2005 American Control Conference
June 8-10, 2005. Portland, OR, USA

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

Siddhartha P. Chakrabarty† and Floyd B. 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. This gives rise to a coupled system of equations with a forward state equation and a backward co-state equation, which are solved using a modified double shot, forward-backward method.**

A numerical procedure based upon the Crank-Nicolson method is used to solve the coupled nonlinear system of six one-dimension partial differential equations, along with a quasi-linear approximation of the nonlinearites using extrapolator-predictor-corrector iteration techniques.

I. Introduction

The growth and control of brain tumors have been the subject of medical and scientific scrutiny for a very long time [15]. Simply speaking a tumor, like most cancerous cells originates from a single 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. 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 [16]. Despite improved diagnostic procedures such as computerized tomography (CT) scan and magnetic resonance imaging (MRI), their benefits have been restricted by the treatment options available. One major impediment to administering the drugs to the brain tumor site is the *blood brain barrier (BBB)* [4], which exists 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. The reader can consult the recent work by Araujo et al.[3], for more details on solid tumors. Websites provide very useful information on *Clinical Trials and Noteworthy Treatments for Brain Tumors* like BCNU [1] and Gliadel wafers [2]. Wang et al. [17], [18] have modeled drug delivery behavior to tumors in three dimensions for drugs like IgG and BCNU.

In this paper, the focus will be mostly on the distribution and optimal control of the drug about the original tumor site. First the mechanisms behind the growth of tumor cells as well as normal tissues and the drug concentration in the tumor are considered. Unlike a lot of other tumors, gliomas can be highly diffusive [15]. Gatenby et al. [5] and Mansuri [13] 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. [19] look at the various types of tumor growth, namely exponential, logistic and Gompertz. Murray's books [14], [15] are excellent references for the study of different types of growth mechanisms. Also, Woodward et al. [20] study a model of glioma growth and the effects of surgical resection. Here, 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* [6] for including constraints in the control problem. Finally a *double shot*, forward-backward iteration method is given to approximate the state, co-state and regular control. For implementing this method for the *PDEs*, a substantial modification is developed from the Crank-Nicolson predictor-corrector method developed by Hanson et al. [9] and Hanson [10] for stochastic dynamic programming of biological control applications. It is important to point out that the model presented is applicable to many cancer and non-cancer model applications.

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.

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, $n_2 = n_2(\mathbf{x}, t)$,

$$\frac{\partial n_1}{\partial t} = D_1 \nabla_x^2 [n_1] + a_1 n_1 g_1(n_1) - (\alpha_{1,2} n_2 + \kappa_{1,3} c)\, n_1, \qquad (1)$$

where the tumor diffusivity is D_1. Let the term $a_1 n_1 g_1(n_1)$ be the growth rate of the tumor cells, where $n_1 g_1(n_1)$ could

*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.

† 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 or hanson@math.uic.edu

be exponential, logistic or Gompertz growth, $g_1(n_1) = 1$, $(1-n_1/k_1)$ or $\ln(k_1/n_1)$, 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 $\kappa_{1,3}c$ be the death rate of tumor cells due to drug treatment, although it could be a nonlinear function.

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

$$\frac{\partial n_2}{\partial t} = D_2 \nabla_x^2[n_2] + a_2 n_2 g_2(n_2) - (\alpha_{2,1} n_1 + \kappa_{2,3} c)\, n_2, \tag{2}$$

where a_2 is the normal cell intrinsic growth rate and the normal cell growth function $n_2 g_2(n_2)$ is either exponential, logistic or Gompertz growth, $g_2(n_2) = 1$, $(1-n_2/k_2)$ or $\ln(k_2/n_2)$, respectively, where k_2 is the normal tissue carrying capacity. Note that the $\kappa_{2,3}c$ term indicates that some normal tissues could die as a result of the treatment.

C. Concentration

The drug exhibits a diffusive behavior and there is a reabsorption at the rate a_3. Let $u = u(\mathbf{x}, t)$ be the rate at which the drug is being injected and is the control variable in an optimal control system. The equation for drug concentration at position $\mathbf{x}$ and time t is,

$$\frac{\partial c}{\partial t} = D_3 \nabla_x^2[c] + a_3 c g_3(c) + u, \tag{3}$$

where $c g_3(c) = -c$ is the reabsorption function [18].

D. Global State Vector

Let the global state vector be

$$\mathbf{Y}(\mathbf{x}, t) = [Y_i(\mathbf{x}, t)]_{3\times 1} = [n_1(\mathbf{x}, t)\ n_2(\mathbf{x}, t)\ c(\mathbf{x}, t)]^\top, \tag{4}$$

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

E. Initial and Boundary Conditions

Let the initial conditions for the state be

$$\mathbf{Y}(\mathbf{x}, 0) \equiv \mathbf{Y}_0(\mathbf{x}), \tag{5}$$

for $\mathbf{x}$ in Ω. Murray [15] recommends using Gaussian distribution for the initial distributions of tumors. The *no flux* boundary conditions are

$$-D\left(\widehat{\mathbf{N}} \cdot \nabla_x\right)[\mathbf{Y}](\mathbf{x}, t) = \Big[-D_i(\widehat{\mathbf{N}} \cdot \nabla_x)[Y_i](\mathbf{x}, t)\Big]_{3\times 1} \tag{6}$$

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, and the diffusion matrix is diagonal, $D(\mathbf{x}) = [D_i \delta_{i,j}]_{3\times 1}$ and could be inhomogeneous depending on the brain matter [16], 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

The objective functional is taken to be a quadratic form of running and terminal costs,

$$J(u) = \frac{1}{2}\int_0^{t_f} dt \int_\Omega d\mathbf{x}\left(r_1 n_1^2(\mathbf{x}, t) + s_3\,(u - u_0)^2\,(\mathbf{x}, t)\right) + \int_\Omega d\mathbf{x}\left(q_1 n_1^2(\mathbf{x}, t_f) + q_3 c^2(\mathbf{x}, t_f)\right). \tag{7}$$

The goal is to minimize this functional with respect to the drug input rate $u(\mathbf{x}, t)$ relative to some threshold rate $u_0(\mathbf{x}, t)$ and the terminal costs at t_f, i.e., $\min_u [J(u)]$. Note that here $r_1 > 0$ is the tumor burden cost coefficient and $s_3 > 0$ is the drug delivery cost coefficient, while $q_1 > 0$ and $q_3 > 0$ are the corresponding final costs. We could have chosen a linear control which would have been less realistic, but would give rise to singular control complications. In addition no assumption is made about the control constraints, but $u_0(\mathbf{x}, t)$ serves as physical restriction on the amount and costs 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 is governed by a nonlinear PDE:

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

where

$$A(\mathbf{Y}) = [a_i g_i(Y_i)\delta_{i,j}]_{3\times 3}, \qquad \mathbf{U}(\mathbf{x}, t) = U_3(\mathbf{x}, t)\mathbf{e}_3, \tag{9}$$

$$B(\mathbf{Y}, t) = -(\alpha_{1,2} n_2 + \kappa_{1,3} c)\mathbf{e}_1\mathbf{e}_1^\top - (\alpha_{2,1} n_1 + \kappa_{2,3} c)\mathbf{e}_2\mathbf{e}_2^\top,$$

$\mathbf{e}_i$ is the ith unit vector and $U_3(\mathbf{x}, t) = u(\mathbf{x}, t)$.

B. Objective Functional

The quadratic objective in vector form is

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

where $R = r_1\mathbf{e}_1\mathbf{e}_1^\top$, $S = s_3\mathbf{e}_3\mathbf{e}_3^\top$, $Q = q_1\mathbf{e}_1\mathbf{e}_1^\top + q_3\mathbf{e}_3\mathbf{e}_3^\top$ and $\mathbf{U}_0 = u_0(\mathbf{x}, t)\mathbf{e}_3$.

V. Defining the Pseudo-Hamiltonian

There are three vector *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),

$$\boldsymbol{\xi}(\mathbf{x}, t) = [\xi_i]_{3\times 1}\ ,\ \boldsymbol{\eta}(\mathbf{x}, t) = [\eta_i]_{3\times 1}\ ,\ \boldsymbol{\chi}(\mathbf{x}) = [\chi_i]_{3\times 1}, \tag{11}$$

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

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

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 [6], [12] 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 terms linear in $\delta\mathbf{Z}$ using (12),

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

Before the critical conditions for first variation in (13) can be applied, the higher order derivatives in time and state of the extended state perturbations must be reduced 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 [7],

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

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

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

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:

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

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.,

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

and with initial conditions on the interior Ω from the coefficient of $(\delta\boldsymbol{\chi})^\top$, i.e.,

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

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. Regular Optimal Control

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

$$U_3^*(\mathbf{x},t) = u_0(\mathbf{x},t) + \xi_3^*(\mathbf{x},t)/s_3, \quad (\mathbf{x},t)\in\Omega\times[0,t_f], \tag{17}$$

provided $s_3 \neq 0$. Note that this control law only requires solving for the 3rd component of the first co-state vector $\boldsymbol{\xi}^*(\mathbf{x},t)$, since $\delta U_1 \equiv 0$ and $\delta U_2 \equiv 0$.

D. Co-State Equations

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

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

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

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

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 $\mathbf{\Gamma}=\partial\Omega$,

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

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. Double Shot, Forward-Backward Computational Iteration Method

It is necessary to solve the system consisting of 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 in time. The method is a *double shot*, forward-backward iteration method, since the model has two vector-valued PDEs and the method consists of one forward shot with (14) followed by one backward shot with (18). This *double shot* method is similar to the *multiple shooting method* of Hackbusch [8] used for solving parabolic equations with *opposite orientations* or to what Gunzberger [6] calls the *one-shot* method. See Gunzberger [6] also for a more rigorous justification with Sobolev spaces in the more general abstract case, but the model here is quite concrete.

At this point one must be cautioned that the numerical method suffers from the *Curse of Dimensionality* for PDEs in higher space dimensions as the number of nodes grows exponentially, which limits the size of the problem that can be numerically computed. However, Hanson [10] has used parallel and other supercomputer processors with related numerical procedures on many control problems for biological applications to reduce the effects of the curse of dimensionality.

In reality, the problem is highly nonlinear as are many problems in biology and we need numerical approximations of the solution as well as modifications of standard linear numerical methods. The main problem here is the fact that we have a forward state equation and a backward co-state equation along with two pairs of bilinear terms due to the intrinsic growth term $A(\mathbf{Y})\mathbf{Y}$ and interaction term $B(\mathbf{Y},t)\mathbf{Y}$, both amenable to quasi-linear approximations. The double shot method is a major modification of the shooting methods [11] 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 $U_3^{(\ell+1)}(\mathbf{x},t)$.

An initial guess for the first ($\ell=1$) forward-backward shot iteration is made for the control $U_3(\mathbf{x},t)=U_3^{(1)}(\mathbf{x},t)$ in (17), where $U_3^{(1)}(\mathbf{x},t)$ is taken as a Gaussian distribution after [15] with weight appropriate to the concentration level. Substituting it into the *state forward PDE* (14) solving for $\mathbf{Y}^{(1)}(\mathbf{x},t)$, using the initial condition $\mathbf{Y}_0(\mathbf{x})$ (16) and boundary condition (15). For each successive double shot for $\ell>1$, the starting control is given from a discrete version of (17).

Next, a predictor-corrector adaptation of the Crank-Nicolson implicit method in one space dimension or alternating directions implicit method in higher dimensions to this forward-backward problem is made in order to correct for a quasi-linear approximation of the nonlinear terms and preserve the tridiagonal properties of the implicit step. Central finite differences for all derivatives are used everywhere, except that appropriate second order accurate forward or backward differences are used at nodes adjacent to the boundary due to the no flux boundary conditions. Here our test results will be for the one-dimensional model.

The three-dimensional space is discretized as follows,

$$\mathbf{x}\to\mathbf{x}_\mathbf{j}=[x_{j_i,1}+(j_i-1)\cdot\Delta x_i]_{3\times1}.$$

Here Δx_i is the mesh size for dimension i and $\mathbf{j}=[j_i]_{3\times1}$ where, $j_i=1:M_i$ nodes per state 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 $t_K=t_f$. Next consider the vector state/co-state PDE system in the convenient general notation:

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

with appropriate vector functions $\mathbf{F}(\mathbf{x},t,\mathbf{y},\mathbf{u})$ for (8) and $\mathbf{G}(\mathbf{x},t,\boldsymbol{\xi},\mathbf{y})$ for (18). In this coupled set of vector equations, the state is discretized at the forward midpoint in time $\mathbf{Y}(\mathbf{x}_\mathbf{j},t_{k+0.5})\simeq\mathbf{Y}_{\mathbf{j},k+0.5}$ and the co-state at the backward midpoint $\boldsymbol{\xi}(\mathbf{x}_\mathbf{j},t_{k-0.5})\simeq\boldsymbol{\xi}_{\mathbf{j},k-0.5}$. The space-time partial derivatives of $\mathbf{Y}^*(\mathbf{x},t)$ and $\boldsymbol{\xi}^*(\mathbf{x},t)$ are discretized by the usual central finite differences of second order at these respective midpoints. Consequently, the forward and backward numerical schemes are given by

$$\begin{aligned}\mathbf{Y}_{\mathbf{j},k+1}^{(\gamma+1,\ell)} &= \mathbf{Y}_{\mathbf{j},k}^{(\ell)}+\Delta t\mathbf{F}_{\mathbf{j},k+0.5}^{(\gamma,\ell)},\\ \boldsymbol{\xi}_{\mathbf{j},k-1}^{(\gamma+1,\ell)} &= \boldsymbol{\xi}_{\mathbf{j},k}^{(\ell)}+\Delta t\mathbf{G}_{\mathbf{j},k-0.5}^{(\gamma,\ell)},\end{aligned}$$

respectively, for $\gamma = 0{:}nc$ corrections with each time step k until

$$\left|\left|Y^{(\gamma+1,\ell)}_{1,\mathbf{j},k+1} - Y^{(\gamma,\ell)}_{1,\mathbf{j},k+1}\right|\right| < \mathrm{tol}_y \left|\left|Y^{(\gamma,\ell)}_{1,\mathbf{j},k+1}\right|\right|$$

is satisfied for all $\mathbf{j}$, for $k = 0{:}K{-}1$, and for $\ell = 1{:}L$ double shots. The Crank-Nicolson temporal mid-point is approximated by average $\mathbf{Y}^{(\gamma,\ell)}_{\mathbf{j},k+0.5} = 0.5\left(\mathbf{Y}^{(\gamma,\ell)}_{\mathbf{j},k+1} + \mathbf{Y}^{(\ell)}_{\mathbf{j},k}\right)$, where $\mathbf{Y}^{(\ell)}_{\mathbf{j},k}$ is the final correction for each time step k given shot ℓ.

Similarly, the backward temporal mid-point is approximated by $\boldsymbol{\xi}^{(\gamma,\ell)}_{\mathbf{j},k-0.5} = 0.5\left(\boldsymbol{\xi}^{(\gamma,\ell)}_{\mathbf{j},k} + \boldsymbol{\xi}^{(\ell)}_{\mathbf{j},k-1}\right)$, where $\boldsymbol{\xi}^{(\ell)}_{\mathbf{j},k}$ is the final correction for each $k-1$ time step given shot ℓ. These averages can be used to construct finite differences for the derivatives, with a similar form for $\mathbf{U}^{(\ell)}_{\mathbf{j},k+0.5}$, for $\ell = 0{:}L$ and other terms. For each double shot ℓ beyond the first, the state starts from $\mathbf{Y}^{(\ell)}_{\mathbf{j},0} = \mathbf{Y}_{0,\mathbf{j},0} = \mathbf{Y}_0(\mathbf{x}_\mathbf{j}, 0)$ using the update

$$U^{(\ell)}_{3,\mathbf{j},k} = u_{0,\mathbf{j},k} + \xi^{(\ell-1)}_{3,\mathbf{j},k}/s_3, \quad \text{for } k = 0{:}K{-}1,$$

except when $\ell = 1$ and the initial guess $U^{(1)}_{3,\mathbf{j},k} = u_{0,\mathbf{j},k}$, is used. For each updated forward state shot is completed, then the backward co-state shot starts from $\boldsymbol{\xi}^{(\ell)}_{\mathbf{j},K} = Q\mathbf{Y}^{(\ell)}_{\mathbf{j},K}$ using the whole state set $\mathbf{Y}^{(\ell)}_{\mathbf{j},k}$ for $k = 0{:}K$.

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,

$$\left|\left|U^{(\ell)}_3(\mathbf{x},t) - U^{(\ell-1)}_3(\mathbf{x},t)\right|\right| < \mathrm{tol}_u \left|\left|U^{(\ell-1)}_3(\mathbf{x},t)\right|\right|,$$

and

$$\left|\left|Y^{(\ell)}_1(\mathbf{x},t) - Y^{(\ell-1)}_1(\mathbf{x},t)\right|\right| < \mathrm{tol}_y \left|\left|Y^{(\ell-1)}_1(\mathbf{x},t)\right|\right|,$$

where the norm is over all $(\mathbf{x}, t)$, for $\ell = 2{:}L$ until satisfied, provided $||U^{(\ell-1)}_3(\mathbf{x},t)|| \neq 0$ and $||Y^{(\ell-1)}_1(\mathbf{x},t)|| \neq 0$, where $\mathrm{tol}_u > 0$ and $\mathrm{tol}_y > 0$ are some prescribed tolerances.

The treatment of the nonlinear terms is to make their approximation compatible with the linear properties of the Crank-Nicolson implicit method, so using an exptrapolator-predictor-corrector technique superimposed with

$$A\left(\mathbf{Y}^{(\ell)}_{\mathbf{j},k+0.5}\right)\mathbf{Y}^{(\ell)}_{\mathbf{j},k+0.5} \simeq A\left(\mathbf{Y}^{(\gamma,\ell)}_{\mathbf{j},k+0.5}\right)\mathbf{Y}^{(\gamma+1,\ell)}_{\mathbf{j},k+0.5}$$
$$B\left(\mathbf{Y}^{(\ell)}_{\mathbf{j},k+0.5}\right)\mathbf{Y}^{(\ell)}_{\mathbf{j},k+0.5} \simeq B\left(\mathbf{Y}^{(\gamma,\ell)}_{\mathbf{j},k+0.5}\right)\mathbf{Y}^{(\gamma+1,\ell)}_{\mathbf{j},k+0.5}$$

for $\gamma = 0{:}nc$ corrections in the case of the state equation where $\gamma = 0$ is the initial value prediction when $k = 0$ but is the extrapolated value (e.g., $\mathbf{Y}^{(\gamma,\ell)}_{\mathbf{j},k+0.5} \simeq 0.5(3\mathbf{Y}^{(\ell)}_{\mathbf{j},k} - \mathbf{Y}^{(\ell)}_{\mathbf{j},k-1})$) from the two prior times (k and $k-1$) when $k > 0$. The number of corrections for each k is made small by selecting Δt sufficiently small to satisfy accurate relative stopping criterion similar to that for the control and state given in the last section. Note that these approximations of the nonlinear functions are explicitly linear in the new value $\mathbf{Y}^{(\gamma+1,\ell)}_{\mathbf{j},k+0.5}$, so the very efficient Thomas tridiagonal elimination algorithm can be used with Crank-Nicolson in the one-dimensional case. Similar quasi-linear approximations with prediction and correction are used for the backward co-state equations.

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

$$\mathbf{0} \simeq -(3\mathbf{Y}^{(\ell)}_{\mathbf{j},k} - 4\mathbf{Y}^{(\ell)}_{\mathbf{j}-\mathbf{N},k} + \mathbf{Y}^{(\ell)}_{\mathbf{j}-2\mathbf{N},k})/(2|\mathbf{N}{\cdot}\Delta\mathbf{x}|),$$
$$\mathbf{0} \simeq (3\boldsymbol{\xi}^{(\ell)}_{\mathbf{j},k} - 4\boldsymbol{\xi}^{(\ell)}_{\mathbf{j}-\mathbf{N},k} + \boldsymbol{\xi}^{(\ell)}_{\mathbf{j}-2\mathbf{N},k})/(2|\mathbf{N}{\cdot}\Delta\mathbf{x}|),$$

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, an extrapolation or prediction and corrections of the state and co-state are used to account for the usual nonlinearities in the biological models [10], stopping when the changes are sufficiently small. The overall 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 $\mathrm{R}\boldsymbol{\xi}^{(\ell)}_{\mathbf{j},k}$ 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).

VIII. 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. [18] and the brain tumor modeling of Swanson [16] and Murray [15], with some difficult to find parameters from Mansuri [13] or from reasonable estimates from other areas. For example, diffusion diagonal vector is $\mathbf{D}$ =[4.2e-3,1.e-15,0.22] cm^2 per day (normal tissue diffusion is assumed to be insignificant), the quadratic cost coefficients are $r_1 = 0.1 = q_1 = q_3$ and $s_3 = 0.2$, the net growth coefficients are a =[1.2e-2,8.6e-7,11.3] 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 $\{\alpha_{1,2}, \alpha_{2,1}, \kappa_{2,3}\}$ are all given the arbitrary value 1.0e-4, but $\kappa_{1,3} = 0.5$. 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 0.02 about a mean of 0.0 with a weight of 1.0e-3. The initial drug concentration has a Gaussian spread of 0.02 about a mean 0.0 with weight

0.15, while the threshold drug control $u_0(x,t)$ is similarly distributed, but with weight 1.0.

The results (obtained using MATLAB™) are plotted only for the tumor density $N_1^* = N_1^*(x,t)$ in Figure 1 on the symmetric 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 29.4% reduction of the total tumor density integral over this simulated five day drug treatment trial. The running time on a 2GHz processor was 168 seconds.

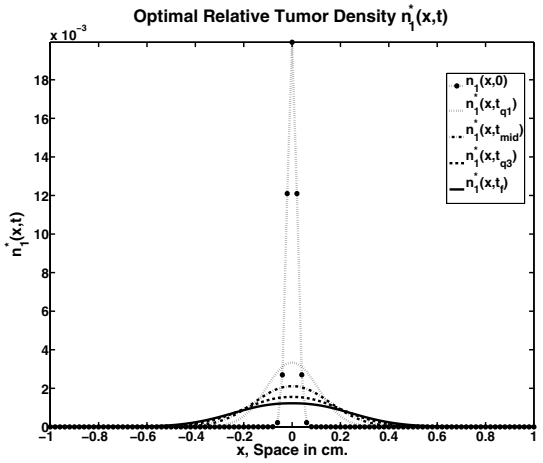

Fig. 1. Tumor density $N_1^*(x,t)$ versus the one-dimensional spatial coordinate x with time t at the rounded quartile values $\{0, t_{q1}{=}0.25t_f, t_{mid}{=}0.5t_f, t_{q3}{=}0.75t_f, t_f\}$, where $t_f = 5$ days. The targeted tumor density rapidly decays in this simulated 5 day trial.

IX. 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. Here, a one-dimensional test case is used. This paper leaves room for many new directions for this extremely complex problem.

A. Future Directions

One such direction would be running a simulation for more than one space dimension to implement the algorithm, perhaps using supercomputing tools. Of course this would require more realistic medical data. Also, the effects of using symmetric initial data needs to be examined by considering non-symmetric initial distributions. The effects of brain geometry and the diversity of the brain structure such as fluid cavities, vascular systems and brain matter needs to be explored. 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. A study of the increase of memory requirements and execution time if the dimension of state-control space increases is very important when using this model with actual real life data. We hope to examine all these aspects of the problem in future works.

X. 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. The authors are grateful to the three reviewers for their comments.

References

[1] D. Acquafredda, "Brain Tumor Treatments: BCNU (Carmustine)," Musella Foundation, `http://virtualtrials.com/bcnu2.cfm`

[2] Anonymous, "Brain Tumor Treatments: Gliadel® Wafer," Musella Foundation, 2004, `http://www.virtualtrials.com/gliadel`

[3] R. P. Araujo and D. L. S. McElwain "A history of the study of solid tumour growth: The contribution of mathematical modelling, " *Bull. Math. Biol.*, vol. 66, pp. 1039-1091,2004.

[4] H.H. Engelhard, "Brain tumors and the blood-brain barrier," *Neuro-Oncology, The Essentials*, pp. 49-53, Thieme Medical Publishers, Inc., New York 2000.

[5] R.A. Gatenby and E.T. Gawlinski, "A reaction-diffusion model of cancer invasion," *Cancer Research*, vol. 56, pp. 5745-5753, December 15, 1996.

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

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

[8] W. Hackbusch, "A numerical method for solving parabolic equations with opposite orientations," *Computing*, vol. 20, pp. 229-240, 1978.

[9] F.B.Hanson and K.Naimipour, "Convergence of numerical method for multistate stochastic dynamic progamming," *Proc. Int. Fed. of Auto. Control 12th World Congress*, vol. 9, pp.501-504, 1993.

[10] 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.

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

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

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

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

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

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

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

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

[19] J. J. Westman, B. R. Fabijonas, D. L. Kern and F. B. Hanson, "Probabilistic rate compartment cancer model: Alternate versus traditional chemotherapy dcheduling," in *Stochastic Theory and Control, Lecture Notes in Control and Information Sciences*, vol. 280, B. Pasik-Duncan (Editor), Springer-Verlag, New York, pp. 491-506, 2002.

[20] 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.