Optimal control of drug delivery to brain tumors for a distributed parameters model

Siddhartha P. Chakrabarty
Department of Mathematics, Statistics and Computer Science
University of Illinois at Chicago
851 S Morgan St.; m/c 249
Chicago, IL 60607–7045, USA
E–Mail: pratim@math.uic.edu
URL: http://www.math.uic.edu/∼pratim/

and

Floyd B. Hanson
Laboratory for Advanced Computing
University of Illinois at Chicago
851 S Morgan St.; m/c 249
Chicago, IL 60607–7045, USA
E–Mail: hanson@math.uic.edu
URL: http://www.math.uic.edu/∼hanson/

***Abstract*— The growth and treatment of brain tumors is mathematically examined using a distributed parameters model. The model is a system of three coupled reaction diffusion equations involving the tumor cells, normal tissue and the drug concentration. An optimal control problem is designed, with the drug delivery rate as the control and solved to obtain the state and co-state equations as well as the regular control using a modified double shot forward-backward method. This gives rise to a coupled system of equations with a forward state equation and a backward co-state equation, which is solved using a double shot method. A numerical procedure based upon the Crank-Nicolson method is used to solve the coupled system of two three-dimension partial differential equations.**

I. Introduction

The growth and control of brain tumors have been the subject of medical and scientific scrutiny for a very long time. Simply speaking a tumor, like most cancerous cells originate from a cell, that proliferates and effects its neighboring normal tissues. As the tumor cells become malignant they become more dangerous for the host. Understanding the mechanism of tumor progression is necessary for its diagnosis and treatment. As mentioned earlier, brain tumors have been studied for centuries, dating back to as early as 2500 BC [11]. The most common and deadly form of brain tumor are the gliomas, which account for more than half of the brain tumor cases. Gliomas are highly invasive and severely infiltrate the surrounding tissues [12]. Despite improved diagnostic procedures such as computerized tomography (CT) scan and magnetic resonance imaging (MRI), the benefits of such modern accessories have been restricted by the treatment options available. One major problem of administering the drugs to the brain tumor site is the *blood brain barrier (BBB)* [1], which exists in the human brain as a protection for the brain cells and as a restriction on the transport of water soluble substances between the blood and the central nervous system. Another problem that arises is the resection of a tumor, after the core mass of the tumor has been surgically removed. One way to deliver the drug is to use the drug conjugated with a *non-toxic* polymer like the BCNU and place it in the brain cavity for controlled release. A more traditional way would be to deliver the drug by an optimal distribution of the drug about the original tumor site. Wang et al. [13], [14] have studied the drug delivery method to tumors in three dimension for drugs like IgG and BCNU.

The paper will mostly focus on the control for the optimal distribution of the drug about the original tumor site. Before that we need to look at the mechanisms behind the growth of tumor cells as well as normal tissues and the drug concentration in the tumor. Unlike a lot of other tumors, gliomas can be highly diffusive [11]. Gatenby et al. [2] and Mansuri [9] study the mechanism of reaction diffusion in the growth of tumors. They also take into account the effects of competition for resources between the cancerous cells and the healthy tissues. Westman et al. [15] look at the various types of tumor growth, namely exponential, logistic and Gompertz. Murray's books [10], [11] are excellent references for the study of different types of growth mechanisms. Also, Woodward et al. [16] study a model of glioma growth and the effects of surgical resection. In this paper we set up a fairly generalized distributed parameters model for the PDE driven system, define an objective functional to minimize drug delivery and tumor burden costs, and use a modified *Lagrange multiplier method* [3] for including constraints in the control problem. Finally a *double shot*, forward and backward, numerical method is given to solve for the state, co-state and regular control. To solve for the *PDEs* we use a Crank-Nicolson predictor-corrector method developed by Hanson et al. [5] and Hanson [6], for stochastic dynamic programming.

II. Mathematical Model

Let $Y_1 = n_1(\mathbf{x}, t)$ be the density of tumor cells, $Y_2 = n_2(\mathbf{x}, t)$ be the density of normal tissue and $Y_3 = c(\mathbf{x}, t)$ be the drug concentration at any vector position $\mathbf{x}$ and time t. This spatio-temporal model is a system of three coupled reaction-diffusion equations.

Work supported in part by the National Science Foundation Grants DMS–99–73231 and DMS–02–07081. The first author was also supported by the *Provost's award for Graduate Research* awarded by the Graduate College at the University of Illinois at Chicago. Submitted to *Proceedings of American Control Conference 2005*.

A. Tumor Cells

It is assumed that the density of tumor cells, $n_1 = n_1(\mathbf{x}, t)$, satisfy a reaction-diffusion equation subject to competition with the normal cells, $n_2 = n_2(\mathbf{x}, t)$,

$$\frac{\partial n_1}{\partial t} = D_1 \nabla_x^2 [n_1] + a_1 f_1(n_1) - (\alpha_{1,2} n_2 + b_1(c))\, n_1, \qquad (1)$$

where the tumor diffusivity is D_1. Let the term $a_1 f_1(n_1)$ be the growth rate of the tumor cells, where f_1 could be exponential, logistic or Gompertz growth ,

$$f_1(n_1) = n_1, \quad n_1 \left(1 - \frac{n_1}{k_1}\right) \quad \text{or} \quad n_1 \ln\left(\frac{k_1}{n_1}\right),$$

respectively, where a_1 is the tumor cell intrinsic growth rate and k_1 is the tumor cell carrying capacity. Let $\alpha_{1,2}$ denote the death rate of the tumor cells due to competition for resources with the normal tissue. Let $b_1(c)$ be the death rate of tumor cells due to drug treatment, which could be a function of the localized drug concentration $c(\mathbf{x}, t)$ at the tumor site, e.g., purely linear function,

$$b_1(c(\mathbf{x}, t)) = \kappa_{1,3} \cdot c(\mathbf{x}, t)$$

respectively.

B. Normal Tissue

Similar assumptions are made for the density of normal cells $n_2 = n_2(\mathbf{x}, t)$ with similar coefficients. Thus, the reaction-diffusion equation for normal tissue evolution is as follows,

$$\frac{\partial n_2}{\partial t} = D_2 \nabla_x^2 [n_2] + a_2 f_2(n_2) - (\alpha_{2,1} n_1 + b_2(c))\, n_2, \qquad (2)$$

where a_2 is the normal cell intrinsic growth rate and the normal cell growth function f_2 is either exponential, logistic or Gompertz growth ,

$$f_2(n_2) = n_2, \quad n_2 \left(1 - \frac{n_2}{k_2}\right) \quad \text{or} \quad n_2 \ln\left(\frac{k_2}{n_2}\right),$$

respectively, where k_2 is the normal tissue carrying capacity. Note that the $b_2(c)$ term indicates that some normal tissues could die as a result of the treatment and $b_2(c)$ must also depend on the local drug concentration $c(\mathbf{x}, t)$, e.g.,

$$b_2(c(\mathbf{x}, t)) = \kappa_{2,3} \cdot c(\mathbf{x}, t).$$

C. Concentration

It is assumed that the drug shows a diffusive behavior and that there is a reabsorption at the rate a_3. Also let $u = u(\mathbf{x}, t)$ be the rate at which the drug is being injected. The choice of the letter u indicates that we will use it as the control when dealing with the optimal control system. The equation for drug concentration at position $\mathbf{x}$ and time t is,

$$\frac{\partial c}{\partial t} = D_3 \nabla_x^2 [c] + a_3 f_3(c) + u, \qquad (3)$$

where $f_3(c) = -c$ is the reabsorption function[14].

D. Global State Vector

Let the global state vector be

$$\mathbf{Y}(\mathbf{x}, t) = [Y_i(\mathbf{x}, t)]_{3\times 1} = \begin{bmatrix} n_1(\mathbf{x}, t) \\ n_2(\mathbf{x}, t) \\ c(\mathbf{x}, t) \end{bmatrix}, \qquad (4)$$

at position $\mathbf{x}$ in the *interior* Ω of the state domain and time t on $[0, t_f]$.

E. Initial and Boundary Conditions

Let the initial conditions for the state be

$$\mathbf{Y}(\mathbf{x}, 0) = \begin{bmatrix} n_1(\mathbf{x}, 0) \\ n_2(\mathbf{x}, 0) \\ c(\mathbf{x}, 0) \end{bmatrix} = \begin{bmatrix} n_{10}(\mathbf{x}) \\ n_{20}(\mathbf{x}) \\ c_0(\mathbf{x}) \end{bmatrix} \equiv \mathbf{Y}_0(\mathbf{x}), \qquad (5)$$

for $\mathbf{x}$ in Ω. Murray [11] recommends using Gaussian distribution for the initial distributions of tumors. The *no flux* boundary conditions are

$$-D(\widehat{\mathbf{N}} \cdot \nabla_x)[\mathbf{Y}](\mathbf{x}, t) = \begin{bmatrix} -D_1(\widehat{\mathbf{N}} \cdot \nabla_x)[n_1] \\ -D_2(\widehat{\mathbf{N}} \cdot \nabla_x)[n_2] \\ -D_3(\widehat{\mathbf{N}} \cdot \nabla_x)[c] \end{bmatrix} (\mathbf{x}, t) = \mathbf{0}, \quad (6)$$

for $\mathbf{x} \in \mathbf{\Gamma} = \partial\Omega$, i.e., on the boundary of the domain, and for $t \in [0, t_f]$, assuming $D_i \neq 0$ or else the D_i would not be used in the condition, where $\widehat{\mathbf{N}}(\mathbf{x}, t)$ is the normal to the boundary, $-D_i \nabla_x [Y_i]$ is the flux of the ith component and the diffusion matrix,

$$D(\mathbf{x}) = \begin{bmatrix} D_1 & 0 & 0 \\ 0 & D_2 & 0 \\ 0 & 0 & D_3 \end{bmatrix} = [D_i \delta_{i,j}]_{3\times 1}$$

is diagonal and inhomogeneous depending on the brain matter [12], where $\delta_{i,j}$ is the Kronecker delta. Note that the no flux condition at the boundary is motivated by the physical reality that the brain is a finite and closed domain.

III. Optimal Control Problem

A possible objective functional is the quadratic form of running and terminal costs,

$$J(u) = \frac{1}{2} \int_0^{t_f} dt \int_\Omega d\mathbf{x} \left(r n_1^2(\mathbf{x}, t) + s\,(u - u_0)^2\,(\mathbf{x}, t) \right) + \int_\Omega d\mathbf{x} \left(q_1 n_1^2(\mathbf{x}, t_f) + q_3 c^2(\mathbf{x}, t_f) \right). \qquad (7)$$

The goal is to minimize this functional with respect to the drug input rate relative to some threshold rate u_0 and the terminal costs at t_f, i.e., $\min_u [J(u)]$. Note that here $r > 0$ is the tumor burden cost coefficient and $s > 0$ is the drug delivery cost coefficient, while $q_1 > 0$ and $q_3 > 0$ are the corresponding final costs. We are trying to minimize the density of tumor cells and the drug delivery quadratic control term $(u(\mathbf{x}, t) - u_0(\mathbf{x}, t))^2$. We could have chosen a linear control which would have been less realistic, and also would give rise to problems like singular control. In addition no assumption is made about the control constraints, even though there might be physical restriction on the amount of drugs that can be administered.

IV. Vector Form

For the sake of *brevity* we put the mathematical model in vector form with vectors in boldface.

A. Governing equations

The vector state satisfies the PDE:

$$\frac{\partial \mathbf{Y}}{\partial t} = D\nabla_x^2[\mathbf{Y}] + A\mathbf{f}(\mathbf{Y}) + B(\mathbf{Y},t)\mathbf{Y} + \mathbf{U}, \tag{8}$$

where

$$A = \begin{bmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{bmatrix} = [a_i\delta_{i,j}]_{3\times 1}, \quad \mathbf{f}(\mathbf{Y}) = \begin{bmatrix} f_1(n_1) \\ f_2(n_2) \\ f_3(c) \end{bmatrix},$$

$$B(\mathbf{Y},t) = -\begin{bmatrix} \alpha_{1,2}n_2 + \kappa_{1,3}c & 0 & 0 \\ 0 & \alpha_{2,1}n_1 + \kappa_{2,3}c & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{U}(\mathbf{x},t) = \begin{bmatrix} 0 \\ 0 \\ u(\mathbf{x},t) \end{bmatrix} = U_3(\mathbf{x},t)\mathbf{e}_3, \tag{9}$$

where $\mathbf{e}_i$ is the ith unit vector.

B. Objective Functional

The objective in vector form is

$$J[\mathbf{Y},\mathbf{U}] = \frac{1}{2}\int_0^{t_f} dt \int_\Omega d\mathbf{x} \left(\mathbf{Y}^\top R\mathbf{Y} + (\mathbf{U}–\mathbf{U}_0)^\top S\,(\mathbf{U}–\mathbf{U}_0) \right) + \frac{1}{2}\int_\Omega d\mathbf{x}\,(\mathbf{Y}^\top Q\mathbf{Y})\Big|_{t=t_f}, \tag{10}$$

where $R = r\mathbf{e}_1\mathbf{e}_1^\top$, $S = s\mathbf{e}_3\mathbf{e}_3^\top$, $Q = q_1\mathbf{e}_1\mathbf{e}_1^\top + q_3\mathbf{e}_3\mathbf{e}_3^\top$ and $\mathbf{U}_0 = u_0(\mathbf{x},t)\mathbf{e}_3$.

V. Defining the Pseudo-Hamiltonian

We have three vectors for the *Lagrange multipliers*, two of which are functions of space and time and one is independent of time, needed to include the optimization constraints in the extended objective for the state PDE (8), the boundary condition (6) and the initial condition (5),

$$\boldsymbol{\xi}(\mathbf{x},t) = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{bmatrix}, \ \boldsymbol{\eta}(\mathbf{x},t) = \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{bmatrix}, \ \boldsymbol{\chi}(\mathbf{x}) = \begin{bmatrix} \chi_1 \\ \chi_2 \\ \chi_3 \end{bmatrix}, \tag{11}$$

i.e., $\xi_i = \xi_i(\mathbf{x},t), \eta_i = \eta_i(\mathbf{x},t)$ and $\chi_i = \chi_i(\mathbf{x})$, for $i = 1:3$. Letting $\mathbf{Z} = (\mathbf{Y},\mathbf{U},\boldsymbol{\xi},\boldsymbol{\eta},\boldsymbol{\chi})$, define the *pseudo-Hamiltonian* as,

$$\begin{aligned} \mathcal{H}(\mathbf{Z}) = {} & \frac{1}{2}\int_0^{t_f} dt \int_\Omega d\mathbf{x} \left(\mathbf{Y}^\top R\mathbf{Y} + (\mathbf{U}–\mathbf{U}_0)^\top S\,(\mathbf{U}–\mathbf{U}_0) \right) \\ & + \frac{1}{2}\int_\Omega d\mathbf{x}\,(\mathbf{Y}^\top Q\mathbf{Y})\Big|_{t=t_f} \\ & + \int_0^{t_f} dt \int_\Omega d\mathbf{x}\ \boldsymbol{\xi}^\top \left(\frac{\partial \mathbf{Y}}{\partial t} - D\nabla_x^2[\mathbf{Y}] - A\mathbf{f}(\mathbf{Y}) \right. \\ & \left. - B(\mathbf{Y},t)\mathbf{Y} - \mathbf{U} \right) \\ & + \int_0^{t_f} dt \int_{\partial\Omega} d\boldsymbol{\Gamma}\ \boldsymbol{\eta}^\top \left(-D\left(\widehat{\mathbf{N}}\cdot\nabla_x \right)[\mathbf{Y}] \right) \\ & + \int_\Omega d\mathbf{x}\ \boldsymbol{\chi}^\top \left(\mathbf{Y}\Big|_{t=0} - \mathbf{Y}_0 \right). \end{aligned} \tag{12}$$

VI. Optimal Control Variational Formulation

The *calculus of variations* is used to find differential equation of optimal control for the control, state and the co-state (adjoint or Lagrange multiplier) by seeking the functional critical point necessary conditions for the first variation [3], [8] of the *pseudo-Hamiltonian* $\mathcal{H}(\mathbf{Z})$.

A. Pseudo-Hamiltonian First Variation

Let the extended state vector be perturbed about the optimal trajectory $\mathbf{Z}^*$, so that $\mathbf{Z} = \mathbf{Z}^* + \delta\mathbf{Z}$, where $\delta\mathbf{Z}$ is the perturbation. Next expand the pseudo-Hamiltonian

$$\mathcal{H}(\mathbf{Z}^* + \delta\mathbf{Z}) = \mathcal{H}(\mathbf{Z}^*) + \delta\mathcal{H}(\mathbf{Z}^*,\delta\mathbf{Z}) + O((\delta\mathbf{Z})^2).$$

Neglecting the quadratic order terms, including the 2nd variation of $\mathcal{H}$, the first variation is given by functional terms linear in $\delta\mathbf{Z}$ using (12),

$$\begin{aligned} \delta\mathcal{H}(\mathbf{Z}^*,\delta\mathbf{Z}) = {} & \int_0^{t_f} dt \int_\Omega d\mathbf{x} \left((\mathbf{Y}^*)^\top R\delta\mathbf{Y} + (\mathbf{U}^*–\mathbf{U}_0)^\top S\delta\mathbf{U} \right) \\ & + \int_\Omega d\mathbf{x}\,((\mathbf{Y}^*)^\top Q\delta\mathbf{Y})\Big|_{t=t_f} \\ & + \int_0^{t_f} dt \int_\Omega d\mathbf{x} \left((\boldsymbol{\xi}^*)^\top (\delta\mathbf{Y}_t - D\nabla_x^2[\delta\mathbf{Y}] \right. \\ & - A(\delta\mathbf{Y}\cdot\nabla_Y)[\mathbf{f}](\mathbf{Y}^*) - B(\mathbf{Y}^*,t)\delta\mathbf{Y} \\ & - (\delta\mathbf{Y}\cdot\nabla_Y)[B](\mathbf{Y}^*,t)\mathbf{Y}^* - \delta\mathbf{U}) \\ & + \delta\boldsymbol{\xi}^\top (\mathbf{Y}_t^* - D\nabla_x^2[\mathbf{Y}^*] - A\mathbf{f}(\mathbf{Y}^*) \\ & \left. - B(\mathbf{Y}^*,t)\mathbf{Y}^* - \mathbf{U}^*) \right) \\ & - \int_0^{t_f} dt \int_{\partial\Omega} d\boldsymbol{\Gamma} \left((\boldsymbol{\eta}^*)^\top D\left(\widehat{\mathbf{N}}\cdot\nabla_x \right)[\delta\mathbf{Y}] \right. \\ & \left. + \delta\boldsymbol{\eta}^\top D\left(\widehat{\mathbf{N}}\cdot\nabla_x \right)[\mathbf{Y}^*] \right) \\ & + \int_\Omega d\mathbf{x} \left((\boldsymbol{\chi}^*)^\top \delta\mathbf{Y} + \delta\boldsymbol{\chi}^\top \left(\mathbf{Y}\Big|_{t=0} - \mathbf{Y}_0 \right) \right). \end{aligned} \tag{13}$$

Before the critical conditions for first variation in (13) can by used to obtain the extended state equations, the functional dependence of the higher derivatives in time and state of the extended state perturbations must be eliminated

on lower order terms by one or two integrations by parts, i.e., by one,

$$\int_0^{t_f} dt(\boldsymbol{\xi}^*)^\top \delta\mathbf{Y}_t = (\boldsymbol{\xi}^*)^\top \delta\mathbf{Y}\Big|_0^{t_f} - \int_0^{t_f} dt \delta\mathbf{Y}^\top \boldsymbol{\xi}_t^*$$

and by two using the Green's formula [4],

$$\begin{aligned}\int_\Omega d\mathbf{x}(\boldsymbol{\xi}^*)^\top D\nabla_x^2[\delta\mathbf{Y}] = &\int_\Omega d\mathbf{x}\delta\mathbf{Y}^\top\nabla_x^2[D\boldsymbol{\xi}^*] \\ &+ \int_{\partial\Omega} d\boldsymbol{\Gamma}\Big((\widehat{\mathbf{N}}\cdot\nabla_x)[\delta\mathbf{Y}^\top]D\boldsymbol{\xi}^* \\ &-\delta\mathbf{Y}^\top(\widehat{\mathbf{N}}\cdot\nabla_x)[D\boldsymbol{\xi}^*]\Big).\end{aligned}$$

Merging these identities with (13), rearranging inner products and collecting terms the extended state equations yields the following intermediate form:

$$\begin{aligned}\delta\mathcal{H}(\mathbf{Z}^*,\delta\mathbf{Z}) = &\int_0^{t_f} dt \int_\Omega d\mathbf{x}\, \delta\mathbf{Y}^\top\big(R\mathbf{Y}^* - \boldsymbol{\xi}_t^* - \nabla_x^2[D\boldsymbol{\xi}^*] \\ &-\nabla_Y[\mathbf{f}](\mathbf{Y}^*)A\boldsymbol{\xi}^* - B(\mathbf{Y}^*,t)\boldsymbol{\xi}^* \\ &-\nabla_Y[B](\mathbf{Y}^*,t)):(\boldsymbol{\xi}^*(\mathbf{Y}^*)^\top\big) \\ &+\int_0^{t_f} dt \int_\Omega d\mathbf{x}\, \delta\mathbf{U}^\top(S\,(\mathbf{U}^* - \mathbf{U}_0) - \boldsymbol{\xi}^*) \\ &+\int_0^{t_f} dt \int_\Omega d\mathbf{x}\, \delta\boldsymbol{\xi}^\top\big(\mathbf{Y}_t^* - D\nabla_x^2[\mathbf{Y}^*] \\ &-A\mathbf{f}(\mathbf{Y}^*) - B(\mathbf{Y}^*,t)\mathbf{Y}^* - \mathbf{U}^*) \\ &-\int_0^{t_f} dt \int_{\partial\Omega} d\boldsymbol{\Gamma}\, \delta\boldsymbol{\eta}^\top D\left(\widehat{\mathbf{N}}\cdot\nabla_x\right)[\mathbf{Y}^*] \\ &+\int_0^{t_f} dt \int_{\partial\Omega} d\boldsymbol{\Gamma}\, \delta\mathbf{Y}^\top\left(\widehat{\mathbf{N}}\cdot\nabla_x\right)[D\boldsymbol{\xi}^*] \\ &-\int_0^{t_f} dt \int_{\partial\Omega} d\boldsymbol{\Gamma}\left(\widehat{\mathbf{N}}\cdot\nabla_x\right)[\delta\mathbf{Y}^\top]D(\boldsymbol{\eta}^* + \boldsymbol{\xi}^*) \\ &+\int_\Omega d\mathbf{x}\left(\delta\boldsymbol{\chi}^\top(\mathbf{Y}^* - \mathbf{Y}_0(\mathbf{x}))\right)\Big|_{t=0} \\ &+\int_\Omega d\mathbf{x}\left(\delta\mathbf{Y}^\top(\boldsymbol{\chi}^* - \boldsymbol{\xi}^*)\right)\Big|_{t=0} \\ &+\int_\Omega d\mathbf{x}\left(\delta\mathbf{Y}^\top(\boldsymbol{\xi}^* + Q\mathbf{Y})\right)\Big|_{t=t_f},\end{aligned}$$

where $A:B$ denotes the trace of the matrix AB or the double-dot product.

B. State Equations

The optimal state equation is recovered by setting the coefficient of $(\delta\boldsymbol{\xi}^*)^\top$ to zero:

$$\frac{\partial\mathbf{Y}^*}{\partial t} = D\nabla_x^2[\mathbf{Y}^*] + A\mathbf{f}(\mathbf{Y}^*) + B(\mathbf{Y}^*,t)\mathbf{Y}^* + \mathbf{U}^* \quad (14)$$

on $\Omega\times(0,t_f]$, with boundary conditions on $\partial\Omega\times[0,t_f]$ from the coefficient of $(\delta\boldsymbol{\eta}^*)^\top$, i.e.,

$$-D(\widehat{\mathbf{N}}\cdot\nabla_x)[\mathbf{Y}^*](\mathbf{x},t) = \mathbf{0}, \quad (15)$$

for $(\mathbf{x},t)\in\partial\Omega\times[0,t_f]$ and with initial conditions on the interior Ω from the coefficient of $(\delta\boldsymbol{\chi}^*)^\top$, i.e.,

$$\mathbf{Y}^*(\mathbf{x},0) = \mathbf{Y}_0(\mathbf{x}) \quad (16)$$

for $\mathbf{x}\in\Omega$. Due to the presence of the functions $\mathbf{f}(\mathbf{Y})$ and $B(\mathbf{Y},t)\mathbf{Y}$ the forward PDE (14) will be nonlinear.

C. Degenerate Regular Optimal Control

Since the control has been defined in (9) has been defined as only having on component, only the coefficient of δU_3 is set to zero giving the regular, but degenerate, control as

$$U_3^*(\mathbf{x},t) = U_{0,3}(\mathbf{x},t) + \frac{1}{s}\xi_3^*(\mathbf{x},t), \quad (17)$$

on $\Omega\times[0,t_f]$, provided $s\neq 0$. Note that the control law with $\delta U_1 = 0 = \delta U_2$ only requires solving for the 3rd component of the first co-state vector $\boldsymbol{\xi}^*(\mathbf{x},t)$.

D. Co-State Equations

Upon setting the functional coefficient of $(\delta\mathbf{Y}^*)^\top$ to zero yields the primary co-state backward PDE:

$$\begin{aligned}\mathbf{0} = &\frac{\partial\boldsymbol{\xi}^*}{\partial t} + \nabla_x^2[D\boldsymbol{\xi}^*] + \nabla_Y[\mathbf{f}](\mathbf{Y}^*)A\boldsymbol{\xi}^* \\ &+B(\mathbf{Y}^*,t)\boldsymbol{\xi}^* + \nabla_Y[B](\mathbf{Y}^*,t)):(\boldsymbol{\xi}^*(\mathbf{Y}^*)^\top) \\ &-R\mathbf{Y}^*, \end{aligned} \quad (18)$$

for $(\mathbf{x},t)\in\Omega\times[0,t_f)$. This PDE (18) is unidirectionally coupled to the state PDE (14), except that only the 3rd component $\xi_3^*(\mathbf{x},t)$ is needed for the regular optimal control input $U_3^*(\mathbf{x},t)$ from (17). The boundary condition follows from setting the functional coefficient of $\delta\mathbf{Y}(\mathbf{x},t)$ for $\mathbf{x}$ on $\boldsymbol{\Gamma}=\partial\Omega$ to zero, so

$$(\widehat{\mathbf{N}}\cdot\nabla_x)[D\boldsymbol{\xi}^*](\mathbf{x},t) = \mathbf{0}, \quad (\mathbf{x},t)\in\partial\Omega\times[0,t_f) \quad (19)$$

and the final condition for this backward PDE follows from forcing the coefficient of $\delta\mathbf{Y}(\mathbf{x},t_f)$ to be zero on Ω,

$$\boldsymbol{\xi}^*(\mathbf{x},t_f) = -Q\mathbf{Y}(\mathbf{x},t_f). \quad (20)$$

The two other co-state vectors should not be needed, but satisfy rather simple equations. The 2nd co-state vector equation follows as the zero coefficient of $(\widehat{\mathbf{N}}\cdot\nabla_x)[\delta\mathbf{Y}^\top]$ on the state boundary $\boldsymbol{\Gamma}=\partial\Omega$,

$$\boldsymbol{\eta}^*(\mathbf{x},t) = -\boldsymbol{\xi}^*(\mathbf{x},t), \quad (\mathbf{x},t)\in\partial\Omega\times[0,t_f].$$

The 3rd co-state vector equation follows as the zero coefficient of state initial condition $\delta\mathbf{Y}(\mathbf{x},0)$,

$$\boldsymbol{\chi}^*(\mathbf{x}) = \boldsymbol{\xi}^*(\mathbf{x},0), \quad \mathbf{x}\in\Omega.$$

VII. Iterative Approximations for Coupled System

We need to solve the system of equations developed in the previous section, namely the *state equations* (14) using the *regular optimal control* (17) and *co-state equations* (18), with the understanding that the state equations are forward equations while the co-state equations are backward equations. The method a *double shot* method, since model has two vector-valued PDEs, so the method consists of one forward shot with (14) and and one backward shot with (18).

The following steps can be used to get approximate solutions:

1) An initial guess for the first forward-backward shot iteration is made for the control $U_3(\mathbf{x},t) = U_3^{(1)}(\mathbf{x},t)$ in (17). Substituting it into the *state forward PDE* (14) solving for $\mathbf{Y}^{(1)}(\mathbf{x},t)$, using the boundary condition (15) and initial condition (16).
2) Next the approximate final condition $\boldsymbol{\xi}^{(1)}(\mathbf{x},t_f) = -Q\mathbf{Y}^{(1)}(\mathbf{x},t_f)$ (20) is used to start the backward *co-state PDE* (18) approximation using boundary conditions (19) .
3) Once $\boldsymbol{\xi}^{(1)}(\mathbf{x},t)$ is determined the *regular optimal control equation* is used to determine the updated value of the control,

$$U_3^{(2)}(\mathbf{x},t) = U_{0,3}(\mathbf{x},t) + \xi_3^{(1)}(\mathbf{x},t)/s.$$

4) This process is repeated for $\ell = 2:L$ double shot iterations until a convergence criterion for sufficiently large L is reached, e.g., the relative criterion for the control,

$$\left|\left|U_3^{(\ell)}(\mathbf{x},t) – U_3^{(\ell-1)}(\mathbf{x},t)\right|\right| < \mathrm{tol}_u \left|\left|U_3^{(\ell-1)}(\mathbf{x},t)\right|\right|,$$

and, say,

$$\left|\left|Y_1^{(\ell)}(\mathbf{x},t) – Y_1^{(\ell-1)}(\mathbf{x},t)\right|\right| < \mathrm{tol}_y \left|\left|Y_1^{(\ell-1)}(\mathbf{x},t)\right|\right|,$$

for $\ell = 2{:}L$ until satisfied, provided $||U_3^{(\ell-1)}(\mathbf{x},t)|| \neq 0$ and $||Y_1^{(\ell-1)}(\mathbf{x},t)|| \neq 0$, where $\mathrm{tol}_u > 0$ and $\mathrm{tol}_y > 0$ are some prescribed tolerances.

VIII. Double Shot Forward-Backward Computational Method

In reality the problem is highly non-linear as are many problems in biology and we need numerical approximations of the solution. The main problem here is the fact that we have a forward state equation and a backward co-state equation. What we really do here is modify shooting methods [7] for initial-final-boundary value problems, where the starting aim is replaced by an estimate of the full control law $U_3^{(\ell)}(\mathbf{x},t)$ for the forward integration of the state PDE (14) whose final approximation $\mathbf{Y}^{(\ell)}(\mathbf{x},t_f)$ serves as the backward aim (20) for the backward integration of the co-state PDE (18) producing an approximation $\boldsymbol{\xi}^{(\ell)}(\mathbf{x},t)$ whose third component is used to update (17) the control law $U_3^{(\ell+1)}(\mathbf{x},t)$.

Before we actually go into the details of the shooting method, we will briefly discuss the Crank-Nicolson, predictor-corrector central finite difference method used to solve the nonlinear PDEs. The space is discretized as follows,

$$\mathbf{x} \to \mathbf{x_j} = [x_{j_i,1} + (j_i - 1)\cdot \Delta x_i]_{3\times 1}.$$

Here Δx_i is the mesh size for state i and $\mathbf{j} = [j_i]_{3\times 1}$ where, $j_i = 1{:}M_i$ nodes per node for states $i = 1{:}3$. For the forward state equation we have the forward time discretization, $t \to t_k = k\Delta t$, for $k = 0{:}K$ time steps where Δt is the forward time step size, $t_0 = 0$ and $t_K = t_f$. Now we consider the vector state/co-state PDE system in general form,

$$\begin{aligned}\mathbf{Y}_t^* &= \mathbf{F}(\mathbf{x},t,\mathbf{Y}^*(\mathbf{x},t),\mathbf{U}^*(\mathbf{x},t)),\\ \mathbf{0} &= \boldsymbol{\xi}_t^* + \mathbf{G}(\mathbf{x},t,\boldsymbol{\xi}^*(\mathbf{x},t),\mathbf{Y}^*(\mathbf{x},t)).\end{aligned}$$

In this equation, using the ordinate discretization $\mathbf{Y}(\mathbf{x_j},t_k) \simeq \mathbf{Y}_{\mathbf{j},k}$ and $\boldsymbol{\xi}(\mathbf{x_j},t_k) \simeq \boldsymbol{\xi}_{\mathbf{j},k}$, the relevant derivatives discretization are

$$(\mathbf{Y}_t^*)_{\mathbf{j},k+1/2} \simeq \frac{\mathbf{Y}_{\mathbf{j},k+1/2+1/2} - \mathbf{Y}_{\mathbf{j},k+1/2-1/2}}{2(\Delta t/2)},$$
$$(\boldsymbol{\xi}_t^*)_{\mathbf{j},k-1/2} \simeq \frac{\boldsymbol{\xi}_{\mathbf{j},k-1/2+1/2} - \boldsymbol{\xi}_{\mathbf{j},k-1/2-1/2}}{2(\Delta t/2)}$$

and

$$\left(\nabla_x^2[\mathbf{Y}^*]\right)_{\mathbf{j},k+1/2} \simeq \frac{\mathbf{Y}_{\mathbf{j}+\mathbf{e}_i,k+1/2} - 2\mathbf{Y}_{\mathbf{j},k+1/2} + \mathbf{Y}_{\mathbf{j}-\mathbf{e}_i,k+1/2}}{(\Delta x_i)^2},$$
$$\left(\nabla_x^2[\boldsymbol{\xi}^*]\right)_{\mathbf{j},k-1/2} \simeq \frac{\boldsymbol{\xi}_{\mathbf{j}+\mathbf{e}_i,k-1/2} - 2\boldsymbol{\xi}_{\mathbf{j},k-1/2} + \boldsymbol{\xi}_{\mathbf{j}-\mathbf{e}_i,k-1/2}}{(\Delta x_i)^2},$$

Consequently the forward and backward numerical schemes are given by

$$\begin{aligned}\mathbf{Y}_{\mathbf{j},k+1}^{(\ell)} &= \mathbf{Y}_{\mathbf{j},k}^{(\ell)} + \Delta t \mathbf{F}_{\mathbf{j},k+\frac{1}{2}}^{(\ell)},\ \text{for } k = 0{:}K{-}1,\\ \boldsymbol{\xi}_{\mathbf{j},k-1}^{(\ell)} &= \boldsymbol{\xi}_{\mathbf{j},k}^{(\ell)} + \Delta t \mathbf{G}_{\mathbf{j},k-\frac{1}{2}}^{(\ell)},\ \text{for } k = \text{-}1{:}1,\end{aligned}$$

respectively, for $\ell = 1{:}L$. For each double shot ℓ the state starts from $\mathbf{Y}_{\mathbf{j},0}^{(\ell)} = \mathbf{Y}_{0,\mathbf{j},0} = \mathbf{Y}_0(\mathbf{x_j},0)$ using

$$U_{3,\mathbf{j},k}^{(\ell)} = U_{0,3,\mathbf{j},k} + \xi_{3,\mathbf{j},k}^{(\ell-1)}/s,\ \text{for } k = 0{:}K{-}1,$$

except when $\ell = 1$ and the initial guess $U_{3,\mathbf{j},k}^{(1)}$, e.g., $U_{0,3,\mathbf{j},k}$, is used. For each updated forward state shot is completed, then the backward co-state shot starts from $\boldsymbol{\xi}_{\mathbf{j},K}^{(\ell)} = Q\mathbf{Y}_{\mathbf{j},K}^{(\ell)}$ using the whole state set $\mathbf{Y}_{\mathbf{j},k}^{(\ell)}$ for $k = 0:K$. The Crank-Nicolson temporal mid-point $\mathbf{F}_{\mathbf{j},k+\frac{1}{2}}^{(\ell)}$ state function is approximated by average with

$$\mathbf{Y}_{\mathbf{j},k+\frac{1}{2}}^{(\ell)} = \frac{1}{2}\left(\mathbf{Y}_{\mathbf{j},k+1}^{(\ell)} + \mathbf{Y}_{\mathbf{j},k}^{(\ell)}\right),$$

which can be used to construct finite differences for the derivatives, with a similar form for $\mathbf{U}_{\mathbf{j},k+\frac{1}{2}}^{(\ell)}$, for $\ell = 0{:}L$.

Similarly, the backward temporal mid-point $\mathbf{G}^{(\ell)}_{\mathbf{j},k-\frac{1}{2}}$ co-state function is approximated with the average,

$$\boldsymbol{\xi}^{(\ell)}_{\mathbf{j},k-\frac{1}{2}} = \frac{1}{2}\left(\boldsymbol{\xi}^{(\ell)}_{\mathbf{j},k} + \boldsymbol{\xi}^{(\ell)}_{\mathbf{j},k-1}\right),$$

and is used for the derivatives as well.

The no flux boundary conditions for both the state and co-state present some extra complexity, since the central differences of Crank-Nicolson are not suitable at the boundary if it is necessary to avoid using artificial external points. External points can be avoided by judicious use of forward and backward differences of second order, matching the accuracy of the Crank-Nicolson central differences. In the simplest case of rectangular grids, the discretized no flux boundary conditions (15,19) with second order accuracy are

$$\begin{aligned}
\mathbf{0} &= -\left((\widehat{\mathbf{N}}\cdot\nabla_x)[\mathbf{Y}^*]\right)^{(\ell)}_{\mathbf{j},k} \\
&\simeq -\frac{3\mathbf{Y}^{(\ell)}_{\mathbf{j},k} - 4\mathbf{Y}^{(\ell)}_{\mathbf{j}-\mathbf{N},k} + \mathbf{Y}^{(\ell)}_{\mathbf{j}-2\mathbf{N},k}}{2|\mathbf{N}\cdot\Delta\mathbf{x}|}, \\
\mathbf{0} &= \left((\widehat{\mathbf{N}}\cdot\nabla_x)[(\boldsymbol{\xi})^*]\right)^{(\ell)}_{\mathbf{j},k} \\
&\simeq \frac{3(\boldsymbol{\xi})^{(\ell)}_{\mathbf{j},k} - 4(\boldsymbol{\xi})^{(\ell)}_{\mathbf{j}-\mathbf{N},k} + (\boldsymbol{\xi})^{(\ell)}_{\mathbf{j}-2\mathbf{N},k}}{2|\mathbf{N}\cdot\Delta\mathbf{x}|},
\end{aligned}$$

respectively, where $\mathbf{N} = \widehat{\mathbf{N}}_{\mathbf{j},k}$, $\Delta\mathbf{x} = [\Delta x_i]_{3\times 1} > \mathbf{0}$, D is not needed, and, e.g., $\mathbf{Y}^{(\ell)}_{\mathbf{j}-\mathbf{N},k} = \mathbf{Y}^{(\ell)}(\mathbf{x}_\mathbf{j} - |\mathbf{N}\cdot\Delta\mathbf{x}|\mathbf{N}, t_k)$. For non-rectangular domains, interpolation would be needed to convert evaluations to defined spatial nodes or else domain compatible grids should be used, e.g., for circular or spherical grid boundaries, $\mathbf{N} = \mathbf{e}_r$, where here r is the radius and $\mathbf{N}\cdot\Delta\mathbf{x} = \Delta r$.

During each ℓth double shot, a prediction and corrections of the state and co-state are used to account for the usual nonlinearities in the biological models [6], stopping when the changes are sufficiently small.

The above mentioned method is a sequential double shot method since one shot is used to get $\mathbf{Y}^{(\ell)}_{\mathbf{j},k}$ and a subsequent shot it used to get $\boldsymbol{\xi}^{(\ell)}_{\mathbf{j},k}$. Alternately, a parallel two shot method could be used to get an approximate solution by integrating for both $\mathbf{Y}^{(\ell)}_{\mathbf{j},k}$ and $\boldsymbol{\xi}^{(\ell)}_{\mathbf{j},k}$ in the forward direction using a guess initial condition for $\boldsymbol{\xi}^{(\ell)}_{\mathbf{j},0}$ at $t_0 = 0$, with several genuine shooting method shots until some ℓ^* shot where $||\boldsymbol{\xi}^{(\ell^*)}_{\mathbf{j},K} + Q\mathbf{Y}^{(\ell^*)}_{\mathbf{j},K}|| < \mathrm{tol}_\xi$, i.e., the final co-state value is small enough using some sufficiently small tolerance tol_ξ to approximate the final condition (20). Interpolation should be used to accelerate the convergence of $\boldsymbol{\xi}^{(\ell)}_{\mathbf{j},0}$ to some sufficiently small value.

IX. Test Results

The double shot forward and backward algorithm outlined in the previous two sections has been tested on one space dimension, x, example with three state dimensions $\{Y_1^*{=}N_1^*, Y_2^*{=}N_2^*, Y_3^*{=}C^*\}$, plus the drug input control U_3^*. The numerical parameter data come from the BCNU drug simulations for the brain of Wang et al. [14] and the brain tumor modeling of Swanson [12] and Murray [11], with some difficult to find parameters from Mansuri [9] or from reasonable estimates from other areas. For example, diffusion diagonal vector is $\mathbf{D} = [$4.2e-3, 1.e-15, 2.2e-1$]$ cm^2 per day (normal tissue diffusion is assumed to be insignificant), the quadratic cost coefficients are $r = 0.1 = s = q_1 = q_3$, the net growth coefficients are $a = [$1.2e-2, 8.6e-7, 1.1e+1$]$ per day, the carrying capacities for tumor k_1 and normal k_2 tissues are scaled to one for the normal value and the interaction coefficients $\{\alpha_{1,2}, \alpha_{2,1}, \kappa_{1,3}, \kappa_{2,3}\}$ are all given the arbitrary value 1.0e-4. The initial states are given to be uniformly one for the normal tissue, while the tumor density was assumed to be a spatial Gaussian with spread 2.0e-2 about a mean of 0.0 with a weight of 1.e-3. The initial drug concentration has a Gaussian spread of 1.0e-2 about a mean 0.0 with weight 1.e-4, while the threshold drug control $u_0(x,t)$ is similarly distributed, but with weight 1.e-5.

The results are given only for the tumor density $N_1^* = N_1^*(x,t)$ in Figure 1 on the symmetric space interval $x \in [-1.0, +1.0]$ in centimeters over a $t_f = 5$ day treatment. For this simple one space dimension test example, we see that the optimal distribution of the tumor using an optimal distribution of the drug delivery results in the 16.2% reduction of the tumor density over this simulated five day drug treatment trial.

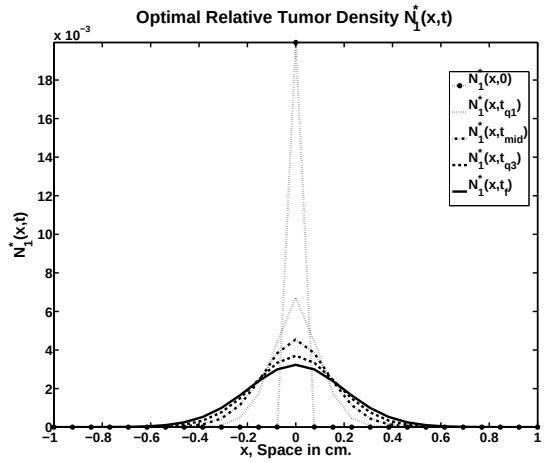

Fig. 1. Tumor density $N_1^*(x,t)$ versus the one-dimensional spatial coordinate x with time t at the rounded parameter values $\{0, t_{q1} = 0.25t_f, t_{mid} = 0.5t_f, t_{q3} = 0.75t_f, t_f\}$, where $t_f = 5$ days. The targeted tumor density rapidly decays in this simulated 5 day trial.

X. Conclusion

The main interest of this paper was to provide the necessary foundation to study the mechanism of drug delivery to the brain. We have set up a fairly realistic distributed parameters model which takes into account the spatial dependence of the state variables. The main focus of the paper was to develop an algorithm to determine the optimal

drug delivery to brain tumors using an optimal distribution of the drug about the original tumor site. This paper leaves room for a lot of new directions of work.

A. Future Directions

One such direction would be running a simulation to implement the algorithm, using supercomputing tools. Of course this would require more realistic medical data. Another important aspect that can be examined is the effect of stochasticity, most notably the Gaussian and Poisson type of noise. The physical basis for such stochasticity would be the phenomenon of metastasis, which gives rise to additional tumor sites and also the side effects produced by the drug. We hope to examine all these aspects of the problem in future works.

XI. Acknowledgments

The authors gratefully acknowledge, Dr. Herbert H. Engelhard of Neurological Surgery at the University of Illinois Medical Center for suggesting the problem and Professor William D. O'Neill of Bioengineering at the University of Illinois at Chicago for putting us in contact.

References

[1] H.H. Engelhard, "Brain Tumors and the Blood-Brain Barrier," *Neuro-Oncology, The Essentials*, pp. 49-53, Thieme Medical Publishers, Inc., New York 2000.

[2] R.A. Gatenby and E.T. Gawlinski, "A Reaction-Diffusion Model of Cancer Invasion," *Cancer Research*, vol. 56, pp. 5745-5753, December 15, 1996.

[3] M.D. Gunzburger, *Perspectives in Flow Control and Optimization*, SIAM Series in Design and Control, Philadelphia, 2003.

[4] R. Haberman, *Elementary Applied Partial Differential Equations with Fourier Series and Boundary Value Problems*, Prentice-Hall, Englewood Cliffs, NJ, 1983.

[5] F.B.Hanson and K.Naimipour, "Convergence of Numerical Method for Multistate Stochastic Dynamic Progamming," *Proceedings of International Federation of Automatic Control 12th World Congress*, vol. 9, pp.501-504, 1993.

[6] F.B.Hanson, "Techniques in Computational Stochastic Dynamic Programming," *Control and Dynamic Systems: Advances in Theory and Applications*, vol. 76, pp.103-162, Academic Press, New York, April 1996.

[7] H. B. Keller, *Numerical Methods for Two-Point Boundary Value Problems*, Gin Blaisdell, Waltham, MA, 1968.

[8] D. E. Kirk, *Optimal Control Theory: An Introduction*, Prentice-Hall, Englewood Cliffs, NJ, 1970.

[9] J. Mansuri, *The Modelling of Tumor Growth Using Reaction-Diffusion Equations*, MSc Thesis, Oxford University, 2002.

[10] J.D.Murray, *Mathematical Biology I: An Introduction*, Springer, New York, 2002.

[11] J.D.Murray, *Mathematical Biology II: Spatial Models and Biomedical Applications*, Springer, New York, 2003.

[12] K.R. Swanson, *Mathematical Modeling of the Growth and Control of Tumors*, PhD Thesis, University of Washington, 1999.

[13] C-H. Wang and J. Li, "Three dimensional simulation of IgG delivery to tumors," *Chemical Engineering Science*, vol. 53, pp. 3579-3600, 1998.

[14] C-H. Wang, J. Li, C.S. Teo and T. Lee, "The delivery of BCNU to brain tumors," *Journal of Controlled Release*, vol. 61, pp. 21-41, 1999.

[15] J. J. Westman, B. R. Fabijonas, D. L. Kern, and F. B. Hanson, "Compartmental Model for Cancer Evolution: Chemotherapy and Drug Resistance," *Mathematical BioSciences*, pp. 1-20, May 2002, under revision.

[16] D.E.Woodward, J.Cook, P.Tracqui, G.C.Cruywagen, J. D Murray, E.C.Alvord, Jr., "A mathematical model of glioma growth: the effect of extent of surgical resection," *Cell Proliferation*, vol. 29, pp. 269-288, 1996.