

Estimating intratumoral heterogeneity from spatiotemporal data

E. M. Rutter · H. T. Banks · K. B. Flores

Received: September 10, 2017/ Accepted: date

Abstract Glioblastoma Multiforme (GBM) is a malignant brain cancer with a tendency to both migrate and proliferate. We propose modeling GBM with heterogeneity in cell phenotypes using a random differential equation version of the reaction-diffusion equation, where the parameters describing diffusion (D) and proliferation (ρ) are random variables. We investigate the ability to perform the inverse problem to recover the probability distributions of D and ρ using the Prohorov metric, for a variety of probability distribution functions. We test the ability to perform the inverse problem for noisy synthetic data. We then examine the predicted effect of treatment, specifically, chemotherapy, when assuming such a heterogeneous population and compare with predictions from a homogeneous cell population model.

Keywords Glioblastoma Multiforme · Random Differential Equation · Parameter Estimation

1 Introduction

The current biological understanding of cancer is that tumors are comprised of a genotypically diverse population of cells with plastic phenotypes and that treatment acts as an evolutionary pressure, selecting for therapy resistant sub-populations. Glioblastoma multiforme (GBM) is an aggressive, inherently heterogeneous brain tumor for which treatment is minimally effective: mean sur-

This research was supported in part by the Air Force Office of Scientific Research under grant number AFOSR FA9550-15-1-0298 and in part by the National Science Foundation under NSF grant number DMS-0946431

E. M. Rutter, H. T. Banks, and K. B. Flores
Center for Research in Scientific Computation
Department of Mathematics
North Carolina State University
E-mail: erutter@ncsu.edu, htbanks@ncsu.edu, kbflors@ncsu.edu

vival time with treatment is 12-15 months [41], nearly all GBM tumors recur [59], and it has a five-year survival rate of approximately 5% [30]. GBM is known for exhibiting both strong migratory and proliferative characteristics.

The simplest spatial model that can be used to describe the spread of GBM is the reaction-diffusion equation (Equation (1)) [40,54]. Although this model is clinically relevant [3,34], its major drawback is that it assumes isotropic diffusion, which does not describe the majority of tumor shapes. Therefore, it is imperative to include some method of accounting for anisotropy in tumor growth models. Methods of incorporating spatial dependence include using spatially-dependent piece-wise constant diffusion, diffusion tensor imaging, and assuming heterogeneous cell sub-populations of the tumor that exhibit different behavior. The presence of this phenotypic heterogeneity is the basis of the so called the “go or grow” hypothesis [33], and the estimation of such heterogeneity from spatiotemporal data is the aim of our investigation in this work.

We propose to perform an inverse problem to estimate the diffusion D , and growth rate ρ in the reaction-diffusion equation as distributions of a random differential equation rather than as point estimates for a deterministic differential equation. The inverse problem is performed using the Prohorov metric framework which has been developed and tested in other biological scenarios with phenotypic heterogeneity [16,17], and which makes no prior assumption of the underlying probability distribution of the parameters. We test our framework by generating simulated data for a reaction-diffusion tumor growth model with heterogenous sub-populations and attempt to recover the underlying probability distributions for the parameters describing the heterogeneity. Specifically, we generate data for the reaction-diffusion equation as a random differential equation, using a variety of distributions such as normal, lognormal, and bigaussian (corresponding to the “go or grow” hypothesis) to describe the growth rate (ρ) and the diffusion (D) parameters. We then use the Prohorov metric framework to estimate, i.e., attempt to recover, the distributions for ρ and D . We examine the cases where only one parameter is estimated as a distribution and then proceed by estimating in the case where both ρ and D are distributed parameters. We introduce observation noise into the solutions to determine how this perturbs the estimated distributions. We then model the effect of chemotherapy and compare predicted tumor burden using our random differential equation (i.e., a heterogeneous cell population) with predictions made by the reaction-diffusion equation.

We discuss the current status of GBM mathematical models and the importance of modeling tumor heterogeneity in Section 2. In Section 3 we introduce the Prohorov metric framework and its use to estimate parameter distributions. We follow with our results for synthetically generated data in Section 4, including the case where we recover both one and two parameter distributions. This is followed by discussing the ability to perform the inverse problem with noisy synthetically generated data in Section 5. We then carry out a comparison of drug treatment predictions with our model versus the standard

homogeneous population model in Section 6. We conclude in Section 7 with a discussion of future applications of this framework.

2 Mathematical Modeling Background

The normalized reaction-diffusion equation

$$\frac{\partial c(t, x)}{\partial t} = D \frac{\partial^2 c(t, x)}{\partial x^2} + \rho c(t, x)(1 - c(t, x)) \quad (1)$$

is frequently used as a parsimonious model for GBM growth and diffusion, where $c(t, x)$ represents the cell density at time t and location x , D is the diffusion coefficient, and ρ is the intrinsic growth rate. However, its simplistic formulation results in isotropic diffusion, which does not often match *in vivo* GBM growth. Several methods of altering Equation (1) have been developed to obtain anisotropic diffusion.

One method is incorporating spatially-dependent diffusion via a piecewise constant function (i.e., $D = D(x)$), where, for example, the diffusion constant differs spatially depending on if the spatial location is in grey matter or white matter [54, 57]. This allows the shape of the tumor to travel along white matter tracts at a faster rate. However, one must know the spatial distribution of grey and white matter within the brain, or approximate the distribution with a human brain atlas, to inform the diffusion function.

A second method of introducing anisotropic diffusion is to use diffusion weighted imaging (DWI), a neurological imaging technique using Magnetic Resonance Imaging (MRI). These data can be collected concurrently with T2-weighted or T1-weighted images used for diagnostic purposes. Diffusion weighted images, from 6 or more directions, can be used to create a diffusion tensor [60]. The eigenvalues of these tensors describe the anisotropic diffusion of water. One lingering question is how to use information, which describes water diffusion, to inform cellular diffusion. Multiple methods have been proposed [27, 28, 35, 38, 45], however this remains an area of active research. Recent work showed superior fits in nine out of ten patient cases as compared to the standard reaction-diffusion equation [53]. Although this may be the most biologically accurate model, DWI data is not always available, although some work has been performed using a human brain atlas [27, 37].

In vitro experimental work showed that the reaction-diffusion equation was insufficient to describe cell migratory patterns [51]. This led to the hypothesis that two distinct phenotypes are present in GBM: a proliferative population with a low diffusion rate and a migratory population that does not reproduce as much. This “go or grow” hypothesis has been present in many forms of biology, yet remains unproven [31]. However, there is experimental evidence that GBM cells move in a bursting fashion [29]. Hatzikirou et al. [33] introduced the concept of “go or grow” to mathematical modeling of GBM, theorizing that phenotypic switching between the two phenotypes was induced by hypoxia, rather than random mutation [44]. Modeling hypoxia-induced phenotypic switching

has been shown to replicate histologic patterns of pseudopalisades present in human GBM [39]. Recent work shows that one equation describing a population with phenotypic plasticity affected by cell density was sufficient to fit *in vitro* experimental data without assuming separate subpopulations [52].

The commonality between the previously mentioned mathematical models is their ability to include a non-constant diffusion rate. In this work, we assume there may exist heterogeneous subpopulations of cells which behave differently from one another, but that we are only able to observe the aggregate (total) population data. We propose to describe GBM cell phenotypic heterogeneity by using parameter distributions for the parameters ρ and D . The random differential equation governing diffusion and growth is:

$$\frac{\partial c(t, x, \mathbf{D}, \rho)}{\partial t} = \nabla \cdot (D \nabla c(t, x, \mathbf{D}, \rho)) + \rho c(t, x, \mathbf{D}, \rho)(1 - c(t, x, \mathbf{D}, \rho)) \quad (2)$$

We assume that the parameters \mathbf{D} and ρ are random variables defined on a compact set $\Omega = \Omega_{\mathbf{D}} \times \Omega_{\rho}$. The distribution of the parameters is given by $P(\mathbf{D}, \rho)$, and $c(t, x)$ represents the aggregate population observable (which is defined as the expectation over subpopulations $c(t, x, \mathbf{D}, \rho)$):

$$c(t, x) = \mathbb{E}[c(t, x, \cdot, \cdot), P] = \int_{\Omega} c(t, x, \mathbf{D}, \rho) dP(\mathbf{D}, \rho) \quad (3)$$

The random differential equation (2) is sufficiently flexible to model the classic reaction-diffusion equation (assuming a point distribution of \mathbf{D} and ρ) and a “go or grow” type equation (assuming, for example, a bi-gaussian distribution of \mathbf{D} and ρ).

3 Prohorov Metric

An inverse problem framework relying on the Prohorov metric, called the Prohorov Metric Framework (PMF) was developed and discussed in [6, 21, 23]). Here, we describe how this framework can be utilized for estimating the continuous probability measure $P(\mathbf{D}, \rho)$ as a finite-dimensional approximation using aggregate population data. Previous authors have used the Prohorov metric to estimate parameter distributions for mathematical models in a wide range of applications. These include estimating the distribution of individual growth rates for size-structured population models (first order hyperbolic partial differential equations) using aggregate or population level data. For example, efforts have involved mosquitofish [8, 11, 15, 16] and shrimp populations [12–14]. Such problems involve equations modeling the number of individuals $v(t, \xi; g)$ at time t and size ξ for a given subpopulation of mosquitofish or shrimp with growth rate g , however, the observable data are *aggregate* measurements of the total population level. These type of aggregate data are common in ecological studies with *catch and release* measurements [24], where one samples at different times from the same population but the change in the structured variable, e.g., size, for each individual is not tracked over time. More recently

these techniques have been used in determining heterogeneous amplification rates in protein aggregate formation in structured population models [17]. In this instance, a first order system of partial differential equations is used with the PMF to model dynamics of a structured variable, i.e., prion aggregates per cell, within a population of cells. Such type of problems are also typical in experiments where the organism or population member being studied is sacrificed in the process of making a single observation (e.g., certain physiologically based pharmacokinetic (PBPK) modeling [26] and whole organism transport models [24]). In these cases one may still have longitudinal models for individuals as in the mosquitofish problem, but no individual level data are available. In more complex systems involving second order hyperbolic systems, the PMF has been used with Maxwell systems for complex polarization models of heterogeneous materials [9, 18]. Second order hyperbolic systems are also the focus of the shear wave propagations studies in [7, 19, 20, 25]. In our efforts here we propose use of the PMF techniques for estimation of random variables representing rates governing a reaction-diffusion system.

Two different methods for approximating the probability measure $P(\mathbf{D}, \boldsymbol{\rho})$ are using either delta functions or spline functions. Although using spline functions are known to yield more accurate convergence in the probability density function (PDF) and cumulative distribution function (CDF), delta functions are able to better approximate CDFs that have discontinuous derivatives. Therefore, we illustrate use of both approximations, since we do not wish to make any assumptions about, or restrictions on, the CDF.

Suppose that the aggregate spatiotemporal data we want to model is given by v_{ji} , representing the data at time j and spatial location i , where $j = 1, \dots, N_t$ and $i = 1, \dots, N_x$. Then, we estimate:

$$\hat{P} = \operatorname{argmin}_{P^M(\Omega)} \sum_{j,i=1}^{N_t, N_x} (v_{ji} - v(t_j, x_i; P))^2 \quad (4)$$

where M represents the number of elements used in the approximation (explained in the sections below). This becomes:

$$\hat{P} = \operatorname{argmin}_{P^M(\Omega)} \sum_{j,i} \left(v_{ji} - \int_{\Omega} c(t_j, x_i; \mathbf{D}, \boldsymbol{\rho}) dP(\mathbf{D}, \boldsymbol{\rho}) \right)^2 \quad (5)$$

where c is the numerical solution. Previous experimental data [51] suggests that it is not unreasonable to assume that data can be collected radially at spatial increments of 40 microns, daily. This is the basis for describing the data in the form of v_{ji} . Two methods are used to approximate the probability measure P : a discrete approximation based on delta functions, and a continuous approximation using spline basis functions.

3.1 Discrete Approximation

One-Parameter Approximation We first describe how to estimate one parameter as a distribution using the delta function approximation. Essentially, we define a discrete mesh for a parameter, in this case ρ , over M nodes, such that $\boldsymbol{\rho}^M = \{\Delta_{\rho_l}, l = 1, \dots, M\}$. Thus, Equation (5) simplifies to:

$$\hat{P} = \operatorname{argmin}_{\mathbb{R}} \sum_{j,i} \left[v_{ji} - \left(\sum_l c(t_j, x_i; D, \rho_l) w_l^M \right) \right]^2 \quad (6)$$

where the numbers $w_l^M \geq 0$ are weights describing a discrete probability density function, i.e., such that $\sum_{l=1}^M w_l^M = 1$. We will refer to a discrete approximation with M nodes as DEL(M).

Similarly, if we wish to estimate only the parameter D as a distribution, we define a mesh for D over M nodes, such that $\boldsymbol{D}^M = \{\Delta_{D_l}, l = 1, \dots, M\}$. We would then solve:

$$\hat{P} = \operatorname{argmin}_{\mathbb{R}} \sum_{j,i} \left[v_{ji} - \left(\sum_l c(t_j, x_i; D_l, \rho) w_l^M \right) \right]^2 \quad (7)$$

under the assumption $\sum_{l=1}^M w_l^M = 1$.

The number of nodes, M , can greatly influence the convergence of the approximating CDFs. By incorporating too few nodes, we may not adequately recover the CDF, but by incorporating too many nodes, we may introduce numerical instability and over-fitting. Therefore, we simulate our results for many choices of M .

Two-Parameter Estimations We next describe how to estimate the two parameters ρ and D as distributions. We make the assumption of independence between the distributions of the random variables $\boldsymbol{\rho}$ and \boldsymbol{D} .

Similarly to the one-parameter case, we generate a uniform mesh for ρ and D over the total number of nodes. Therefore we have $\boldsymbol{D}^{M_D} = \{\Delta_{D_l}, l = 1, \dots, M_D\}$, and $\boldsymbol{\rho}^{M_\rho} = \{\Delta_{\rho_k}, k = 1, \dots, M_\rho\}$. We do not require $M_D = M_\rho$. We then independently find the weights for each parameter by minimizing:

$$\hat{P} = \operatorname{argmin}_{\mathbb{R}} \sum_{j,i} \left[v_{ji} - \left(\sum_{l,k} c(t_j, x_i; D_l, \rho_k) w_l^{M_D} w_k^{M_\rho} \right) \right]^2 \quad (8)$$

under the assumption $\sum_{l=1}^{M_D} w_l^{M_D} = 1$ and $\sum_{k=1}^{M_\rho} w_k^{M_\rho} = 1$. We refer to a discrete approximation of two variables as DEL(M_D, M_ρ).

3.2 Spline-Based Approximations

An alternative method to using delta functions involves using spline basis functions to approximate a probability density function. Similar to the discrete approximations, we choose a certain number of nodes, but we also choose the number of integrative nodes – that is, the number of nodes we use to approximate the integral. The splines are “hat functions” defined as:

$$s_l(\xi) = \begin{cases} \frac{\xi - \xi_{l-1}}{\xi_l - \xi_{l-1}} & \text{if } \xi \in [\xi_{l-1}, \xi_l] \\ \frac{\xi_{l+1} - \xi}{\xi_{l+1} - \xi_l} & \text{if } \xi \in [\xi_l, \xi_{l+1}] \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

One-Parameter Approximation We first describe how to estimate one parameter as a continuous distribution. Essentially, we generate a mesh for a random variable describing the parameter, in this case $\boldsymbol{\rho}$, over M nodes, such that $\boldsymbol{\rho}^M = \{s_l(\boldsymbol{\rho}), l = 1, \dots, M\}$. Thus Equation (5) simplifies to:

$$\hat{P} = \underset{\mathbb{R}}{\operatorname{argmin}} \sum_{j,i} \left[v_{ji} - \left(\sum_l a_l \int_{\Omega_\rho} c(t_j, x_i; D, \boldsymbol{\rho}) s_l(\boldsymbol{\rho}) d\rho \right) \right]^2 \quad (10)$$

where $p_l = a_l s_l(\boldsymbol{\rho})$ represents the probability density. Therefore, we require that $\sum_{l=1}^M a_l \int_{\Omega_\rho} s_l(\boldsymbol{\rho}) d\rho = 1$. The $s_l(\boldsymbol{\rho})$ are calculated as in Equation (9). In order to calculate the integral, we use the composite trapezoid rule using Q quadrature nodes. We will refer to a spline approximation with M nodes and Q quadrature nodes as SPL(M,Q).

A similar formulation is used for the diffusion coefficient. We partition D over M nodes, such that $\mathbf{D}^M = \{s_l(\mathbf{D}), l = 1, \dots, M\}$. We then find the probability distribution by minimizing the following:

$$\hat{P} = \underset{\mathbb{R}}{\operatorname{argmin}} \sum_{i,j} \left[v_{ij} - \left(\sum_l a_l \int_{\Omega_D} c(t_j, x_i; \mathbf{D}, \rho) s_l(\mathbf{D}) dD \right) \right]^2 \quad (11)$$

Two-Parameter Approximation As in the discrete case, we assume that $\boldsymbol{\rho}$ and \mathbf{D} are independent random variables. Because of this assumption, we can individually estimate the distributions. We generate a uniform mesh for ρ and D over the total number of nodes. Therefore we have $\mathbf{D}^{M_D} = \{s_l(\mathbf{D}), l = 1, \dots, M_D\}$, and $\boldsymbol{\rho}^{M_\rho} = \{s_k(\boldsymbol{\rho}), k = 1, \dots, M_\rho\}$. As in the discrete case, we do not require $M_D = M_\rho$. We then minimize:

$$\hat{P} = \underset{\mathbb{R}}{\operatorname{argmin}} \sum_{i,j} \left[v_{ij} - \sum_l a_l \int_{\Omega_\rho} \left(\sum_k b_k \int_{\Omega_D} c(t_j, x_i; \mathbf{D}, \boldsymbol{\rho}) s_l(\mathbf{D}) dD \right) s_k(\boldsymbol{\rho}) d\rho \right]^2 \quad (12)$$

We require that $\sum_{l=1}^{M_D} a_l \int_{\Omega_D} s_l(\mathbf{D}) d\mathbf{D} = 1$ and $\sum_{k=1}^{M_\rho} b_k \int_{\Omega_\rho} s_k(\boldsymbol{\rho}) d\rho = 1$. As with the 1D case, we perform the approximation of each of the integrals using the composite trapezoid rule with Q_D and Q_ρ quadrature nodes, respectively. The number of quadrature nodes Q_D and Q_ρ do not need to be the same. We will refer to this approximation as $\text{SPL}(M_D, M_\rho, Q_D, Q_\rho)$.

3.3 Convergence and Consistency Theory

The Prohorov metric represents the weak* convergence of measures when the space of probability measures $P(\Omega)$ is imbedded in the dual $C^*(\Omega)$ of the space of bounded continuous functions on Ω . We discuss briefly the convergence and consistency theory here, assuming that we are only estimating \mathbf{D} as a distribution, but we note the theory extends to two parameters. We assume a family of permissible probability functions for our diffusion and growth rates on $P(\Omega)$ as defined for Equation (2). Thus we attempt to perform the following estimation in a least-squares framework

$$\hat{P} = \underset{P \in P(\Omega)}{\operatorname{argmin}} \sum_{j,i} (v_{ji} - c(t_j, x_i; P))^2 \quad (13)$$

to obtain a best fit for a nominal or “true” parameter P_0 . To illustrate for our special 1-D case, we let $\Omega = \Omega_{\mathbf{D}}$ be the continuum of values on $[0, D_{\max}]$. This implies that the family of probability functions $P(\Omega)$ is a compact metric space in the Prohorov metric. The minimizer function used in Equation (13) is continuous in P , and therefore there exists a (not necessarily unique) minimizer \hat{P} .

In order to approximate this minimizer, we replace the infinite dimensional optimization problem by a sequence of finite-dimensional optimization problems with Dirac distributions. Thus, we set $\Omega^M = \{\Delta_{D_k}, k = 1, \dots, M\}$, where M represents the number of nodes, or elements, used in the approximation. Our family of approximating probability functions becomes

$$P^M(\Omega^M) = \left\{ P^M = \sum_{k=1}^M w_k \Delta_{D_k} \mid w_k \geq 0 \text{ and } \sum_{k=1}^M w_k = 1 \right\},$$

where Δ_{D_k} represent the Dirac delta functions at the point D_k and w_k are the weights and/or probabilities. Again it has been previously proven [6, 21] that there exists a minimizer for the discrete approximation problem

$$\hat{P}^M = \underset{P \in P^M(\Omega^M)}{\operatorname{argmin}} \sum_{j,i=1}^{N_t, N_x} (v_{ji} - c(t_j, x_i; P))^2. \quad (14)$$

There are a number of questions that arise immediately in the class of problems we have defined. Perhaps the most obvious are questions of convergence (what happens as $M \rightarrow \infty$ in the Dirac or spline approximations?) and

Table 1 Information for the various parameter distributions of \mathbf{D} and $\boldsymbol{\rho}$. When applicable, values are reported as mean (standard deviation).

Distribution Name	\mathbf{D}	$\boldsymbol{\rho}$
Gaussian	1e-4 (2e-5)	1 (0.25)
Bigaussian	[5e-5 (5e-6), 1.25e-4 (1e-5)]	[0.25 (0.05), 1.25 (0.2)]
Point	1e-4	1
Two-Point	[1e-5, 1.5e-4]	[0.5, 1.5]
Lognormal	-10.6484 (0.3782)	-0.5 (0.3)

consistency (what happens as $N = (N_t, N_x) \rightarrow \infty$?) These questions have been successfully investigated both theoretically ([21,25] and the references therein) and computationally ([10,11] and the references). A further issue (as mentioned above) involves the partial differential equation approximations $c_{\tilde{N}}$ to the solution c of Equation (2). Again, the necessary convergence issues have been successfully addressed in [21,24]. In summary we can assert that the approximations $\hat{P}_{N,\tilde{N}}^M$ converge to a true distribution P_0 as the number of elements used in the approximations increase (i.e., $M, N, \tilde{N} \rightarrow \infty$).

4 Results for Distributed Parameters

In the following section we describe results showing the Prohorov metric can correctly estimate parameter distributions for both \mathbf{D} and $\boldsymbol{\rho}$ individually, as well as estimating a joint distribution for \mathbf{D} and $\boldsymbol{\rho}$. Our goal is to ensure that we can adequately recover these parameter distributions.

Within each distribution, we use both the delta function approximations as well as the spline approximations. We also investigate the number of nodes necessary to generate a good fit.

Generation of Synthetic Data. To first prove the viability of recovering parameter distributions, we generate synthetic data by solving the RDE (Equation (2)) for a variety of different parameter distributions. To ensure a strong connection to the biological processes, we examine a wide family of biologically-relevant distributions such as normal, lognormal, bigaussian, point, two point, and a distribution made up of a point and a normal distribution.

Table 1 displays the associated information with the parameter distributions for both \mathbf{D} and $\boldsymbol{\rho}$. The synthetic data was generated by finely meshing over each parameter, $D \in [0, 2e-4]$ and $\rho \in [0, 2]$, and then creating the probability density function (pdf). For discrete probability distributions (point, 2 point), a probability mass function (pmf) was created. We then solved the following:

$$\frac{\partial c(t, x, D_t, \rho_k)}{\partial t} = D_t \frac{\partial^2 c}{\partial x^2} + \rho_k c(1 - c) \quad (15)$$

We then weight this solution to calculate our synthetic data:

$$v_{ij} = \sum_{l,k} c(t, x, D_l, \rho_k) p(D_l, \rho_k) \quad (16)$$

where $p(D_l, \rho_k)$ represents the probability (pdf) of $D = D_l$ and $\rho = \rho_k$. We finely mesh over \mathbf{D} and $\boldsymbol{\rho}$.

Numerical Setup. We use MATLAB's built-in pde solver, PDEPE to solve the partial differential equations. We perform the minimization (Equation (5)) using constrained optimization (FMINCON), requiring that all probabilities are non-negative and that the sum of probabilities equal 1: $\sum_{k=1}^M w_k = 1$ in the discrete case and $\sum_{k=1}^M a_k \int_{\Omega_\rho} s_k(\boldsymbol{\rho}) d\rho = 1$ in the spline case.

To ensure that the amount of data necessary to recover the distributions are reasonable, we discretize the time as $t = 1, 2, \dots, 20$ days (i.e., experimental data are collected once per day). We solve the pde on a fine spatial structure, but we assume data are collected every 40 microns (as previously mentioned), so our comparison occurs at these nodes.

For each of these cases, we need to quantify the necessary hyperparameters (nodes, quadrature nodes) that generate the best fit. This inverse problem requires us to consider each *node* as a separate parameter that we are estimating. Therefore, we need to assess model fit versus number of parameters used. We will use the Akaike Information Criteria (AIC) [2] in order to compare the various models. In the context of least-squares framework, the AIC [22] is given by:

$$AIC = N\nu \ln \left(\frac{RSS}{N\nu} \right) + N\nu(1 + \ln(2\pi)) + 2(p + 1) \quad (17)$$

where N is the number of data points, ν is the number of observables, RSS represents the residual sum of squares between our solution $c(t, x)$ and the synthetically generated solution, and p is the number of parameters (i.e., nodes) being estimated. Note that the number of quadrature nodes in the spline case are not considered free parameters since these are not estimated during the inverse problem.

4.1 Estimating \mathbf{D}

We assume that we already know the value of ρ and are only estimating the distribution for \mathbf{D} . Figure 1 depicts the AIC scores versus number of nodes for the discrete approximation (left) and spline approximation (right). For the spline approximations, solid lines represent 50 quadrature nodes, while the dashed lines are 100 quadrature nodes. For the spline approximations, it appears that 30 nodes is ideal for estimating the parameter distribution for

D. For the discrete approximations, it appears that more nodes are needed to ensure a good fit. Since the various parameter distributions use different data for estimation, we cannot compare the AIC scores from one distribution to the next.

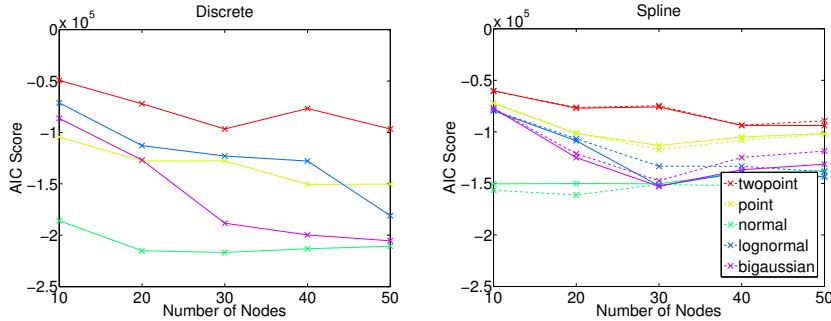


Fig. 1 AIC scores with respect to number of nodes for approximating \mathbf{D} using discrete approximations (left) and spline approximations (right). Dashed lines represent 100 quadrature nodes, while solid lines are 50 quadrature nodes.

Although we cannot compare the AIC scores between distributions, we can compare the AIC scores for the same distribution (i.e., we can compare the normal discrete approximation of \mathbf{D} with the normal spline approximations for \mathbf{D}). We can see that the discrete approximation appears to perform better for almost all distributions. There also appears to be little difference between using a spline approximation with 50 quadrature nodes and using a spline approximation with 100 quadrature nodes.

4.2 Estimating ρ

The AIC scores with respect to number of nodes for estimating ρ are shown in Figure 2. For the discrete approximation, it appears that fewer than 20 nodes does not provide a good approximation, however, increasing the number of nodes past 30 does not appear to provide much benefit. Therefore, due to the computational costs associated with increasing the number of nodes, we would choose to estimate ρ with 30 nodes. For the spline approximation, it appears that 20 nodes is sufficient to provide (in general) an accurate solution.

We see that, especially for the discrete distributions (point and two point) the discrete approximation performs better than the spline approximations. The spline approximations with 50 quadrature nodes and 100 quadrature nodes are essentially the same in most cases, corroborating what we found for estimates of \mathbf{D} in the previous section.

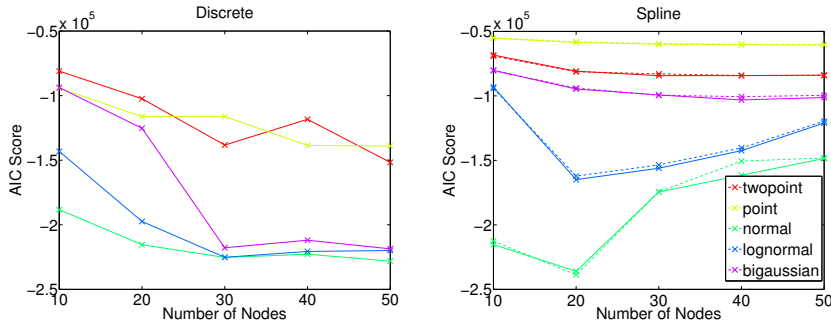


Fig. 2 AIC scores with respect to number of nodes for approximating ρ using discrete approximations (left) and spline approximations (right). Dashed lines represent 100 quadrature nodes, while solid lines are 50 quadrature nodes.

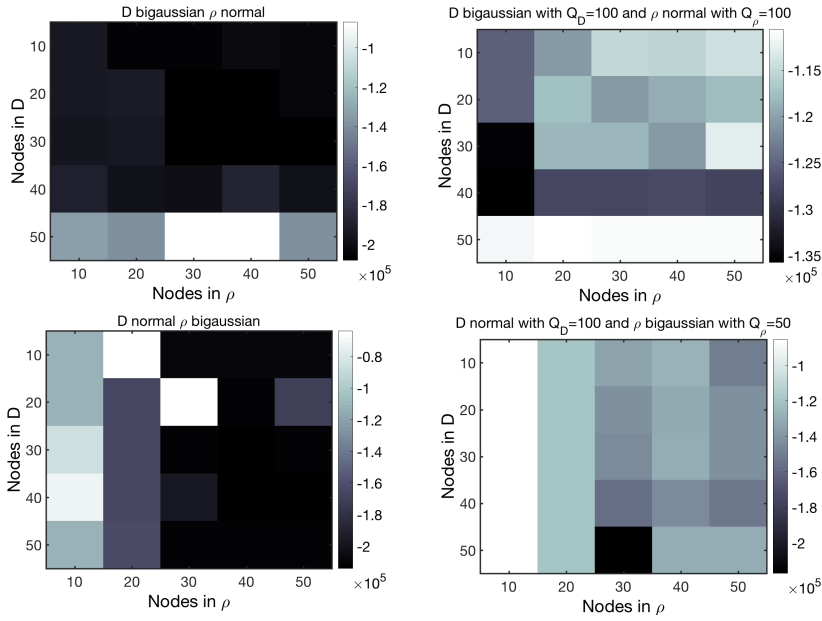


Fig. 3 Representative examples of the AIC scores with respect to number of nodes for approximating D and ρ using discrete approximations (left) and spline approximations (right) for the case where D is bigaussian and ρ is normal (top) and the case when D is normal and ρ is bigaussian (bottom).

4.3 Estimating D and ρ

We extend the optimization problem to estimate both D and ρ as distributions for both the discrete case and the continuous case. We focus on two representative examples (D is bigaussian and ρ is normal and vice versa). See Table 1 for the exact values of the distributions. We present the AIC score surfaces for the discrete case and the spline cases in Figure 3. The optimal number of

nodes for a bigaussian and normal distribution for \mathbf{D} and ρ , respectively, using a discrete approximation is 30 (for both \mathbf{D} and ρ), while the optimal number of nodes for the continuous case is 10 for ρ and 20 for \mathbf{D} . We note in this case the minimum AIC score occurs using the discrete approximation rather than the continuous spline approximation. For the case when \mathbf{D} normal and ρ is bigaussian, the optimal number of nodes in the discrete case is 40 for ρ and 20 for \mathbf{D} , while the optimal number of nodes for the continuous case is 30 for ρ and 10 for \mathbf{D} . In this case, the minimal AIC score is comparable for the spline and discrete approximations.

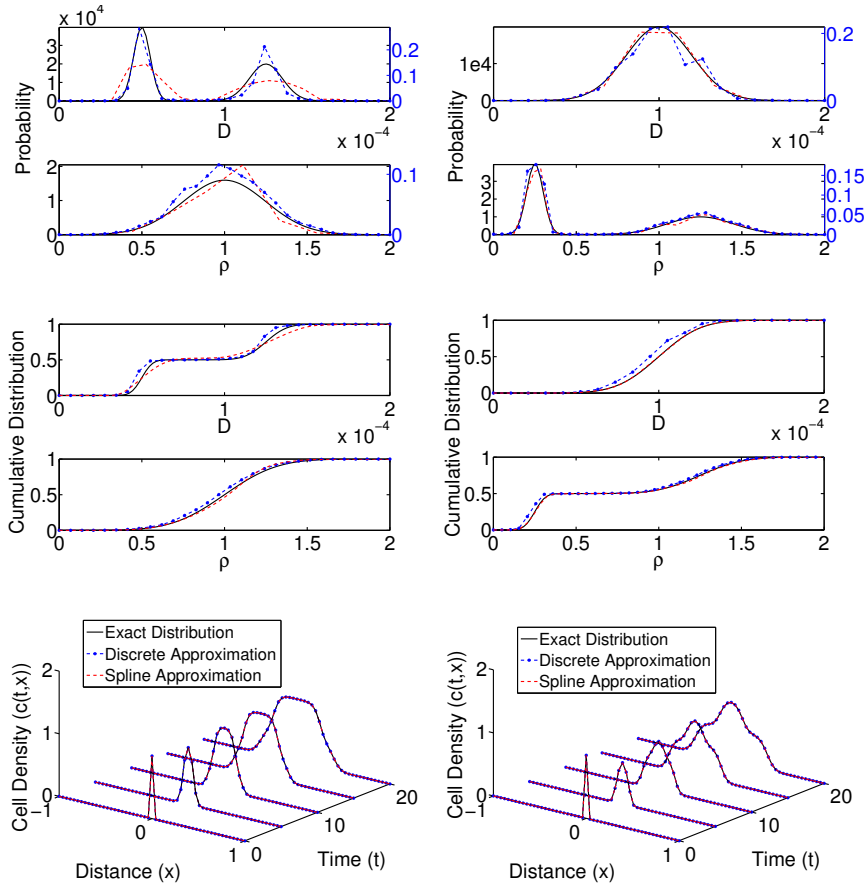


Fig. 4 Fits using the optimal number of nodes (determined by AIC scores) when ρ is bigaussian and \mathbf{D} is normally distributed (left) and when \mathbf{D} is bigaussian and ρ is normally distributed (right). The top panels display the pdf comparisons for the actual distribution (black), the spline approximation (red) and the discrete approximation (blue). The axis scale for the discrete approximation in the top panel is on the right side of the plot. The middle panel depicts the cdf comparisons. The bottom panel reveals the solutions of the RDE.

Figure 4 exhibits the best fits determined by AIC scores when \mathbf{D} is normally distributed and ρ is bigaussian (Figure 4, left), and when \mathbf{D} is bigaussian and ρ is normally distributed (Figure 4, right). The top panel reveals the estimated probability density functions, the middle panel has the cumulative density functions, and the bottom panel depicts the solutions to the RDE. We found that the solutions to the RDE match for both choices of joint distributions for ρ and \mathbf{D} . Similarly, the fits for the cdf and pdf appear to match well. Note that the axis scale for the probability density functions are different for the discrete node approximation. This is because we are not guaranteed convergence in pdf for the discrete node approximation, rather we are guaranteed convergence in cdf for the discrete node approximation.

For the case where \mathbf{D} is bigaussian and ρ is normally distributed (Figure 4, left), we found that the pdf and cdf agree best with the true distributions when using the discrete node approximation. The discrete approximation also had much lower AIC scores than the spline approximations. In the case where \mathbf{D} is normally distributed and ρ is bigaussian (Figure 4, right), both the spline approximations and the discrete approximations estimate the pdf and cdf well. This is unsurprising because the AIC scores were very similar between the discrete node approximation and the spline approximation.

One question of interest is what are the spatiotemporal dynamics of the distribution of the relative subpopulations. Figure 5 shows the subpopulations in the case where diffusion is bigaussian and growth is normally distributed (Figure 5, left) and the case where diffusion is normally distributed and growth is bigaussian (Figure 5, right). In these simulations, we used discrete approximations for the parameter distributions, based on our findings above that this was more accurate than using spline approximations.

Subpopulations for the \mathbf{D} bigaussian and ρ normally distributed were calculated by splitting the computed probability density function for \mathbf{D} into two halves and independently simulating their solutions. “Slow-diffusing” represents those cells with the smaller half of the pdf (the first part of the bigaussian) and “Fast-diffusing” represents the second half of the pdf. Similar computations were carried out for the case where ρ is bigaussian and \mathbf{D} was normally distributed. As can be seen in Figure 5, these subpopulations exhibit different behavior that is not recovered by assuming static parameters ρ and D , shown in magenta. The reaction-diffusion solution was obtained by estimating the parameters D and ρ for Equation (1) to the synthetic data using FMINCON under the same computation setup defined for the RDE inverse problem. Especially in the case where ρ is bigaussian (Figure 5, right), there is a large difference in the solutions when assuming a reaction-diffusion type model.

5 Noisy Data

We repeat the work performed in Section 4, except assuming the data (exact solution) are noisy. To generate these noisy data, we perturb the system

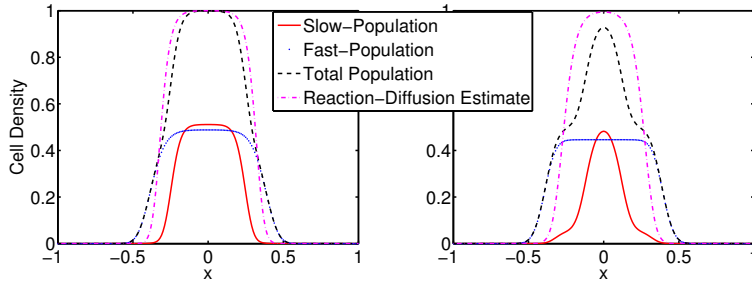


Fig. 5 Spatiotemporal distributions for cell subpopulations in the case where \mathbf{D} bigaussian and $\boldsymbol{\rho}$ normally distributed (left), and the case where \mathbf{D} normally distributed and $\boldsymbol{\rho}$ bigaussian (right).

under the assumption of a proportional error model. The proportional error model assumes that the noise level is proportional to the size of the model solution, which commonly occurs in biological scenarios in general (see [4, 1] for examples). In order to generate these errors, we modify Equation (16) to incorporate noise:

$$\bar{v}_{ij} = v_{ij} + \epsilon_{ij}v_{ij}; \quad (18)$$

where the residuals ϵ_{ij} are distributed as $\epsilon \sim 0.05\mathcal{N}(0, 1)$.

In order to perform the inverse problem, we use a generalized least squares framework. The generalized least squares framework assumes a proportional error model:

$$C(t_j, x_i) = \mathbb{E}[c(t_j, x_i; \mathbf{D}, \boldsymbol{\rho})] + \mathbb{E}[c(t_j, x_i; \mathbf{D}, \boldsymbol{\rho})]^\gamma \epsilon_{ij} \quad (19)$$

where ϵ_{ij} is assumed to be independent and identically distributed with mean 0. We then estimate the parameter distributions by performing:

$$\hat{P} = \operatorname{argmin}_{P^M(\Omega)} \sum_{i,j} w_{ij} \left(\bar{v}_{ij} - \int_{\Omega} c(t_j, x_i; \mathbf{D}, \boldsymbol{\rho}) dP(\mathbf{D}, \boldsymbol{\rho}) \right)^2 \quad (20)$$

where w_{ij} are weights for the least-squares calculation and M represents the number of elements used in approximating \mathbf{D} and $\boldsymbol{\rho}$.

Although we are mainly concerned in this paper with the mathematical model rather than the statistical error models, we mention that similar convergence and consistency results discussed in Section 3.3 can be considered in the context of both [21] an uncertain mathematical model (assumed a major question here) and a statistical model (assumed here as in Equation (20)).

We must perform an iterative algorithm in order to update the weights with the updated solution. We present the iterative algorithm for the discrete approximation case:

1. Initialize weights $w_{ij}^0 = 1$. Set iteration number, $\text{iter} = 0$.

2. Estimate parameters in equation (20):

$$\hat{p}^{\text{iter}} = \underset{p \in \mathbb{R}}{\operatorname{argmin}} \sum_{i,j} w_{ij}^{\text{iter}} \left(\bar{v}_{ij} - \sum_{l,k} c(t_j, x_i; \mathbf{D}^{M_D}, \boldsymbol{\rho}^{M_\rho}, p) \right)^2 \quad (21)$$

where \mathbf{D}^{M_D} and $\boldsymbol{\rho}^{M_\rho}$ are defined as in previous sections ($\mathbf{D}^{M_D} = \{\Delta_{D_l}, l = 1, \dots, M_D\}$, and $\boldsymbol{\rho}^{M_\rho} = \{\Delta_{\rho_k}, k = 1, \dots, M_\rho\}$), and $p = [p_D, p_\rho]$, where p_D and p_ρ are vectors of the estimated probability densities of \mathbf{D} and $\boldsymbol{\rho}$ respectively (meshed over \mathbf{D}^{M_D} and $\boldsymbol{\rho}^{M_\rho}$).

3. Determine the weights. Since many of the data points that we are estimating are essentially zero (i.e., no tumor cell density), we use a truncated version of our function as a weight, as introduced in [5], to avoid numerical instability:

$$w_{ij}^{\text{iter}+1} = \begin{cases} 0, & c(t_j, x_i; \mathbf{D}^{M_D}, \boldsymbol{\rho}^{M_\rho}, \hat{p}^{\text{iter}}) < C^* \\ \frac{1}{c(t_j, x_i; \mathbf{D}^{M_D}, \boldsymbol{\rho}^{M_\rho}, \hat{p}^{\text{iter}})^2}, & c(t_j, x_i; \mathbf{D}^{M_D}, \boldsymbol{\rho}^{M_\rho}, \hat{p}^{\text{iter}}) \geq C^* \end{cases} \quad (22)$$

4. Repeat steps 2-3 for at least 10 iterations (iter=10) and until we obtain convergence defined as:

$$\frac{1}{(M_D + M_\rho)} \sum_{\text{index}=1}^{M_D+M_\rho} \frac{|\hat{p}_{\text{index}}^{\text{iter}} - \hat{p}_{\text{index}}^{\text{iter}-1}|}{\hat{p}_{\text{index}}^{\text{iter}}} < \delta \quad (23)$$

In our simulations, we discovered that $C^* = 1e - 4$ still resulted in converged results. We adaptively changed our choice of δ : we initially set $\delta = 0.1$, however, if the number of iterations increased to greater than 200, we set $\delta = 0.5$. In the case of the spline approximations, we only alter Equation (21) to the spline approximations defined in Section 3.2.

Estimating \mathbf{D} and $\boldsymbol{\rho}$. We examine only the ability of estimating the distributions for \mathbf{D} and $\boldsymbol{\rho}$ jointly with the assumption that if we are able to recover both distributions at the same time, we would be able to recover the distributions independently.

Figure 6 exhibits the pdf, cdf, and best-fit solutions for the noisy case where \mathbf{D} is bigaussian and $\boldsymbol{\rho}$ is normally distributed (Figure 6, left). The AIC scores, calculated as in Section 4, revealed that the discrete approximation with 10 nodes in $\boldsymbol{\rho}$ and 20 nodes in \mathbf{D} resulted in the best approximation. The optimal AIC score for the spline case occurred with 10 nodes in $\boldsymbol{\rho}$ and 20 nodes in \mathbf{D} with 100 quadrature nodes in each $\boldsymbol{\rho}$ and \mathbf{D} . The RDE solutions using estimated parameters appear to match the noisy solution well and both the pdf and cdfs appear to match the true distributions. We note that (results not shown), in the case of spline approximations, if we choose to use a larger number of nodes, we quickly introduce numerical instability.

Figure 6 also portrays the pdf, cdf, and best-fit solutions for the noisy case where \mathbf{D} is normally distributed and $\boldsymbol{\rho}$ is bigaussian (Figure 6, right).

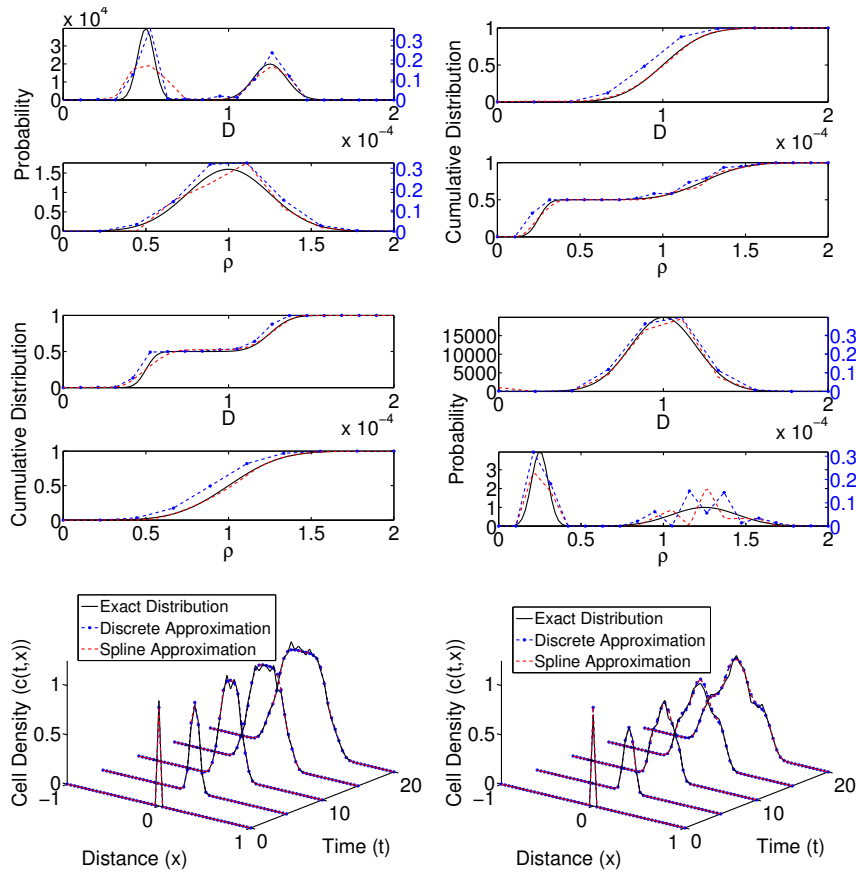


Fig. 6 Fits using the optimal number of nodes (determined by AIC scores) when D is bigaussian and ρ is normally distributed (left) and the case where ρ is bigaussian and D is normally distributed (right) for solutions which have added noise. The top panels display the pdf comparisons for the actual solution, the spline approximation and the discrete approximation. In the middle, we compare the cdfs. On the bottom we present the solutions of the RDE.

The AIC scores determined the best fit occurs when there are 10 nodes in D and 20 nodes in ρ for both the discrete and spline case. For the spline approximations, quadrature nodes of 100 for both Q_D and Q_ρ resulted in the best fit. We see that, in this case, the exact pdfs are not fully recovered (there are some oscillations in the second mode of the bigaussian). However, the overall behavior, and specifically the means of each mode, are accurate. We hypothesize that, in order to fit the mode of the bigaussian with small variance, we need a larger number of nodes, but that this increased number of nodes overfits the portion of the bigaussian with wide variance. We were able to show better agreement between the estimated and true distributions when choosing 16 nodes for ρ (results not shown). This implies for a multimodal distribution with differing variances fitting to noisy data, the number

of nodes and discretization mesh may need to be adaptive, i.e., non-uniform. The investigation of using an adaptive mesh for performing the parameter distribution estimation is left for future work.

6 Effect of Drug Administration

Mathematical models are often used to predict the change in growth of gliomas in response to treatment such as chemotherapy [46], radiotherapy [48], resection [55], viralththerapy [47], or combinations of the above [32,58]. However, most of these models assume a homogeneous population of tumor cells with static values of D and ρ . It is known that chemotherapy selectively targets proliferative tumor cells, i.e., cells with larger values of ρ . Popular methods of mathematically incorporating chemotherapy response includes the log-kill hypothesis [50] and the Norton-Simon hypothesis [42,43], among others [36]. In this section, we compare the predicted total tumor size under Equation (1) as compared with the various distributed parameter solutions, i.e., using random differential equations.

We model treatment assuming a variation on the log-kill hypothesis, which assumes that the number of cells killed by chemotherapy is proportional to the total number of cells. This is modeled by the following equation:

$$\frac{\partial c(t, x)}{\partial t} = \mathbf{D} \frac{\partial^2 c(t, x)}{\partial x^2} + \boldsymbol{\rho} c(t, x)(1 - c(t, x)) - r \frac{\boldsymbol{\rho}}{\bar{\rho}} c(t, x) \quad (24)$$

where $\bar{\rho}$ is the mean value of $\boldsymbol{\rho}$, and r is the coefficient regarding the drug treatment strength (it is assumed that treatment is administered at a constant rate for simplicity). This is an extension of the log-kill hypothesis [50], which simplifies to the original log-kill hypothesis for the basic reaction-diffusion model, Equation (1), when $\boldsymbol{\rho} = \bar{\rho}$. We keep the death due to drug administration constant, as previously modeled in [56], for simplicity.

We integrate this solution over space to obtain the total tumor burden (tumor size) as a function of time. We wish to see how predictions of tumor burden differ when assuming homogeneous populations *if* the population is inherently heterogeneous (modeled as the 'exact solution' calculated over the finely meshed D and ρ as described in Section 4). We also use the solutions and parameters generated by the noisy solutions computed in Section 5, although we note that the conclusions drawn from this noisy data set also hold for the noise-free data.

Figure 7 exhibits the comparison between the predicted tumor burden with respect to time for both the RDE model (Equation (24)), in red dashed lines with circles, the RD model (Equation (1)), in blue solid lines with plusses, and the "exact solution" (described above), shown in black solid line, which mostly overlaps the RDE solution. These simulations of tumor burden are for the case where \mathbf{D} is bigaussian and $\boldsymbol{\rho}$ is normal (Figure 7, left) and the case where \mathbf{D} is normally distributed and $\boldsymbol{\rho}$ is bigaussian (Figure 7, right) for varying values of the chemotherapy-induced death rate, r . We observed that

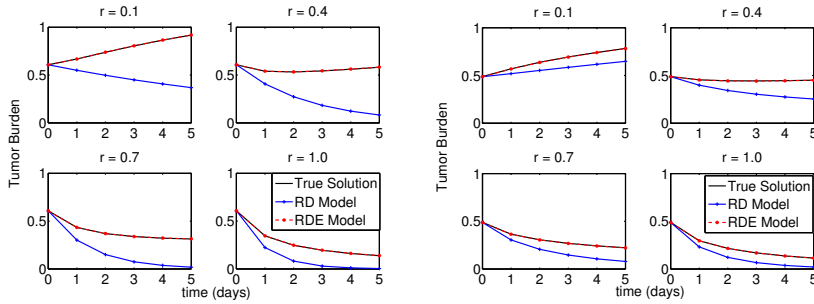


Fig. 7 Simulations of log-kill chemotherapy treatment (Equation (24)) for the case where \mathbf{D} is bigaussian and ρ is normal (left) and the case where \mathbf{D} is normally distributed and ρ is bigaussian for varying values of r (right). The reaction-diffusion equation (Equation (1)), shown in blue solid line with pluses, vastly overestimates the efficacy of chemotherapy treatment. The RDE model (red dashes) with parameter distributions estimated from noisy data almost exactly match the true solution.

when tumor heterogeneity is modeled by parameter distributions we defined, the reaction-diffusion equation will always underestimate the remaining tumor burden size. In fact, in the case where $r = 0.4$, the reaction-diffusion equation would predict tumor shrinkage while the heterogeneous model accurately predicts slight tumor growth.

7 Discussion and Conclusions

We investigated the ability to recover distributions for tumor diffusion and growth parameters using an extension of the reaction-diffusion equation to random differential equations. Under assumptions of noise up to a value of 5%, we were still able to obtain accurate recovery of various combinations of parameter distributions for diffusion and growth. In addition, we investigated the sensitivity of tumor burden with respect to drug administration when modeling homogeneous and heterogeneous populations. We found that we overestimate the efficacy of chemotherapy treatment if we assume a homogeneous population when tumors are actually composed of heterogeneous cell subpopulations with distributions of diffusion and growth.

This initial work has raised many questions which subsequent work should address. Initially, it would be interesting to extend the simulations to incorporate 2D and 3D spatial structure. Mathematically, we should quantify how much information is needed in order to recover the parameter distributions. Since we are, in effect, estimating many parameters to describe a joint distribution (instead of just 2 parameters for the constant-coefficient reaction-diffusion equation), it is imperative that we understand how much data is necessary to avoid the issue of parameter identifiability (which is a known issue in estimating parameters for the reaction-diffusion equation using only one type of imaging data [49]). Quantifying the uncertainty for the distributions estimated using the Prohorov metric framework remains to be investigated.

Our assumption of independence between the parameter distributions should be relaxed in future work. A “go or grow” type distribution assumes that one cell phenotype exhibits slow growth *and* high diffusion and another cell phenotype with fast growth and low diffusion. Modeling this scenario with random differential equations would rely on assuming dependence between the parameter distributions for the diffusion and growth rates. In future work, we will extend the estimation framework described here to allow for estimation of the joint probability density function of \mathbf{D} and $\boldsymbol{\rho}$ without assumptions about parameter independence.

In addition to mathematical questions, a key interest in future work will be to determine how effectively the Prohorov metric framework performs when estimating parameter distributions from *in vitro* and *in vivo* experimental data. This will allow us to validate the application of this estimation framework to realistic spatiotemporal experimental data from heterogeneous tumors.

References

1. Adoteye, K., Banks, H., Cross, K., Etchyson, S., Flores, K., LeBlanc, G., Nguyen, T., Ross, C., Smith, E., Stemkovski, M., Stokley, S.: Statistical validation of structured population models for *Daphnia magna*. *Mathematical Biosciences* **266**, 73–84 (2015)
2. Akaike, H.: A new look at the statistical model identification. *IEEE Transactions on Automatic Control* **19**(6), 716–723 (1974)
3. Baldock, A.L., Ahn, S., Rockne, R., Johnston, S., Neal, M., Corwin, D., Clark-Swanson, K., Sterin, G., Trister, A.D., Malone, H., et al.: Patient-specific metrics of invasiveness reveal significant prognostic benefit of resection in a predictable subset of gliomas. *PLoS One* **9**(10), e99057 (2014)
4. Banks, H., Collins, E., Flores, K., Pershad, P., Stemkovski, M., Stephenson, L.: Statistical error model comparison for logistic growth of green algae (*Raphidocelis subcapitata*). *Applied Math Letters* **64**, 213–222 (2017)
5. Banks, H., Sutton, K.L., Thompson, W.C., Bocharov, G., Roose, D., Schenkel, T., Meyerhans, A.: Estimation of cell proliferation dynamics using CFSE data. *Bulletin of Mathematical Biology* **73**(1), 116–150 (2011)
6. Banks, H.T.: *A Functional Analysis Framework for Modeling, Estimation and Control in Science and Engineering*. CRC Press (2012)
7. Banks, H.T., Barnes, J.H., Eberhardt, A., Tran, H., Wynne, S.: Modeling and computation of propagating waves from coronary stenosis. *Computational Applied Mathematics* **21**, 767–788 (2002)
8. Banks, H.T., Botsford, L.W., Kappel, F., Wang, C.: Modeling and estimation in size structured population models. LCDS-CCS Report 87-13, Brown University; Proc. 2nd Course on Mathematical Ecology, (Trieste, December 8–12, 1986) pp. 521–541 (1988)
9. Banks, H.T., Buksas, M.W., Lin, T.: *Electromagnetic Material Interrogation Using Conductive Interfaces and Acoustic Wavefronts*. SIAM FR 21 Philadelphia (2002)
10. Banks, H.T., Davis, J.L.: A comparison of approximation methods for the estimation of probability distributions on parameters. *Applied Numerical Mathematics* **57**(5-7), 753–777 (2007)
11. Banks, H.T., Davis, J.L.: Quantifying uncertainty in the estimation of probability distributions with confidence bands. CRSC-TR07-21, December, 2007; *Mathematical Biosciences and Engineering* **5**, 647–667 (2008)
12. Banks, H.T., Davis, J.L., Ernstberger, S.L., Hu, S., Artimovich, E., Dhar, A.K., Browdy, C.L.: A comparison of probabilistic and stochastic formulations in modeling growth uncertainty and variability. CRSC-TR08-03, February, 2008; *Journal of Biological Dynamics* **3**, 130–148 (2009)

13. Banks, H.T., Davis, J.L., Ernstberger, S.L., Hu, S., E, A., Dhar, A.K.: Experimental design and estimation of growth rate distributions in size-structured shrimp populations. CRSC-TR08-20, November, 2008; *Inverse Problems* 095003 **25**, (095,003–28pages) (2009)
14. Banks, H.T., Davis, J.L., Hu, S.: A computational comparison of alternatives to including uncertainty in structured population models. CRSC-TR09-14, June, 2009; in *Three Decades of Progress in Systems and Control* (X. Hu, et al., eds.) pp. 19–33 (2010)
15. Banks, H.T., Fitzpatrick, B.G.: Estimation of growth rate distributions in size structured population models. *Quarterly of Applied Mathematics* **49**, 215–235 (1991)
16. Banks, H.T., Fitzpatrick, B.G., Potter, L.K., Zhang, Y.: Estimation of probability distributions for individual parameters using aggregate population data. CRSC-TR98-06, January 1998; in *Stochastic Analysis, Control, Optimization, and Applications*, (W. McEneaney, G. Yin, and Q. Zhang, eds.) pp. 353–371 (1998)
17. Banks, H.T., Flores, K.B., Langlois, C.R., Serio, T.R., Sindi, S.S.: Estimating the rate of prion aggregate amplification in yeast with a generation and structured population model. *Inverse Problems in Science and Engineering* pp. 1–23 (2017)
18. Banks, H.T., Gibson, N.L.: Electromagnetic inverse problems involving distributions of dielectric mechanisms and parameters. CRSC-TR05-29, August, 2005; *Quarterly of Applied Mathematics* **64**, 749–795 (2006)
19. Banks, H.T., Hu, S., Kenz, Z.R., Kruse, C., Shaw, S., Whiteman, J.R., Brewin, M.P., Greenwald, S.E., Birch, M.J.: Material parameter estimation and hypothesis testing on a 1d viscoelastic stenosis model: methodology. CRSC-TR12-09, April, 2012; *J. Inverse and Ill-posed Problems* **21**, 25–57 (2013)
20. Banks, H.T., Hu, S., Kenz, Z.R., Kruse, C., Shaw, S., Whiteman, J.R., Brewin, M.P., Greenwald, S.E., Birch, M.J.: Model validation for a noninvasive arterial stenosis detection problem. CRSC-TR12-22, December, 2012; *Mathematical Biosciences and Engr.* **11**, 427–448 (2013)
21. Banks, H.T., Hu, S., Thompson, W.C.: *Modeling and Inverse Problems in the Presence of Uncertainty*. CRC Press (2014)
22. Banks, H.T., Joyner, M.L.: AIC under the framework of least squares estimation. *Applied Mathematics Letters* **74**, 33–45 (2017)
23. Banks, H.T., Kenz, Z.R., Thompson, W.C.: A review of selected techniques in inverse problem nonparametric probability distribution estimation. CRSC-TR12-13, May 2012; *J. Inverse and Ill-Posed Problems* pp. 429–460 (2012)
24. Banks, H.T., Kunisch, K.: *Estimation Techniques for Distributed Parameter Systems*. Birkhausen, Boston (1989)
25. Banks, H.T., Pinter, G.A.: A probabilistic multiscale approach to hysteresis in shear wave propagation in biotissue. CRSC-TR04-03, January, 2004; *SIAM J. Multiscale Modeling and Simulation* **3**, 395–412 (2005)
26. Banks, H.T., Potter, L.K.: Probabilistic methods for addressing uncertainty and variability in biological models: Application to a toxicokinetic model. CRSC-TR02-27, September, 2002; *Mathematical Biosciences* **192**, 193–22 (2004)
27. Bondiau, P.Y., Clatz, O., Sermesant, M., Marcy, P.Y., Delingette, H., Frenay, M., Ayache, N.: Biocomputing: numerical simulation of glioblastoma growth using diffusion tensor imaging. *Physics in Medicine and Biology* **53**(4), 879 (2008)
28. Clatz, O., Sermesant, M., Bondiau, P.Y., Delingette, H., Warfield, S.K., Malandain, G., Ayache, N.: Realistic simulation of the 3-d growth of brain tumors in mr images coupling diffusion with biomechanical deformation. *IEEE Transactions on Medical Imaging* **24**(10), 1334–1346 (2005)
29. Farin, A., Suzuki, S.O., Weiker, M., Goldman, J.E., Bruce, J.N., Canoll, P.: Transplanted glioma cells migrate and proliferate on host brain vasculature: a dynamic analysis. *Glia* **53**(8), 799–808 (2006)
30. Gallego, O.: Nonsurgical treatment of recurrent glioblastoma. *Current Oncology* **22**(4), e273 (2015)
31. Garay, T., Juhász, É., Molnár, E., Eisenbauer, M., Czirók, A., Dekan, B., László, V., Hoda, M.A., Döme, B., Tímár, J., et al.: Cell migration or cytokinesis and proliferation?—revisiting the go or grow hypothesis in cancer cells *in vitro*. *Experimental Cell Research* **319**(20), 3094–3103 (2013)

32. Hathout, L., Ellingson, B., Pope, W.: Modeling the efficacy of the extent of surgical resection in the setting of radiation therapy for glioblastoma. *Cancer Science* **107**(8), 1110–1116 (2016)
33. Hatzikirou, H., Basanta, D., Simon, M., Schaller, K., Deutsch, A.: ‘Go or grow’: the key to the emergence of invasion in tumour progression? *Mathematical Medicine and Biology* **29**(1), 49–65 (2012)
34. Jackson, P.R., Juliano, J., Hawkins-Daarud, A., Rockne, R.C., Swanson, K.R.: Patient-specific mathematical neuro-oncology: using a simple proliferation and invasion tumor model to inform clinical practice. *Bulletin of Mathematical Biology* **77**(5), 846–856 (2015)
35. Jbabdi, S., Mandonnet, E., Duffau, H., Capelle, L., Swanson, K.R., Pélégri-issac, M., Guillemin, R., Benali, H.: Simulation of anisotropic growth of low-grade gliomas using diffusion tensor imaging. *Magnetic Resonance in Medicine* **54**(3), 616–624 (2005)
36. Kohandel, M., Sivaloganathan, S., Oza, A.: Mathematical modeling of ovarian cancer treatments: sequencing of surgery and chemotherapy. *Journal of Theoretical Biology* **242**(1), 62–68 (2006)
37. Konukoglu, E., Clatz, O., Bondiau, P.Y., Delingette, H., Ayache, N.: Extrapolating glioma invasion margin in brain magnetic resonance images: Suggesting new irradiation margins. *Medical Image Analysis* **14**, 111–125 (2010)
38. Konukoglu, E., Clatz, O., Bondiau, P.Y., Sermesant, M., Delingette, H., Ayache, N.: Towards an identification of tumor growth parameters from time series of images. In: *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2007*, pp. 549–556. Springer (2007)
39. Martínez-González, A., Calvo, G.F., Pérez Romasanta, L.A., Pérez-García, V.M.: Hypoxic cell waves around necrotic cores in glioblastoma: a biomathematical model and its therapeutic implications. *Bulletin of Mathematical Biology* pp. 1–22 (2012)
40. Murray, J.D.: *Mathematical Biology II: Spatial Models and Biomedical Applications*, 3 edn. Springer-Verlag New York Incorporated (2003)
41. Norden, A.D., Wen, P.Y.: Glioma therapy in adults. *Neurologist* **12**(6), 279–292 (2006)
42. Norton, L., Simon, R.: Tumor size, sensitivity to therapy, and design of treatment schedules. *Cancer Treatment Reports* **61**(7), 1307 (1977)
43. Norton, L., Simon, R.: The Norton-Simon hypothesis revisited. *Cancer Treatment Reports* **70**(1), 163 (1986)
44. Onishi, M., Ichikawa, T., Kurozumi, K., Date, I.: Angiogenesis and invasion in glioma. *Brain Tumor Pathology* **28**(1), 13–24 (2011)
45. Painter, K., Hillen, T.: Mathematical modelling of glioma growth: the use of diffusion tensor imaging (dti) data to predict the anisotropic pathways of cancer invasion. *Journal of Theoretical Biology* **323**, 25–39 (2013)
46. Powathil, G., Kohandel, M., Sivaloganathan, S., Oza, A., Milosevic, M.: Mathematical modeling of brain tumors: effects of radiotherapy and chemotherapy. *Physics in Medicine and Biology* **52**(11), 3291 (2007)
47. de Rioja, V.L., Isern, N., Fort, J.: A mathematical approach to virus therapy of glioblastomas. *Biology Direct* **11**(1), 1 (2016)
48. Rockne, R., Rockhill, J., Mrugala, M., Spence, A., Kalet, I., Hendrickson, K., Lai, A., Cloughesy, T., Alvord Jr, E., Swanson, K.: Predicting the efficacy of radiotherapy in individual glioblastoma patients *in vivo*: a mathematical modeling approach. *Physics in Medicine and Biology* **55**(12), 3271 (2010)
49. Rutter, E.M., Stepien, T.L., Anderies, B.J., Plasencia, J.D., Woolf, E.C., Scheck, A.C., Turner, G.H., Liu, Q., Frakes, D., Kodibagkar, V., Kuang, Y., Preul, M.C., Kostelich, E.J.: Mathematical analysis of glioma growth in a murine model. *Scientific Reports* **7**, 2508 (2017). DOI 10.1038/s41598-017-02462-0
50. Skipper, H.E.: Experimental evaluation of potential anticancer agents. xiii. on the criteria and kinetics associated with” curability” of experimental leukemia. *Cancer Chemother Rep* **35**, 1–111 (1964)
51. Stein, A.M., Demuth, T., Mobley, D., Berens, M., Sander, L.M.: A mathematical model of glioblastoma tumor spheroid invasion in a three-dimensional *in vitro* experiment. *Biophysical Journal* **92**(1), 356–365 (2007)

52. Stepien, T.L., Rutter, E.M., Kuang, Y.: A data-motivated density-dependent diffusion model of *in vitro* glioblastoma growth. *Mathematical Biosciences and Engineering* **12**(6), 1157–1172 (2015)
53. Swan, A., Hillen, T., Bowman, J.C., Murtha, A.D.: A patient-specific anisotropic diffusion model for brain tumour spread. *Bulletin of Mathematical Biology* (2017). DOI 10.1007/s11538-017-0271-8. URL <http://dx.doi.org/10.1007/s11538-017-0271-8>
54. Swanson, K., Alvord, E., Murray, J.: A quantitative model for differential motility of gliomas in grey and white matter. *Cell Proliferation* **33**(5), 317–330 (2000)
55. Swanson, K., Alvord, E., Murray, J.: Virtual resection of gliomas: effect of extent of resection on recurrence. *Mathematical and Computer Modelling* **37**(11), 1177–1190 (2003)
56. Swanson, K.R., Alvord, E., Murray, J.: Virtual brain tumours (gliomas) enhance the reality of medical imaging and highlight inadequacies of current therapy. *British Journal of Cancer* **86**(1), 14–18 (2002)
57. Swanson, K.R., Bridge, C., Murray, J., Alvord, E.C.: Virtual and real brain tumors: using mathematical modeling to quantify glioma growth and invasion. *Journal of the Neurological Sciences* **216**(1), 1–10 (2003)
58. Tian, J.P., Friedman, A., Wang, J., Chiocca, E.A.: Modeling the effects of resection, radiation and chemotherapy in glioblastoma. *Journal of Neuro-oncology* **91**(3), 287 (2009)
59. Wen, P.Y., Kesari, S.: Malignant gliomas in adults. *New England Journal of Medicine* **359**(5), 492–507 (2008)
60. Westin, C.F., Maier, S.E., Mamata, H., Nabavi, A., Jolesz, F.A., Kikinis, R.: Processing and visualization for diffusion tensor mri. *Medical Image Analysis* **6**(2), 93–108 (2002)