Optimal and Suboptimal Protocols for a Class of Mathematical Models of Tumor Anti-Angiogenesis

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

Heinz Schättler (Corresponding author)
Dept. of Electrical and Systems Engineering, Campus Box 1127
Washington University,
One Brookings Drive,
St. Louis, Missouri, 63130-4899,
(314) 935-6019; fax: (314) 935-7500
hms@wustl.edu

February 2, 2008

Abstract

Tumor anti-angiogenesis is a cancer treatment approach that aims at preventing the primary tumor from developing its own vascular network needed for further growth. In this paper the problem of how to schedule an a priori given amount of angiogenic inhibitors in order to minimize the tumor volume is considered for three related mathematical formulations of a biologically validated model developed by Hahnfeldt, Panigrahy, Folkman and Hlatky [Cancer Research, 1999]. Easily implementable piecewise constant protocols are compared with the mathematically optimal solutions. It is shown that a constant dosage protocol with rate given by the averaged optimal control is an excellent suboptimal protocol for the original model that achieves tumor values that lie within 1% of the theoretically optimal values. It is also observed that the averaged optimal dose is decreasing as a function of the initial tumor volume.

Keywords: Cancer therapy, anti-angiogenic inhibitor, drug dosage, dynamical system, optimal control

1 Introduction

Tumor anti-angiogenesis is a cancer treatment approach targeted at the vasculature of a growing tumor. Its biological foundation was first introduced by J. Folkman (1971, 1972): A primary solid tumor, after going through a state of avascular growth, at the size of about 2mm in diameter, starts the process of *angiogenesis* to recruit surrounding, mature, host blood vessels in order to develop its own blood vessel capillaries needed for supply of nutrients. The lining of these newly developing blood vessels consist of endothelial cells and the tumor produces vascular endothelial growth factor (VEGF) to stimulate their growth (Klagsburn and Soker, 1993) as well as inhibitors to suppress it (Folkman, 1995). Overall this process is based on a bi-directional signaling that can be viewed as a complex balance of tightly regulated stimulatory and inhibitory mechanisms (Folkman and Klagsburn, 1987; Davis and Yancopoulos, 1999). Anti-angiogenic treatments rely on these mechanisms by bringing in external angiogenic inhibitors (e.g., endostatin) targeting the endothelial cells and thus blocking their growth. This indirectly effects the tumor which, ideally, deprived of necessary nutrition, would regress. Since contrary to traditional chemotherapy this treatment targets normal, not cancer cells, it was observed that no resistance to the angiogenic inhibitors developed in experimental cancer (Boehm et al., 1997). For this reason tumor anti-angiogenesis has been called a therapy "resistant to resistance" that provides a new hope for the treatment of tumor type cancers (Kerbel, 1997) and as such in the last ten years became an active area of research not only in medicine (e.g., Davis and Yancopoulos, 1999; Hahnfeldt et al., 1999; see also, Kerbel, 2000), but also in other disciplines including mathematical biology, (e.g., Anderson and Chaplain, 1998; Sachs et al., 2001; Ergun et al., 2003; Hahnfeldt et al., 2003; d'Onofrio and Gandolfi, 2004; Agur et al., 2004; Forys et al., 2005).

Specifically in mathematical modelling several models describing the dynamics of angiogenesis have been proposed. Some of these aim at fully reflecting the complexity of the biological processes, (e.g., Anderson and Chaplain, 1998; Arakelyan et al., 2003), and allow for large scale simulations, but are less amenable to a mathematical analysis. Most theoretical techniques from such fields as dynamical systems or optimal control theory can only effectively be used in low dimensional systems. Hahnfeldt, Panigrahy, Folkman and Hlatky, (1999), a group of researchers then at Harvard Medical School, developed and biologically validated a two-dimensional model of ordinary differential equations for the interactions between the tumor volume, p, and the carrying capacity of the endothelial cells, q. The latter is defined as the maximum tumor volume sustainable by the vascular network. Henceforth we also refer to this as the endothelial support of the tumor for short. Based on this model and the underlying spatial analysis carried out in that research two main modifications of the original model have been formulated since then, one

by d'Onofrio and Gandolfi (2004) at the European Institute of Oncology in Milan, the other by Ergun, Camphausen and Wein (2003) at the Cancer Research Institute at NIH. In each formulation a Gompertzian model with variable carrying capacity q is chosen for tumor growth, but the equations for the endothelial support differ in their inhibition and stimulation terms, $I(p,q)$ and $S(p,q)$. The dynamics of these systems as well as the important problem of how to schedule an a priori given amount of angiogenic inhibitors in such a way as to realize the maximum tumor reduction possible was analyzed in several papers (Ergun et al., 2003; d'Onofrio and Gandolfi, 2004; Ledzewicz and Schättler, 2005, 2006a,b, 2007a,b; Swierniak et al., 2006a,b). Here we draw on this research to investigate easily implementable protocols.

For each of these three models we gave a complete mathematical analysis of the structure of optimal controls (Ledzewicz and Schättler, 2005, 2006b, 2007a,b). While optimal controls only contain one interval where all available inhibitors are given at maximum dose for the model in (d'Onofrio and Gandolfi, 2004), mathematically these are called *bang-bang* controls, the other two models have optimal solutions that contain so-called *singular* controls. These are specific, time-varying, state dependent, feedback controls at less than maximum dose that are not medically or biologically realizable. Thus the question about suboptimal, but realizable protocols arises. In this paper we compare the effectiveness of simple protocols that give all available inhibitors at constant dosages from the beginning of the therapy with those of the optimal protocols. We consider two ad hoc strategies of giving full dose and half dose and a third protocol when the actual dosage is computed as the average dose of the optimal control, that is

$$\bar{u} \equiv \frac{1}{T_{\text{opt}}} \int_0^{T_{\text{opt}}} u_{\text{opt}}(t)dt = \frac{A}{T_{\text{opt}}},$$

where u_{opt} denotes the optimal control as a function of time, T_{opt} is the time when all inhibitors are exhausted along this optimal control, and A denotes the a priori specified overall amount of inhibitors to be given. Optimal controls naturally depend on the initial conditions for the problem, that is the initial tumor volume p_0 and its endothelial support q_0, and thus also the averaged optimal dose $\bar{u}$ becomes a function of these initial data. The constant ad hoc dosage protocols at maximum or half the maximum dose have the advantage of not requiring precise information, but these ad hoc strategies tend to be inferior to the average of the optimal controls. This strategy indeed *defines an excellent suboptimal protocol* for the model considered in (Hahnfeldt et al., 1999) that generally comes within 1% of the theoretically optimal value for the model. Naturally, the type of suboptimal protocol that is best depends on the model under consideration, but the reasons for the superiority of specific simple strategies can be understood from the properties of the dynamics and knowledge of the optimal solution.

Applications of optimal control to mathematical models arising in biomedical problems have

a long history going back to Eisen's monograph (Eisen, 1979) and some of the classical papers by Swan (Swan, 1988, 1990). The early focus was on models in connection with cancer chemotherapy and these efforts have continued to the present day (e.g., Swierniak, 1995; Fister and Panetta, 2000; Ledzewicz and Schättler, 2002a,b). But methods of optimal control have also been applied in new directions of research to models of HIV-infection (e.g., Kirschner, Lenhart, and Serbin, 1997; Butler, Kirschner and Lenhart, 1997) and novel treatment approaches to cancer. These include the research on the models for tumor anti-angiogenesis described in this paper, but also a strong and increasing recent interest in applications to models for immunotherapy (e.g., de Pillis and Radunskaya, 2001; Castiglione and Piccoli, 2006). While there exist some mathematical models in which optimization methods are utilized in the context of combination therapies such as the combination of radiotherapy and anti-angiogenic treatment considered by Ergun, Camphausen and Wein (2003), these approaches are very recent and mathematical models are only in the stage of development. In most of the literature that applies optimal control methods to biomedical problems it is common to only analyze the first order necessary conditions and then proceed with numerical simulations. We want to emphasize that contrary to this approach our research presented here is based on a fully worked out theoretical solution of the problem considered and numerical computations are only used to illustrate these results.

The paper is organized as follows: In section 2 we review the biological/medical background and formulate the three mathematical models that will be considered. Section 3 then gives a full description of the *optimal solutions* for arbitrary initial conditions, a so-called *regular synthesis*, for each of these models. We refer to our previous work (Ledzewicz and Schättler, 2005, 2006b, 2007a,b) for the derivation and proofs of these results, but include a brief presentation of the concept of a regular synthesis and its sufficiency for optimality in an appendix. The optimal protocols resulting from these theoretical solutions provide a benchmark with which the other more realistic strategies can be compared. Then in section 4 an extensive study of constant dosage protocols, easily implementable, realistic strategies, is done. It is shown that, depending on the initial conditions (tumor volume and size of endothelial support at the beginning of therapy) different types of constant dosage suboptimal protocols provide the best approximation of the optimal controls. An interesting observation is that a lower constant dosage generally does better for high initial tumor volumes. The reason for this lies in the strongly differential-algebraic character of the overall dynamics and the fact that for a high initial tumor volume the so-called slow manifolds for constant dosage systems better approximate the optimal singular arc if the dosage is lower. We briefly comment on multi-period protocols in the conclusion.

2 Mathematical Models for the Dynamics of Tumor Anti-Angiogenesis

The underlying mathematical model on which all the models considered here are based was developed and biologically validated by Hahnfeldt, Panigrahy, Folkman and Hlatky (1999) and, as already mentioned, has the primary tumor volume, p, and the carrying capacity of the vasculature, q, as its principal variables. Tumor growth is described as Gompertzian with a variable carrying capacity represented by q. Consequently the rate of change in the volume of primary tumor cells is given by

$$\dot{p} = -\xi p \ln\left(\frac{p}{q}\right) \tag{1}$$

where ξ denotes a tumor growth parameter. Other growth models, like, for example, logistic growth considered by d'Onofrio and Gandolfi, (2004), or general growth functions considered by Forys et al., (2005), are equally realistic, but lead to different computations and the corresponding optimal control problems would need to be analyzed separately. We thus retain the original modelling. The models we consider here differ in the equations for the dynamics of the endothelial support. Overall this dynamics is a balance between stimulatory and inhibitory effects and its basic structure can be written in the form

$$\dot{q} = -\mu q + S(p,q) - I(p,q) - Guq \tag{2}$$

where I and S denote endogenous inhibition and stimulation terms and the terms μq and Guq that have been separated from the general terms describe, respectively, loss to the endothelial cells through natural causes (death etc.), and loss of endothelial cells due to additional outside inhibition. Generally μ is small and often this term is negligible compared to the other factors and thus in the literature sometimes μ is set to 0 in this equation. The variable u represents the control in the system and corresponds to the concentration of the inhibitors while G is a constant that represents the anti-angiogenic killing parameter. In the model by Hahnfeldt et al. (Hahnfeldt, 1999), making "the usual pharmacokinetic assumptions," the concentration is still linked with the angiogenic dose rate v by a first order linear ODE,

$$\dot{u} = -\gamma u + v, \tag{3}$$

where γ is the clearing rate for the inhibitors. In the model analyzed here this relation has been dropped or, in other words, we only consider a zero-order pharmacokinetic model that identifies concentration with dosage. While clearly a simplification, this only neglects the transient behavior in (3) and, as has been argued by us in connection with mathematical models for chemotherapy (Ledzewicz and Schättler, 2005b) a linear PK-model does not change the

qualitative structure of solutions and also quantitatively the changes are minor, especially if the dynamics (3) is fast. Here we thus identify the concentration of inhibitors with the angiogenic dose rate administered.

The problem of how to administer a given amount of inhibitors to achieve the "best possible" effect arises naturally. One possible formulation, considered first in (Ergun et al., 2003) and then taken up by us in (Ledzewicz and Schättler, 2005a, 2006b, 2007a,b), is to solve the following optimal control problem: for a free terminal time T, minimize the value $p(T)$ subject to the dynamics (1) and (2) over all piecewise continuous (more generally, Lebesgue measurable) functions $u : [0, T] \to [0, a]$ that satisfy a constraint on the total amount of anti-angiogenic inhibitors to be administered,

$$\int_0^T u(t)dt \leq A. \tag{4}$$

The upper limit a in the definition of the control set $U = [0, a]$ is a previously determined maximum dose at which inhibitors can be given. Note that the time T does not correspond to a therapy period in this formulation. Instead the solution to this problem gives the maximum tumor reduction achievable with an overall amount A of inhibitors available and T is the time when this minimum tumor volume is being realized. Mathematically it is more convenient to adjoin the constraint as third variable and define the problem in $\mathbb{R}^3$. Hence we consider the following optimal control problem:

[OC] For a free terminal time T, minimize the value $p(T)$ subject to the dynamics

$$\dot{p} = -\xi p \ln\left(\frac{p}{q}\right), \qquad p(0) = p_0, \tag{5}$$

$$\dot{q} = -\mu q + S(p, q) - I(p, q) - Guq, \qquad q(0) = q_0, \tag{6}$$

$$\dot{y} = u, \qquad y(0) = 0, \tag{7}$$

over all piecewise continuous functions $u : [0, T] \to [0, a]$ for which the corresponding trajectory satisfies $y(T) \leq A$.

While some of the analysis can be done for the general model [OC], the stimulation and inhibition terms need to be specified further in order to obtain more than just qualitative statements. In this paper we consider three specifications of this model that all are based on the paper by Hahnfeldt et al. (1999). In that paper a spatial analysis of the underlying consumption-diffusion model was carried out that led to the following two principal conclusions:

1. *The inhibitor will impact endothelial cells in a way that grows like volume of cancer cells to the power $\frac{2}{3}$. (The exponent $\frac{2}{3}$ arises through the interplay of the surface of the tumor through which the inhibitor needs to be released with the volume of endothelial cells.)*

Thus in (Hahnfeldt et al., 1999) the inhibitor term is taken in the form

$$I(p,q) = dp^{\frac{2}{3}}q \tag{8}$$

with d a constant, mnemonically labelled the "death" rate. The second implication of the analysis in (Hahnfeldt et al., 1999) is that:

2. *The inhibitor term will tend to grow at a rate of* $q^\alpha p^\beta$ *faster than the stimulator term with* $\alpha + \beta = \frac{2}{3}$.

However, the choice of α and β in this analysis allows some freedom and in fact is the main source of variability for the other models considered in the literature. Hahnfeldt et al. (1999) select $\alpha = 1$ and $\beta = -\frac{1}{3}$ resulting in the simple stimulation term

$$S(p,q) = bp \tag{9}$$

with b a constant, the "birth" rate. However, other choices are possible and, for example, taking $\alpha = 0$ and $\beta = \frac{2}{3}$ results in the equally simple form

$$S(p,q) = bq \tag{10}$$

chosen by d'Onofrio and Gandolfi (2004) and resulting in a much simpler q-dynamics for the model. It is a feature of either system (see below) that the q-dynamics is much faster than the p-dynamics and the systems exhibit a behavior characteristic of differential-algebraic models. In fact, it is argued by Ergun, Camphausen and Wein (2003) that the systems tend to reach their steady state too fast. Since p and q tend to move together in steady state, ideally $p = q$, there is some freedom in selecting the terms for inhibition and stimulation, and Ergun et al. (2003) modify the $\dot{q}$ equation to

$$\dot{q} = -\mu q + bq^{\frac{2}{3}} - dq^{\frac{4}{3}} - Guq. \tag{11}$$

Their justification for this change or approximation lies in a different balance in the dynamics for the substitution of stimulation and inhibition, but compared with the underlying model of Hahnfeldt et al. (1999) the inhibitor term in this model is now only proportional to tumor radius and thus the premises of this model are not fully consistent with the implications of the analysis in (Hahnfeldt et al., 1999). As another justification for the choice $bq^{\frac{2}{3}}$ for the stimulation term, it could be argued that the necrotic core of the tumor does not interact with endothelial cells and thus the power 2/3 could also be interpreted as scaling down the interactions from the tumor volume p to the surface area $p^{\frac{2}{3}}$ of the tumor and then interchanging p and q for the steady-state analysis, as it is done in (Ergun et al., 2003). The mathematical advantage

of this approach is that the dynamics becomes a tremendous simplification in the sense that it eliminates a direct link between tumor cells p and endothelial cells q. As we shall see in section 3, with all these simplifications, this model nevertheless retains the essential features of the problem and its optimal solution is qualitatively identical with the one for the original model, while the modification by d'Onofrio and Gandolfi (2004) that is fully consistent with the modelling implications of (Hahnfeldt et al., 1999) leads to a qualitatively different structure. We summarize the differences between the three models in Table 1 and henceforth refer to these models as models (A), (B) and (C).

Before analyzing the optimal controls for problem [OC] for these models, we first describe the dynamics for the two standard situations: (a) the uncontrolled system, that is with $u \equiv 0$, and (b) the system under constant, maximum dosage, that is for the control $u \equiv a$. Figs. 1(a)-3(a) give simulations of the phase portraits of these systems with $u \equiv 0$ and Figs. 1(b)-3(b) show the dynamics with $u \equiv a$. In all our figures we plot p vertically and q horizontally since this easier visualizes tumor reductions. For the simulations in this paper we use the following parameter values that are taken from (Hahnfeldt et al., 1999): The variables p and q are volumes measured in mm^3; $\xi = \frac{0.192}{\ln 10} = 0.084$ *per day* (adjusted to the natural logarithm), $b = 5.85$ *per day*, $d = 0.00873$ *per* mm^2 *per day*, $G = 0.15$ *kg per mg of dose per day* with concentration in *mg of dose per kg*, and for illustrative purposes we chose a small positive value for μ, $\mu = 0.02$ *per day*. Since we identify dosage and concentration, both a and A are in units of concentration. Since the dynamics of the systems considered are different, and this does lead to different quantitative results, to allow for these variations we picked $a = 75$ and $A = 300$ for models (A) and (B), but reduced these values to $a = 15$ and $A = 45$ for model (C). These specific values chosen are purely for illustrative purposes and are not based on biological data. The values of all the parameters used in our numerical calculations are summarized in Table 2.

It is shown by d'Onofrio and Gandolfi (2004) that the uncontrolled models (A) and (B) have the same and unique equilibrium point given by $\bar{p} = \bar{q} = \left(\frac{b-\mu}{d}\right)^{\frac{3}{2}}$, equal to $17,258$ mm^3 for our parameter values. This equilibrium is globally asymptotically stable, a node for model (A) (Fig. 1(a)), a focus for model (B), (Fig. 2(a)). The uncontrolled model (C) also has a globally asymptotically stable node, but now given by $\bar{p} = \bar{q} = \left(\frac{\sqrt{\mu^2+4bd}-\mu}{2d}\right)^3$ and equal to $15,191$ mm^3 in our simulation (Fig. 3(a)), (Ledzewicz and Schättler, 2005). Note that these equilibria coincide for $\mu = 0$. These values are too high to be acceptable and it is the aim to lower these set points through anti-angiogenic therapy. The biologically most relevant region for all these models is therefore contained in the square

$$\mathcal{D} = \{(p,q) : 0 < p \leq \bar{p},\ 0 < q \leq \bar{q}\} \tag{12}$$

and in order to exclude discussions about the structure of optimal controls in regions where the models do not represent the underlying biological problem, we thus restrict our discussions to this domain $\mathcal{D}$. Adding the control term Guq and making the reasonable assumption that

$$Ga > b - \mu > 0, \tag{13}$$

i.e., the parameters related to outside inhibition are able to overcome the net effect of "birth" minus "death", one can show that this globally asymptotically stable node is preserved for model (C), (Fig. 3(b)), but can be shifted to a much lower value, 17 mm^3 for our parameter values. For models (A) and (B), however, the equilibrium ceases to exist and all trajectories converge to the origin in infinite time, (Figs. 1(b) and 2(b), respectively). This, in principle, would be the desired situation since, at least theoretically, it allows eradication of the tumor using a constant dose $u = a$ for all time. But clearly this is not a feasible strategy due to limits on the total amount of inhibitors and potential side effects. This leads to the formulation and analysis of optimal control problems like problem [OC] which we pursue in this paper. We also do assume in our descriptions here that the overall amount A of inhibitors is large enough to make a reduction of the initial tumor volume p_0 possible since otherwise the mathematically optimal solution comes out to be $T = 0$. This merely excludes some unrealistic initial conditions for which q_0 is much larger than p_0 (see, (Ledzewicz and Schättler, 2007a)) and it streamlines the presentation of the results.

The following statement about the dynamical behavior of all three models guarantees the existence and positivity of solutions for all times and arbitrary controls. Clearly, if the state variables would not remain positive, the model would not make sense. For models (A) and (B) this is an easy corollary of the results proven in (d'Onofrio and Gandolfi, 2004) and for model (C) it was verified in (Ledzewicz and Schättler, 2005).

Proposition 2.1 *For each of the models (A), (B) or (C) and any admissible control u and arbitrary positive initial conditions p_0 and q_0 the corresponding solution (p, q) exists for all times $t \geq 0$ and both p and q remain positive.* □

3 Optimal Protocols for the Models

We now present the structure of optimal protocols for the three models (A), (B) and (C). These protocols will be given in mathematical form as a *regular synthesis* of optimal controls (e.g., Boltyansky, 1966; Fleming and Rishel, 1975; Bressan and Piccoli, 2007). Such a synthesis provides a full "road map" for how optimal protocols look like for all possble initial conditions in the problem, both qualitatively and quantitatively, and establishes the optimality of all the

controls and trajectories in the synthesis. (A brief exposition of this theory is given in the Appendix.) The results presented in this paper have been proven mathematically in (Ledzewicz and Schättler, 2007a) for model (A), in (Ledzewicz and Schättler, 2007b) for model (B) and an outline of the proofs for model (C) is given in (Ledzewicz and Schättler, 2005). It follows from these results that optimal controls have a common qualitative behavior for models (A) and (C), but follow a different pattern for model (B). Some general and intuitively clear properties are shared for all models:

- Optimal trajectories satisfy $y(T) = A$, i.e., all available inhibitors are exhausted by the end of the therapy.
- Optimal trajectories end with $p(T) = q(T)$. This is a property generally valid for any realistic growth model for the tumor, not just the Gompertzian model considered here, and is a consequence of using the carrying capacity of the vasculature as a modelling variable.

We first summarize the results of our earlier research in Theorem 3.1 below and then proceed to a more precise description of the optimal controls.

Theorem 3.1 *For both models (A) and (C), given any initial condition $(p_0, q_0) \in \mathcal{D}$, optimal controls are at most concatenations of the form* **0asa0** *where* **0** *denotes an interval along which the optimal control is given by the constant control $u = 0$, that is, no inhibitors are given,* **a** *denotes an interval along which the optimal control is given by the constant control $u = a$ at full dose, and* **s** *denotes an interval along which the optimal control follows a time-varying feedback control (that will be specified below), the so-called singular control. This control is only optimal while the system follows a particular curve $\mathcal{S}$ in the (p, q)-space, the optimal singular arc. However, depending on the initial condition (p_0, q_0), not all of these intervals need to be present in a specific solution. For the biologically most relevant initial conditions typically optimal controls have the form* **bs0** *with* **b** *denoting an interval along which the optimal control is given by either* **a** *or* **0** *depending on the initial condition. For model (B) singular controls do not exist and optimal controls are bang-bang of at most the form* **0a0**.

Despite their name, which has its origin in classical literature on optimal control (e.g., (Bryson and Ho, 1975; Krener, 1977)), singular controls and the corresponding singular curves are to be expected in a synthesis of optimal controls for a problem of the type [OC] for nonlinear models (Bonnard and Chyba, 2003). If they exist - they do for model (A) and (C), but not for model (B) - they typically will be either locally maximizing or minimizing for the objective and higher order conditions for optimality, like the Legendre-Clebsch conditions (Bryson and Ho,

1975; Krener, 1977), allow to determine their optimality status. In fact, for the problem [OC] optimal singular trajectories can only lie on *one* specific curve in (p,q)-space, the singular curve $\mathcal{S}$. If singular controls are locally minimizing, as it is the case here, then this curve becomes the center piece to the optimal synthesis. Thus for models (A) and (C) singular controls and the geometry of the singular curve $\mathcal{S}$ are an essential part of the design of the optimal protocols and in order to construct a full synthesis of solutions, the formulas for singular controls and corresponding singular trajectories need to be determined. The proposition below presents these formulas for model (A). All mathematical arguments leading to the full derivation of these formulas can be found in (Ledzewicz and Schättler, 2007a).

Proposition 3.1 (Ledzewicz and Schättler, 2007a) *For model (A) the singular curve $\mathcal{S}$ entirely lies in the sector $\{(p,q) : x_1^* q < p < x_2^* q\}$ where x_1^* and x_2^* are the unique zeroes of the equation*

$$\varphi(x) = \frac{b}{d}x(\ln x - 1) + \frac{\mu}{d} = 0 \tag{14}$$

and satisfy $0 \leq x_1^* < 1 < x_2^* \leq e$. In new variables (p,x) with $x = \frac{p}{q}$ the singular curve $\mathcal{S}$ can be parameterized in the form

$$p^2 + \varphi(x)^3 = 0 \qquad \text{for} \quad x_1^* < x < x_2^*. \tag{15}$$

The singular control keeps the system on the singular curve and is given as a feedback function of x in the form

$$u_{\sin}(x) = \frac{1}{G}\left[\left(\frac{1}{3}\xi + bx\right)\ln x + \frac{2}{3}\xi\left(1 - \frac{\mu}{bx}\right)\right]. \tag{16}$$

There exists exactly one connected arc on the singular curve $\mathcal{S}$ along which the singular control is admissible, i.e., satisfies the bounds $0 \leq u_{\sin}(x) \leq a$. This arc is defined over an interval $[x_\ell^*, x_u^*]$ where x_ℓ^* and x_u^* are the unique solutions to the equations $u_{\sin}(x_\ell^*) = 0$ and $u_{\sin}(x_u^*) = a$ and these values satisfy $x_1^* < x_\ell^* < 1 < x_u^* < x_2^*$.

The two graphs given in Fig. 4 illustrate the proposition for the parameter values from (Hahnfeldt et al., 1999) specified earlier. Fig. 4(a) shows the plot for the singular control defined by (16) also indicating the values x_ℓ^* and x_u^* where the control saturates at $u_{\sin}(x) = 0$ and $u_{\sin}(x) = a$. Fig. 4(b) shows the graph of the singular curve given by formula (15). Saturation of the singular control at x_ℓ^* and x_u^* restricts the admissible part of this petal-like curve to the portion lying between the lines $p = x_l^* q$ and $p = x_u^* q$. This portion is marked with a solid line in Fig. 4(b). The qualitative structures shown in Fig. 4 are generally valid for arbitrary parameter values both for the control and the singular curve. With decreasing values for the upper control limit a the admissible portion shrinks, but it is always preserved.

For model (C) the analysis of singular controls and corresponding singular arcs was pursued in (Ledzewicz and Schättler, 2005).

Proposition 3.2 (Ledzewicz and Schättler, 2005) *For model (C) the singular curve $\mathcal{S}$ is defined in (p,q)-space by*

$$p = p(q) = q\exp\left(3\frac{b-\mu q^{\frac{1}{3}}-dq^{\frac{2}{3}}}{b+dq^{\frac{2}{3}}}\right) \tag{17}$$

and the corresponding singular control is given in feedback form as

$$u_{\sin}(q) = \frac{1}{G}\left(\frac{b-dq^{\frac{2}{3}}}{q^{\frac{1}{3}}} + 3\xi\frac{b+dq^{\frac{2}{3}}}{b-dq^{\frac{2}{3}}} - \mu\right). \tag{18}$$

This control is admissible over an interval $q_\ell^* \le q \le q_u^*$ where the values q_ℓ^* and q_u^* are the unique solutions to the equation $u_{\sin}(q) = a$ in $(0, \sqrt{\frac{b}{d}})$. ■

The two graphs in Fig. 5 illustrate this singular control and singular curve. Again we use the numerical values given earlier, but for this model we take $a = 15$ as upper limit on the controls. The graph shown in Fig. 5(a) represents the singular control given by (18). For the numerical values used here the optimal control saturates at $q_\ell^* = 23.69$ and $q_u^* = 12,319$. The lower value is irrelevant for all practical purpose and thus the corresponding singular curve given by (17) shown in Fig. 5(b) has its admissible part (the solid portion of the graph) starting almost at the origin. The function $p = p(q)$ that defines the singular arc is strictly increasing over the interval where the singular control is admissible. One difference to model (A) is that here the singular curve will become inadmissible if the upper limit a on the allowable dosage lies below the minimum of the function defined by (18). In (Ledzewicz and Schättler, 2005) our analysis was done for $\mu = 0$ while here we use $\mu = 0.02$ to allow for comparison of the syntheses of the two models (A) and (C). The equations in Proposition 3.2 reflect this change. Furthermore, in (Ledzewicz and Schättler, 2005) these and also the figures for the synthesis below were presented in the variables (p,x) where $q = x^3$ which simplified the analysis, but also generated distortions along the horizontal axis.

The admissible singular arcs for models (A) and (C) become an essential part of the synthesis of solutions for both models which are depicted in Figs. 6 and 7. The top of each figure gives a representation of the synthesis as a whole and the bottom gives an example of one particular optimal trajectory and its corresponding optimal control. The important curves for the synthesis are the admissible portions of the singular curve (solid blue curve), portions of trajectories corresponding to the constant controls $u = 0$ (dash-dotted green curves) and $u = a$ (solid green curves), and the line $p = q$ (dotted black line) where the trajectories achieve the maximum tumor reduction. These diagrams represent the optimal trajectories as a whole and each of the different curves gives a different optimal trajectory depending on the actual initial condition. The thick lines in the graphs mark one specific such trajectory. In each case the initial value p_0

for the tumor volume and q_0 for the endothelial support are high and require to immediately start with the treatment. The optimal trajectory therefore initially follows the curve corresponding to the control $u = a$. Note that, although inhibitors are given at full dose along this curve, this shows very little effect on the number of the cancer cells in a sense of decrease. Once the trajectory corresponding to the full dose hits the singular arc $\mathcal{S}$, according to our analysis it is no longer optimal to give full dose and the optimal controls here switch to the singular control and the optimal trajectory follows the singular arc. Ignoring some special cases that are due to saturation of the singular control along this arc and are described in (Ledzewicz and Schättler, 2007a), the optimal control will now follow the singular arc until all inhibitors are exhausted according to the condition that $y(T) = A$. It is clear from the top graphs in Figs. 6 and 7 that this is the part where most of the shrinkage of the tumor occurs. When the inhibitors have been exhausted, therapy is over and the optimal trajectory now follows a trajectory for the control $u = 0$. The reason for this lies in the fact that due to aftereffects in the dynamics the minimum tumor volume is only realized along this trajectory when it crosses the diagonal $p = q$. The corresponding time T then is the limit of the horizon considered in the problem formulation [OC].

We close this section with giving one particular example of an optimal trajectory and its corresponding control for each of models (A) (Figs. 6 (bottom)) and (C) (Figs. 7 (bottom)). The initial conditions for each run are chosen as $(p_0, q_0) = (12,000\ mm^3; 15,000\ mm^3)$. For model (A) the dynamics along the constant controls $u = 0$ and $u = a$ is very fast resulting in short initial and terminal pieces along which the control is constant while the bulk of time is spent along the singular arc. In both cases the optimal singular control administers the inhibitors first at lower levels and then the dosage intensifies along the singular arc, an observation already made by Ergun et al., (2003), for model (C). But for this model the times along the constant controls are significantly longer. Protocols of this type correspond to the most typical structure $as0$, but depending on the initial condition other scenarios are possible. Clearly, if the initial condition already would lie on the admissible part of the singular arc, the portion with $u = a$ will not be present. On the other hand, if the overall amount of inhibitors is so large that the lower saturation point on the singular arc would be reached with inhibitors remaining, then it is actually not optimal to wait until this point is reached, but optimal trajectories need to leave the singular arc earlier and switch to the control $u = a$ until all inhibitors are being used up. The behavior of the optimal solutions around the saturation point on the singular arc is actually a complex mathematical problem and more details can be found in (Ledzewicz and Schättler, 2007a).

Note that, although there are clear quantitative differences in the solutions to both models shown in Figs. 6 (bottom) and 7 (bottom), qualitatively the optimal syntheses for models (A) and (C) are identical (Figs. 6 (top) and 7 (top)). On the other hand, for model (B), as stated in Theorem 3.1, no singular arcs exist (Swierniak et al., 2006b; Ledzewicz and Schättler, 2007b) and the synthesis has the simple form $0a0$. This makes the protocols resulting from model (B) easy to implement. The optimal protocols for models (A) and (C) on the other hand are not realistically implementable, but they define the benchmarks for realistic and implementable, but suboptimal strategies.

4 Suboptimal Protocols

Singular controls are time varying feedback controls that depend on the current values of the state of the system - tumor volume and endothelial support - and are only optimal along one specific curve. Clearly these are not implementable strategies since the required information generally is not available in continuous-time. Even at the initial time a reliable estimate of these values may not be known. It is therefore of interest to formulate simple, easily implementable, but also robust strategies that could be employed with great uncertainty in the state of the system. The significance of the optimal solution lies in providing a theoretical benchmark to which these practical schemes can be compared to assess their efficiency. In this section we explore several protocols that *apply all available inhibitors in one session at a constant dosage at the beginning of therapy* and compare their effectiveness with the theoretically optimal protocols. As it was done implicitly in the formulation of the optimization model [OC] described earlier, we do assume here that the overall amount A corresponds to what is considered an acceptable amount for one treatment period. We shall comment on multi-period treatments briefly in the conclusion.

A simple ad hoc strategy is to give all available inhibitors at maximum rate from the beginning of therapy and we call this the *full dose protocol*. Again, because of after effects the minimum tumor volume that is generated by this strategy is not the value of the system when the inhibitors are being exhausted, but it is realized when the subsequent trajectory of the system corresponding to $u = 0$ crosses the diagonal. This applies to any strategy considered. Thus mathematically this strategy corresponds to a simple bang-bang control of the type **a0**. In fact, for both models (A) and (C), if the initial tumor volume p_0 is close to the p-value for the lower saturation point on the singular arc, then these $a0$-strategies are indeed the optimal solutions for all realistic values of endothelial support q_0. Naturally thus the **a0**-protocol is an excellent sub-optimal strategy for small tumor volumes that is almost as good as the optimal solution.

However, as our simulations show, for initial conditions with higher initial values p_0 protocols that use a reduced dose do better. We therefore also consider for comparison *half dose protocols* that give the full amount of inhibitors at half the maximum dose for twice the time. It will be seen that this is another very good ad-hoc strategy that generally does well for high initial values of tumor volume p_0. A third protocol that we consider is based on the optimal control and gives as dose the averaged dose of the optimal control, that is

$$\bar{u} = \frac{1}{T}\int_0^T u(t)dt = \frac{A}{T} \tag{19}$$

where T is the time when all inhibitors are exhausted along the optimal control. The interval when the tumor volume still decreases due to after effects is not included in this computation. We call this the *averaged optimal dose protocol*. Clearly this protocol depends on the initial condition while the ad hoc full and half dose protocols do not. We now compare these three suboptimal strategies to the optimal one for models (A) and (C).

Suboptimal protocols for model (A). For a general constant dose protocol $u \equiv v$ (see Fig. 1 for the phase portraits for $u = 0$ and $u = a$), the dynamics has a strong differential-algebraic character: it is fast in q and slow in p. Essentially the systems follows an almost "horizontal" line until the algebraic constraint manifold determined by the $\dot{q} = 0$ nullcline,

$$q = \frac{bp}{\mu + Gv + dp^{\frac{2}{3}}}, \tag{20}$$

is reached and then the dynamics evolves along this curve. In fact, this nullcline very much plays the same role for constant dose protocols as the singular curve does for the optimal solutions. Fig. 8 shows both the singular curve (in red) and the $\dot{q} = 0$ nullclines for the full dose $u = a$ (in blue) and half dose $u = \frac{a}{2}$ (in green). The a-nullcline intersects the singular curve in the saturation point and it is clear that the two curves are almost identical around and below this saturation point. Hence in this area the full dose protocols come very close to the optimal protocols. But for higher values of p and q these curves separate and now the singular curve is better approximated by the $\frac{a}{2}$-nullcline. Hence the half-dose protocol does better there.

Fig. 9 (left column) gives three graphs of the minimum tumor volume realized by the various protocols for a fixed initial condition p_0 and varying initial conditions q_0, i.e., slices through the graphs of the associated value functions for $p_0 = const.$ In all diagrams the red curve gives the optimal values, the blue curve corresponds to giving the maximum dosage a and the green curve to giving half dosage $\frac{a}{2}$ while the black curve gives the values for the averaged optimal dose protocol.

For small tumor volumes ($p_0 = 6000\ mm^3$), there is no discernable difference between the optimal, averaged and full dose protocols. The saturation value of the singular control lies at

$p_{sat} = 4,122\ mm^3$ and in this range, as the value of the singular control is close to $a = 75$, the optimal, averaged and full dose protocols give almost identical values. The half dose protocol does noticeable worse for these values. In fact, it is so far off that we did not include this curve in the range for Fig. 9 (top left). Naturally, the realizable minimum values increase with growing initial endothelial support q_0. But as the initial tumor volume p_0 increases, the full dose protocol starts to perform worse while the half dose protocol improves. For $p_0 = 12,000\ mm^3$ the full dose protocol does considerably worse while the half dose protocol starts to perform better. But in all cases *the averaged optimal dose protocol does best* among these three constant infusion protocols. Actually, for all these values the averaged optimal dose protocol stayed within 0.5% of the optimal value. But even for the full dose protocols the differences to the optimal value barely exceed 2% for high initial values of q_0 if the initial condition (p_0, q_0) lies to the right of the singular curve (the blue curve in Fig. 8) since the optimal control starts with $u = a$ in this region. If, however, the initial condition lies to the left of the singular curve, then the discrepancies become larger, exceeding the 5% range for the full dose protocol. The reason lies in an argument we presented in (Ledzewicz and Schättler, 2007a) where it was shown that it is optimal in this region to first wait (i.e., start with the control $u = 0$) until the level of endothelial support reaches the singular curve upon which treatment ensues. In this case it simply is unnecessary waste if the control $u = a$ is applied immediately. This feature, that the optimal control starts with a segment where $u = 0$ if the initial condition (p_0, q_0) lies to the left of the singular curve also is responsible for the fact that the averaged optimal dosages in Fig. 9 increase for small values of q_0 and then tend to level off. Fig. 9 gives the graphs for the averaged optimal dosages as function of q_0 in the right column. These values tend to increase in q_0 since the initial interval when $u \equiv a$ is being used, however brief it is, increases with q_0. However, note from these three diagrams that the values of the averaged controls for fixed q_0 *decrease* with increasing initial tumor volumes p_0. The reason for this lies in the fact that the optimal singular control has this property of dose-intensification already noted for model (C) in (Ergun et al., 2003) - the dosage increases in time as the tumor volume decreases. This is inherited by the averaged values.

Since anti-angiogenic therapies do not kill the cancer, but only delay its further growth, *one of the positive effects of the optimal solution is also that it delays the time when this minimum is reached*. For the averaged optimal dose protocol by construction the time when the inhibitors are exhausted is the same as for the optimal protocols and thus the times when the minimum tumor reductions are achieved are almost identical. But since the optimal singular arc applies the inhibitors at time-varying lower doses, the time T_{opt} when the minimum is realized along the optimal solution is larger, at times significantly, than the time T_{full} for the full dose protocol.

In Table 3 these times are compared for the optimal and the full dose protocols. Given the data, with a full dose protocol all inhibitors are exhausted in $4 = \frac{A}{a}$ days. For **a0**-protocols the minimum tumor volumes are being realized almost immediately afterward and this does not change much with the initial condition (p_0, q_0) (see Table 3, left portion). However, if the initial tumor volume p_0 is high, then inhibitors are given at a much lower rate initially along the optimal control and this leads to significantly larger times T when the minimum is realized (Table 3, right portion). For example, if $p_0 = 6,000\ mm^3$ and $q_0 = 15,000\ mm^3$, then the time T for the optimal control (9.138 days) is almost double the time for the straightforward **a0**-protocol (4.228 days). These times are compared in Table 3 for a wide range of initial data. The differences become especially pronounced for initial conditions to the left of the singular curve when the times until the minimum is reached more than double for the optimal protocol. Clearly this indicates that in these cases it was not such a good strategy to give all inhibitors in one session at the beginning and that they might have better been applied at lower doses like the singular control does. In fact, not only in this region, but generally for high initial values p_0 of the tumor volume, for the value a of the upper limit on the control or dosage used in these simulations, a control that applies all available inhibitors at half the maximum dose, $\frac{a}{2}$, over twice the time interval does significantly better than the straightforward full dose protocol.

The reason for this is easily understood from the geometry of the trajectories involved. Figure 10 shows an example of the relevant trajectories for initial conditions $(p_0, q_0) = (15,000\ mm^3; 6,000\ mm^3)$, the optimal trajectory shown in red, the averaged optimal dose trajectory in black, and the responses to the full and half dose protocols in blue and green, respectively. It is clear that the $\dot{q} = 0$ nullcline for $u = \frac{a}{2}$ is a much better approximation of the optimal singular curve for high initial values of p_0 than the $\dot{q} = 0$ nullcline for $u = a$ is and thus the half-dose strategy is a superior sub-optimal control for large tumor volumes. In fact, if we were to reduce the upper limit a defining the control set to $\frac{a}{2}$ in the optimal control problem, the saturation point of the singular arc is given by $11,902\ mm^3$ and thus for almost all initial conditions the optimal controls will be given by bang-bang controls that give the new "full" dose $\frac{a}{2}$ from the beginning, thus explaining the superior performance of the half-dose protocols for this range of initial conditions. Lowering the upper limit of the dose further to $\frac{a}{4}$ no longer improves the value. In fact, these protocols generally do quite worse, since the $\dot{q} = 0$ nullcline for $u = \frac{a}{4}$ now becomes a poor approximation of the singular curve. For the initial condition $(15,000\ mm^3; 6,000\ mm^3)$ the realized minimal value for the quarter dose strategy is only $p(T) = 12,316\ mm^3$, almost 20% worse than the value realized with the half-dose protocol.

Naturally, the tumor reduction realized for a constant dose protocol depends on the dosage given. Our calculations show that *a higher dose is not necessarily better if the same overall*

amount is administered, and that *optimal protocols give excellent guidance on how to choose this dosage*. In conclusion, *for model (A) the averaged optimal dose protocol is the best of the three constant dose protocols considered here and consistently comes within* 1% *of the optimal value. The full dose protocols are close to optimal for smaller initial tumor volumes* p_0 *while the half dose protocols do better for higher values of* p_0. Overall these two ad-hoc protocols stay within a reasonable range of the optimal values and from a practical side they have the advantage of not requiring any knowledge about the initial conditions.

Suboptimal a0-protocols for model (C). For this model the q-dynamics is slower and this results in a different behavior of the suboptimal protocols. Fig. 11 gives the graphs of the minimum tumor volume realized for varying initial endothelial support q_0 and fixed initial tumor volume p_0 for $p_0 = 6,000\ mm^3$ and $p_0 = 15,000\ mm^3$. For $p_0 = 6,000\ mm^3$ (more generally for smaller p_0) and small values of q_0 the half dose protocol (green curve) does almost as good as the optimal protocol, but then becomes inferior as q_0 increases; for high values of q_0 the full dose protocol is the best of the three suboptimal protocols. The reason for this lies in the fact that for high q_0 the first portion of the optimal solution is a trajectory with $u = a$ and there is a significant time period spent along this curve due to the now slower q-dynamics (see also Figure 1). This shifts the balance towards the full dose strategy. On the other hand, for small values of q_0 this portion is small and the values of the singular control are low, in about the half dose range, and thus the half dose protocol is better. For model (A) similar effects were not seen because its q-dynamics was considerably faster. For high values of p_0, like $p_0 = 15,000\ mm^3$, the four graphs separate and the half-dose protocol is the best of the suboptimal approximations we consider and the full dose does the worse. For this model the averaged optimal control is not as effective as for model (A), and again the reason for this lies in the slower q-dynamics.

Figure 12 shows the four trajectories corresponding to the optimal (red), the averaged optimal control (black), full (blue) and half dose (green) protocols for initial conditions $(p_0, q_0) = (15,000\ mm^3; 2,000\ mm^3)$. Since the point (p_0, q_0) lies to the left to the singular curve, the optimal control has an initial segment with $u = 0$ until the corresponding trajectory hits the singular arc. On the other hand, all the suboptimal protocols due to their construction immediately give inhibitors, although at different doses. All four protocols result in minimal tumor volumes that lie within 11% of the optimal minimal value. The blow-up of the final segments of these trajectories exhibits the ordering of the values.

5 Conclusion

In this paper we showed the role analytically obtained optimal protocols, although they may not be practically realizable, can still play in designing practical protocols for anti-angiogenic treatments. The optimal protocol in this process serves as a benchmark for other simpler and implementable protocols by providing theoretically calculated optimal values to which these protocols can be compared, thus determining a measure for how close to optimal a given general protocol is. Here we considered three suboptimal protocols for anti-angiogenic treatments that applied all available inhibitors in one session from the beginning of therapy, but at different dosages, and evaluated their overall efficiency by comparing the minimum tumor volumes that these protocol achieve with the optimal solution for problem [OC]. Besides this obvious application, optimal protocols also provide a set of data that can be used to design simple implementable protocols like the *averaged optimal dose protocol* we introduced in this paper. For the model of Hahnfeldt et al., (1999), these are actually the best of the suboptimal protocols considered here and they come exceptionally close to the optimal values, within 1%. Without knowing the data coming from these theoretically derived optimal protocols it would be extremely difficult, if not impossible, to come up with such a good suboptimal protocol even pursuing an exhausting trial and error search. While the choice of the best suboptimal protocol naturally depends on the specific model considered, the initial tumor volume, and the size of the endothelial support, it was observed in general that the averaged optimal dose decreased with increasing initial tumor volume. Here an important fact also is that the averaged optimal dosage was quite robust in the sense that it was not very sensitive to the initial data. Protocols that have strong robustness properties in this sense are of medical importance if these data are difficult to get. In this context we have also shown for the models considered here that the simple full and half-dose protocols are good sub-optimal strategies (generally within the $5-15\%$ range of the optimal values). Also, in all these cases it was observed that the full dose protocol was better for low initial values of p_0, but the half dose protocol performed better for higher initial tumor volumes.

An implicit assumption in the problem formulation considered in this paper is that the amount of inhibitors, A, is a priori specified to be given over one time or therapy period. However, since anti-angiogenic therapies do not kill the cancer cells, but starve the tumor by depriving it of its endothelial support, if left to itself, then at the end of therapy, this support will simply redevelop. In the absence of further treatment it follows from the dynamics of the uncontrolled system that the tumor volume will again start to increase after the time when the minimum tumor volume is realized and the system will converge to the medically non-viable globally asymptotically stable equilibrium point $(\bar{p}, \bar{q})$. It is therefore clear that repeated applications are necessary if one wants to control the tumor volume in the long run to lie below

a certain acceptable level. A simple practical scheme would be to schedule periodic sessions over some predetermined horizon $[0, T]$ that includes both the period of application of angiogenic inhibitors and a subsequent rest period. Mathematically this leads to straightforward control strategies given by periodic controls. Depending on the dynamics the overall effectiveness of such a protocol depends on the length of the predetermined therapy horizon T_{th} and adjustments in T_{th} may become necessary as treatment progresses (see also (Ledzewicz and Schättler, 2007b) in the context of model (B)). But this will be considered elsewhere.

6 Appendix: Regular Synthesis

The top diagrams in Figures 6 and 7 illustrate the structure of a regular synthesis of optimal controls and trajectories projected in (p, q)-space. Here we briefly outline why such a structure indeed implies that all controls and trajectories of this family are optimal. The mathematical arguments are not necessarily difficult, but they are quite technical and lengthy and we therefore refer the interested reader to the literature, e.g., the original article by Boltyansky (1965), a treatment of the subject in a classical textbook on optimal control (Fleming and Rishel, 1975, Chapter IV, section 6) or in a more modern style in (Bressan and Piccoli, 2007, Section 7.6).

Essentially, a *regular synthesis* for an optimal control problem is a decomposition of the relevant region in the state space into a finite or possibly countably infinite collection of embedded submanifolds, $\mathcal{M} = \{M_i\}_{i\in\mathbb{N}}$, sometimes called *strata*, together with (i) a well-defined flow of trajectories corresponding to admissible controls on each stratum and (ii) regular transitions between the strata that generate (iii) a memoryless flow of extremal trajectories; that is, forward in time there exist unique solutions and the resulting controlled trajectories satisfy the conditions of the Pontryagin Maximum Principle. If such a decomposition exists and if some rather mild technical regularity conditions are satisfied, then it can be shown that the cost evaluated along the trajectories in the synthesis, as function of the initial conditions, is a solution to a certain partial differential equation, the so-called Hamilton-Jacobi-Bellman equation, and that it has enough regularity properties to conclude that it indeed is the value function of the problem. This then verifies the optimality of the selection of controls that goes into the synthesis.

In our case the relevant subset of the state space that is covered in a $1-1$ way by a family of extremals is the set

$$\mathcal{X} = \{(p, q, y) :\ 0 < p < \bar{p},\ 0 < q < \bar{q},\ 0 \leq y \leq A\} \subset \mathbb{R}^3$$

and the strata of our decomposition are open subsets away from the singular curve (and on all of these the admissible control is simply given by the constant control $u \equiv a$ or $u \equiv 0$) and lower

dimensional strata (surfaces, curves and points) that lie at the boundaries between these higher dimensional strata. For example,

$$M_1 = \{(p,q,y) : \ (p,q) \in \mathcal{S}, \ 0 < u_{\sin}\left(\frac{p}{q}\right) < a, \ y > 0\} \tag{21}$$

is the 2-dimensional embedded analytic submanifold of the state space where the optimal control is given by the singular control. Recall that $\mathcal{S}$ denotes the singular curve; the restriction on $u_{\sin}$ ensures that it is admissible and not at saturation and the inequality $y > 0$ says that there still are inhibitors available to give. Since the singular control also is real analytic this generates a well-defined analytic flow of trajectories on M_1 that remains on M_1. There are various submanifolds lying in the boundary of M_1 corresponding to all possible entry and exit strategies defining the decompositon of the state space. For example,

$$M_2 = \{(p,q,y) : \ (p,q) \in \mathcal{S}, \ 0 < u_{\sin}\left(\frac{p}{q}\right) < a, \ y = 0\}, \tag{22}$$

consists of the points when the inhibitors run out ($y = 0$). This particular submanifold does not support trajectories, but only acts as a transient stratum into one of the open subsets of the decomposition where the control is defined by $u \equiv 0$. In this sense the 'flow' of trajectories on M_2 is trivial since it only lasts for length zero. It is not that difficult, albeit lengthy and somewhat technical to write down a complete description of all the strata in the decomposition and the flows of trajectories defined on them. By construction the resulting piecewise analytic flow of trajectories is memoryless and covers the state space $1-1$. Essentially this is a consequence of the uniqueness of solutions to ordinary differential equations and having an appropriate set of switching rules between the strata. This precisely is guaranteed by the analysis of optimal controls and trajectories carried out in (Ledzewicz and Schättler, 2007a). The sufficiency of this construction, i.e., that all the controls in this synthesis are optimal, then is a direct consequence of the results on regular synthesis (Boltyansky, 1965; Fleming and Rishel, 1975; Bressan and Piccoli, 2007).

Acknowledgement. The authors would like to thank two anonymous reviewers for valuable comments that were incorporated in this paper and helped to improve the presentation of our results. This material is based upon research supported by the National Science Foundation under collaborative research grants DMS 0707404/0707410. U. Ledzewicz's research also was partially supported by an SIUE 2007 Summer Research Fellowship. The views expressed in this paper are those of the authors and do not necessarily represent those of the funding agencies and institutions.

REFERENCES

Agur, Z., Arakelyan, L., Daugulis, P., and Ginosar, Y., (2004). Hopf point analysis for angiogenesis models, *Discr. and Cont. Dyn. Syst., Series B*, **4**, 29-38

Anderson, A., and Chaplain, M., (1998). Continuous and discrete mathematical models of tumor-induced angiogenesis, *Bull. Math. Biol.*, **60**, 857ff

Arakelyan, L., Vainstain, V., and Agur, Z., (2003). A computer algorithm describing the process of vessel formation and maturation, and its use for predicting the effects of anti-angiogenic and anti-maturation therapy on vascular tumour growth, *Angiogenesis*, **5**, 203ff

Boehm, T., Folkman, J., Browder, T., and O'Reilly, M. S., (1997). Antiangiogenic therapy of experimental cancer does not induce acquired drug resistance, *Nature*, **390**, 404ff

Bonnard, B., and Chyba, M., *Singular Trajectories and their Role in Control Theory*, Springer Verlag Paris, 2003

Boltyansky, V.G., (1966), Sufficient conditions for optimality and the justification of the dynamic programming method, *SIAM J. Control*, **4**, 326-361

Bressan, A., and Piccoli, B., *Introduction to the Mathematical Theory of Control*, American Institute of Mathematical Sciences, 2007

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

Butler, S., Kirschner, D., and Lenhart, S., (1997). Optimal control of Chemotherapy Affecting the Infectivity of HIV, in: *Advances in Mathematical Population Dynamics: Molecules, Cells, Man*, (Arino O., Axelrod D., Kimmel M., and Langlais M., Eds.), World Scientific Publishing, 104-120

Castiglione, F., and Piccoli,. B., (2006) Optimal control in a model of dendritic cell transfection cancer immunotherapy, *Bul. Math. Biol.*, **68**, 255-274

Davis, S., and Yancopoulos, G.D., (1999). The angiopoietins: Yin and Yang in angiogenesis, *Curr. Top. Microbiol. Immunol.*, **237**, 173-185

De Pillis, L.G., and Radunskaya, A.E., (2001). A mathematical tumor model with immune resistance and drug therapy: an optimal control approach, *J. of Theoret. Med.*, **3**, 79-100

d'Onofrio, A., and Gandolfi, A., (2004). Tumour eradication by antiangiogenic therapy: analysis and extensions of the model by Hahnfeldt et al., *Math. Biosci.*, **191**, 159-184

Eisen, M., *Mathematical Models in Cell Biology and Cancer Chemotherapy*, Springer Verlag,

Berlin, 1979

Ergun, A., Camphausen, K., and Wein, L.M., (2003). Optimal scheduling of radiotherapy and angiogenic inhibitors, *Bull. Math. Biol.*, **65**, 407-424

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

Fleming, W.H., and Rishel, R.W., *Deterministic and Stochastic Optimal Control*, Springer Verlag New York, 1975

Folkman, J., (1971). Tumor angiogenesis: therapeutic implications, *N. Engl. J. Med.*, **295**, 1182-1186

Folkman, J., (1972). Antiangiogenesis: new concept for therapy of solid tumors, *Ann. Surg.*, **175**, 409-416

Folkman, J., (1995). Angiogenesis inhibitors generated by tumors, *Mol. Med.*, **1**, 120-122

Folkman, J., and Klagsburn, M., (1987). Angiogenic factors, *Science*, **235**, 442-447

Forys, U., Keifetz, Y., and Kogan, Y., (2005). Critical-point analysis for three-variable cancer angiogenesis models, *Math. Biosci. Engr.*, **2**, 511-525

Hahnfeldt, P., Panigrahy, D., Folkman, J., and Hlatky, L., (1999). Tumor development under angiogenic signaling: a dynamical theory of tumor growth, treatment response, and postvascular dormancy, *Cancer Res.*, **59**, 4770-4775

Hahnfeldt, P., Folkman, J., and Hlatky, L., (2003). Minimizing long-term burden: the logic for metronomic chemotherapeutic dosing and its angiogenic basis, *J. Theor. Biol.*, **220**, 545-554

Kerbel, R.S., (1997). A cancer therapy resistant to resistance, *Nature*, **390**, 335-336

Kerbel, R.S., (2000). Tumor angiogenesis: past, present and near future, *Carcinogenesis*, **21**, 505-515

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

Klagsburn, M., and Soker, S., (1993). VEGF/VPF: the angiogenesis factor found?, *Curr. Biol.*, **3**, 699-702

Krener, A., (1977). The high-order maximal principle and its application to singular controls,

SIAM J. Contr. Optim., **15**, 256-293

Ledzewicz, U., and Schättler, H., (2002). Optimal bang-bang controls for a 2-compartment model in cancer chemotherapy, *J. of Optim. Theory and Appl. - JOTA*, **114**, 609-637

Ledzewicz, U., and Schättler, H., (2002). Analysis of a cell-cycle specific model for cancer chemotherapy, *J. of Biol. Syst.*, **10**, 183-206

Ledzewicz, U., and Schättler, H., (2005). A synthesis of optimal controls for a model of tumor growth under angiogenic inhibitors, *Proc. 44th IEEE Conf. on Dec. and Contr.*, 945-950

Ledzewicz, U., and Schättler, H., (2005). The influence of PK/PD on the structure of optimal controls in cancer chemotherapy models, *Math. Biosci. and Engr.*, **2**, 561-578

Ledzewicz, U., and Schättler, H., (2006). Optimal control for a system modelling tumor anti-angiogenesis, *ICGST-ACSE J.*, **6**, 33-39

Ledzewicz, U., and Schättler, H., (2006). Application of optimal control to a system describing tumor anti-angiogenesis, *Proc. 17th Int. Symp. MTNS*, 478-484

Ledzewicz, U., and Schättler, H., (2007). Anti-Angiogenic therapy in cancer treatment as an optimal control problem, *SIAM J. Contr. Optim.*, **46**, 1052-1079

Ledzewicz, U., and Schättler, H., (2007). Analysis of a mathematical model for tumor anti-angiogenesis, *Opt. Contr. Appl. Meth.,* published online July 2007, to appear

Sachs, R.K., Hlatky, L.R., and Hahnfeldt, P., (2001). Simple ODE models of tumor growth and anti-angiogenic or radiation treatment, *Math. Comput. Mod.*, **33**, 1297ff

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

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

Swierniak, A., (1995). Cell cycle as an object of control, *J. of Biol. Syst.*, **3**, 41-54

Swierniak, A., Gala, G., Gandolfi, A., and d'Onofrio, A., (2006). Optimization of angiogenic therapy as optimal control problem, *Proc. 4th IASTED Conf. Biomech.*, Acta Press, (Ed. M. Doblare), 56-60

Swierniak, A., d'Onofrio, A., and Gandolfi, A., (2006). Control problems related to tumor angiogenesis, *Proc. IEEE-IECON'2006,* 667-681

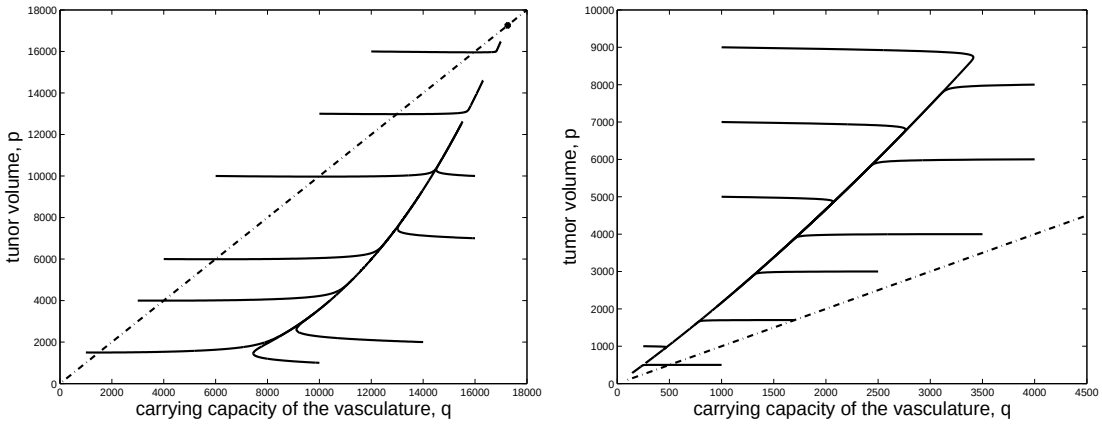

Figure 1: Model (A): Phase portraits for the uncontrolled dynamics, $u = 0$, (a, left), and with a constant control, $u = a = 75$, (b, right). The system has a strong differential-algebraic character and trajectories follow the slow manifold (i.e., the $\dot{q} = 0$ nullcline) into the globally asymptotically stable node $(\bar{p}, \bar{q})$ for $u = 0$ (left) respectively converge to the origin $(0, 0)$ for $u = a$ (right).

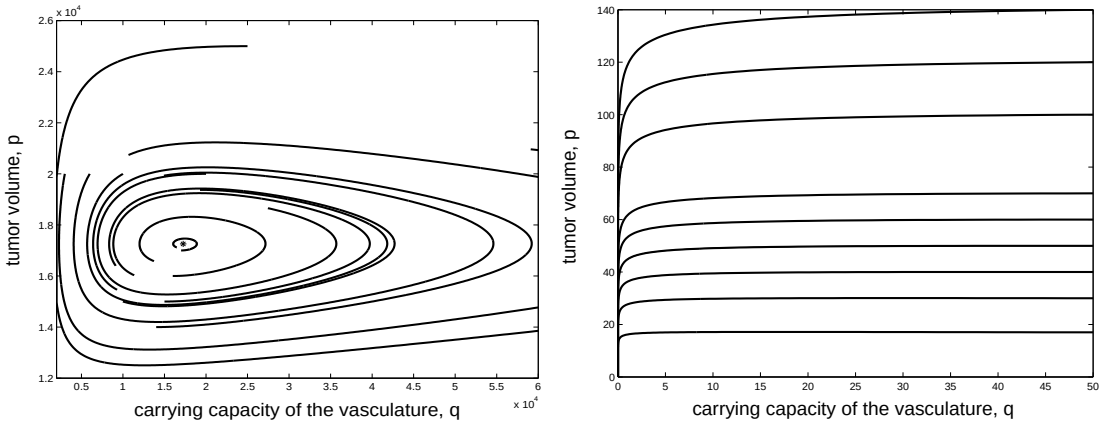

Figure 2: Model (B): Phase portraits for the uncontrolled dynamics, $u = 0$, (a, left), and with a constant control, $u = a = 75$, (b, right). The uncontrolled system has a globally asymptotically stable focus $(\bar{p}, \bar{q})$ (left); all trajectories of the controlled system converge to the origin $(0, 0)$ asymptotic with the p-axis (right).

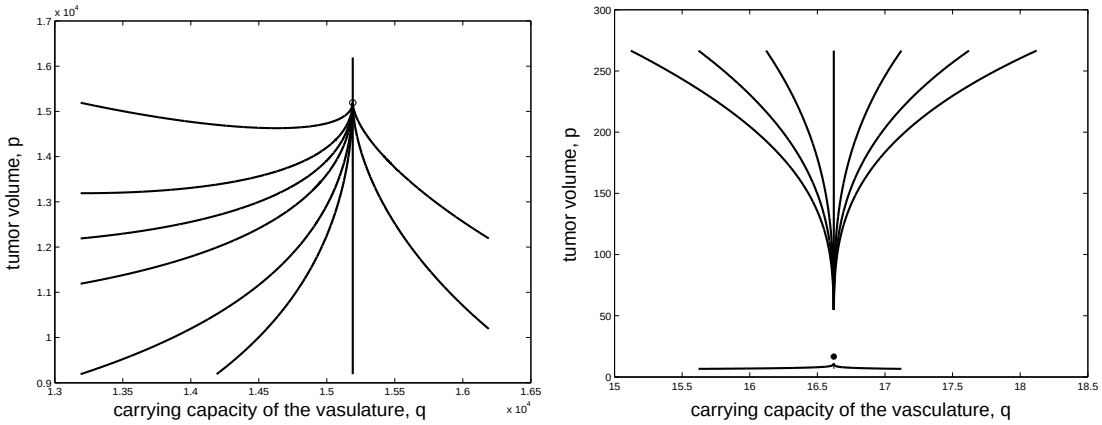

Figure 3: Model (C): Phase portraits for the uncontrolled dynamics, $u = 0$, (a, left) and with a constant control, $u = a = 15$, (b, right). Here the equilibrium point, a globally asymptotically stable node, is preserved and the $\dot{q} = 0$ nullclines are given by the vertical lines $q = \bar{q}$.

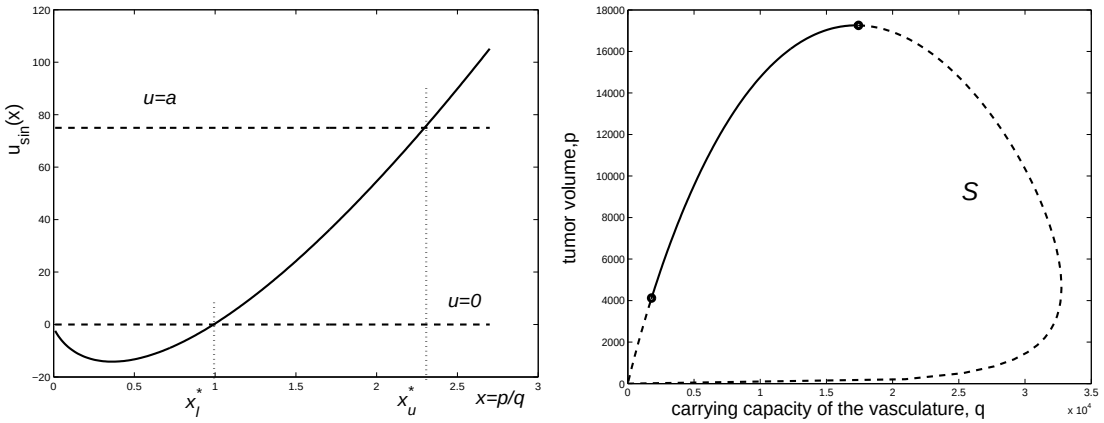

Figure 4: Model (A): (a, left) The singular control $u_{sin}(x)$ is plotted as a function of the quotient $x = \frac{p}{q}$ = tumor volume/carrying capacity of the vasculature and (b, right) the singular curve $\mathcal{S}$ is plotted in the (q, p)-plane with the admissible part (where the singular control takes values in the interval $[0, a]$) marked by the solid portion of the curve. Away from this solid segment the singular control is either negative or exceeds the maximum allowable limit $a = 75$.

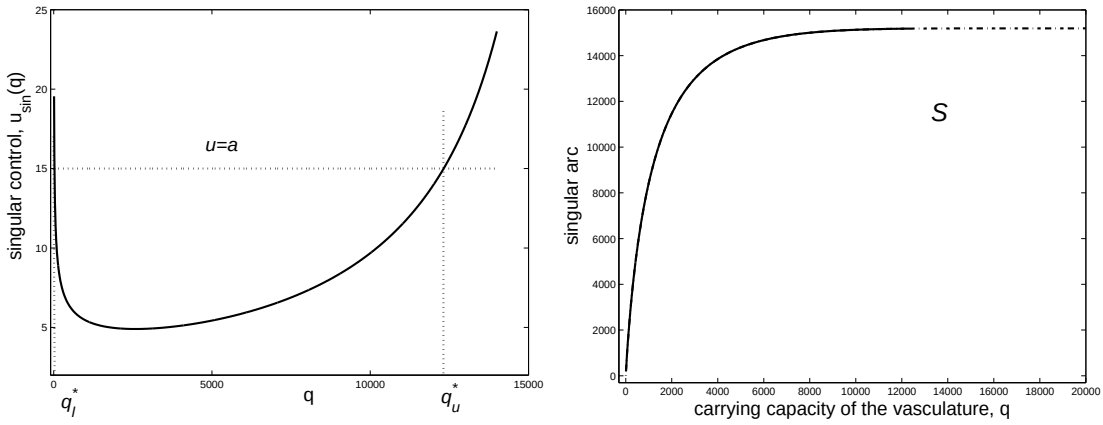

Figure 5: Model (C): (a, left) The singular control $u_{sin}(q)$ is plotted as a function of the carrying capacity of the vasculature, q, and (b, right) the singular curve $\mathcal{S}$ is plotted in the (q, p)-plane with the admissible part (where the singular control takes values in the interval $[0, a]$) marked by the solid portion of the curve. Away from this solid segment the singular control exceeds the maximum allowable limit $a = 15$.

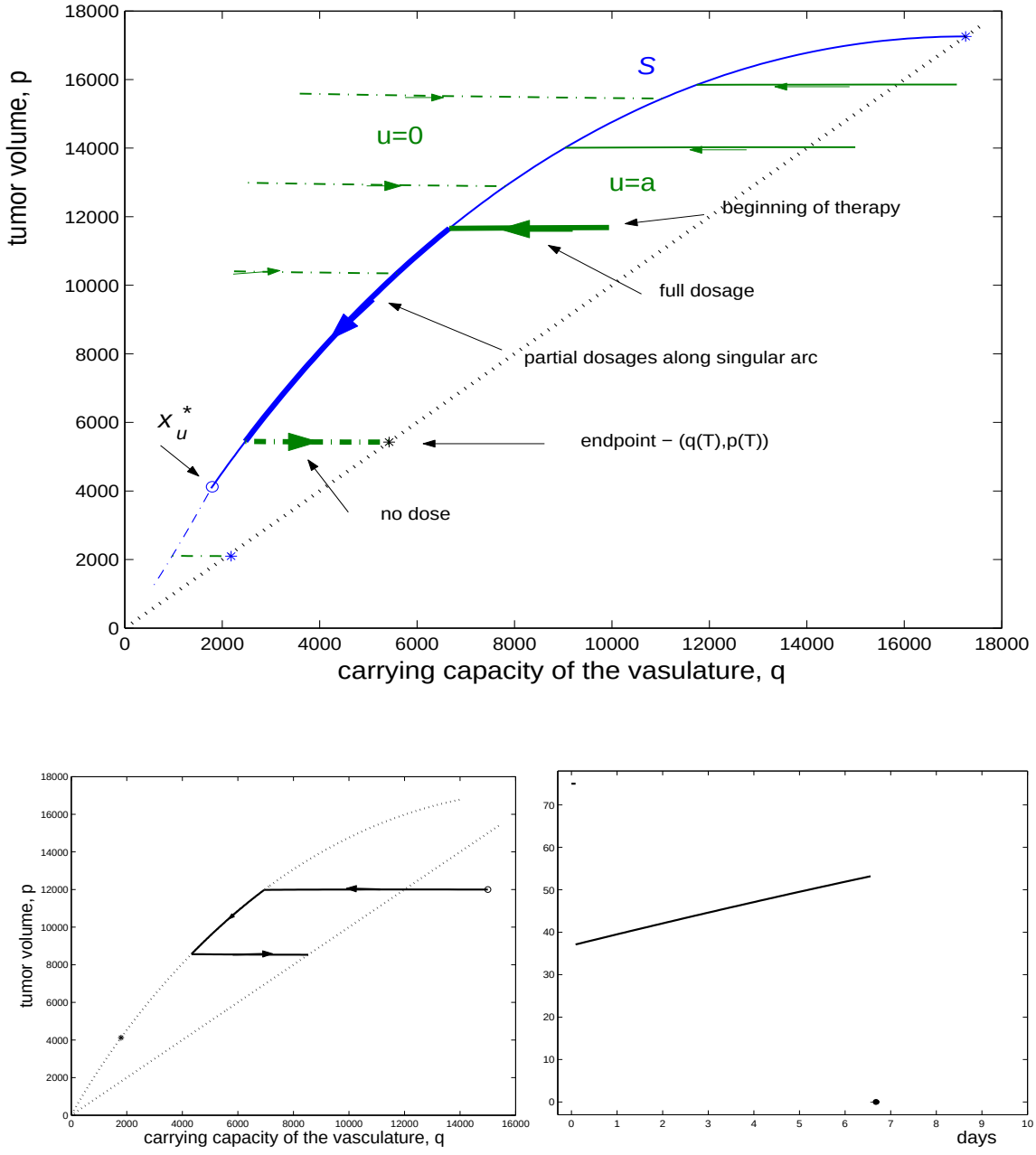

Figure 6: Model (A): Synthesis of optimal trajectories (top) and one example of an optimal trajectory (bottom, left) and corresponding optimal control (bottom, right) for initial condition $(p_0, q_0) = (12,000\,mm^3; 15,000\,mm^3)$. For this particular example first the optimal control is given at full dosage until the singular curve $\mathcal{S}$ is reached ($t_1 = 0.09$ days); then administration follows the time-varying singular control until inhibitors are exhausted ($t_2 = 6.56$ days) and due to after effects the maximum tumor reduction is realized along a trajectory for control $u = 0$ at time $T = 6.73$ days when the trajectory reaches the diagonal $p = q$.

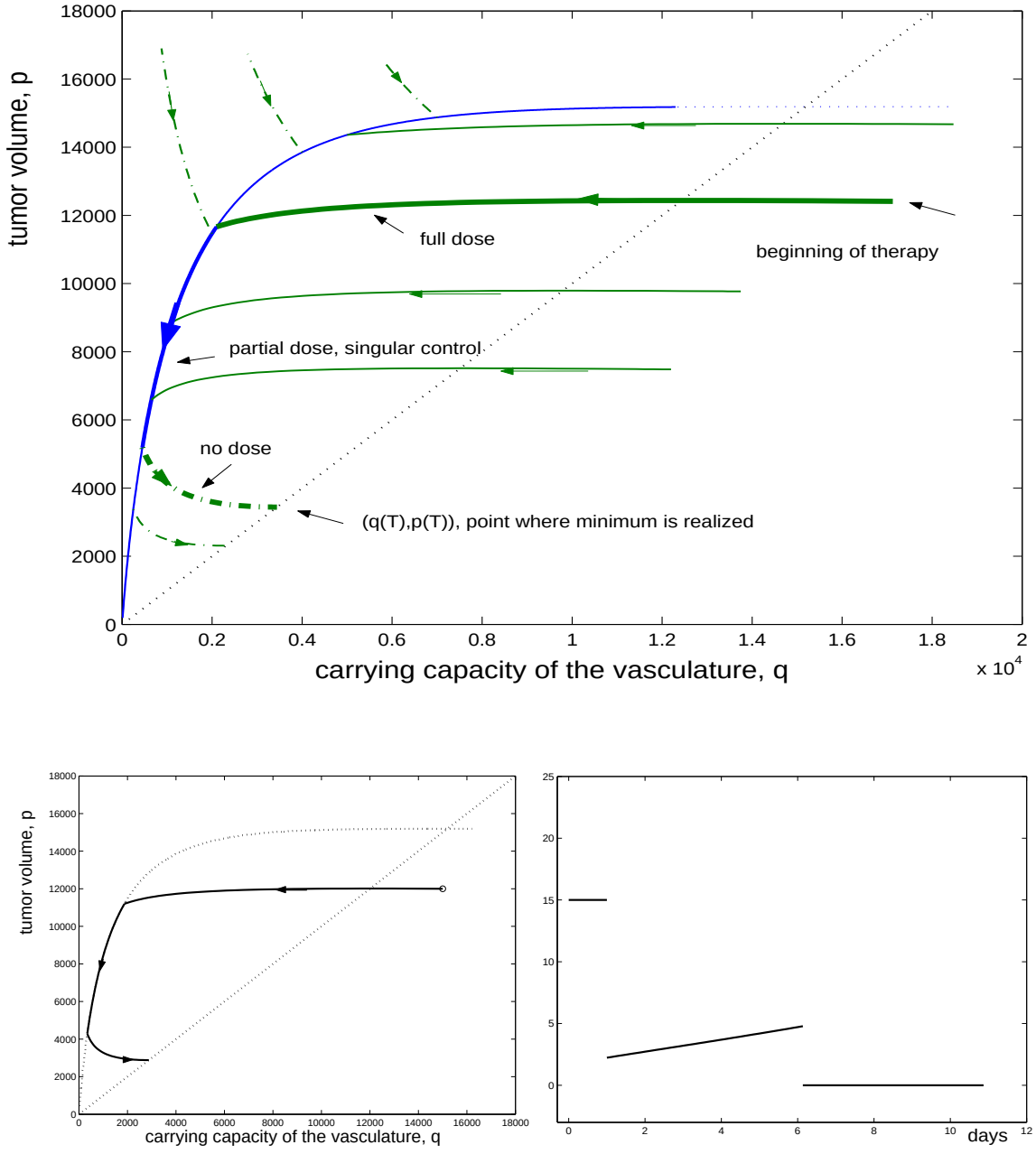

Figure 7: Model (C): Synthesis of optimal trajectories (top) and one example of an optimal trajectory (bottom, left) and corresponding optimal control (bottom, right) for initial condition $(p_0, q_0) = (12,000\, mm^3; 15,000\, mm^3)$. As for model (A), first the optimal control is given at full dosage until the singular curve $\mathcal{S}$ is reached ($t_1 = 1.00$ days); then administration follows the time-varying singular control until inhibitors are exhausted ($t_2 = 6.13$ days) and due to after effects the maximum tumor reduction is realized along a trajectory for control $u = 0$ at time $T = 10.87$ days when the trajectory reaches the diagonal $p = q$.

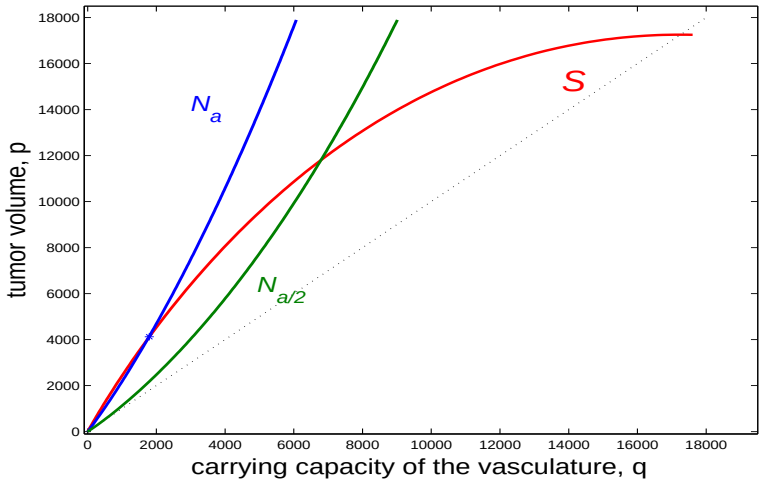

Figure 8: Model (A): The singular curve $\mathcal{S}$ and nullclines $\mathcal{N}_a$ and $\mathcal{N}_{\frac{a}{2}}$ for the constant dose protocols for $u = a = 75$ and $u = \frac{a}{2} = 37.5$, respectively.

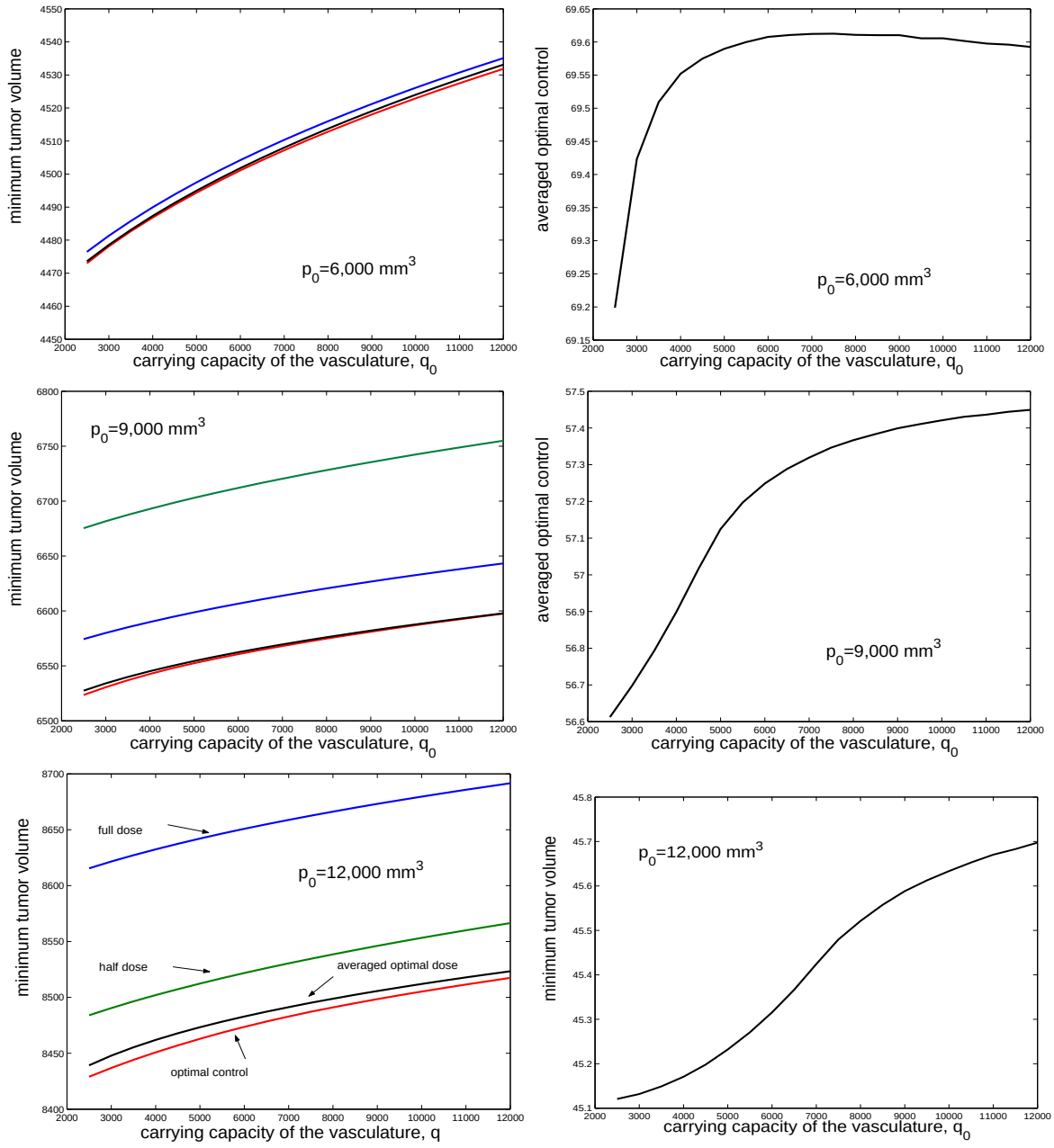

Figure 9: Model (A): Graphs of the minimum tumor volumes for a given initial tumor volume p_0 as function of the initial carrying capacity q_0 realized by the optimal control (red curve), the full-dose protocol (blue), the half-dose protocol (green) and the averaged optimal control protocol (black) for $p_0 = 6,000\,mm^3$ (top), $p_0 = 9,000\,mm^3$ (middle) and $p_0 = 12,000\,mm^3$ (bottom). The graphs on the right give the dosages of the corresponding averaged optimal controls.

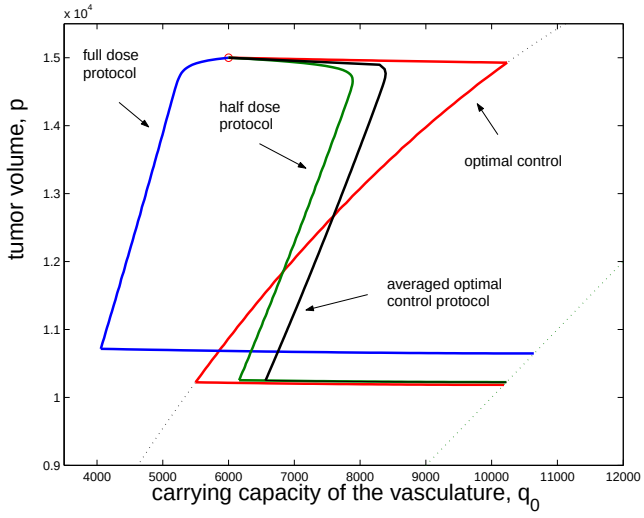

Figure 10: Model (A): A comparison of the optimal trajectory (red) with trajectories for the averaged optimal control protocol (black), the full dose protocol (blue) and half dose protocol (green) for initial condition $(p_0, q_0) = (15,000\, mm^3; 6,000\, mm^3)$.

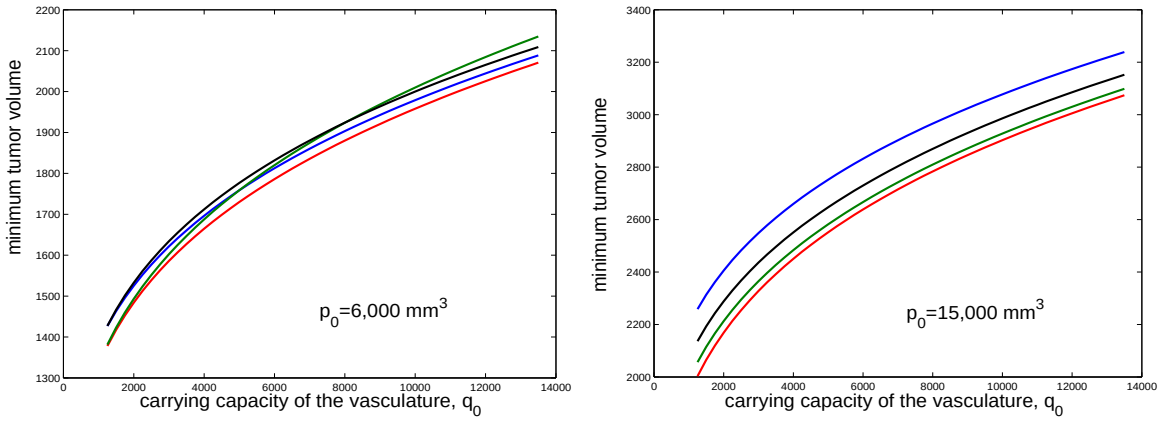

Figure 11: Model (C): Graphs of the minimum tumor volumes for a given initial tumor volume p_0 as function of the initial carrying capacity q_0 realized by the optimal control (red curve), the full-dose protocol (blue), the half-dose protocol (green) and the averaged optimal control protocol (black) for $p_0 = 6,000\, mm^3$ (left) and $p_0 = 15,000\, mm^3$ (right).

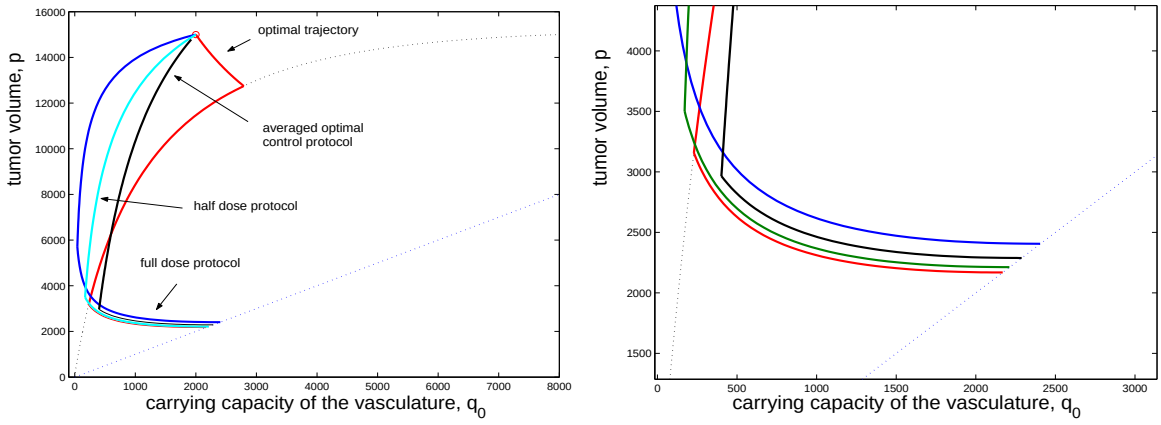

Figure 12: Model (C): A comparison of the optimal trajectory (red) with trajectories for the averaged optimal control protocol (black), the full dose protocol (blue) and half dose protocol (green) for initial condition $(p_0, q_0) = (15,000\, mm^3; 2,000\, mm^3)$. A blow-up of the final segments is included on the right.

Model	inhibition term $I(p,q)$	stimulation term $S(p,q)$	Reference
(A)	$dp^{\frac{2}{3}}q$	bp	Hahnfeldt et al., 1999
(B)	$dp^{\frac{2}{3}}q$	bq	d'Onofrio and Gandolfi, 2004
(C)	$dq^{\frac{4}{3}}$	$bq^{\frac{2}{3}}$	Ergun et al., 2003

Table 1: Models for inhibition and stimulation terms

variable/coefficients	interpretation	numerical value	dimension
p	tumor volume		mm^3
q	carrying capacity of the vasulature		mm^3
y	remaining inhibitors		mg
ξ	tumor growth parameter	0.084	day^{-1}
μ	natural loss of endothelial support	0.02	day^{-1}
b	stimulation parameter, 'birth'	5.85	day^{-1}
d	inhibition parameter, 'death'	0.00873	$mm^{-2}\ day^{-1}$
G	anti-angiogenic killing parameter	0.15	$conc^{-1}\ day^{-1}$
Model (A) a	upper limit on concentration	75	$conc$
A	available amount of inhibitors	300	mg
Model (C) a	upper limit on concentration	15	$conc$
A	available amount of inhibitors	45	mg

Table 2: Variables and parameter values

$q_0 \diagdown p_0$	6000	9000	12000	15000
3000	4.138	4.163	4.192	4.227
6000	4.139	4.164	4.192	4.228
9000	4.139	4.164	4.192	4.228
12000	4.139	4.164	4.193	4.229
15000	4.139	4.164	4.193	4.229
18000	4.140	4.165	4.193	4.229

$q_0 \diagdown p_0$	6000	9000	12000	15000
3000	4.457	5.441	6.810	9.104
6000	4.446	5.390	6.784	9.138
9000	4.446	5.380	6.745	9.134
12000	4.447	5.371	6.729	9.103
15000	4.449	5.370	6.722	9.085
18000	4.451	5.369	6.718	9.077

Table 3: Model (A): Comparison of the times in days when the minimum tumor volumes for the optimal (right) and **a0** protocol (left) are realized.

List of Captions

Figure 1: Model (A): Phase portraits for the uncontrolled dynamics, $u = 0$, (a, left), and with a constant control, $u = a = 75$, (b, right). The system has a strong differential-algebraic character and trajectories follow the slow manifold (i.e., the $\dot{q} = 0$ nullcline) into the globally asymptotically stable node $(\bar{p}, \bar{q})$ for $u = 0$ (left) respectively converge to the origin $(0, 0)$ for $u = a$ (right).

Figure 2: Model (B): Phase portraits for the uncontrolled dynamics, $u = 0$, (a, left), and with a constant control, $u = a = 75$, (b, right). The uncontrolled system has a globally asymptotically stable focus $(\bar{p}, \bar{q})$ (left); all trajectories of the controlled system converge to the origin $(0, 0)$ asymptotic with the p-axis (right).

Figure 3: Model (C): Phase portraits for the uncontrolled dynamics, $u = 0$, (a, left) and with a constant control, $u = a = 15$, (b, right). Here the equilibrium point, a globally asymptotically stable node, is preserved and the $\dot{q} = 0$ nullclines are given by the vertical lines $q = \bar{q}$.

Figure 4: Model (A): (a, left) The singular control $u_{sin}(x)$ is plotted as a function of the quotient $x = \frac{p}{q}$ = tumor volume/carrying capacity of the vasculature and (b, right) the singular curve $\mathcal{S}$ is plotted in the (q, p)-plane with the admissible part (where the singular control takes values in the interval $[0, a]$) marked by the solid portion of the curve. Away from this solid segment the singular control is either negative or exceeds the maximum allowable limit $a = 75$.

Figure 5: Model (C): (a, left) The singular control $u_{sin}(q)$ is plotted as a function of the carrying capacity of the vasculature, q, and (b, right) the singular curve $\mathcal{S}$ is plotted in the (q, p)-plane with the admissible part (where the singular control takes values in the interval $[0, a]$) marked by the solid portion of the curve. Away from this solid segment the singular control exceeds the maximum allowable limit $a = 15$.

Figure 6: Model (A): Synthesis of optimal trajectories (top) and one example of an optimal trajectory (bottom, left) and corresponding optimal control (bottom, right) for initial condition $(p_0, q_0) = (12,000\, mm^3; 15,000\, mm^3)$. For this particular example first the optimal control is given at full dosage until the singular curve $\mathcal{S}$ is reached ($t_1 = 0.09$ days); then administration follows the time-varying singular control until inhibitors are exhausted ($t_2 = 6.56$ days) and due to after effects the maximum tumor reduction is realized along a trajectory for control $u = 0$ at time $T = 6.73$ days when the trajectory reaches the diagonal $p = q$.

Figure 7: Model (C): Synthesis of optimal trajectories (top) and one example of an optimal trajectory (bottom, left) and corresponding optimal control (bottom, right) for initial condition $(p_0, q_0) = (12,000\, mm^3; 15,000\, mm^3)$. As for model (A), first the optimal control is given at full

dosage until the singular curve $\mathcal{S}$ is reached ($t_1 = 1.00$ days); then administration follows the time-varying singular control until inhibitors are exhausted ($t_2 = 6.13$ days) and due to after effects the maximum tumor reduction is realized along a trajectory for control $u = 0$ at time $T = 10.87$ days when the trajectory reaches the diagonal $p = q$.

Figure 8: Model (A): The singular curve $\mathcal{S}$ and nullclines $\mathcal{N}_a$ and $\mathcal{N}_{\frac{a}{2}}$ for the constant dose protocols for $u = a = 75$ and $u = \frac{a}{2} = 37.5$, respectively.

Figure 9: Model (A): Graphs of the minimum tumor volumes for a given initial tumor volume p_0 as function of the initial carrying capacity q_0 realized by the optimal control (red curve), the full-dose protocol (blue), the half-dose protocol (green) and the averaged optimal control protocol (black) for $p_0 = 6,000\,mm^3$ (top), $p_0 = 9,000\,mm^3$ (middle) and $p_0 = 12,000\,mm^3$ (bottom). The graphs on the right give the dosages of the corresponding averaged optimal controls.

Figure 10: Model (A): A comparison of the optimal trajectory (red) with trajectories for the averaged optimal control protocol (black), the full dose protocol (blue) and half dose protocol (green) for initial condition $(p_0, q_0) = (15,000\,mm^3; 6,000\,mm^3)$.

Figure 11: Model (C): Graphs of the minimum tumor volumes for a given initial tumor volume p_0 as function of the initial carrying capacity q_0 realized by the optimal control (red curve), the full-dose protocol (blue), the half-dose protocol (green) and the averaged optimal control protocol (black) for $p_0 = 6,000\,mm^3$ (left) and $p_0 = 15,000\,mm^3$ (right).

Figure 12: Model (C): A comparison of the optimal trajectory (red) with trajectories for the averaged optimal control protocol (black), the full dose protocol (blue) and half dose protocol (green) for initial condition $(p_0, q_0) = (15,000\,mm^3; 2,000\,mm^3)$. A blow-up of the final segments is included on the right.

Table 1: Models for inhibition and stimulation terms

Table 2: Variables and parameter values

Table 3: Comparison of the times in days when the minimum tumor volumes for the optimal (right) and **a0** protocol (left) are realized.