that the flow σ restricted to $D_\ell = \{(t, p : t_\ell < t \leq T, p \in W\}$ defines a field of strictly bang-bang extremals without simultaneous switchings and (u_*, v_*) is optimal relative to any other control whose trajectory lies in the image $R_\ell = \sigma(D_\ell)$, but (u_*, v_*) is no longer optimal for initial times $t \leq t_\ell$.

Essentially, this result is proven by verifying that the ℓ-th switching surface is an envelope in the sense of Sussmann [27] consisting of conjugate points where local optimality ceases. Since we are assuming that only one of the controls switches at t_ℓ this proof is identical to the proof of [18, Thm 5.3].

Figures 5.1 and 5.3 below show the graphs of two sample switching functions for our problem. The parameters chosen for the calculation were $a_1 = 0.197$, $a_2 = 0.395$, $a_3 = .107$, $r_1 = 1$, $r_2 = .5$, $r_3 = 1$, and the time horizon is $T = 7$. The terminal values for the states, respectively the parameter values are $p_1 = p_2 = 10$ and $p_3 = 17$. The control u (see Fig. 5.2) switches once at $t_2 = 1.5279$, the zero of the switching function $\Phi_1(t,p)$ for u (Fig. 5.1) while v (see Fig. 5.4) has two switchings at $t_3 = .7439$ and $t_1 = 3.5583$, the zeroes of its switching function $\Phi_2(t,p)$ (Fig. 5.3). Notice that since the derivative of Φ_1 involves v, but not u, Φ_1 is differentiable at $t_2 \in (1,2)$, but its graph has corners at $t_3 \in (0,1)$ and $t_1 \in (3,4)$ where Φ_1 is not differentiable. Analogously, since the derivative of Φ_2 involves u, but not v, Φ_2 is differentiable at the switchings t_1 and t_3 in v, but its graph has a corner at the time t_2 of the switching in u. Notice also how the discontinuities in the controls affect the graphs of the components $N_i(t,p)$, $i = 1,2,3$, of the state (Figs. 5.5-5.7). Since the control u only enters the equation of $\dot{N}_1$ directly, the discontinuity in u causes a corner in the graph of $N_1(t,p)$ at $t_2 \in (1,2)$ (Fig. 5.5), but is not noticable in the graphs of $N_2(t,p)$ and $N_3(t,p)$. Similarly, the discontinuities in v show as corners in the graphs of $N_2(t,p)$ (Fig. 5.6) and $N_3(t,p)$ (Fig. 5.7) since v enters the equations for $\dot{N}_2$ and $\dot{N}_3$, but are not visible in the graph of N_1. Using the algorithm given in Theorem 5.2, it can be verified that the transversality condition is satisfied at each switching and therefore the pair (u,v) is locally optimal.

Table 1. Data for the switchings

switching time	switch in control	transversality condition
$t_1 = 3.5583$	v	.3081
$t_2 = 1.5279$	u	.5171
$t_3 = 0.7439$	v	.1977

Figs. 5.5 to 5.7 give the graphs of the trajectories verifying in particular that the states are positive. Note that this protocol allows the numbers N_2 and N_3 of cancer cells in the second and third compartment to rise while it drastically reduces the number N_1 of cancer cells in the first compartment. Overall, the weighted average $rN(t)$ decreases from 42.5 at time $t = 0$ to 32 at time $T = 7$. From a medical point of view the shift of higher number of cancer cells from the first to the third compartment makes perfect sense since this is where the killing agent is active, i.e. more cells are exposed to the drug, while smaller numbers of cancer cells in the first compartment imply that less cells will move through the cell-cycle in the future.

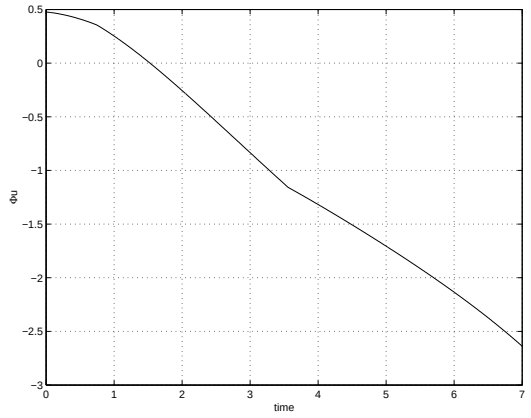

Figure 5.1: Switching function $\Phi_1(t, p)$ for the control $u(t, p)$

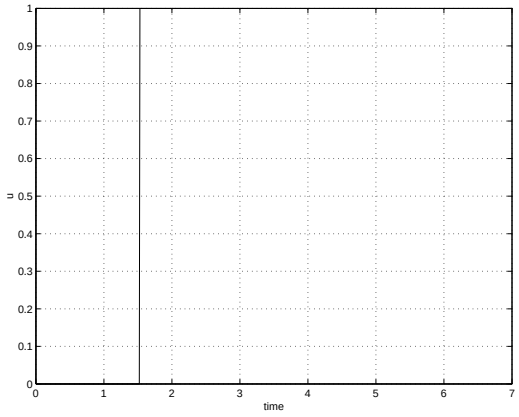

Figure 5.2: Control $u(t,p)$

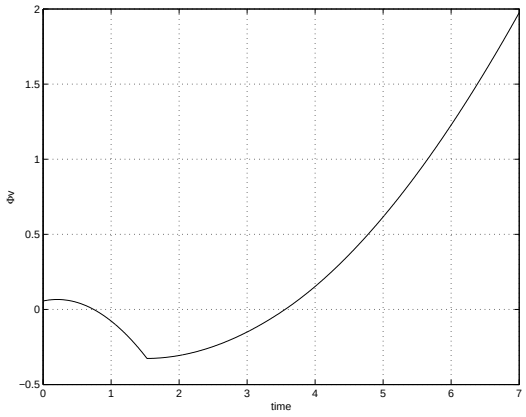

Figure 5.3: Switching function $\Phi_2(t,p)$ for the control $v(t,p)$

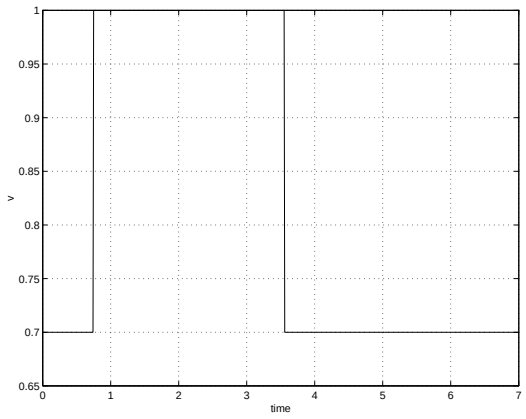

Figure 5.4: Control $v(t,p)$

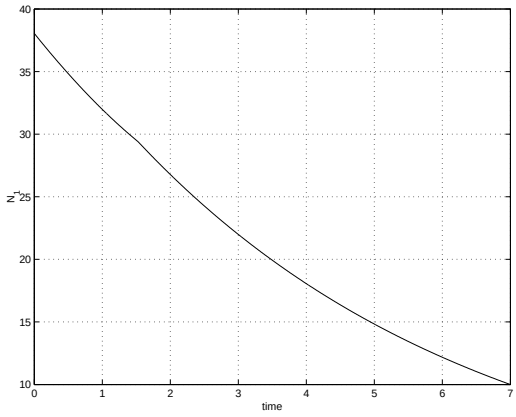

Figure 5.5: First component of the state, $N_1(t,p)$

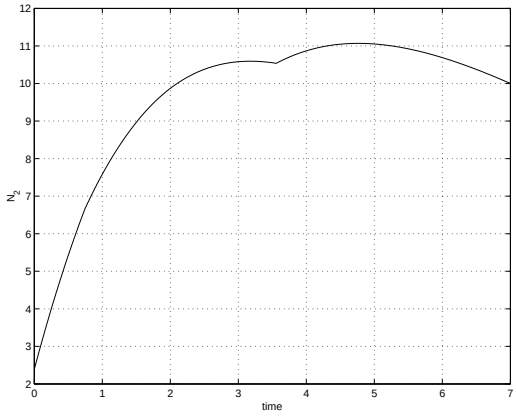

Figure 5.6: Second component of the state, $N_2(t, p)$

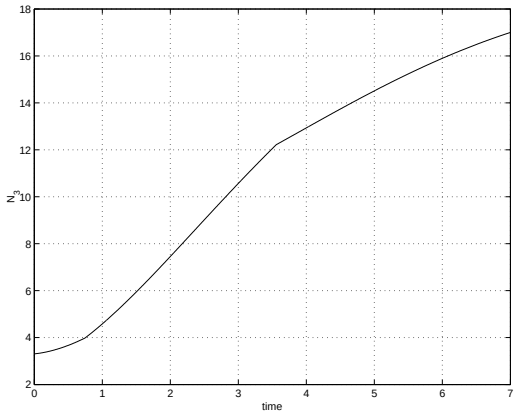

Figure 5.7: Third component of the state, $N_3(t, p)$

6 Conclusion

In the paper we pursued a further analysis, beyond the standardly used Maximum Principle, of optimal controls of the 3-compartment model for cancer chemotherapy by [32]. After eliminating singular controls from the candidates for optimality, a more detailed analysis of bang-bang controls and corresponding trajectories was performed. Sufficient conditions for optimality in the form of conditions for transversal crossings of the flow at switching surfaces were derived. In particular, it follows from our results as a simple corollary that bang-bang controls with at most one switching are locally optimal (assuming that the derivative of the switching function does not vanish at the switching). It was also shown that a violation of the transversal crossings condition causes loss of optimality. The arguments used in this paper are general and apply to arbitrary systems with bang-bang controls. Some numerical calculations were given to support and illustrate the applicability of the theory.

Interpreting the results for the model, it follows that it is not optimal to administer a drug, neither the killing nor the blocking agent, at a level less than their maximum dose allowable. Optimal protocols consist of periods of fully administering one or both of the drugs with restperiods in between as, indeed, it is the common practice in medicine. However, although some of these protocols may satisfy the necessary conditions for optimality, they still need not be optimal. Our results in Theorems 5.2 and 5.3 provide easily verifiable conditions to determine whether a given protocol is at least locally optimal and to eliminate non-optimal protocols.

7 Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant DMS-9971747. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. Urszula Ledzewicz was also supported in part by 2001 SIUE Summer Research Fellowship and SIUE Research Scholar Award.

References

[1] V.G. Boltyanskii, Sufficient conditions for optimality and the justification of the dynamic programming method, *SIAM J. on Control*, **4**, (1966), pp. 326-361

[2] A.E. Bryson and Y.C. Ho, *Applied Optimal Control*, Hemisphere Publishing, 1975

[3] M.I.S. Costa, J.L. Boldrini and R.C. Bassanezi, Drug kinetics and drug resistance in optimal chemotherapy, *Mathematical Biosciences*, **125**, (1995), pp. 191-209

[4] Z. Duda, A gradient method for application of chemotherapy models, *J. of Biological Systems*, **3**, (1995), pp. 3-11

[5] Z. Duda, Numerical solutions to bilinear models arising in cancer chemotherapy, *Nonlinear World*, **4**, (1997), pp. 53-72

[6] M. Eisen, *Mathematical Models in Cell Biology and Cancer Chemotherapy*, Lecture Notes in Biomathematics, Vol. 30, Springer Verlag, (1979)

[7] K.R. Fister and J.C. Panetta, Optimal control applied to cell-cycle-specific cancer chemotherapy, *SIAM J. Appl. Math.*, **60**, (2000), pp. 1059-1072

[8] T. Kailath, *Linear Systems*, Prentice Hall, 1980

[9] M. Kim, K.B. Woo and S. Perry, Quantitative approach to the design of antitumour drug dosage schedule via cell cycle kinetics and systems theory, *Ann. Biomed. Eng.*, **5**, (1977), pp. 12-33.

[10] M. Kimmel and A. Swierniak, An optimal control problem related to leukemia chemotherapy, *Scientific Bulletins of the Silesian Technical University*, **65**, (1983), pp. 120-130

[11] D. Kirschner, Using mathematics to understand HIV immune dynamics, *Notices of the American Mathematical Society*, **43**, (1996), pp. 191-202

[12] D. Kirschner, S. Lenhart, and S. Serbin, Optimal control of chemotherapy of HIV, *J. Math. Biol.*, **35**, (1997), pp. 775-792

[13] D.E. Kirschner and G.F. Webb, A mathematical model of combined drug therapy of HIV infection, *J. of Theoretical Medicine*, **1**, (1997), pp. 25-34

[14] A. Krener, The high-order maximal principle and its application to singular controls, *SIAM J. Control and Optimization*, **15**, (1977), pp. 256-293

[15] U. Ledzewicz and H. Schättler, Optimal control for a mathematical model of cancer chemotherapy, in: Control and Applications, Proceed-

ings of the IASTED International Conference, (M.H. Hamza, Ed.), IASTED/ACTA Press, Anaheim, (2000), pp. 8-14

[16] U. Ledzewicz and H. Schättler, On a synthesis of controls for a mathematical model of cancer chemotherapy, Proceedings of the 39th IEEE Conference on Decision and Control, Sydney, Australia, (2000), pp. 4845-4850

[17] U. Ledzewicz and H. Schättler, On optimal controls for a 3-compartment model of cancer chemotherapy, in: Control and Applications, Proceedings of the IASTED International Conference, (M.H. Hamza, Ed.), IASTED/ACTA Press, Anaheim, (2001), pp. 109-114

[18] U. Ledzewicz and H. Schättler, Optimal bang-bang controls for a 2-compartment model in cancer chemotherapy, *Journal of Optimization Theory and Applications - JOTA*, **114**, No. 3, (2002), to appear

[19] A. Matveev and A. Savkin, Optimal control applied to drug administration in cancer chemotherapy: the case of several toxicity constraints, Proceedings of the 39th IEEE Conference on Decision and Control, Sydney, Australia, (2000), pp. 4833-4838

[20] R.B. Martin, Optimal control drug scheduling of cancer chemotherapy, *Automatica*, **28**, (1992), pp. 1113-1123

[21] R.R. Mohler, *Bilinear Control Systems*, Academic Press, (1973)

[22] J.M. Murray, Some optimal control problems in cancer chemotherapy with a toxicity limit, *Mathematical Biosciences*, **100**, (1990), pp. 49-67

[23] J. Noble and H. Schättler, On the optimality of parametrized families of bang-bang trajectories, *Nonlinear Analysis*, **47**, (2001), pp. 351-362

[24] J. Noble and H. Schättler, Sufficient conditions for relative minima of broken extremals in optimal control theory, *J. of Mathematical Analysis and Applications*, accepted for publication, to appear

[25] L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelidze and E.F. Mishchenko, *The Mathematical Theory of Optimal Processes*, MacMillan, New York, (1964)

[26] E.O. Roxin, Control Theory and its Applications, Gordon and Breach Science Publ., (1997)

[27] H. Sussmann, Envelopes, high-order optimality conditions and Lie brackets, Proceedings of the 28th IEEE Conference on Decision and Control, Tampa, Florida, 1989, pp. 1107-1112

[28] G.W. Swan, General applications of optimal control theory in cancer chemotherapy, *IMA J. Math. Appl. Med. Biol.*, **5**, (1988), pp. 303-316

[29] G.W. Swan, Role of optimal control in cancer chemotherapy, *Math. Biosci.*, **101**, (1990), pp. 237-284

[30] A. Swierniak, Some control problems for simplest differential models of proliferation cycle, *Appl. Math. and Comp. Sci.*, **4**, (1994), pp. 223-232

[31] A. Swierniak, Cell cycle as an object of control, *Journal of Biological Systems*, **3**, (1995), pp. 41-54.

[32] A. Swierniak, A. Polanski and M. Kimmel, Optimal control problems arising in cell-cycle-specific cancer chemotherapy, *Cell prolif.*, **29**, (1996), pp. 117-139

[33] A. Swierniak and Z. Duda, Bilinear models of cancer chemotherapy-singularity of optimal solutions, in: Mathematical Population Dynamics, **2**, (1995), pp. 347-358

[34] A. Swierniak, A. Polanski and Z. Duda, "Strange" phenomena in simulation of optimal control problems arising in cancer chemotherapy, Procceedings of the 8th Prague Symposium on Computer Simulation in Biology, Ecology and Medicine, (1992), pp. 58-62

[35] I. Tannock, Cell kinetics and chemotherapy: a critical review, *Cancer Treat. Rep.*, **62**, (1978)

[36] J. Velasco-Hernández, J. Garcia and D. Kirschner, Remarks on modeling host-viral dynamics and treatment, in: Vol. 125, Mathematical approaches for emerging and reemerging infectious diseases, Part I: An introduction to models, Methods and Theory, (Eds. Carlos Castillo-Chavez with S. Blower, P. van den Driessche, D. Kirschner and A. Yakubu), (2001), to appear

[37] T.E. Wheldon, *Mathematical models in cancer research*, Bristol: Medical Sci. Series, Hilger, (1988)

List of Figures.

Fig. 5.1: Switching function $\Phi_1(t,p)$ for the control $u(t,p)$

Fig. 5.2: Control $u(t,p)$

Fig. 5.3: Switching function $\Phi_2(t,p)$ for the control $v(t,p)$

Fig. 5.4: Control $v(t,p)$

Fig. 5.5: First component of the state, $N_1(t,p)$

Fig. 5.6: Second component of the state, $N_2(t,p)$

Fig. 5.7: Third component of the state, $N_3(t,p)$