Cancer Drug Delivery in Three Dimensions For a Distributed Parameter Control Model Using Finite Elements*

Siddhartha P. Chakrabarty† and Floyd B. Hanson†

***Abstract*— The Galerkin finite element method is used to develop procedures for the optimal drug delivery to brain tumors. The mathematical model comprises of a system of three coupled reaction diffusion models, involving the density of tumor cells and normal tissue as also the drug concentration. An optimal control problem is formulated with the goal of minimizing the tumor cell density and reducing the side effects of the drug. A distributed parameter method based on the application of variational calculus is used on a pseudo-Hamiltonian, which is then used to obtain a coupled system of forward state equations and backward co-state equations. The Galerkin finite element method is used to realistically represent the brain structure. Finally a three dimensional test case is considered and partitioned into a set of brick finite elements in spherical coordinates, with tri-linear basis functions. Non-uniqueness of nodes in spherical coordinates is removed by combining like nodes, such as at the origin, at the poles and at the polar angle discontinuity. The Galerkin ODEs are solved by a combination of Crank-Nicolson and predictor-corrector methods.**

I. Introduction

Varius kinds of cancerous growth have been studied not only from the medical perspective, but also from the mathematical point of view. One such kind is the brain tumor, which like most cancerous cells originates from a single cell and proliferates into the neighboring normal cells. Understanding the mechanism behind the growth of tumors is necessary for designing an optimal treatment. *Gliomas*, a very deadly form of brain tumors, account for a majority of the cases [10], [12]. Despite the advanced disgnostic procedures available, their benefits have been limited due to various impedimets like the existence of the *blood brain barrier (BBB)* [3]. The commonly used forms of drug delivery are drugs congugated with polymer and delivery by optimal distribution of drugs about the original tumor site. Wang et al. [13], [14] have worked extensively on drug delivery to tumors in three dimensions. In this paper we focus on post-operative treatment of the resection that occurs after the bulk of the tumor has been removed.

The tumor-drug model here has been taken from Chakrabarty and Hanson [1], [2], which in turn was influenced by the works of Gatenby et al. [4] and Mansuri [8]. Murray's fine text [9] and the paper by Westman et al. [15] are good sources of different types of growth mechanishms.

The paper is organized in the following manner. In section II, we have the mathematical model. In section III, the optimal control problem is formulated, a *pseudo-Hamiltonian* is defined and then calculus of variations [5] is applied to it to obtain a forward state equation and a backward co-state equation. In the next two sections we present the Galerkin finite element method and finite element test configuration using the spherical co-ordinates. Finally we present the computational results of numerical implementation of this finite element formulation.

II. Mathematical Model

In the *distributed parameter control* model of Chakrabarty and Hanson [1], [2], the tumor cell and normal cell density and the drug concentration at any position vector $\mathbf{x}$ and time $t \in [0, t_f]$, in the *interior* Ω of the domain, denoted by $n_1(\mathbf{x}, t)$, $n_2(\mathbf{x}, t)$ and $c(\mathbf{x}, t)$ respectively, are taken as the state variables. Defining the global state vector as

$$\mathbf{Y}(\mathbf{x}, t) \equiv \begin{bmatrix} n_1(\mathbf{x}, t) & n_2(\mathbf{x}, t) & c(\mathbf{x}, t) \end{bmatrix}^\top, \tag{1}$$

the governing nonlinear vector PDE is given by

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

where the nonlinearities are given by

$$A(\mathbf{Y}) = a_1(1 - Y_1/k_1)\mathbf{e}_1\mathbf{e}_1^\top + a_2(1 - Y_2/k_2)\mathbf{e}_2\mathbf{e}_2^\top - a_3\mathbf{e}_3\mathbf{e}_3^\top,$$
$$B(\mathbf{Y}) = -(\alpha_{1,2}Y_2 + \kappa_{1,3}Y_3)\mathbf{e}_1\mathbf{e}_1^\top - (\alpha_{2,1}Y_1 + \kappa_{2,3}Y_3)\mathbf{e}_2\mathbf{e}_2^\top,$$

$$\mathbf{U}(\mathbf{x}, t) = U_3(\mathbf{x}, t)\mathbf{e}_3. \tag{3}$$

Here, $D_i > 0$ is the ith component of the diagonal diffusion coefficient D (could be inhomogeneous depending on the brain matter [12]), $A_{i,i}(\mathbf{Y})Y_i$ is the ith growth rate (logistic for $i = 1:2$ and exponentially decaying for $i = 3$ for our test case) $\alpha_{i,j}$ are death rates due to competition, $\kappa_{i,j}$ are the death rates due to treatment and $u = U_3(\mathbf{x}, t)$ is the rate at which the drug is being delivered and will be the control variable in an optimal control system. Also, $\mathbf{e}_i$ is the ith unit vector. The initial conditions and the no-flux boundary conditions on the boundary $\partial\Omega$ are, respectively,

$$\mathbf{Y}(\mathbf{x}, 0) = \mathbf{Y}_0(\mathbf{x}), \tag{4}$$

$$-D(\widehat{\mathbf{N}} \cdot \nabla_x)[\mathbf{Y}](\mathbf{x}, t) = \mathbf{0}. \tag{5}$$

*The work is supported by the National Science Foundation under Grant DMS-02-07081. The content of this material is that of the authors and does not necessarily reflect the views of the National Science Foundation.. The second 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 the *Proceedings of the Conference in Decision and Control 2006*.

†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: hanson@math.uic.edu and pratim@math.uic.edu

III. The Optimal Control Problem

The quadratic objective functional for running and terminal costs is given by,

$$J[\mathbf{Y},\mathbf{U}] = \frac{1}{2}\int_0^{t_f} dt \int_\Omega d\mathbf{x}\left(\mathbf{Y}^\top R\mathbf{Y} + (\mathbf{U}-\mathbf{U}_0)^\top S\,(\mathbf{U}-\mathbf{U}_0)\right) + \frac{1}{2}\int_\Omega d\mathbf{x}\left(\mathbf{Y}^\top Q\mathbf{Y}\right)(\mathbf{x},t_f), \tag{6}$$

where coefficients $R = r_1\mathbf{e}_1\mathbf{e}_1^\top$, $S = s_3\mathbf{e}_3\mathbf{e}_3^\top$ and $Q = q_1\mathbf{e}_1\mathbf{e}_1^\top + q_3\mathbf{e}_3\mathbf{e}_3^\top$, while $\mathbf{U}_0 = U_{0,3}(\mathbf{x},t)\mathbf{e}_3$. The goal is to minimize this functional with respect to the drug input rate relative to some threshold rate $U_{0,3}$ and the terminal costs at t_f, i.e., $\min_u[J(u)]$ also reduces the effects of toxicity. 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. In addition, no assumption is made about the control constraints like physical restriction on the amount of drugs that can be administered.

Using three *Lagrange multiplier* vectors, two of which are functions of space and time and one is independent of time, and letting $\mathbf{Z} = (\mathbf{Y},\mathbf{U},\boldsymbol{\xi},\boldsymbol{\eta},\boldsymbol{\chi})$ be an extended state vector, we define the *pseudo-Hamiltonian* as,

$$\begin{aligned}\mathcal{H}(\mathbf{Z}) \equiv\ & \frac{1}{2}\int_0^{t_f} dt \int_\Omega d\mathbf{x}\left(\mathbf{Y}^\top R\mathbf{Y} + (\mathbf{U}-\mathbf{U}_0)^\top S\,(\mathbf{U}-\mathbf{U}_0)\right)\\ &+\frac{1}{2}\int_\Omega d\mathbf{x}\left(\mathbf{Y}^\top Q\mathbf{Y}\right)(\mathbf{x},t_f)\\ &+\int_0^{t_f} dt\int_\Omega d\mathbf{x}\ \boldsymbol{\xi}^\top\left(\mathbf{Y}_t - D\nabla_x^2[\mathbf{Y}]\right.\\ &\left.-(A+B)(\mathbf{Y})\mathbf{Y}-\mathbf{U}\right)\\ &+\int_0^{t_f} dt\int_{\partial\Omega} d\boldsymbol{\Gamma}\ \boldsymbol{\eta}^\top\left(-D\left(\widehat{\mathbf{N}}\cdot\nabla_x\right)[\mathbf{Y}]\right)\\ &+\int_\Omega d\mathbf{x}\left(\boldsymbol{\chi}^\top(\mathbf{Y}-\mathbf{Y}_0)\right)(\mathbf{x},0).\end{aligned} \tag{7}$$

The *calculus of variations* is used to determine the critical point necessary condition for the first variation [5] of the *pseudo-Hamiltonian* $\mathcal{H}(\mathbf{Z})$. Let the perturbation $\delta\mathbf{Z}$ about the optimal trajectory $\mathbf{Z}^*$, be defined as $\delta\mathbf{Z} = \mathbf{Z} - \mathbf{Z}^*$. The pseudo-Hamiltonian is expanded as follows,

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

The quadratic order terms, including the 2nd variation of $\mathcal{H}$ are neglected. In addition 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, (using Green's formula [6]). Merging these identities, rearranging inner products and collecting terms, the extended state equations yields an intermediate form (for details see [1] for one space dimension and [2] for two space dimensions formulations). The coefficients of independent variations in this intermediate form are set to zero to obtain the state, control and the co-state equations.

A. State Equations

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

$$\mathbf{Y}_t^* = D\nabla_x^2[\mathbf{Y}^*] + (A+B)(\mathbf{Y}^*)\mathbf{Y}^* + \mathbf{U}^* \tag{8}$$

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.,

$$-D(\widehat{\mathbf{N}}\cdot\nabla_x)[\mathbf{Y}^*](\mathbf{x},t) = \mathbf{0}, \tag{9}$$

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.,

$$\mathbf{Y}^*(\mathbf{x},0) = \mathbf{Y}_0(\mathbf{x}) \tag{10}$$

for $\mathbf{x}\in\Omega$. Due to the presence of the functions $A(\mathbf{Y})\mathbf{Y}$ and $B(\mathbf{Y})\mathbf{Y}$ the forward PDE (8) will be nonlinear.

B. Regular Optimal Control

Since the control has been defined in (3) with only one component, only the coefficient of δU_3 is set to zero giving the corresponding drug regular control

$$U_3^*(\mathbf{x},t) = U_{0,3}(\mathbf{x},t) + \xi_3^*(\mathbf{x},t)/s_3, \tag{11}$$

on $\Omega\times[0,t_f]$,provided $s_3\neq 0$. Note that this control law only requires solving for the 3rd component of the 1st co-state vector $\boldsymbol{\xi}^*(\mathbf{x},t)$, since $\delta U_1\equiv 0$ and $\delta U_2\equiv 0$.

C. Co-State Equations

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

$$\begin{aligned}\mathbf{0} = \boldsymbol{\xi}_t^* + \nabla_x^2[D\boldsymbol{\xi}^*] + (A+B)(\mathbf{Y}^*)\boldsymbol{\xi}^*\\ +\nabla_Y[A+B](\mathbf{Y}^*):(\boldsymbol{\xi}^*(\mathbf{Y}^*)^\top) - R\mathbf{Y}^*,\end{aligned} \tag{12}$$

for $(\mathbf{x},t)\in\Omega\times[0,t_f)$. This PDE (12) is *unidirectionally coupled* to the state PDE (8), 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 (11). 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

$$(\widehat{\mathbf{N}}\cdot\nabla_x)[D\boldsymbol{\xi}^*](\mathbf{x},t) = \mathbf{0},\quad (\mathbf{x},t)\in\partial\Omega\times[0,t_f) \tag{13}$$

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,

$$\boldsymbol{\xi}^*(\mathbf{x},t_f) = -Q\mathbf{Y}^*(\mathbf{x},t_f). \tag{14}$$

The two other co-state vectors should not be needed, but satisfy rather simple equations [1], [2].

IV. Galerkin Finite Element Method

In [1], [2] the *Crank-Nicolson implicit method combined with predictor-corrector methods* are used to study the problem numerically. However, using finite difference methods like *Crank-Nicolson implicit method* alone and *alternating directions implicit method* have serious drawbacks. Finite difference techniques are more likely to have higher computational requirements, *i.e,* they suffer from the *curse of dimensionality*. Finite element methods require a relatively

smaller number of nodes as compared to the finite difference methods while maintaining the same level of accuracy. Also, the finite element method can better handle irregular structure, such as a brain tumor or ventricles. Hanson [7] has worked extensively in this area and has made a comparative study of different numerical methods for stochastic dynamic programming. For the problem under consideration, we use the *Galerkin finite element method* so as to reduce the number of state nodes. The following steps can be used to get an approximate numerical solution. Note that the assumption made in this paper is that the growth is *logistic* for the tumor and normal cells.

1) The first *double shot* step ($\delta = 1$) of forward-backward iteration shots is to guess the control $U_3^*(\mathbf{x},t) \simeq U_3^{(1)}(\mathbf{x},t)$, substitute it into the forward state equations and use the finite element method to solve for the state $\mathbf{Y}^*(\mathbf{x},t) \simeq \mathbf{Y}^{(1)}(\mathbf{x},t)$ for $t > 0$. Initially, $\mathbf{Y}^*(\mathbf{x},0) = \mathbf{Y}_0(\mathbf{x})$. Let the Galerkin approximation for the state vector be

$$\mathbf{Y}^*(\mathbf{x},t) \simeq \widehat{\mathbf{Y}}(\mathbf{x},t) \equiv \sum_{\hat{k}=1}^{\widehat{M}} \widehat{\mathbf{Y}}_{\hat{k}}(t) \cdot \phi_{\hat{k}}(\mathbf{x}), \qquad (15)$$

along with a similar approximation for the optimal control,

$$\mathbf{U}^*(\mathbf{x},t) \simeq \widehat{\mathbf{U}}(\mathbf{x},t) \equiv \sum_{\hat{k}=1}^{\widehat{M}} \widehat{\mathbf{U}}_{\hat{k}}(t) \cdot \phi_{\hat{k}}(\mathbf{x}), \qquad (16)$$

where, $[\phi_i(\mathbf{x})]_{\widehat{M}\times 1}$, is a set of $\widehat{M}$ *linearly independent* continuous basis functions, with the interpolation condition $\phi_{\hat{k}}(\mathbf{x}_{\hat{j}}) = \delta_{\hat{j},\hat{k}}$, at the element node $\mathbf{x}_{\hat{j}}$, implying that $\mathbf{Y}^*(\mathbf{x}_{\hat{j}},t) = \widehat{\mathbf{Y}}_{\hat{j}}(t)$ for $\hat{j} = 1{:}\widehat{M}$ finite element nodes.

2) Before applying the Galerkin approximation (15) to the state equation (8), the equation must be put into integral form on Ω with respect to a test function $\phi_{\hat{j}}(\mathbf{x})$ taken from the basis and then further prepared for low order basis function by reducing the 2nd order derivatives to 1st order derivatives by integration by parts (Green's formula [6]), so

$$\begin{aligned}
\mathbf{0} &= \int_\Omega d\mathbf{x}\phi_{\hat{j}}(\mathbf{x}) \left(\mathbf{Y}_t^* - D\nabla_x^2[\mathbf{Y}^*]\right. \\
&\quad \left. -(A+B)(\mathbf{Y}^*)\mathbf{Y}^* - \mathbf{U}^*\right) \\
&= \int_\Omega d\mathbf{x} \left(\phi_{\hat{j}}\mathbf{Y}_t^* + D\nabla_x^\top[\phi_{\hat{j}}]\nabla_x[\mathbf{Y}^*]\right. \\
&\quad \left. -\phi_{\hat{j}}\left((A+B)(\mathbf{Y}^*)\mathbf{Y}^* + \mathbf{U}^*\right)\right) \\
&\quad -\int_{\partial\Omega} d\mathbf{\Gamma}\phi_{\hat{j}} D\Big(\widehat{\mathbf{N}}\cdot\nabla_x\Big)\,[\mathbf{Y}^*] \\
&= \int_\Omega d\mathbf{x} \left(\phi_{\hat{j}}\mathbf{Y}_t^* + D\nabla_x^\top[\phi_{\hat{j}}]\nabla_x[\mathbf{Y}^*]\right. \\
&\quad \left. -\phi_{\hat{j}}\left((A+B)(\mathbf{Y}^*)\mathbf{Y}^* + \mathbf{U}^*\right)\right),
\end{aligned}$$

for $\hat{j} = 1:\widehat{M}$, where the exact no-flux boundary condition has been used.

3) Now, the Galerkin approximation (15) can be applied

$$\begin{aligned}
\mathbf{0} \simeq \sum_{\hat{k}=1}^{\widehat{M}} \int_\Omega d\mathbf{x} &\left(\widehat{\mathbf{Y}}'_{\hat{k}}\phi_{\hat{j}}\phi_{\hat{k}} + D\widehat{\mathbf{Y}}_{\hat{k}}\left(\nabla_x^\top[\phi_{\hat{j}}]\nabla_x[\phi_{\hat{k}}]\right)\right. \\
&\left. -\Big((A+B)\left(\widehat{\mathbf{Y}}\right)\widehat{\mathbf{Y}}_{\hat{k}} + \widehat{\mathbf{U}}_{\hat{k}}\Big)\phi_{\hat{j}}\phi_{\hat{k}}\right),
\end{aligned}$$

for $\hat{j} = 1:\widehat{M}$. Futher reduction to finite element integrals is accomplished by letting

$$\widehat{\mathcal{M}}_{\hat{j},\hat{k}} \equiv \int_\Omega d\mathbf{x}\phi_{\hat{j}}(\mathbf{x})\phi_{\hat{k}}(\mathbf{x}) \qquad (17)$$

be an element mass integral for $\hat{j},\hat{k} = 1{:}\widehat{M}$,

$$\widehat{\mathcal{K}}_{\hat{j},\hat{k}} \equiv \int_\Omega d\mathbf{x}\nabla_x^\top[\phi_{\hat{j}}]\nabla_x[\phi_{\hat{k}}] \qquad (18)$$

be an element stiffness integral for $\hat{j},\hat{k} = 1{:}\widehat{M}$, and

$$\widehat{\mathcal{T}}_{\hat{j},\hat{k},\hat{l}} \equiv \int_\Omega d\mathbf{x}\phi_{\hat{j}}(\mathbf{x})\phi_{\hat{k}}(\mathbf{x})\phi_{\hat{l}}(\mathbf{x}) \qquad (19)$$

be a triple basis element integral for $\hat{j},\hat{k},\hat{l} = 1:\widehat{M}$ arising from the purely bilinear terms in $A(\mathbf{Y})\mathbf{Y}$ and $B(\mathbf{Y})\mathbf{Y}$. Thus, the Galerkin equation becomes,

$$\begin{aligned}
\mathbf{0} = \sum_{\hat{k}=1}^{\widehat{M}} &\Big(\widehat{\mathcal{M}}_{\hat{j},\hat{k}}\Big(\widehat{\mathbf{Y}}'_{\hat{k}}(t) - \Big(a_1\mathbf{e}_1\mathbf{e}_1^\top + a_2\mathbf{e}_2\mathbf{e}_2^\top \\
&- a_3\mathbf{e}_3\mathbf{e}_3^\top\Big)\widehat{\mathbf{Y}}_{\hat{k}}(t) - \widehat{\mathbf{U}}_{\hat{k}}(t)\Big) + D\widehat{\mathcal{K}}_{\hat{j},\hat{k}}\widehat{\mathbf{Y}}_{\hat{k}}(t) \\
&+ \sum_{\hat{l}=1}^{\widehat{M}} \widehat{\mathcal{T}}_{\hat{j},\hat{k},\hat{l}}\Big(\frac{a_1}{k_1}\widehat{Y}_{1,\hat{k}}(t)\widehat{Y}_{1,\hat{l}}(t)\mathbf{e}_1 \\
&+ \frac{a_2}{k_2}\widehat{Y}_{2,\hat{k}}(t)\widehat{Y}_{2,\hat{l}}(t)\mathbf{e}_2 \\
&+ \Big(\alpha_{1,2}\widehat{Y}_{2,\hat{l}}(t) + \kappa_{1,3}\widehat{Y}_{3,\hat{l}}(t)\Big)\widehat{Y}_{1,\hat{k}}(t)\mathbf{e}_1 \\
&+ \Big(\alpha_{2,1}\widehat{Y}_{1,\hat{l}}(t) + \kappa_{2,3}\widehat{Y}_{3,\hat{l}}(t)\Big)\widehat{Y}_{2,\hat{k}}(t)\mathbf{e}_2\Big)\Big),
\end{aligned} \qquad (20)$$

for $\hat{j} = 1:\widehat{M}$. This Galerkin ODE can be solved by approximating the Galerkin basis integral coefficients $(\widehat{\mathcal{M}}_{\hat{j},\hat{k}}, \widehat{\mathcal{K}}_{\hat{j},\hat{k}}, \widehat{\mathcal{T}}_{\hat{j},\hat{k},\hat{l}})$ using symbolic methods and two-point Gaussian quadrature in case of singularities, and then the ODE system is solved by the combined Crank-Nicolson and predictor-corrector methods. The coefficients can be computed for all double shots for fixed finite elements off-line since they will be fixed. These coefficients can be calculated on an element-by-element decomposition and element results can later be reassembled to form the global solution [11].

4) In the second, backward shot of the double shot algorithm [1], the final condition (14),

$$\begin{aligned}
\boldsymbol{\xi}^{(\delta)}(\mathbf{x},t_f) &\simeq -Q\widehat{\mathbf{Y}}^{(\delta)}(\mathbf{x},t_f) \\
&= -Q\sum_{\hat{k}=1}^{\widehat{M}} \widehat{\mathbf{Y}}_{\hat{k}}^{(\delta)}(t_f)\phi_{\hat{k}}(\mathbf{x}),
\end{aligned}$$

for $\delta = 1{:}L$ double shots, is used to start the backward co-state solution. Similar to the state equation, a Galerkin approximation for the co-state equation (after

dropping the (δ) subscript) using the same basis is given by,

$$\boldsymbol{\xi}^*(\mathbf{x},t) \simeq \widehat{\boldsymbol{\xi}}(\mathbf{x},t) \equiv \sum_{\hat{k}=1}^{\widehat{M}} \widehat{\boldsymbol{\xi}}_{\hat{k}}(t)\cdot\phi_{\hat{k}}(\mathbf{x}) \tag{21}$$

for $t < t_f$. As with the state Galerkin variational formulation, the variation formulation for the co-state equation (12) is

$$\begin{aligned}\mathbf{0} &= \int_\Omega d\mathbf{x}\phi_{\hat{j}}(\mathbf{x})\left(\boldsymbol{\xi}^*_t + \nabla_x^2[D\boldsymbol{\xi}^*] + (A+B)(\mathbf{Y}^*)\boldsymbol{\xi}^*\right.\\ &\quad \left. + \nabla_Y[A+B](\mathbf{Y}^*):(\boldsymbol{\xi}^*(\mathbf{Y}^*)^\top) - R\mathbf{Y}^*\right)\\ &= \int_\Omega d\mathbf{x}\left(\phi_{\hat{j}}\left(\boldsymbol{\xi}^*_t + (A+B)(\mathbf{Y}^*)\boldsymbol{\xi}^*\right.\right.\\ &\quad \left. + \nabla_Y[A+B](\mathbf{Y}^*):(\boldsymbol{\xi}^*(\mathbf{Y}^*)^\top) - R\mathbf{Y}^*\right)\\ &\quad \left. - \nabla_x^\top[\phi_{\hat{j}}]\nabla_x[D\boldsymbol{\xi}^*]\right) + \int_{\partial\Omega} d\Gamma\phi_{\hat{j}}\left(\widehat{\mathbf{N}}\cdot\nabla_x\right)[D\boldsymbol{\xi}^*].\end{aligned}$$

A form with reduced order derivatives is derived by eliminating the boundary integral by the no-flux condition (13) and then the Galerkin approximations are substituted for the state and co-state, thus producing

$$\begin{aligned}\mathbf{0} \simeq \sum_{\hat{k}=1}^{\widehat{M}} \int_\Omega d\mathbf{x}\left(\left(\widehat{\boldsymbol{\xi}}'_{\hat{k}}(t) + (A+B)(\widehat{\mathbf{Y}})\widehat{\boldsymbol{\xi}}_{\hat{k}}\right.\right.\\ \left. + \nabla_Y[A+B](\widehat{\mathbf{Y}}):(\widehat{\boldsymbol{\xi}}_{\hat{k}}(\widehat{\mathbf{Y}})^\top) - R\widehat{\mathbf{Y}}_{\hat{k}}\right)\phi_{\hat{j}}\phi_{\hat{k}}\\ \left. - D\widehat{\boldsymbol{\xi}}_{\hat{k}}\nabla_x^\top[\phi_{\hat{j}}]\nabla_x[\phi_{\hat{k}}]\right),\end{aligned}$$

except that the bilinear terms are only symbolically designated by $\widehat{\mathbf{Y}}$, for $\hat{j} = 1:\widehat{M}$. Next by substituting the Galerkin approximation for $\widehat{\mathbf{Y}}$ in the nonlinear terms, using the element Galerkin integral notation for the mass $\widehat{\mathcal{M}}_{\hat{j},\hat{k}}$ (17), stiffness $\widehat{\mathcal{K}}_{\hat{j},\hat{k}}$ (18) and the pure bilinear triple $\widehat{\mathcal{T}}_{\hat{j},\hat{k},\hat{l}}$ (19), the compact Galerkin ODEs are obtained:

$$\begin{aligned}\mathbf{0} = \sum_{\hat{k}=1}^{\widehat{M}}\left(\widehat{\mathcal{M}}_{\hat{j},\hat{k}}\left(\widehat{\boldsymbol{\xi}}'_{\hat{k}}(t) + a_1\widehat{\xi}_{1,\hat{k}}\mathbf{e}_1 + a_2\widehat{\xi}_{2,\hat{k}}\mathbf{e}_2\right.\right.\\ \left. - a_3\widehat{\xi}_{3,\hat{k}}\mathbf{e}_3 - R\widehat{\mathbf{Y}}_{\hat{k}}(t)\right) - D\widehat{\mathcal{K}}_{\hat{j},\hat{k}}\widehat{\boldsymbol{\xi}}_{\hat{k}}(t)\\ - \sum_{\hat{l}=1}^{\widehat{M}}\widehat{\mathcal{T}}_{\hat{j},\hat{k},\hat{l}}\left(\frac{2a_1}{k_1}\widehat{Y}_{1,\hat{l}}(t)\widehat{\xi}_{1,\hat{k}}(t)\mathbf{e}_1\right.\\ + \frac{2a_2}{k_2}\widehat{Y}_{2,\hat{l}}(t)\widehat{\xi}_{2,\hat{k}}(t)\mathbf{e}_2\\ + \alpha_{1,2}\left(\widehat{Y}_{2,\hat{l}}(t)\mathbf{e}_1 + \widehat{Y}_{1,\hat{l}}(t)\mathbf{e}_2\right)\widehat{\xi}_{1,\hat{k}}(t)\\ + \kappa_{1,3}\left(\widehat{Y}_{3,\hat{l}}(t)\mathbf{e}_1 + \widehat{Y}_{1,\hat{l}}(t)\mathbf{e}_3\right)\widehat{\xi}_{1,\hat{k}}(t)\\ + \alpha_{2,1}\left(\widehat{Y}_{2,\hat{l}}(t)\mathbf{e}_1 + \widehat{Y}_{1,\hat{l}}(t)\mathbf{e}_2\right)\widehat{\xi}_{1,\hat{k}}(t)\\ \left.\left. + \kappa_{2,3}\left(\widehat{Y}_{3,\hat{l}}(t)\mathbf{e}_2 + \widehat{Y}_{2,\hat{l}}(t)\mathbf{e}_3\right)\widehat{\xi}_{2,\hat{k}}(t)\right)\right),\end{aligned} \tag{22}$$

for $\hat{j} = 1:\widehat{M}$. This Galerkin ODE (22) may be computed by the appropriate numerical methods using the same Galerkin integral basis coefficients.

5) For each completed double shot for $\delta = 1:L$, the co-state approximation $\widehat{\boldsymbol{\xi}}^{(\delta)}(\mathbf{x},t) = \sum_{\hat{k}=1}^{\widehat{M}}\widehat{\boldsymbol{\xi}}_{\hat{k}}^{(\delta)}(t)\phi_{\hat{k}}(\mathbf{x})$ is used to determine the *regular optimal control* (11) updated value third component

$$\widehat{U}_3^{(\delta+1)}(\mathbf{x},t) = U_{0,3}(\mathbf{x},t) + \widehat{\xi}_3^{(\delta)}(\mathbf{x},t)/s_3. \tag{23}$$

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

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

and say,

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

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

V. Spherical Finite Element Test Configuration

Consider a three dimensional test configuration that is a sphere of radius R_r. Transforming the spherical coordinates in space as usual,

$$(x,y,z) = r(\cos(\theta)\sin(\psi), \sin(\theta)\sin(\psi), \cos(\psi)) \tag{24}$$

where r, θ, ψ are the radius, polar angle and azimuthal angle, respectively. Also

$$0 \le r \le R_r, \quad 0 \le \theta \le 2\pi \quad \text{and} \quad 0 \le \psi \le \pi.$$

The brick element grid in spherical coordinates is constructed of M_r radial sectors of width $\Delta r = R_r/M_r$, M_θ polar sectors of width $\Delta\theta = 2\pi/M_\theta$ and M_ψ azimuthal sectors of width $\Delta\psi = \pi/M_\psi$. The nodal values are given by

$$(\theta_{i_e}, \psi_{j_e}, r_{k_e}) = ((i_e-1)\Delta\theta, (j_e-1)\Delta\psi, (k_e-1)\Delta r)$$

$$\text{for } i_e = 1{:}M_\theta + 1, j_e = 1{:}M_\psi + 1 \text{ and } k_e = 1{:}M_r + 1.$$

The elements are numbered in (θ,ψ,r) linear priority order like the nodal values,

$$e_{i_e,j_e,k_e} = i_e + (j_e-1)\cdot M_\theta + (k_e-1)\cdot M_\theta\cdot M_\psi,$$

$$\text{for } i_e = 1{:}M_\theta, j_e = 1{:}M_\psi \text{ and } k_e = 1{:}M_r.$$

Within element e_{i_e,j_e,k_e}, the element primary node with local node number $i = 1$ has the same global node number $\bar{k}_e = \{i_e, j_e, k_e\}$ as the element, i.e.,

$$n_{\bar{k}_e,1} = e_{i_e,j_e,k_e} \text{ for } i_e = 1{:}M_\psi, j_e = 1{:}M_\theta, k_e = 1{:}M_r.$$

The element local node numbering is $i = 1{:}8$ as shown in the Figure 1.

For simplicity, *trilinear basis functions* are used for all brick elements in spherical coordinates. The trilinear bases

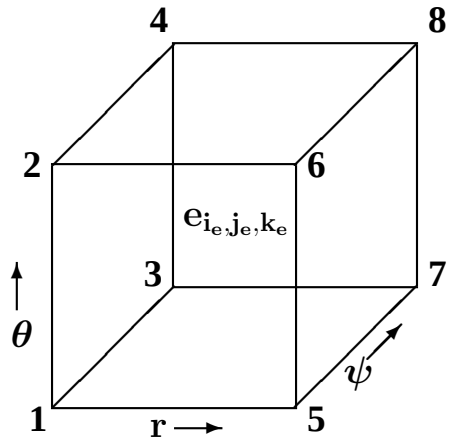

Fig. 1. Local node numbering with $i = 1:8$ for general element e_{i_e,j_e,k_e}.

are constructed from the more basic one-dimensional linear bases with only two nodes labeled 1 and 2:

$$\phi_{1r}^{(k_e)}(r) = \left(\frac{r_{k_e+1} - r}{\Delta r}\right), \quad \phi_{2r}^{(k_e)}(r) = \left(\frac{r - r_{k_e}}{\Delta r}\right); \tag{25}$$

$$\phi_{1\theta}^{(i_e)}(\theta) = \left(\frac{\theta_{i_e+1} - \theta}{\Delta\theta}\right), \quad \phi_{2\theta}^{(i_e)}(\theta) = \left(\frac{\theta - \theta_{i_e}}{\Delta\theta}\right); \tag{26}$$

$$\phi_{1\psi}^{(j_e)}(\psi) = \left(\frac{\psi_{j_e+1} - \psi}{\Delta\psi}\right), \quad \phi_{2\psi}^{(j_e)}(\psi) = \left(\frac{\psi - \psi_{j_e}}{\Delta\psi}\right); \tag{27}$$

For the general element

$$e_{i_e,j_e,k_e} \text{ on } [\theta_{i_e}, \theta_{i_e+1}] \times [\psi_{j_e}, \psi_{j_e+1}] \times [r_{k_e}, r_{k_e+1}]$$
$$\text{for } i_e = 1:M_\theta,\ j_e = 1:M_\psi \ \text{ and } k_e = 2:M_r$$

in the $i = 1:8$ element node numbering,

$$\begin{aligned}
\phi_1^{(i_e,j_e,k_e)}(r,\theta,\psi) &= \phi_{1r}^{(k_e)}(r)\cdot\phi_{1\theta}^{(i_e)}(\theta)\cdot\phi_{1\psi}^{(j_e)}(\psi); && (28)\\
\cdot\phi_2^{(i_e,j_e,k_e)}(r,\theta,\psi) &= \phi_{1r}^{(k_e)}(r)\cdot\phi_{2\theta}^{(i_e)}(\theta)\cdot\phi_{1\psi}^{(j_e)}(\psi); && (29)\\
\phi_3^{(i_e,j_e,k_e)}(r,\theta,\psi) &= \phi_{1r}^{(k_e)}(r)\cdot\phi_{1\theta}^{(i_e)}(\theta)\cdot\phi_{2\psi}^{(j_e)}(\psi); && (30)\\
\phi_4^{(i_e,j_e,k_e)}(r,\theta,\psi) &= \phi_{1r}^{(k_e)}(r)\cdot\phi_{2\theta}^{(i_e)}(\theta)\cdot\phi_{2\psi}^{(j_e)}(\psi); && (31)\\
\phi_5^{(i_e,j_e,k_e)}(r,\theta,\psi) &= \phi_{2r}^{(k_e)}(r)\cdot\phi_{1\theta}^{(i_e)}(\theta)\cdot\phi_{1\psi}^{(j_e)}(\psi); && (32)\\
\phi_6^{(i_e,j_e,k_e)}(r,\theta,\psi) &= \phi_{2r}^{(k_e)}(r)\cdot\phi_{2\theta}^{(i_e)}(\theta)\cdot\phi_{1\psi}^{(j_e)}(\psi); && (33)\\
\phi_7^{(i_e,j_e,k_e)}(r,\theta,\psi) &= \phi_{2r}^{(k_e)}(r)\cdot\phi_{1\theta}^{(i_e)}(\theta)\cdot\phi_{2\psi}^{(j_e)}(\psi); && (34)\\
\phi_8^{(i_e,j_e,k_e)}(r,\theta,\psi) &= \phi_{2r}^{(k_e)}(r)\cdot\phi_{2\theta}^{(i_e)}(\theta)\cdot\phi_{2\psi}^{(j_e)}(\psi)\,. && (35)
\end{aligned}$$

The mapping (24) from spherical to cartesian coordinates is not unique since the mapping is a many-to-one. The non-unique nodes arise at the *origin* for $r = r_1 = 0$ and any (θ,ψ) or at the *poles* for $\psi = 0,\pi$ and any (r,θ) or at the polar angle line of discontinuity for $\theta = 2\pi$ for any (r,ψ) (really a periodic boundary condition). The non-uniqueness or resulting over-determinism can simply be removed by adding together the appropriate bases in (28)-(31) corresponding to the same non-unique nodes, using the identities for the one-dimensional bases and their derivatives,

$$\phi_{1\rho}^{(\ell_e)}(\rho) + \phi_{2\rho}^{(\ell_e)}(\rho) = 1, \quad \phi_{1\rho}^{(\ell_e)\,\prime}(\rho) + \phi_{2\rho}^{(\ell_e)\,\prime}(\rho) = 0; \tag{36}$$

for $\rho = r,\ \theta$ or ψ and $\ell_e = k_e,\ i_e$ or j_e. While it may appear awkward to have to make this adjustment, the disadvantage is out-weighed by the ease of deforming a sphere into a brain geometry than deforming a brick or rectangular solid into a brain geometry. Another advantage of spherical coordinates is the ease of imposing the *no-flux boundary condition (BC)* at $r = R_r$, since on the element e_{i_e,j_e,k_e} the i_sth state solution for $i_s = 1:3$ is expressed as a preliminary Galerkin approximation,

$$Y_{i_s}^{(i_e,j_e,k_e)}(r,\theta,\psi,t) \simeq \sum_{j=1}^{8} \widetilde{Y}_{i_s,j}^{(i_e,j_e,k_e)}(t)\cdot\phi_j^{(i_e,j_e,k_e)}(r,\theta,\psi), \tag{37}$$

so the normal gradient, at boundary element $ke = M_r$ with $r_{M_r+1} = R_r$ for local nodes $j = 5:8$, reduces to

$$\begin{aligned}
Y_{i_s,r}^{(i_e,j_e,M_r)}(R_r,\theta,\psi,t) \simeq \sum_{j=5}^{8} &\left(\widetilde{Y}_{i_s,j}^{(i_e,j_e,M_r)}(t) - \widetilde{Y}_{i_s,j-4}^{(i_e,j_e,M_r)}(t)\right)\\
&\cdot\phi_{j,r}^{(i_e,j_e,M_r)}(R_r,\theta,\psi),
\end{aligned}$$

and then

$$\widetilde{Y}_{i_s,j}^{(i_e,j_e,M_r)}(t) = \widetilde{Y}_{i_s,j-4}^{(i_e,j_e,M_r)}(t) \tag{38}$$

for $j = 5:8$ and arbitrary (θ,ψ) if $Y_{i_s,r}^{(i_e,j_e,M_r)}(R_r,\theta,\psi,t) = 0$ with no-flux, where the symmetries and asymmetries of the bases functions (25-35) and derivatives have been used. This version of the no-flux condition is much better and simpler to use than that in [2], even when dealing with a deformed sphere in the form of a brain case.

The element matrices for local node numbers $i,j,k = 1:8$ are,

$$\begin{aligned}
\mathcal{M}_{i,j}^{(i_e,j_e,k_e)} &= \int_{\theta_{i_e}}^{\theta_{i_e+1}} d\theta \int_{\psi_{i_e}}^{\psi_{j_e+1}} d\psi \int_{r_{k_e}}^{r_{k_e+1}} dr r^2\sin(\psi)\\
&\quad \left(\phi_i^{(i_e,j_e,k_e)}\phi_j^{(i_e,j_e,k_e)}\right)(r,\theta,\psi), && (39)\\
\mathcal{K}_{i,j}^{(i_e,j_e,k_e)} &= \int_{\theta_{i_e}}^{\theta_{i_e+1}} d\theta \int_{\psi_{i_e}}^{\psi_{j_e+1}} d\psi \int_{r_{k_e}}^{r_{k_e+1}} dr r^2\sin(\psi)\\
&\quad \Big(\phi_{i,r}^{(i_e,j_e,k_e)}\phi_{j,r}^{(i_e,j_e,k_e)}\\
&\quad + \frac{1}{r^2}\phi_{i,\psi}^{(i_e,j_e,k_e)}\phi_{j,\psi}^{(i_e,j_e,k_e)}\\
&\quad + \frac{1}{r^2\sin^2(\psi)}\phi_{i,\theta}^{(i_e,j_e,k_e)}\phi_{j,\theta}^{(i_e,j_e,k_e)}\Big)(r,\theta,\psi), && (40)\\
\mathcal{T}_{i,j,k}^{(i_e,j_e,k_e)} &= \int_{\theta_{i_e}}^{\theta_{i_e+1}} d\theta \int_{\psi_{i_e}}^{\psi_{j_e+1}} d\psi \int_{r_{k_e}}^{r_{k_e+1}} dr r^2\sin(\psi)\\
&\quad \left(\phi_i^{(i_e,j_e,k_e)}\phi_j^{(i_e,j_e,k_e)}\phi_k^{(i_e,j_e,k_e)}\right)(r,\theta,\psi), && (41)
\end{aligned}$$

where in (40) $\phi_{i,\rho}^{(i_e,j_e,k_e)}(\rho)$ denotes the partial derivative of $\phi_i^{(i_e,j_e,k_e)}(\rho)$ with respect to generic spherical coordinate ρ. In the stiffness matrix (40), the mapping singularities of the gradient lead to reciprocal factors in r and $\sin(\psi)$, but the r factors are simply cancelled by the Jacobian $r^2\sin(\psi)$ and leave uncanceled $\sin(\psi)$-denominators in the θ-derivative term. However, these $\sin(\psi)$-denominators are completely eliminated in later analysis upon eliminating non-unique nodes by combining terms and associated ODEs.

VI. Computational Results

The double shot, forward-backward iteration algorithm using the finite element method outlined in Section IV is implemented on three-dimensional space with the three states and the drug input control. The numerical implementation of

the algorithm is similar to our one-dimensional application in [1], except there the finite difference version of *Crank Nicolson's* method was used and is too costly to extend to another dimension. The implementation is more similar to the two-dimensional problem treated in [2], except that the complexity of the mapping from spherical to cartesian is much greater than the polar mapping due to the degeneracy at the poles. Once the non-unique degeneracies of states and co-states due to aliases and boundary conditions are eliminated so that there are only $\widehat{M}$ linearly independent Galerkin coefficients, $\widehat{\mathbf{Y}}_{\hat{k}}(t)$ in ODE (20), $\widehat{\boldsymbol{\xi}}_{\hat{k}}(t)$ in ODE (22) and control $\widehat{\mathbf{U}}_{\hat{k}}(t)$. This non-uniqueness elimination keeps the system of ODEs from being over-determined, preserving the symmetry of the mass and other coefficient arrays, and eliminating the $1/\sin(\psi)$ singularity in the stiffness integrals by virtue of identities (36). A summary of the degeneracy removal by combining unknowns and the corresponding equations follows;

1) *Origin*: $r = 0$, $k_e = 1$, $(x, y, z) = (0, 0, 0)$ for $i_e = 1\!:\! M_\theta$, $j_e = 1\!:\! M_\psi$, $i = 1\!:\!4$ (see Fig. 1 to see why only local nodes $i = 1\!:\!4$ are involved), so
$$Y^{(i_e,j_e,1)}_{i_s,i}(t) = Y^{(1,1,1)}_{i_s,1}(t) = \widehat{Y}_{i_s,\hat{j}}(t),$$
where the initial count is $\hat{j} = 1$ of an independent set of unknowns; then the ODEs must be combined corresponding to the combined aliased unknowns, but this is only illustrated on the central derivative terms as
$$\widehat{M}_{\hat{j},\hat{j}} \cdot \widehat{Y}'_{i_s,\hat{j}}(t) \equiv Y^{(1,1,1)\,\prime}_{i_s,1}(t) \cdot \sum_{i_e=1}^{M_\theta} \sum_{j_e=1}^{M_\psi} \sum_{i=1}^{4} \sum_{j=1}^{4} M^{(i_e,j_e,1)}_{i,j}.$$
2) *Higher Pole*: $\psi = 0$, $(x, y, z) = r(0, 0, +1)$ for elements $i_e = 1 : M_\theta$, $j_e = 1$, $ke = 2 : M_r - 1$ for $i = 1\!:\!2$ (see Fig. 1),
$$Y^{(i_e,1,k_e)}_{i_s,i}(t) = Y^{(1,1,k_e)}_{i_s,1}(t) = \widehat{Y}_{i_s,\hat{j}}(t)$$
where $\hat{j} = 2+(k_e-2)(M_\theta(M_\psi-1)+2)$, but also in the $k_e - 1$ neighboring element $Y^{(i_e,1,k_e-1)}_{i_s,i}(t) = \widehat{Y}_{i_s,\hat{j}}(t)$ for $i = 5\!:\!6$, then
$$\widehat{M}_{\hat{j},\hat{j}} \cdot \widehat{Y}'_{i_s,\hat{j}}(t), \equiv Y^{(1,1,k_e)\,\prime}_{i_s,1}(t) \cdot \sum_{i_e=1}^{M_\theta} \sum_{k=0}^{1} \sum_{i=\bar{i}_k}^{\bar{i}_k+1} \sum_{j=\bar{i}_k}^{\bar{i}_k+1} M^{(i_e,1,k_e-k)}_{i,j},$$
where $\bar{i}_0 = 1$ and $\bar{i}_1 = 5$.
If $k_e = M_r$ when $j_e = 1$, then the no-flux BC (38) holds so these terms must be added,
$$\widehat{M}_{\hat{j},\hat{j}} \cdot \widehat{Y}'_{i_s,\hat{j}}(t) \equiv Y^{(1,1,M_r)\,\prime}_{i_s,1}(t) \cdot \sum_{i_e=1}^{M_\theta} \left(\sum_{k=0}^{1} \sum_{i=\bar{i}_k}^{\bar{i}_k+1} \sum_{j=\bar{i}_k}^{\bar{i}_k+1} M^{(i_e,1,M_r-k)}_{i,j} + \sum_{i=5}^{6} \sum_{j=5}^{6} M^{(i_e,1,M_r)}_{i,j} \right),$$
where $\hat{j} = 2 + (M_r - 2)(M_\theta(M_\psi - 1) + 2)$.
3) *Lower Pole*: $\psi = \pi$, $(x, y, z) = r(0, 0, -1)$ for elements $i_e = 1 : M_\theta$, $j_e = M_\psi$, $ke = 2 : M_r - 1$ for $i = 3\!:\!4$ (see Fig. 1),
$$Y^{(i_e,M_\psi,k_e)}_{i_s,i}(t) = Y^{(1,M_\psi,k_e)}_{i_s,3}(t) = \widehat{Y}_{i_s,\hat{j}}(t)$$
where $\hat{j} = 1 + (k_e - 1)(M_\theta(M_\psi - 1) + 2)$, but also in the $k_e - 1$ neighboring element $Y^{(i_e,M_\psi,k_e-1)}_{i_s,i}(t) = \widehat{Y}_{i_s,\hat{j}}(t)$ for $i = 7\!:\!8$, then
$$\widehat{M}_{\hat{j},\hat{j}} \cdot \widehat{Y}'_{i_s,\hat{j}}(t) \equiv Y^{(1,M_\psi,k_e)\,\prime}_{i_s,3}(t) \cdot \sum_{i_e=1}^{M_\theta} \sum_{k=0}^{1} \sum_{i=\bar{i}_k}^{\bar{i}_k+1} \sum_{j=\bar{i}_k}^{\bar{i}_k+1} M^{(i_e,M_\psi,k_e-k)}_{i,j},$$
where $\bar{i}_0 = 3$ and $\bar{i}_1 = 7$.
If $k_e = M_r$ when $j_e = M_\psi$, then the no-flux BC (38) holds so these terms must be added,
$$\widehat{M}_{\hat{j},\hat{j}} \cdot \widehat{Y}'_{i_s,\hat{j}}(t) \equiv Y^{(1,M_\psi,M_r)\,\prime}_{i_s,m}(t) \cdot \sum_{i_e=1}^{M_\theta} \left(\sum_{k=0}^{1} \sum_{i=\bar{i}_k}^{\bar{i}_k+1} \sum_{j=\bar{i}_k}^{\bar{i}_k+1} M^{(i_e,M_\psi,M_r-k)}_{i,j} + \sum_{i=7}^{8} \sum_{j=7}^{8} M^{(i_e,M_\psi,M_r)}_{i,j} \right),$$
where $\hat{j} = 2 + (M_r - 1)(M_\theta(M_\psi - 1) + 2)$.
4) *Periodic BC*: $\theta = 2\pi$, $i_e = M_\theta$, $(x, y, z) = r(\sin(\psi), 0, \cos(\psi))$, the same if $\theta = 0$ or $\theta = 2\pi$, for element $i_e = M_\theta$, $j_e = 1 : M_\psi - 1$, $k_e = 2 : M_r - 1$ for local node $i = 4$ (see Fig. 1), so
$$Y^{(M_\theta,j_e,k_e)}_{i_s,4}(t) = Y^{(1,j_e,k_e)}_{i_s,3}(t) = \widehat{Y}_{i_s,\hat{j}}(t)$$
where $\hat{j} = 3+(k_e-2)(M_\theta(M_\psi-1)+2)+(j_e-1)M_\theta$, but also $Y^{(M_\theta,j_e+\bar{j}_m,k_e-\bar{k}_m)}_{i_s,2m}(t) = \widehat{Y}_{i_s,\hat{j}}(t)$ when $\bar{j}_m = 1, 0, 1, 0$ and $\bar{k}_m = 0, 0, 1, 1$, respectively for $m = 1\!:\!4$, then
$$\widehat{M}_{\hat{j},\hat{j}} \cdot \widehat{Y}'_{i_s,\hat{j}}(t) \equiv \left(M^{(1,j_e,k_e)}_{3,3} + \sum_{m=1}^{4} M^{(M_\theta,j_e+\bar{j}_m,k_e-\bar{k}_m)}_{2m,2m} \right) \cdot Y^{(1,j_e,k_e)\,\prime}_{i_s,3}(t).$$
If $k_e = M_r$ when $i_e = M_\theta$ and $j_e = 1 : M_\psi - 1$, then the no-flux BC must be added, so
$$\widehat{M}_{\hat{j},\hat{j}} \cdot \widehat{Y}'_{i_s,\hat{j}}(t) \equiv Y^{(1,j_e,M_r)\,\prime}_{i_s,3}(t) \cdot \left(M^{(1,j_e,M_r)}_{3,3} + \sum_{m=1}^{4} M^{(M_\theta,j_e+\bar{j}_m,M_r-\bar{k}_m)}_{2m,2m} + M^{(M_\theta,j_e+1,M_r)}_{6,6} + M^{(M_\theta,j_e,M_r)}_{8,8} \right),$$
where $\hat{j} = 3+(M_r-2)(M_\theta(M_\psi-1)+2)+(j_e-1)M_\theta$.

The actual implementation uses subscripted subscripts to gather the aliased unknowns and equation into their unique locations.

The general method uses a combination of Crank-Nicolson and prediction-correction methods developed in [7] for solving high dimensional stochastic control problems on supercomputers. The general method can handle both implicit and nonlinear terms. For simplicity, the forward ODE (20) for the degeneracy removed i_sth-state $\widehat{Y}_{i_s,\hat{j}}$ at node $\hat{j}$ can be written symbolically, for $i_s = 1\!:\!3$ and nodes $\hat{j} = 1\!:\!\widehat{M}$, as
$$\sum_{\hat{k}=1}^{\widehat{M}} \mathcal{M}_{\hat{j},\hat{k}} \widehat{Y}'_{i_s,\hat{k}}(t) = \sum_{j_s=1}^{3} \sum_{\hat{k}=1}^{\widehat{M}} \mathcal{A}_{i_s,j_s,\hat{j},\hat{k}}(\widehat{\mathcal{Y}}(t)) \cdot \widehat{Y}_{j_s,\hat{k}}(t) - \sum_{\hat{k}=1}^{\widehat{M}} \mathcal{M}_{\hat{j},\hat{k}} \widehat{U}_{\hat{k}}(t),$$
where $\mathcal{A}_{i_s,j_s,\hat{j},\hat{k}}(\widehat{\mathcal{Y}}(t))$ symbolically represents theright-hand-side matrices including nonlinear terms and $\widehat{\mathcal{Y}}(t)$ represents the combined states and nodes array. A similar backward ODE can be written for the co-state $\widehat{\xi}_{i_s\hat{k}}(t)$ with the state replacing the role of the control in the source while with remaining in the general nonlinear coefficient $\mathcal{B}_{i_s,j_s,\hat{j},\hat{k}}(\widehat{\mathcal{Y}}(t))$as in (22), i.e.,
$$\sum_{\hat{k}=1}^{\widehat{M}} \mathcal{M}_{\hat{j},\hat{k}} \widehat{\xi}'_{i_s,\hat{k}}(t) = -\sum_{j_s=1}^{3} \sum_{\hat{k}=1}^{\widehat{M}} \mathcal{B}_{i_s,j_s,\hat{j},\hat{k}}(\widehat{\mathcal{Y}}(t)) \cdot \widehat{\xi}_{j_s,\hat{k}}(t) + \sum_{\hat{k}=1}^{\widehat{M}} \mathcal{M}_{\hat{j},\hat{k}} R \widehat{Y}_{i_s\hat{k}}(t).$$
The essential setup for a general Crank-Nicolson method is to use the midpoint approximation on the integral form of the differential equations followed by an average approximation

of the midpoint values, producing from the state ODE for $\widehat{Y}_{i_s,\hat{k},\ell+1}$ at time $t_\ell = (\ell-1)*\Delta t$ with $\ell = 1:N_t$,

$$\sum_{\hat{k}=1}^{\widehat{M}} \mathcal{M}_{\hat{j},\hat{k}} \left(\widehat{Y}_{i_s,\hat{k},\ell+1} - \widehat{Y}_{i_s,\hat{k},\ell}\right) =$$
$$+\Delta t \sum_{j_s=1}^{3} \sum_{\hat{k}=1}^{\widehat{M}} \mathcal{A}_{i_s,j_s,\hat{j},\hat{k},\ell+0.5} \cdot \widehat{Y}_{j_s,\hat{k},\ell+0.5}$$
$$-\Delta t \sum_{\hat{k}=1}^{\widehat{M}} \mathcal{M}_{\hat{j},\hat{k}} \widehat{U}_{\hat{k},\ell+0.5},$$

where the average approximation at the midpoint is

$$\widehat{Y}_{j_s,\hat{k},\ell+0.5} \simeq 0.5 * \left(\widehat{Y}_{i_s,\hat{k},\ell+1} + \widehat{Y}_{i_s,\hat{k},\ell}\right),$$
$$\widehat{U}_{\hat{k},\ell+0.5} \simeq 0.5 * \left(\widehat{U}_{\hat{k},\ell+1} + \widehat{U}_{\hat{k},\ell}\right)$$

and is compatible with the midpoint approximation in accuracy. The approximation $\mathcal{A}_{i_s,j_s,\hat{j},\hat{k},\ell+0.5}$ is similar computed.Similarly, except for backward integration, the co-state $\widehat{\xi}_{i_s,\hat{k},\ell-1}$ satisfies

$$\sum_{\hat{k}=1}^{\widehat{M}} \mathcal{M}_{\hat{j},\hat{k}} \left(\widehat{\xi}_{i_s,\hat{k},\ell-1} - \widehat{\xi}_{i_s,\hat{k},\ell}\right) =$$
$$+\Delta t \sum_{j_s=1}^{3} \sum_{\hat{k}=1}^{\widehat{M}} \mathcal{B}_{i_s,j_s,\hat{j},\hat{k},\ell-0.5} \cdot \widehat{\xi}_{j_s,\hat{k},\ell-0.5}$$
$$-\Delta t \sum_{\hat{k}=1}^{\widehat{M}} \mathcal{M}_{\hat{j},\hat{k}} R\widehat{Y}_{\hat{k},\ell-0.5},$$

where thee average approximation at the midpoint is

$$\widehat{\xi}_{j_s,\hat{k},\ell-0.5} \simeq 0.5 * \left(\widehat{\xi}_{i_s,\hat{k},\ell-1} + \widehat{\xi}_{i_s,\hat{k},\ell}\right),$$

Next the *predictor-corrector* procedure is used to handle the remaining implicit and nonlinear terms. The zeroth corrector, given the final correction $\widehat{Y}_{i_s,\hat{j},\ell}$ at time stage i is the predictor,

$$\mathrm{YC}^{(0)}_{i_s,\hat{j},\ell+1} = \widehat{Y}_{i_s,\hat{j},\ell}.$$

This initialization permits finding the $(\gamma+1)$th correction $\mathrm{YC}^{(\gamma+1)}_{\hat{j},\ell+1}$ from

$$\sum_{\hat{k}=1}^{\widehat{M}} \mathcal{M}_{\hat{j},\hat{k}} \left(\mathrm{YC}^{(\gamma+1)}_{i_s,\hat{k},\ell+1} - \mathrm{YC}^{(\gamma)}_{i_s,\hat{k},\ell}\right) =$$
$$+\Delta t \sum_{j_s=1}^{3} \sum_{\hat{k}=1}^{\widehat{M}} \mathrm{AC}^{(\gamma)}_{i_s,j_s,\hat{j},\hat{k},\ell+0.5} \cdot \mathrm{YC}^{(\gamma)}_{j_s,\hat{k},\ell+0.5}$$
$$-\Delta t \sum_{\hat{k}=1}^{\widehat{M}} \mathcal{M}_{\hat{j},\hat{k}} \mathrm{UC}^{(\gamma)}_{\hat{k},\ell+0.5},$$

where $YC^{(\gamma)}_{j_s,\hat{k},\ell+0.5}$ and other midpoint terms are evaluated as before by averaging. The final correction at the final time of the *state-shot forward iteration*, $\widehat{Y}_{\hat{j},N_t+1}$ yields the starting or final-time condition for $\widehat{\xi}_{i_s,\hat{k},Nt+1}$ using (14). Then the final correction at the initial time of the *costate-shot backward iteration* produces the initial control condition (23) when $i_s = 3$ which is used to begin another double for $\delta = 2:L$. The algorithm is implemented in MATLAB™ on a desktop computer.

The data for the numerical parameters are drawn from various sources including Wang et al. [13], [14], Swanson [12] and Murray [10], while unavailable parameters were estimated. The diffusion diagonal vector is D = [4.2e-3, 1.e-15, 0.216] cm^2 per day. The quadratic cost coefficients are $r_1 = 0.1 = q_1 = q_3$ and $s_3 = 0.2$. The net growth coefficient is a =[1.2e-2, 8.64e-7, 11.3] per day. The other coefficients are $k_i = 1, i = 1:2$, $\alpha_{1,2} = \alpha_{2,1} = \kappa_{2,3}$ =1.e-4 and $\kappa_{1,3} = 0.5$. Following Murray the initial state for the normal tissue is assumed to be 1. The initial tumor spread and drug concentration are assumed to be Gaussian with scale and state dependent means, spread and weights. Maple™ 9.5 was used to exactly evaluate the integrals of the element matrices off-line.

A sample history of the optimal relative tumor density $Y_1^*(r,\theta,\psi,t)$ for r over [0, 5] in centimeters at fixed (θ,ψ) = (0, 0) in radians and at quartiles in time of a 5 day treatment schedule is given in Figure 2. The initial tumor peak is at $(r,\theta,\psi) = (0.5, 0, 0)$. This test case shows significant reduction of the tumor density over the treatment schedule.

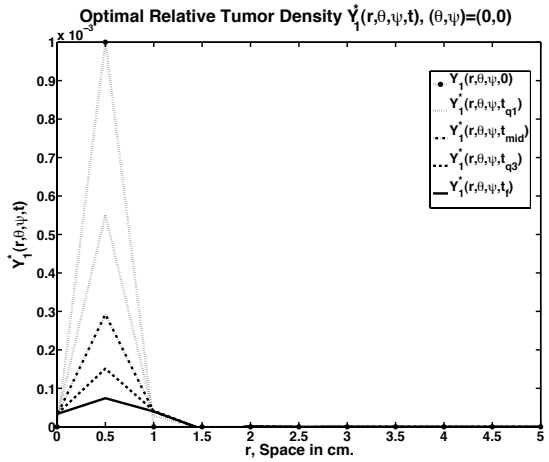

Fig. 2. The optimal, relative tumor density $Y_1^*(r,\theta,\psi,t)$ versus r at time, $t = 0$ days, as a cross-section at $(\theta,\psi) = (0,0)$ radians, with the initial tumor density peak location with $(r,\theta,\psi) = (0.5, 0, 0)$. The grid size is $(M_r, M_\theta, M_\psi) = (10, 3, 3)$.

A more detailed presentation of the initial to final tumor density from $t{=}t_0{=}0$ to $t{=}t_{mid}{=}2.5$ to $t{=}t_f{=}5$ days is given in Figs. 3–5 over the larger (r,ψ) plane section with fixed $\theta = 0$ showing that the final peak value is small and the tumor has not spread significantly through the rest of the plane in spherical coordinates, although somewhat in the ψ–direction but not much in the r–direction.

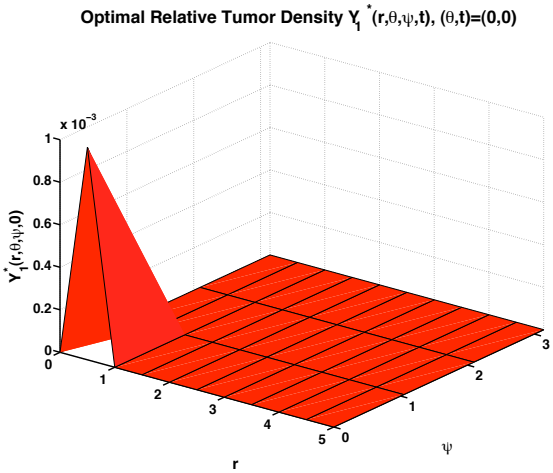

Fig. 3. The initial relative tumor density $Y_1^*(r,\theta,\psi,t)$ over (r,ψ) plane $\theta = 0$ and at time $t = 0$ days. The FEM grid size is $(M_r, M_\theta, M_\psi) = (10, 3, 3)$, with the mid-time and final distributions given in Figs. 4–5, respectively.

VII. Conclusion and Future Directions

The theory of Galerkin finite elements is used to develop approximations to the distributed parameter optimal control

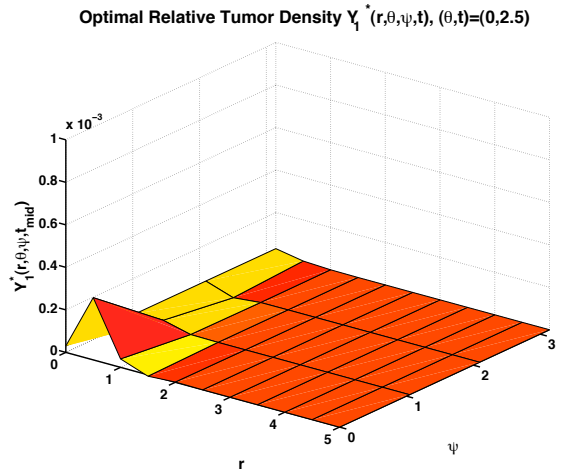

Fig. 4. The mid-time optimal, relative tumor density $Y_1^*(r, \theta, \psi, t)$ over (r, θ) plane at $\theta = 0$ and time $t = t_{mid} = 2.5$ days. The FEM grid size is $(M_r, M_\theta, M_\psi) = (10, 3, 3)$, starting from the distribution displayed in Fig. 3.

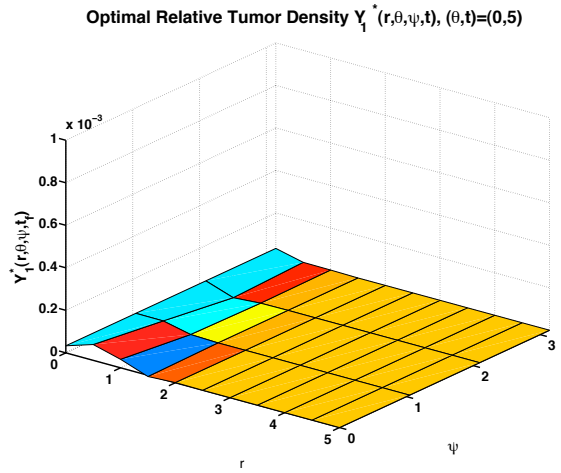

Fig. 5. The final optimal, relative tumor density $Y_1^*(r, \theta, \psi, t)$ over (r, θ) plane at $\theta = 0$ and time $t = T = 5.0$ days. The FEM grid size is $(M_r, M_\theta, M_\psi) = (10, 3, 3)$, starting from the initial distribution displayed in Fig. 3 through mid-time distribution in Fig; 4. The tumor has shrunk down to very small values over this plane section.

problem of cancer drug delivery to the brain governed by a coupled set of three reaction diffusion PDEs in three space dimensions. The three state variables are the tumor cell density, the normal cell density and the cancer drug concentration. While the tumor and normal cells are highly coupled through intrinsic and competitive interactions, the concentration is directly controlled by the drug delivery control rate. The optimally controlled distributed parameter system is derived by a straight-forward calculus of variations technique without resort to an extremely abstract formulation, and that should be useful in other similar scientific or engineering applications.

The system of optimal PDEs in three state and three co-state dimensions is reduced by Galerkin approximations of the state, co-state and control vectors to a system of six ODEs in time with three fundamental element integral coefficient forms: the mass, the stiffness and nonlinear coefficients. The finite element configuration is given for a spherical geometry that can be used to test the optimal drug delivery computations. This finite element configuration will be more amenable to complex brain structures and three-dimensional geometries than the finite difference method and low dimension of our earlier work.

Future directions include:

- Application to general curvilinear coordinates for general brain geometries;
- Application to heterogeneous brain structures such as spinal fluid cavities, variable brain matter, vascular system and the blood brain barrier.
- For the final paper, there will be more results with more quantitative descriptions.

VIII. 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] S. P. Chakrabarty and F. B. Hanson,"Optimal control of drug delivery to brain tumors for a distributed parameters model," in *Proceedings of the American Control Conference 2005*, pp. 973-978, 2005.

[2] S. P. Chakrabarty and F. B. Hanson," Optimal control of drug delivery to brain tumors for a test of PDE driven models using the Galerkin finite element Method", in *Proceedings of the Conference in Decision and Control and European Control Conference 2005*, pp. 1613-1618, 2005.

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

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

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

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

[7] 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.

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

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

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

[11] G. Sewell, *The Numerical Solution of Ordinary and Partial Differential Equations*, Academic Press, New York, 1988.

[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, "Probabilistic Rate Compartment Cancer Model: Alternate versus Traditional Chemotherapy Scheduling," in *Stochastic Theory and Control, Proceedings of a Workshop held in Lawrence, Kansas, October 18-20, 2001, Lecture Notes in Control and Information Sciences*, vol. 280, B. Pasik-Duncan (Editor), Springer-Verlag, New York, pp. 491-506, 2002.