# **Mathematical Biology**

CrossMark

# Estimating intratumoral heterogeneity from spatiotemporal data

E. M. Rutter<sup>1</sup> · H. T. Banks<sup>1</sup> · K. B. Flores<sup>1</sup>

Received: 11 September 2017 / Revised: 13 April 2018 / Published online: 8 May 2018 © Springer-Verlag GmbH Germany, part of Springer Nature 2018

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 ( $\rho$ ) are random variables. We investigate the ability to perform the inverse problem to recover the probability distributions of *D* and  $\rho$  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.

In memory of Karl Hadeler, for his inspiring research and teaching, and dedication to mentoring.

**Electronic supplementary material** The online version of this article (https://doi.org/10.1007/s00285-018-1238-6) contains supplementary material, which is available to authorized users.

⊠ K. B. Flores kbflores@ncsu.edu

> E. M. Rutter erutter@ncsu.edu

H. T. Banks htbanks@ncsu.edu

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.

<sup>&</sup>lt;sup>1</sup> Department of Mathematics, Center for Research in Scientific Computation, North Carolina State University, Raleigh, USA

**Keywords** Glioblastoma multiforme · Random differential equation · Parameter estimation

#### Mathematics Subject Classification 92B05 · 62G07

# 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 subpopulations. Glibolastoma multiforme (GBM) is an aggressive, inherently heterogeneous brain tumor for which treatment is minimally effective: mean survival time with treatment is 12–15 months Norden and Wen (2006), nearly all GBM tumors recur Wen and Kesari (2008), and it has a five-year survival rate of approximately 5% Gallego (2015). 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 (Eq. (1)) (Murray 2003; Swanson et al. 2000). Although this model is clinically relevant (Baldock et al. 2014; Jackson et al. 2015), it's 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 "go or grow" hypothesis Hatzikirou et al. (2012), 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 coefficient D, and growth rate  $\rho$  in the reaction–diffusion equation as distributions of a random differential equation rather than as point estimates for a deterministic differential equation. Differential equations with random parameters have been studied previously (Soong and Chuang 1973; Soong 1973), including applications to biologically-relevant situations (Chu et al. 2009; Hall and Gandar 1995). However, much of the previous literature concerns forward solutions of the random differential equation. Here, we are interested mainly in the recovery of the distributions of the random variables via the inverse problem. The inverse problem is performed using the Prohorov metric framework which has been developed and tested in other biological scenarios with phenotypic heterogeneity (Banks et al. 1998, 2017), 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 ( $\rho$ ) and the diffusion (D) parameters. We then use the Prohorov metric framework to estimate, i.e., attempt to recover, the distributions for  $\rho$  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  $\rho$  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 Sect. 2. In Sect. 3 we introduce the Prohorov metric framework and it's use to estimate parameter distributions. We follow with our results for synthetically generated data in Sect. 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 Sect. 5. We then carry out a comparison of drug treatment predictions with our model versus the standard homogeneous population model in Sect. 6. We conclude in Sect. 7 with a discussion of future applications of this framework.

#### 2 Mathematical modeling background

The normalized reaction-diffusion equation

$$\frac{\partial c(t,x)}{\partial t} = D \frac{\partial^2 c(t,x)}{\partial x^2} + \rho c(t,x)(1-c(t,x))$$
(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 spatial location x, D is the diffusion coefficient, and  $\rho$  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 Eq. (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 the tissue type, either grey or white matter (Swanson et al. 2000, 2003). 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 Westin et al. (2002). The eigenvalues of these tensors describe the anisotropic diffusion of water. One lingering question is how to use this information, which describes water diffusion, to inform cellular diffusion. Multiple methods have been proposed (Bondiau et al. 2008; Clatz et al.

2005; Jbabdi et al. 2005; Konukoglu et al. 2007; Painter and Hillen 2013), 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 Swan et al. (2017). 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 (Bondiau et al. 2008; Konukoglu et al. 2010).

In vitro experimental work showed that the reaction–diffusion equation was insufficient to describe cell migratory patterns Stein et al. (2007). 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 Garay et al. (2013). However, there is experimental evidence that GBM cells move in a bursting fashion Farin et al. (2006). Hatzikirou et al. (2012) 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 Onishi et al. (2011). Modeling hypoxia-induced phenotypic switching has been shown to replicate histologic patterns of pseudopalisades present in human GBM Martínez-González et al. (2012). 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 Stepien et al. (2015).

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  $\rho$  and D. The random differential equation governing diffusion and growth is:

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

We assume that the parameters D and  $\rho$  are random variables defined on a probability space, denoted by  $\Omega$ . In this case, we assume  $\Omega$  to be a compact set and defined by  $\Omega = \Omega_D \times \Omega_\rho$ . The distribution of the parameters is given by  $P(D, \rho)$ , and c(t, x)represents the aggregate population observable (which is defined as the expectation over subpopulations  $c(t, x, D, \rho)$ :

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

The random differential Eq. (2) is sufficiently flexible to model the classic reaction– diffusion equation (assuming a point distribution of D and  $\rho$ ) and a "go or grow" type equation (assuming, for example, a bi-gaussian distribution of D and  $\rho$ ).

# **3** Prohorov metric

An inverse problem framework relying on the Prohorov metric, called the Prohorov Metric Framework (PMF) was developed and discussed in (Banks 2012; Banks et al. 2014, 2012). Here, we describe how this framework can be utilized for estimating the continuous probability measure  $P(\mathbf{D}, \boldsymbol{\rho})$  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 (Banks et al. 1988; Banks and Davis 2008; Banks and Fitzpatrick 1991; Banks et al. 1998) and shrimp populations (Banks et al. 2009a, b, 2010). Such problems involve equations modeling the number of individuals  $v(t, \xi; g)$  at time t and size  $\xi$  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 Banks and Kunisch (1989), 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 Banks et al. (2017). 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 Banks and Potter (2004) and whole organism transport models Banks and Kunisch (1989)). 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 (Banks et al. 2002; Banks and Gibson 2006). Second order hyperbolic systems are also the focus of the shear wave propagations studies in (Banks et al. 2002, 2013a, b; Banks and Pinter 2005). 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  $P(D, \rho)$  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  $j = 1, ..., N_t$  and  $i = 1, ..., N_x$ . Then, we estimate:

$$\hat{P} = \underset{P^{M}(\Omega)}{\operatorname{argmin}} \sum_{j,i=1}^{N_{t},N_{x}} (v_{ji} - v(t_{j}, x_{i}; P))^{2}$$
(4)

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

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

where *c* is the numerical solution. Previous experimental data Stein et al. (2007) 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  $\rho$ , over M nodes, such that  $\rho^M = \{\Delta_{\rho_l}, l = 1, \dots, M\}$ . Thus, Eq. (5) simplifies to:

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

where the numbers  $w_l^M \ge 0$  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  $\mathbf{D}^M = \{\Delta_{D_l}, l = 1, ..., M\}$ . We would then solve:

$$\hat{P} = \underset{\mathbb{R}}{\operatorname{argmin}} \sum_{j,i} \left[ v_{ji} - \left( \sum_{l} c(t_j, x_i; D_l, \rho) w_l^M \right) \right]^2$$
(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  $\rho$  and D as distributions. We make the assumption of independence between the distributions of the random variables  $\rho$  and D.

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

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

#### 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:

$$s_{l}(\xi) = \begin{cases} \frac{\xi - \xi_{l-1}}{\xi_{l} - \xi_{l-1}} & \text{if } \xi \in [\xi_{l-1}, \xi_{l}] \\ \frac{\xi_{l+1} - \xi_{l}}{\xi_{l+1} - \xi_{l}} & \text{if } \xi \in [\xi_{l}, \xi_{l+1}] \\ 0 & \text{otherwise} \end{cases}$$
(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  $\rho$ , over M nodes, such that  $\rho^M = \{s_l(\rho), l = 1, ..., M\}$ . Thus Eq. (5) simplifies to:

$$\hat{P} = \underset{\mathbb{R}}{\operatorname{argmin}} \sum_{j,i} \left[ v_{ji} - \left( \sum_{l} a_{l} \int_{\Omega_{\rho}} c(t_{j}, x_{i}; D, \boldsymbol{\rho}) s_{l}(\boldsymbol{\rho}) d\rho \right) \right]^{2}$$
(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 Eq. (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  $D^M = \{s_l(D), l = 1, ..., M\}$ . We then find the probability distribution by minimizing the following:

~

$$\hat{P} = \underset{\mathbb{R}}{\operatorname{argmin}} \sum_{j,i} \left[ v_{ji} - \left( \sum_{l} a_{l} \int_{\Omega_{D}} c(t_{j}, x_{i}; \boldsymbol{D}, \rho) s_{l}(\boldsymbol{D}) dD \right) \right]^{2}$$
(11)

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

$$\hat{P} = \underset{\mathbb{R}}{\operatorname{argmin}} \sum_{j,i} \left[ v_{ji} - \sum_{l} a_{l} \int_{\Omega_{\rho}} \left( \sum_{k} b_{k} \int_{\Omega_{D}} c(t_{j}, x_{i}; \boldsymbol{D}, \boldsymbol{\rho}) s_{l}(\boldsymbol{D}) dD \right) s_{k}(\boldsymbol{\rho}) d\rho \right]^{2}$$
(12)

We require that  $\sum_{l=1}^{M_D} a_l \int_{\Omega_D} s_l(\mathbf{D}) dD = 1$  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  $Q_\rho$  quadrature nodes, respectively. The number of quadrature nodes  $Q_D$  and  $Q_\rho$  do not need to be the same. We will refer to this approximation as 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  $C^*(\Omega)$ , the dual of the space of bounded continuous functions on  $\Omega$ . We discuss briefly the convergence and consistency theory here, assuming that we are only estimating **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 Eq. (2). Thus we attempt to perform the following estimation in a least-squares framework

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

to obtain a best fit for a nominal or "true" parameter  $P_0$ . To illustrate for our special 1-D case, we let  $\Omega = \Omega_D$  be the continuum of values on  $[0, D_{\text{max}}]$ . This implies that the family of probability functions  $P(\Omega)$  is a compact metric space in the Prohorov metric. The minimizer function used in Eq. (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  $\Omega^M = \{\Delta_{D_k}, k = 1, ..., M\}$ , where *M* represents the number of nodes, or elements, used in the approximation. Our family of approximating probability functions becomes

$$P^{M}(\Omega^{M}) = \Big\{ P^{M} = \sum_{k=1}^{M} w_{k} \Delta_{D_{k}} | w_{k} \ge 0 \text{ and } \sum_{k=1}^{M} w_{k} = 1 \Big\},\$$

where  $\Delta_{D_k}$  represents the Dirac delta function at the point  $D_k$  and  $w_k$  are the weights and/or probabilities. Again it has been previously proven (Banks 2012; Banks et al. 2014) that there exists a minimizer for the discrete approximation problem

$$\hat{P}^{M} = \underset{P \in P^{M}(\Omega^{M})}{\operatorname{argmin}} \sum_{j,i=1}^{N_{t},N_{x}} (v_{ji} - c(t_{j}, x_{i}; P))^{2}.$$
(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  $M \to \infty$  in the Dirac or spline approximations?) and consistency (what happens as  $N = (N_t, N_x) \to \infty$ ?) These questions have been successfully investigated both theoretically ((Banks et al. 2014; Banks and Pinter 2005) and the references therein) and computationally ((Banks and Davis 2007, 2008) and the references). A further issue involves the partial differential equation approximations  $c_{\tilde{N}}$  to the solution *c* of Eq. (2). Again, the necessary convergence issues have been successfully addressed in Banks et al. (2014); Banks and Kunisch (1989). 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.,  $M, N, \tilde{N} \to \infty$ ).

#### **4** Results for distributed parameters

In the following section we describe results showing the Prohorov metric can correctly estimate parameter distributions for both D and  $\rho$  individually, as well as estimating a joint distribution for D and  $\rho$ . 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 (Eq. (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, and two point.

Table 1 displays the associated information with the parameter distributions for both D and  $\rho$ . The synthetic data were 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:

Distribution name	D	ρ
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)

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

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

We then weight this solution to calculate our synthetic data:

$$v_{ji} = \sum_{l,k} c(t_j, x_i, D_l, \rho_k) p(D_l, \rho_k)$$
(16)

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

*Numerical Setup.* We use MATLAB's built-in pde solver, PDEPE to solve the partial differential equations. We perform the minimization (Eq. (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  $\sum_{k=1}^{M} a_k \int_{\Omega_{\rho}} s_k(\rho) d\rho = 1$  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, ... 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) Akaike (1974) in order to compare the various models. In the context of least-squares framework, the AIC Banks and Joyner (2017) is given by:

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

where *N* is the number of data points, v 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.



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

## 4.1 Estimating **D**

We assume that we already know the value of  $\rho$  and are only estimating the distribution for **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.

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 D with the normal spline approximations for 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 $\rho$

The AIC scores with respect to number of nodes for estimating  $\rho$  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  $\rho$  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 approxima-



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

tions with 50 quadrature nodes and 100 quadrature nodes are essentially the same in most cases, corroborating what we found for estimates of D in the previous section.

#### 4.3 Estimating **D** and $\rho$

We extend the optimization problem to estimate both D and  $\rho$  as distributions for both the discrete case and the continuous case. We focus on two representative examples (D is bigaussian and  $\rho$  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 D and  $\rho$ , respectively, using a discrete approximation is 30 (for both D and  $\rho$ ), while the optimal number of nodes for the continuous case is 10 for  $\rho$ and 20 for 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 Dnormal and  $\rho$  is bigaussian, the optimal number of nodes in the discrete case is 40 for  $\rho$  and 20 for D, while the optimal number of nodes for the continuous case is 30 (for p and 10 for D. In this case, the minimum AIC score is comparable for the spline and discrete approximations.

Figure 4 exhibits the best fits determined by AIC scores when D is normally distributed and  $\rho$  is bigaussian (Fig. 4, left), and when D is bigaussian and  $\rho$  is normally distributed (Fig. 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  $\rho$  and 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.



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

For the case where **D** is bigaussian and  $\rho$  is normally distributed (Fig. 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 **D** is normally distributed and  $\rho$  is bigaussian (Fig. 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. We hypothesize that the reason that the case where **D** is bigaussian and  $\rho$  is normally distributed is not well explained by the spline nodes is that the grid choices over the nodes (10, 20, 30, 40, 50) only allowed choices which either overfit the wide peak or underfit the narrow peak.

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 (Fig. 5, left) and the case where diffusion is normally distributed and growth is bigaussian (Fig. 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 **D** bigaussian and  $\rho$  normally distributed were calculated by splitting the computed probability density function for **D** into two halves and independently simulating their solutions. "Slow-diffusing" represents those cells with the



Fig. 4 Fits using the optimal number of nodes (determined by AIC scores) when  $\rho$  is bigaussian and D is normally distributed (left) and when D is bigaussian and  $\rho$  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

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  $\rho$ is bigaussian and D is normally distributed. As can be seen in Figure 5, these subpopulations exhibit different behavior that is not recovered by assuming static parameters  $\rho$  and D, shown in magenta. The reaction–diffusion solution was obtained by estimating the parameters D and  $\rho$  for Eq. (1) to the synthetic data using FMINCON under the same computation setup defined for the RDE inverse problem. Especially in the case where  $\rho$  is bigaussian (Fig. 5, right), there is a large difference in the solutions when assuming a reaction–diffusion type model. When comparing the parameter results between the basic reaction–diffusion and our methodology, the parameters found by the reaction–diffusion did not correspond to the mean of the distribution.



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

#### 4.4 Computational cost

Here we detail the computational costs associated with estimating the parameter distributions of *D* and  $\rho$  using the Prohorov metric framework. Essentially, the computational method can be considered in two separate parts: the computation of the solution for  $c(t, x, D_l, \rho_k)$  and the optimization to find the weights  $w_l$ ,  $w_k$  (in the discrete case) or  $a_l$ ,  $a_k$  (in the spline case).

To perform our methodology, we precompute the solutions to  $c(t, x, D_l, \rho_k)$ . Essentially, we are computing c(t, x) for each of the node values. This requires  $M_D \times M_\rho$  computations of c(t, x) (in the discrete case) and  $Q_D \times Q_\rho$  computations of c(t, x) (in the spline case). This portion of the computation can easily be parallelized to decrease computational cost.

In the second portion, we use the computed solutions to find the weights  $w_l$ ,  $w_k$  (in the discrete case) or  $a_l$ ,  $a_k$  (in the spline case). For the discrete case, typical times to find the  $w_l$  and  $w_k$  ranged from 10 seconds (for  $M_D = 10$  and  $M_\rho = 10$ ) to 40 minutes (for  $M_D = 50$  and  $M_\rho = 50$ ) using a 2.5GHz processor with 16GB RAM. For the spline case, typical times to find  $a_l$  and  $a_k$  ranged from 14 minutes (for  $M_D = 10$  and  $M_\rho = 10$ ) and  $M_\rho = 50$  and  $Q_\rho = 50$ ) to 7 hours (for  $M_D = 50$  and  $M_\rho = 50$ , with  $Q_D = 100$  and  $Q_\rho = 50$ ). In comparison, performing the optimization assuming static parameters D and  $\rho$  took 2 minutes.

For a full spreadsheet detailing the computational costs associated with the full set of computations for the preceding sections results (assuming *D* is bigaussian and  $\rho$  is normally distributed), please see the supplementary material.

#### 5 Noisy data

We repeat the work performed in Sect. 4, except assuming the data (exact solution) are noisy. To generate these noisy data, we perturb the system 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 (Banks et al. 2017; Adoteye et al. 2015) for examples). In order to generate these errors, we modify Eq. (16) to incorporate noise:

$$\bar{v}_{ji} = v_{ji} + \epsilon_{ji} v_{ji}; \tag{18}$$

where the residuals  $\epsilon_{ji}$  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:

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

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

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

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

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 Sect. 3.3 can be considered in the context of both Banks et al. (2014) an uncertain mathematical model (assumed a major question here) and a statistical model (assumed here as in Eq. (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  $w_{ii}^0 = 1$ . Set iteration number, iter = 0.

2. Estimate parameters in Eq. (20):

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

where  $\mathbf{D}^{M_D}$  and  $\boldsymbol{\rho}^{M_\rho}$  are defined as in previous sections ( $\mathbf{D}^{M_D} = \{\Delta_{D_l}, l = 1, \ldots, M_D\}$ , and  $\boldsymbol{\rho}^{M_\rho} = \{\Delta_{\rho_k}, k = 1, \ldots, M_\rho\}$ ), and  $p = [p_D, p_\rho]$ , where  $p_D$  and  $p_\rho$  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 Banks et al. (2011), to avoid numerical instability:

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

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

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

In our simulations, we discovered that  $C^* = 1e - 4$  still resulted in converged results. We adaptively changed our choice of  $\delta$ : 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 Eq. (21) to the spline approximations defined in Sect. 3.2.

*Estimating* **D** and  $\rho$ . We examine only the ability of estimating the distributions for **D** and  $\rho$  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 **D** is bigaussian and  $\rho$  is normally distributed (Fig. 6, left). The AIC scores, calculated as in Sect. 4, revealed that the discrete approximation with 10 nodes in  $\rho$  and 20 nodes in **D** resulted in the best approximation. The optimal AIC score for the spline case occurred with 10 nodes in  $\rho$  and 20 nodes in **D** with 100 quadrature nodes in each  $\rho$  and **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 D is normally distributed and  $\rho$  is bigaussian (Fig. 6, right). The AIC scores determined the best fit occurs when there are 10 nodes in D and 20 nodes in  $\rho$  for both the discrete and spline case. For the spline approximations, quadrature nodes of 100 for both  $Q_D$  and  $Q_\rho$  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  $\rho$  (results not shown). This, in conjunction with similar results in the previous section implies for a multi-modal distribution with differing variances, 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.



Fig. 6 Fits using the optimal number of nodes (determined by AIC scores) when D is bigaussian and  $\rho$  is normally distributed (left) and the case where  $\rho$  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

In comparison, we also fit the noisy data assuming the basic reaction–diffusion equation. As found in the previous section, the estimated parameters D and  $\rho$  did not correspond to the mean of the distribution.

# 6 Effect of drug administration

Although we are not able to directly compare our methods with classical parameter estimation techniques, we instead propose to compare the ability of the two methods performance in an important clinical application: treatment. In this section, we compare the predicted total tumor size under Eq. (1) as compared with the various distributed parameter solutions, i.e., using random differential equations. Mathemat-



Fig. 7 Simulations of log-kill chemotherapy treatment (Eq. (24)) for the case where **D** is bigaussian and  $\rho$  is normal (left) and the case where **D** is normally distributed and  $\rho$  is bigaussian for varying values of *r* (right). The reaction–diffusion equation (Eq. (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

ical models are often used to predict the change in growth of gliomas in response to treatment such as chemotherapy Powathil et al. (2007), radiotherapy Rockne et al. (2010), resection Swanson et al. (2003), viraltherapy Rioja et al. (2016), or combinations of the above (Hathout et al. 2016; Tian et al. 2009). However, most of these models assume a homogeneous population of tumor cells with static values of D and  $\rho$ . It is known that chemotherapy selectively targets proliferative tumor cells, i.e., cells with larger values of  $\rho$ . Popular methods of mathematically incorporating chemotherapy response includes the log-kill hypothesis Skipper (1964) and the Norton-Simon hypothesis (Norton and Simon 1977, 1986), among others Kohandel et al. (2006).

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:

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

where  $\bar{\rho}$  is the mean value of  $\rho$ , 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 Skipper (1964), which simplifies to the original log-kill hypothesis for the basic reaction–diffusion model, Eq. (1), when  $\rho = \bar{\rho}$ . We keep the death due to drug administration constant, as previously modeled in Swanson et al. (2002), 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  $\rho$  as described in Sect. 4). We also use the solutions and parameters generated by the noisy solutions computed in Sect. 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 (Eq. (24)), in red dashed lines with circles, the RD model (Eq. (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 D is bigaussian and  $\rho$  is normal (Fig. 7, left) and the case where D is normally distributed and  $\rho$  is bigaussian (Fig. 7, right) for varying values of the chemotherapy-induced death rate, r. We observed that 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 introduced a reaction diffusion model for GBM growth that assume the growth and diffusion parameters are random variables. 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. One advantage our method has over classical methods is that our assumption that the parameters are random variables with some underlying distribution does not preclude the existence of point estimates. Thus, our method is able to not only estimate parameter distributions, but also point estimates of parameters.

This initial work has raised many questions which subsequent work should address. Initially, this method should be extended to perform simulations in 2D and 3D in space. Efforts on other more complicated applications of the Prohorov metric framework (Banks et al. 2012; Sirlanci et al. 2018, 2017; Sirlanci and Rosen 2018) including spatially varying coefficient transport problems (as studied in the context of GBM in Engwer et al. (2015); Painter and Hillen (2013); Swan et al. (2017)) have demonstrated the feasibility of greatly expanding the efforts introduced here to include more general and more detailed models of GBM growth.

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. An initial attempt to use *in vitro* data (from Stein et al. (2007)) was unsuccessful due to the lack of information content in the dataset (for full details, see Supplementary Material 2).

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 Rutter et al. (2017)). Quantifying the uncertainty for the distributions estimated using the Prohorov metric framework remains to be investigated. Approaches involving bootstrapping or polynomial chaos may prove useful in estimating the uncertainty of the parameter distributions.

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 **D** and  $\rho$  without assumptions about parameter independence.

#### References

- Adoteye K, Banks H, Cross K, Etchyson S, Flores K, LeBlanc G, Nguyen T, Ross C, Smith E, Stemkovski M, Stokley S (2015) Statistical validation of structured population models for Daphnia magna. Math Biosci 266:73–84
- Akaike H (1974) A new look at the statistical model identification. IEEE Trans Autom Control 19(6):716– 723
- Baldock AL, Ahn S, Rockne R, Johnston S, Neal M, Corwin D, Clark-Swanson K, Sterin G, Trister AD, Malone H et al (2014) Patient-specific metrics of invasiveness reveal significant prognostic benefit of resection in a predictable subset of gliomas. PLoS ONE 9(10):e99057
- Banks H, Collins E, Flores K, Pershad P, Stemkovski M, Stephenson L (2017) Statistical error model comparison for logistic growth of green algae (Raphidocelis subcapitata). Appl Math Lett 64:213– 222
- Banks H, Sutton KL, Thompson WC, Bocharov G, Roose D, Schenkel T, Meyerhans A (2011) Estimation of cell proliferation dynamics using CFSE data. Bull Math Biol 73(1):116–150
- Banks HT (2012) A functional analysis framework for modeling, estimation and control in science and engineering. CRC Press, Boca Raton
- Banks HT, Barnes JH, Eberhardt A, Tran H, Wynne S (2002) Modeling and computation of propagating waves from coronary stenosis. Comput Appl Math 21:767–788
- Banks HT, Botsford LW, Kappel F, Wang C (1988) Modeling and estimation in size structured population models. LCDS-CCS Report 87-13, Brown University. In: Proceedings 2nd course on mathematical ecology, (Trieste, December 8–12, 1986) pp 521–541
- Banks HT, Buksas MW, Lin T (2002) Electromagnetic material interrogation using conductive interfaces and acoustic wavefronts. SIAM FR 21 Philadelphia
- Banks HT, Davis JL (2007) A comparison of approximation methods for the estimation of probability distributions on parameters. Appl Numer Math 57(5–7):753–777
- Banks HT, Davis JL (2008) Quantifying uncertainty in the estimation of probability distributions with confidence bands. CRSC-TR07-21, December 2007. Math Biosci Eng 5:647–667
- Banks HT, Davis JL, Ernstberger SL, Hu S, Artimovich E, Dhar AK, Browdy CL (2009a) A comparison of probabilistic and stochastic formulations in modeling growth uncertainty and variability. CRSC-TR08-03, February 2008. J Biol Dyn 3:130–148
- Banks HT, Davis JL, Ernstberger SL, Hu S, E, A, Dhar AK (2009b) Experimental design and estimation of growth rate distributions in size-structured shrimp populations. CRSC-TR08-20, November, 2008; Inverse Problems 095003 25, (095,003–28 pages)

- Banks HT, Davis JL, Hu S (2010) 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 contro (X. Hu, et al., eds.) pp 19–33
- Banks HT, Fitzpatrick BG (1991) Estimation of growth rate distributions in size structured population models. Q Appl Math 49:215–235
- Banks HT, Fitzpatrick BG, Potter LK, Zhang Y (1998) 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
- Banks HT, Flores KB, Langlois CR, Serio TR, Sindi SS (2017) Estimating the rate of prion aggregate amplification in yeast with a generation and structured population model. Inverse Probl Sci Eng 1–23
- Banks HT, Gibson NL (2006) Electromagnetic inverse problems involving distributions of dielectric mechanisms and parameters. CRSC-TR05-29, August 2005. Q Appl Math 64:749–795
- Banks HT, Hu S, Kenz ZR, Kruse C, Shaw S, Whiteman JR, Brewin MP, Greenwald SE, Birch MJ (2013a) Material parameter estimation and hypothesis testing on a 1d viscoelastic stenosis model: methodology. CRSC-TR12-09, April, 2012; J Inverse Ill-posed Probl 21:25–57
- Banks HT, Hu S, Kenz ZR, Kruse C, Shaw S, Whiteman JR, Brewin MP, Greenwald SE, Birch MJ (2013b) Model validation for a noninvasive arterial stenosis detection problem. CRSC-TR12-22, December, 2012. Math Biosci Eng 11:427–448
- Banks HT, Hu S, Thompson WC (2014) Modeling and inverse problems in the presence of uncertainty. CRC Press, Boca Raton
- Banks HT, Joyner ML (2017) AIC under the framework of least squares estimation. Appl Math Lett 74:33– 45
- Banks HT, Kenz ZR, Thompson WC (2012) A review of selected techniques in inverse problem nonparametric probability distribution estimation. CRSC-TR12-13, May 2012; J Inverse Ill-Posed Probl 429–460
- Banks HT, Kunisch K (1989) Estimation techniques for distributed parameter systems. Birkhausen, Boston
- Banks HT, Pinter GA (2005) A probabilistic multiscale approach to hysteresis in shear wave propagation in biotissue. CRSC-TR04-03, January 2004; SIAM. J Multiscale Model Simul 3:395–412
- Banks HT, Potter LK (2004) 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
- Banks HT, Thompson WC, Peligero GS, Argilaguet J, Meyerhans A (2012) A division-dependent compartmental model for computing cell numbers in cfse-based lymphocyte proliferation assays. CRSC-TR12-03, January 2012. Math Biosci Eng 9:699–736
- Bondiau PY, Clatz O, Sermesant M, Marcy PY, Delingette H, Frenay M, Ayache N (2008) Biocomputing: numerical simulation of glioblastoma growth using diffusion tensor imaging. Phys Med Biol 53(4):879
- Chu J, Ducrot A, Magal P, Ruan S (2009) Hopf bifurcation in a size-structured population dynamic model with random growth. J Differ Eq 247(3):956–1000
- Clatz O, Sermesant M, Bondiau PY, Delingette H, Warfield SK, Malandain G, Ayache N (2005) Realistic simulation of the 3-d growth of brain tumors in mr images coupling diffusion with biomechanical deformation. IEEE Trans Med Imaging 24(10):1334–1346
- de Rioja VL, Isern N, Fort J (2016) A mathematical approach to virus therapy of glioblastomas. Biol Direct 11(1):1
- Engwer C, Hillen T, Knappitsch M, Surulescu C (2015) Glioma follow white matter tracts: a multiscale dti-based model. J Math Biol 71(3):551–582
- Farin A, Suzuki SO, Weiker M, Goldman JE, Bruce JN, Canoll P (2006) Transplanted glioma cells migrate and proliferate on host brain vasculature: a dynamic analysis. Glia 53(8):799–808
- Gallego O (2015) Nonsurgical treatment of recurrent glioblastoma. Curr Oncol 22(4):e273
- Garay T, Juhász É, Molnár E, Eisenbauer M, Czirók A, Dekan B, László V, Hoda MA, Döme B, Tímár J et al (2013) Cell migration or cytokinesis and proliferation? Revisiting the go or grow hypothesis in cancer cells in vitro. Exp Cell Res 319(20):3094–3103
- Hall A, Gandar P (1995) Stochastic models for fruit growth. In: IV International symposium on computer modelling in fruit research and orchard management 416:113–120
- Hathout L, Ellingson B, Pope W (2016) Modeling the efficacy of the extent of surgical resection in the setting of radiation therapy for glioblastoma. Cancer Sci 107(8):1110–1116
- Hatzikirou H, Basanta D, Simon M, Schaller K, Deutsch A (2012) 'Go or grow': the key to the emergence of invasion in tumour progression? Math Med Biol 29(1):49–65

- Jackson PR, Juliano J, Hawkins-Daarud A, Rockne RC, Swanson KR (2015) Patient-specific mathematical neuro-oncology: using a simple proliferation and invasion tumor model to inform clinical practice. Bull Math Biol 77(5):846–856
- Jbabdi S, Mandonnet E, Duffau H, Capelle L, Swanson KR, Pélégrini-Issac M, Guillevin R, Benali H (2005) Simulation of anisotropic growth of low-grade gliomas using diffusion tensor imaging. Magn Reson Med 54(3):616–624
- Kohandel M, Sivaloganathan S, Oza A (2006) Mathematical modeling of ovarian cancer treatments: sequencing of surgery and chemotherapy. J Theor Biol 242(1):62–68
- Konukoglu E, Clatz O, Bondiau PY, Delingette H, Ayache N (2010) Extrapolating glioma invasion margin in brain magnetic resonance images: suggesting new irradiation margins. Med Image Anal 14:111–125
- Konukoglu E, Clatz O, Bondiau PY, Sermesant M, Delingette H, Ayache N (2007) Towards an identification of tumor growth parameters from time series of images. In: Medical image computing and computerassisted intervention–MICCAI 2007, Springer, pp 549–556
- Martínez-González A, Calvo GF, Pérez Romasanta LA, Pérez-García VM (2012) Hypoxic cell waves around necrotic cores in glioblastoma: a biomathematical model and its therapeutic implications. Bull Math Biol 1–22
- Murray JD (2003) Mathematical biology II: spatial models and biomedical applications, 3rd edn. Springer, New York
- Norden AD, Wen PY (2006) Glioma therapy in adults. Neurologist 12(6):279-292
- Norton L, Simon R (1977) Tumor size, sensitivity to therapy, and design of treatment schedules. Cancer Treat Rep 61(7):1307
- Norton L, Simon R (1986) The Norton-Simon hypothesis revisited. Cancer Treat Rep 70(1):163
- Onishi M, Ichikawa T, Kurozumi K, Date I (2011) Angiogenesis and invasion in glioma. Brain Tumor Pathology 28(1):13–24
- Painter K, Hillen T (2013) Mathematical modelling of glioma growth: the use of diffusion tensor imaging (dti) data to predict the anisotropic pathways of cancer invasion. J Theor Biol 323:25–39
- Powathil G, Kohandel M, Sivaloganathan S, Oza A, Milosevic M (2007) Mathematical modeling of brain tumors: effects of radiotherapy and chemotherapy. Phys Med Biol 52(11):3291
- Rockne R, Rockhill J, Mrugala M, Spence A, Kalet I, Hendrickson K, Lai A, Cloughesy T, Alvord E Jr, Swanson K (2010) Predicting the efficacy of radiotherapy in individual glioblastoma patients in vivo: a mathematical modeling approach. Phys Med Biol 55(12):3271
- Rutter EM, Stepien TL, Anderies BJ, Plasencia JD, Woolf EC, Scheck AC, Turner GH, Liu Q, Frakes D, Kodibagkar V, Kuang Y, Preul MC, Kostelich EJ (2017) Mathematical analysis of glioma growth in a murine model. Sci Rep 7:2508. https://doi.org/10.1038/s41598-017-02462-0
- Sirlanci M, Kang D, Fairbairn CE, Luczak SE, Pan R, Yu X, Rosen IG (2018) Estimating the distribution of random parameters in, and deconvolving the input signal to, a diffusion equation forward model for a transdermal alcohol biosensor. Submitted for Publication
- Sirlanci M, Luczak S, Rosen IG (2017) Approximation and convergence in the estimation of random parameters in linear holomorphic semigroups generated by regularly dissipative operators. In: American control conference (ACC), 2017, pp 3171–3176. IEEE
- Sirlanci M, Rosen IG (2018) Estimation of the distribution of random parameters in discrete time abstract parabolic systems with unbounded input and output: Approximation and convergence. Submitted for Publication
- Skipper HE (1964) Experimental evaluation of potential anticancer agents. xiii. on the criteria and kinetics associated with" curability" of experimental leukemia. Cancer Chemother Rep 35:1–111
- Soong T, Chuang S (1973) Solutions of a class of random differential equations. SIAM J Appl Math 24(4):449–459
- Soong TT (1973) Random differential equations in science and engineering. Elsevier, Amsterdam
- Stein AM, Demuth T, Mobley D, Berens M, Sander LM (2007) A mathematical model of glioblastoma tumor spheroid invasion in a three-dimensional in vitro experiment. Biophys J 92(1):356–365
- Stepien TL, Rutter EM, Kuang Y (2015) A data-motivated density-dependent diffusion model of in vitro glioblastoma growth. Math Biosci Eng 12(6):1157–1172
- Swan A, Hillen T, Bowman JC, Murtha AD (2017) A patient-specific anisotropic diffusion model for brain tumour spread. Bull Math Biol. https://doi.org/10.1007/s11538-017-0271-8
- Swanson K, Alvord E, Murray J (2000) A quantitative model for differential motility of gliomas in grey and white matter. Cell Prolif 33(5):317–330

- Swanson K, Alvord E, Murray J (2003) Virtual resection of gliomas: effect of extent of resection on recurrence. Math Comput Modell 37(11):1177–1190
- Swanson KR, Alvord E, Murray J (2002) Virtual brain tumours (gliomas) enhance the reality of medical imaging and highlight inadequacies of current therapy. Br J Cancer 86(1):14–18
- Swanson KR, Bridge C, Murray J, Alvord EC (2003) Virtual and real brain tumors: using mathematical modeling to quantify glioma growth and invasion. J Neurol Sci 216(1):1–10
- Tian JP, Friedman A, Wang J, Chiocca EA (2009) Modeling the effects of resection, radiation and chemotherapy in glioblastoma. J Neurooncol 91(3):287

Wen PY, Kesari S (2008) Malignant gliomas in adults. N Engl J Med 359(5):492-507

Westin CF, Maier SE, Mamata H, Nabavi A, Jolesz FA, Kikinis R (2002) Processing and visualization for diffusion tensor mri. Med Image Anal 6(2):93–108