Mathematical Models and Methods in Applied Sciences
Vol. 23, No. 7 (2013) 1309–1338

DOI: 10.1142/S0218202513500103

SELECTION AND ASSESSMENT OF PHENOMENOLOGICAL MODELS OF TUMOR GROWTH

J. TINSLEY ODEN* and ERNESTO E. PRUDENCIO†
Institute for Computational Engineering and Sciences,
The University of Texas at Austin,
1 University Station C0200, Austin TX 78712, USA
**oden@ices.utexas.edu*
†*prudenci@ices.utexas.edu*

ANDREA HAWKINS-DAARUD
Division of Neuropathology,
University of Washington, 1959 NE Pacific St.,
Box 357470, Seattle WA 98195, USA
andjhd@uw.edu

Received 12 June 2012
Revised 4 August 2012
Accepted 2 September 2012
Published 5 November 2012
Communicated by N. Bellomo

"...the progress of science requires comparison of different conceivable models, a false premise built into a model that is never questioned cannot be removed by any amount of new data"
E. T. Jaynes, Probability Theory: The Logic of Science, 2003

We address general approaches to the rational selection and validation of mathematical and computational models of tumor growth using methods of Bayesian inference. The model classes are derived from a general diffuse-interface, continuum mixture theory and focus on mass conservation of mixtures with up to four species. Synthetic data are generated using higher-order base models. We discuss general approaches to model calibration, validation, plausibility, and selection based on Bayesian-based methods, information theory, and maximum information entropy. We also address computational issues and provide numerical experiments based on Markov chain Monte Carlo algorithms and high performance computing implementations.

Keywords: Bayesian statistics; model validation; diffuse-interface models; model selection; Markov chain Monte Carlo methods.

AMS Subject Classification: 62F07, 62F15, 65C40, 65M60, 93A30

1. Introduction

The heart of predictive science is the development of mathematical and computational models, guided by observational data, but mainly as the result of inductive

arguments on how various physical systems behave. A remarkable development over the last two decades has been the growing number of models that have been proposed for predicting the behavior of biological systems, including the functioning of living tissue. Predictive science for such systems would obviously have a significant impact on biomedicine, various therapies, drug design, and drug delivery. Of particular interest is the increasing number of mathematical models proposed for describing the growth and behavior of tumors as illustrated by the reviews[1,19] and books.[11,34,5] The variety of techniques discussed and used in these references is reflective of the fact that capturing the behavior of living matter is much different than capturing the behavior of inert matter. Quoting Hartwell *et al.*,[18] "*Although living systems obey the laws of physics and chemistry, the notion of function or purpose differentiates biology from other natural sciences.*" Biological systems exist to reproduce and do not exist in thermal equilibrium. In tumor growth, key events dictating the observed tumor kinetics occur on many different scales: genetic mutations within cell nuclei, on the scale of angstroms, generate different proteins in various concentrations in the cell, on the order of nanometers, causing different cellular behavior on the microscale, ultimately giving rise to various behaviors in the host organism on the macroscale.[9,17,16]

In spite of this, there is increasing evidence that many of the complex processes involved in the growth of tumors, including angiogenesis, nutrient transport, instability and progression to metastasis etc. can be captured by phenomenological models, and by careful characterizations of mechanical, chemical, and thermodynamical properties of tissue species. References 6, 5, 14, 15, 45, 40 and 44, the models described in Refs. 1, 4 and 19, and many others contain results that make a convincing case that the prediction of tumor growth and, ultimately, the treatment of certain types of cancer may be made possible through advances in computational modeling. Many of the models in the references already mentioned focus on events at a single level, though an effort is being made to create multiscale models bridging at least two scales. The review by Bellomo *et al.*[4] highlights this effort and goes through the types of analysis required for finding the asymptotic limiting behavior of the larger scales from the smaller scales.

A major issue in modeling tumor growth is choosing the appropriate modeling framework for capturing the desired tumor behavior and then applying such general theories to specific problems. Selecting the appropriate subclass of models for understanding specific events is confounded by the fact that tumors are genetically unstable and thus the appropriate model parameters may change over time. Thus, questions of how to calibrate such models using available data, in what sense are the models valid given the observational data, and how can uncertainties in key quantities of interest, such as net tumor volume, be quantified, are of paramount importance. We believe that answers to these key questions fall within the general domain of Bayesian methods of inference and the powerful ideas of logical probability and information theory which are often associated with them. This paper explores how these methods can be used with examples at continuum scale events,

though a full understanding of the system will require information from all scales. We believe the techniques described are generalizable and could be and should be used for modeling events at any scale. Indeed, the general approach should be applicable to all models of physical and biological events.

As examples of classes of tumor growth models that can be derived from phase-field (diffused interface) theories of mixtures, we lay down four general models of mass balance in the next section of this paper. These are derived as special cases of a general mixture-theory formulation derived in Ref. 30 which generalizes models of similar structure found in earlier literature. Additionally, we review briefly a framework for statistical model validation discussed also in Ref. 30. This provides the setting for a discussion of the general and fundamental question: which of the models in a given set of candidate competing models is the most plausible, for a well-defined physical event and given relevant data? We also address related concepts of information content in probability densities as a means of selecting priors and of quantifying uncertainty to output quantities of interest. Finally, we present the results of several numerical calculations of statistical properties of solutions of two-dimensional models of tumor growth based on the models introduced in Sec. 2.

2. Diffuse-Interface Models of Tumor Growth

The continuum theory of mixtures began with the classical works of Darcy[12] and Fick,[13] and was generalized in Refs. 42 and 43, as well as in the pioneering papers of Bowen[8] and Rajagopal.[38] These provide a powerful framework for modeling the complex behavior of heterogeneous media with many interacting constituents, and has been the foundation of several models of tumor growth, e.g. Refs. 1, 10, 11, 15, 19, 20 and 30.

This multiphase theory represents the presence of each constituent as a volume fraction or a mass concentration, allowing multiple constituents to be present at the same point at the same time. Further, the interface between phases is handled as a feature of the solution by representing surface energies of interfaces between constituents through the appearance of gradients of volume fractions in the Helmholtz free energy, leading to so-called diffuse-interface models of the Cahn–Hilliard type. The theory is a rich source of new largely untried models of the complex behavior of interacting media.

The fundamental idea underlying mixture theory is that a material body $\mathcal{B}$ can be composed of N constituent species $\mathcal{B}_1, \ldots, \mathcal{B}_N$, each occupying a common portion of space at the same time. Each spatial position is then allowed to be occupied by N constituents and each constituent is assigned a mass density $\hat{\rho}_\alpha$, regarded as a function of position and time, $(\mathbf{x}, t)$, which represents the mass of the αth constituent per unit volume of the mixture, a volume fraction φ_α, with $\sum_{i=1}^{N} \varphi_\alpha = 1$, and a mass density ρ_α (so $\hat{\rho}_\alpha = \varphi_\alpha \rho_\alpha$). Each of the N constituents must satisfy its own balance laws which differ from those of classical continuum

mechanics due to the presence of interaction terms representing the exchange of mass, momentum, and energy between constituents.

Here we focus on the principal mechanisms of mass exchange and conservation of mass among species, assuming that momentum and energy transfer are negligible.

For a general mixture occupying an open region Ω in $\mathbb{R}^3$ over a time interval $(0, T]$, the volume fractions, φ_α, and other independent field variables must satisfy the following mass balance law for all α, $1 \leq \alpha \leq N$, all $\mathbf{x} \in \Omega$, and $t \in (0, T]$:

$$\frac{\partial \rho_\alpha \varphi_\alpha}{\partial t} + \nabla \cdot (\rho_\alpha \varphi_\alpha \mathbf{v}_\alpha) = \gamma_\alpha - \nabla \cdot \mathbf{j}_\alpha, \tag{2.1}$$

where γ_α is the mass supplied to constituent α by other constituents and $\mathbf{j}_\alpha$ is the mass flux due to changes in the chemical potential defined in terms of gradients in concentrations and changes in nutrient concentrations, and $\mathbf{v}_\alpha$ is the species velocity field.

Equation (2.1) is closed by the appropriate choices of constitutive equations, which must be consistent with both the second law of thermodynamics for the mixture and the classical balance laws for the entire mixture. By introducing the Helmholtz free energy per unit mass ψ_α for each constituent, or the free energy Ψ_α per unit volume, the classical Coleman–Noll method makes use of the constraints from classical balance laws and the second law of thermodynamics to supply sufficient conditions on forms of key constitutive equations in terms of appropriate derivatives of the free energy (see Refs. 19, 10 and 30). In the diffuse-interface theories, with negligible thermomechanical effects, the Helmholtz free energy is assumed to be of the form

$$\Psi_\alpha = \Psi_\alpha(\varphi_1, \ldots, \varphi_N, \nabla\varphi_1, \ldots, \nabla\varphi_N, m_1, \ldots, m_L). \tag{2.2}$$

The values m_l, $1 \leq l \leq L$, represent concentrations of chemicals or small polypeptides which interact with the species φ_α but are so much smaller in size than the constituents that it is not necessary to consider them directly within the mixture, e.g. oxygen or glucose. It can be argued (see e.g. Ref. 30) that a resulting consistent constitutive equation for $\mathbf{j}_\alpha$ is of the form

$$\mathbf{j}_\alpha = -\sum_{\beta=1}^{N} M_{\alpha\beta}(\boldsymbol{\varphi}, \mathbf{m}_\alpha) \nabla \left(\frac{\mu_{\beta\alpha}}{\rho_\alpha} \right), \tag{2.3}$$

where $M_{\alpha\beta}$ is a positive semi-definite, second-order mobility tensor and $\mu_{\alpha\beta}$ is interpreted as the chemical potential, defined as

$$\mu_{\alpha\beta} = \frac{\partial \Psi_\alpha}{\partial \varphi_\beta} - \nabla \cdot \frac{\partial \Psi_\alpha}{\partial \nabla \varphi_\beta}. \tag{2.4}$$

Other details on this theory are given in Refs. 19 and 30.

Next we consider an organ or gland that occupies a region Ω in $\mathbb{R}^3$ at time $t = 0$. Oxygenated blood and other nutrients may flow into and out of Ω over a

time interval of interest $[0, T]$. Within this domain, we assume there is a subdomain Ω_{tumor}, possibly not connected, which is comprised of tumorous tissue with material properties different than its surrounding tissue. We begin, for simplicity, with the case of an isothermal mixture consisting of at least a tumor species u and non-tumor n (i.e. healthy tissue, extracellular proteins, and fluid). We may include convective velocity, or may assume that it is negligible. Additionally, we may include in our model a representative nutrient, c, say oxygen. It is always assumed that the mixture is saturated (i.e. $\sum_\alpha \varphi_\alpha = 1$). With these conventions in mind, we postulate a form of the free energy functional for the tumor phase similar to that used in the Cahn–Hilliard theory, augmented with a term corresponding to a lower energy when the tumor is able to interact with the chemical species, such as oxygen:

$$\Psi_u(u, \nabla u, c) = \gamma u^2 (u-1)^2 + \frac{\epsilon^2}{2|\nabla u|^2} - \tau(u, c), \tag{2.5}$$

where γ, τ, and ϵ are positive constants. Here $u = \varphi_1$, the volume fraction of cancer cells, and c is the oxygen concentration per unit volume. The first term is a double-well potential with minima at $u = 0$ and $u = 1$, the second term can be thought of as energy due to surface tension, and the final term is energy due to reactions with the nutrient and accounts for cell chemotaxis.

Introducing (2.5), (2.4), and (2.3) into (2.1) produces a general family of tumor growth models in which several specific subfamilies of models can be obtained. We focus on three models, as follows:

A_1: Simple Proliferation

$$\begin{aligned} u_t &= \nabla \cdot (\zeta \nabla \mu) + Pu && \text{in } (0,T) \times \Omega, \\ \mu &= f'(u) - \epsilon^2 \Delta u && \text{in } (0,T) \times \Omega, \\ \nabla u \cdot \mathbf{n} &= \nabla \mu \cdot \mathbf{n} = 0 && \text{on } (0,T) \times \partial\Omega, \\ u(0, \mathbf{x}) &= u_0 && \text{in } \{0\} \times \Omega, \\ f'(u) &= \gamma(4u^3 - 6u^2 + 2u). \end{aligned} \tag{2.6}$$

A_2: Proliferation/Apoptosis, with Degenerate Mobility and Oxygen Dependence

$$\begin{aligned} u_t &= \nabla \cdot (\zeta u^2 \nabla \mu) + Pcu - Au && \text{in } (0,T) \times \Omega, \\ \mu &= f'(u) - \epsilon^2 \Delta u - \epsilon \tau c && \text{in } (0,T) \times \Omega, \\ 0 &= \nabla \cdot (\xi \nabla c) - cu && \text{in } (0,T) \times \Omega, \\ \nabla u \cdot \mathbf{n} &= \nabla \mu \cdot \mathbf{n} = 0 \quad c = 1 && \text{on } (0,T) \times \partial\Omega, \\ u(0, \mathbf{x}) &= u_0 && \text{in } \{0\} \times \Omega, \\ f'(u) &= \gamma(4u^3 - 6u^2 + 2u). \end{aligned} \tag{2.7}$$

A_3: Four Species Proliferation, Including Nutrient Rich Extracellular Water

$$
\begin{aligned}
u_t &= \nabla \cdot (\zeta u^2 \nabla \mu_1) + \delta P(u)(\mu_2 - \mu_1) && \text{in } (0,T)\times\Omega,\\
\mu_1 &= f'(u) - \epsilon^2 \Delta u + \tau_{,u}(u,c) && \text{in } (0,T)\times\Omega,\\
0 &= \nabla \cdot (\delta \xi \nabla \mu_2) - \delta P(u)(\mu_2 - \mu_1) && \text{in } (0,T)\times\Omega,\\
\mu_2 &= \frac{1}{\delta} c + \tau_{,c}(u,c) && \text{in } (0,T)\times\Omega,\\
\nabla u \cdot \mathbf{n} &= \nabla \mu_1 \cdot \mathbf{n} = \nabla \mu_2 \cdot \mathbf{n} = 0, \quad c = 1 && \text{on } (0,T)\times\partial\Omega,\\
u(0,\mathbf{x}) &= u_0 && \text{in } \{0\}\times\Omega,\\
f'(u) &= \gamma(4u^3 - 6u^2 + 2u),\\
\tau(u,c) &= -\tau_0 uc,\\
P(u) &= \hat{P}(u - 0.1)(1-u).
\end{aligned}
\tag{2.8}
$$

We use the terminology "Parametric Model Class", or PMC, to describe the models (2.6)–(2.8), and to emphasize that each of these models is in fact an infinite class of models parametrized by the physical parameters of each system for given Ω and $[0,T]$. PMCs also include the prior information in such parameters. A more complete definition of a PMC is given in subsequent sections.

In deriving PMC M_1 (related to A_1), it is assumed that there are only two species, u and n. Convective velocity is ignored, the term $\tau(u,c)$ in the free energy is assumed zero and γ_u is taken as a simple, linear proliferation wherever there is species u.

There is also assumed to be only two species, u and n, in PMC M_2 (related to A_2), but the effects of oxygen are now considered. Specifically, $\tau(u,c) = -\tau uc$ and the oxygen concentration is modeled via a separate reaction diffusion–reaction equation. Additionally, the mass exchange term γ_u also includes effects of oxygen, that is the tumor will only proliferate significantly if there is enough oxygen present.

Finally, in PMC M_3 (related to A_3) four species are considered. This model differs from M_2 in that it includes the oxygen within the mixture. That is, the four species are tumor cells u, healthy cells n, nutrient-rich extracellular water c, and nutrient-poor extracellular water w. Additional saturation constraints are placed on this model as well; in particular, (1) $u + n = C$ and (2) $c + w = 1 - C$. Again, equations can be renormalized such that $u + n = 1$ and then $c + w = 1$, and still only u and c need be taken as unknowns to determine the entire system. The term $\tau(u,c)$ is taken as the function given in (2.8). Additionally, the mass exchange terms are now chosen such that the equations are guaranteed to be thermodynamically consistent. More details on this model are given in Ref. 21.

3. Abstractions: General Parametric Model Classes

We now digress to the consideration of an abstract class of deterministic or stochastic mathematical models which contains those described in the previous section.

This development represents a generalization of that given in Ref. 30. We begin by considering a set $\mathcal{M} = \{M_1, M_2, \ldots, M_m\}$ of PMCs. In PMC M_j, we have the central forward problem of finding a solution u_j to the problem,

$$A_j(\boldsymbol{\theta_j}, s; u_j(\boldsymbol{\theta_j}, s)) = 0, \tag{3.1}$$

where A_j is a collection of operators (such as the partial differential operators in (2.6)–(2.8)), $\boldsymbol{\theta_j}$ is a vector of model parameters (such as coefficients, moduli, etc. in (2.6), (2.7), or (2.8)), and s is a scenario under which the operations A_j are to function (such as the solution domain Ω, the boundary data, source terms, etc.). The following remarks are given to clarify the intended meaning of (3.1):

- The model parameters $\boldsymbol{\theta_j}$ belong to a parameter manifold $V_{\boldsymbol{\theta_j}}$ and run over a range of physical values. Each specific value $\boldsymbol{\theta_j}^*$ defines a different model in the PMC M_j. The determination of which values $\boldsymbol{\theta_j}$ pertain to a given physical situation is generally a problem of model calibration, and depends upon the scenario s.
- The model equation (3.1) can typically be applied to many different scenarios, ranging from one-dimensional laboratory tests, *in vitro* experiments, multiple components tests in two-dimensional setting produced by tomographical modalities, etc. The initial-condition data can be assumed to be given as part of the scenario, but it can also be regarded as part of the vector $\boldsymbol{\theta_j}$. The most general scenario is that in which the target predictions of the model are to be made, and that is the prediction scenario $s = s_p$. The notion of what constitutes the model parameters $\boldsymbol{\theta_j}$ and what constitutes a scenario is a somewhat subjective decision. For simplicity, we will regard the scenario as a well-defined deterministic feature of the physical setting in which the model is to be implemented, although generalizations to stochastic settings can easily be accommodated.
- The solution $u_j(s, \boldsymbol{\theta_j})$ of problem (3.1) is assumed to exist.
- Upon solving (3.1), one can compute theoretical values Y_j of observables that are to eventually be experimentally measured, for each scenario s:

$$Y_j = Y_j(\boldsymbol{\theta_j}, s, u_j(\boldsymbol{\theta_j}, s)). \tag{3.2}$$

  In general, Y_j will differ from the actual experimentally observational data $D(s)$, and this difference is called noise. In the case of additive noise ϵ_j,

$$\epsilon_j = Y_j(\boldsymbol{\theta_j}, s, u_j(\boldsymbol{\theta_j}, s)) - D(s). \tag{3.3}$$

- The observational data $D(s)$ belong to a data manifold V_D and is determined experimentally for the scenario s.

This framework is designed to identify all the sources of information and uncertainty that affect the use of mathematical (and computational) models to make predictions of physical quantities of interest (as discussed in Sec. 4), and to ultimately quantify the uncertainty in those predictions, which will be used as basis

for decisions. These sources are:

(1) The model parameters $\boldsymbol{\theta_j}$, the initial uncertainty of which we assume is fully characterized by a prior probability density function $\pi^{(j)}_{\text{prior}}(\boldsymbol{\theta_j})$. We may include in these parameters those defining the scenario s, which may also involve uncertainties.
(2) The model equations themselves provide an imperfect map of parameters $\boldsymbol{\theta_j}$ into observations Y_j that do not necessarily match the experimental observations $D = D(s)$. This uncertainty is characterized by a conditional likelihood PDF $\pi^{(j)}_{\text{like}}(D|\boldsymbol{\theta_j})$. For instance, if the additive noise ϵ_j of (3.3) is characterized by a PDF $\pi_{\text{noise}}(\epsilon_j)$, then

$$\pi^{(j)}_{\text{like}}(D|\boldsymbol{\theta_j}) = \pi_{\text{noise}}[Y_j(\boldsymbol{\theta_j}, s, u(\boldsymbol{\theta_j}, s)) - D(s)].$$

(3) The observational data D have prior uncertainties themselves due to experimental error characterized by a prior $\pi_{\text{data}}(D)$.

All of the information available that affects the predictability of the model and the uncertainty of outputs are connected in the Bayes' formula

$$\pi^{(j)}_{\text{post}}(\boldsymbol{\theta_j}|D) = \frac{\pi^{(j)}_{\text{like}}(D|\boldsymbol{\theta_j}) \cdot \pi^{(j)}_{\text{prior}}(\boldsymbol{\theta_j})}{\pi_{\text{data}}(D)}, \tag{3.4}$$

where $\pi^{(j)}_{\text{post}}(\boldsymbol{\theta_j}|D)$ is the posterior PDF defining the Bayesian update of the prior information embodied in $\pi^{(j)}_{\text{prior}}(\boldsymbol{\theta_j})$. Equation (3.4) is valid for each M_j separately, a fact that will be made clear in Eqs. (5.2) and (5.5) below. From now on, for simplification, we will suppress the superscripts "(j)".

4. Model Validation and Uncertainty Quantification

In Sec. 3 we fixed the modeling scenario and explored the ideas of comparing different PMCs competing to explain experimentally observed data $D(s)$. Here we explore the situation of fixing the PMC but having multiple scenarios $s_1, s_2, \ldots$, generally of increasing complexity, that form conceptually a prediction pyramid, with lowest level calibration experiments corresponding to the base of the pyramid (see Fig. 1). The peak is the full prediction scenario s_p, where the extrapolation of lower level information on parameters is performed to compute the target quantity of interest (QoI) Q, the ultimate goal of a modeling program. The QoI is not an observable quantity, else it would not be a prediction. We assume that all possible values of Q belong to a manifold V_Q, which is equipped with a measure μ_Q. We also denote by $V_{\boldsymbol{\theta}}$ the manifold of the input parameters $\boldsymbol{\theta}$, and assume it is equipped with measure $\mu_{\boldsymbol{\theta}}$.

The notion of model validation embodies processes designed to establish if a given model can depict (or predict) physical realities as determined by experimental observations, with acceptable accuracy. Validation processes that make use of a

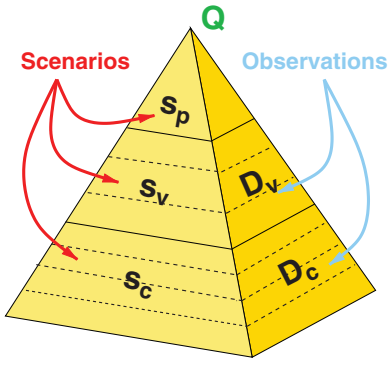

Fig. 1. The prediction pyramid of model calibration, validation, and prediction (cf. Refs. 30 and 31).

Bayesian framework are discussed, e.g. in Refs. 2, 30–32. Here we provide a brief review and slight extension of the validation process described in Ref. 30.

The process begins with statistical calibration of parameters in a lower tier scenario s_1. This is done through an application of Bayes' rule (3.4) for calibration scenario s_1:

$$\pi_{\text{post}}(\boldsymbol{\theta}|D_1) = \frac{\pi_{\text{like}}(D_1|\boldsymbol{\theta}) \cdot \pi_{\text{prior}}(\boldsymbol{\theta})}{\pi_{\text{data}}(D_1)}.$$

Here D_1 denotes experimental observational data at the lowest level of the prediction pyramid. One can presumably improve the distribution for $\boldsymbol{\theta}$ by performing additional calibration experiments at higher level scenarios, yielding for example, an improved posterior PDF

$$\pi_{\text{post}}(\boldsymbol{\theta}|D_1, D_2) = \frac{\pi_{\text{like}}(D_2|\boldsymbol{\theta}) \cdot \pi_{\text{post}}(\boldsymbol{\theta}|D_1)}{\pi_{\text{data}}(D_2)},$$

assuming that D_1 and D_2 are independent, i.e. $\pi_{\text{data}}(D_1, D_2) = \pi_{\text{data}}(D_1) \cdot \pi_{\text{data}}(D_2)$. More data sets $D_3, D_4, \ldots$ might be available from calibration experiments. Let us denote by D_c all the data available from calibration experiments. The updated PDF for $\boldsymbol{\theta}$ is now used to construct a stochastic forward problem in the prediction scenario s_p:

$$A(\pi_{\text{post}}(\boldsymbol{\theta}|D_c), s_p; u(\pi_{\text{post}}(\boldsymbol{\theta}|D_c), s_p)) = 0. \tag{4.1}$$

The notation used here is extended to signify that the mathematical problem of solving for u_c defined in (4.1) is now replaced by a stochastic system in which the parameters $\boldsymbol{\theta}$ are available not as real numbers but as PDFs determined by one or more calibration processes. This so-called forward problem is solved (using standard methods such as Monte Carlo, polynomial chaos, etc.) for

$$u_c \equiv u(\pi_{\text{post}}(\boldsymbol{\theta}|D_c), s_p),$$

and the solution is used to compute the final QoI

$$Q_c \equiv Q(u(\pi_{\text{post}}(\boldsymbol{\theta}|D_c), s_p)),$$

which has PDF

$$\pi_{\text{pred}}(Q|D_c) = \int_{V_{\boldsymbol{\theta}}} \pi_{\text{pred}}(Q|\boldsymbol{\theta}, D_c) \cdot \pi_{\text{post}}(\boldsymbol{\theta}|D_c)\ \mu_{\boldsymbol{\theta}}(d\boldsymbol{\theta}). \tag{4.2}$$

In addition to the calibration tests, we construct validation experiments at as high a level of sophistication practical up the prediction pyramid. Ideally, the validation scenario s_{v} is designed to depict as many features of the final prediction problem that will influence the QoIs Q based on prior information in the hands of the modelers. If D_{v} denotes the observation data measured in the validation experiments, the knowledge gained on the parameters is embodied in

$$\pi_{\text{post}}(\boldsymbol{\theta}|D_c, D_{\text{v}}) = \frac{\pi_{\text{like}}(D_{\text{v}}|\boldsymbol{\theta}) \cdot \pi_{\text{post}}(\boldsymbol{\theta}|D_c)}{\pi_{\text{data}}(D_{\text{v}})},$$

again assuming independence between D_c and D_{v}. With data D_{v}, the stochastic forward problem is

$$A(\pi_{\text{post}}(\boldsymbol{\theta}|D_c, D_{\text{v}}), s_p; u(\pi_{\text{post}}(\boldsymbol{\theta}|D_c, D_{\text{v}}), s_p)) = 0,$$

the corresponding solution is

$$u_{\text{v}} \equiv u(\pi_{\text{post}}(\boldsymbol{\theta}|D_c, D_{\text{v}}), s_p),$$

and the QoI is

$$Q_{\text{v}} \equiv Q(u(\pi_{\text{post}}(\boldsymbol{\theta}|D_c, D_{\text{v}}), s_p)),$$

which has PDF

$$\pi_{\text{pred}}(Q|D_c, D_{\text{v}}) = \int_{V_{\boldsymbol{\theta}}} \pi_{\text{pred}}(Q|\boldsymbol{\theta}, D_c, D_{\text{v}}) \cdot \pi_{\text{post}}(\boldsymbol{\theta}|D_c, D_{\text{v}})\ \mu_{\boldsymbol{\theta}}(d\boldsymbol{\theta}). \tag{4.3}$$

The underlying PMC, denoted simply "M", can be made explicit in formulas (4.2) and (4.3), as will be shown in Eq. (5.11) below. As already mentioned at the end of Sec. 3, such explicit presentation of M will also happen in Eqs. (5.2) and (5.5) below, in comparison to Eq. (3.4).

We now have two probabilistic characterizations of the QoI Q, the calibration-based PDF $\pi_{\text{pred}}(Q|D_c)$ and the validation-based PDF $\pi_{\text{pred}}(Q|D_c, D_{\text{v}})$, and pose two fundamental questions:

(i) Validation: Is the calibrated model valid for predicting the QoI Q?
(ii) Uncertainty quantification: What is the "amount of uncertainty" in the predictions $\pi_{\text{pred}}(Q|D_c)$ and $\pi_{\text{pred}}(Q|D_c, D_{\text{v}})$?

In answer to (i), we first note that no model can ever be claimed to be valid; at best, a model is not invalid when measured against some (subjective) tolerance,

the idea being that it is always possible that new data may be acquired that may invalidate the model by establishing that the model cannot predict the QoIs with sufficient accuracy. We shall declare the model not invalid when $\pi_{\text{pred}}(Q|D_c)$ is close to $\pi_{\text{pred}}(Q|D_c, D_{\text{v}})$ in some meaningful sense. We must first answer (ii) to address how these PDFs can be compared.

According to Jaynes,[23] the only reasonable function (satisfying certain mild conditions laid down in Ref. 23) for quantifying "the amount of uncertainty" in a probability distribution is the information entropy, defined by

$$H(p) = -\sum_{i=1}^{n} p_i \cdot \ln(p_i)$$

for a discrete probability distribution $p = \{p_1, p_2, \ldots, p_n\}$, and by

$$H(p) = -\int p(x) \cdot \ln\left[\frac{p(x)}{m(x)}\right] dx$$

for the continuous case, $m(x)$ being a reference PDF. The *relative entropy* of two PDFs w and z, with the support of z contained in the support of w, is defined by the Kullback–Leibler divergence,[28]

$$D_{\text{KL}}(z\|w) = \int z(x) \cdot \ln\left(\frac{z(x)}{w(x)}\right) dx. \tag{4.4}$$

Thus, the *relative uncertainty between the calibration-based and validation-based* PDFs is

$$\begin{aligned} &D_{\text{KL}}[\pi_{\text{pred}}(Q|D_c, D_{\text{v}})\|\pi_{\text{pred}}(Q|D_c)] \\ &\quad = \int_{V_Q} \pi_{\text{pred}}(Q|D_c, D_{\text{v}}) \cdot \ln\left[\frac{\pi_{\text{pred}}(Q|D_c, D_{\text{v}})}{\pi_{\text{pred}}(Q|D_c)}\right] \mu_Q(dQ). \end{aligned}$$

This can also be written

$$\begin{aligned} &D_{\text{KL}}[\pi_{\text{pred}}(Q|D_c, D_{\text{v}})\|\pi_{\text{pred}}(Q|D_c)] \\ &\quad = H[\pi_{\text{pred}}(Q|D_c, D_{\text{v}}), \pi_{\text{pred}}(Q|D_c)] - H[\pi_{\text{pred}}(Q|D_c, D_{\text{v}})], \end{aligned}$$

where $H[\pi_{\text{pred}}(Q|D_c, D_{\text{v}}), \pi_{\text{pred}}(Q|D_c)]$ is the cross entropy and $H[\pi_{\text{pred}}(Q|D_c, D_{\text{v}})]$ is the entropy of the validation-based QoI PDF. Finally, choosing a tolerance $\eta \in \mathbb{R}_+$, we shall say that the calibrated forward model (4.1) is not invalid if the KL measure

$$D_{\text{KL}}[\pi_{\text{pred}}(Q|D_c, D_{\text{v}})\|\pi_{\text{pred}}(Q|D_c)] < \eta, \tag{4.5}$$

i.e. if the information provided by the validation data on the QoI does not appreciably differ from that corresponding to the calibration data.

We shall show in the next section that the unsymmetric KL measure also occurs naturally in Bayesian model selection.

5. Model Selection and the Concept of Parametric Model Class

The idea of developing criteria for selecting the most plausible model for a prediction among several possible models for given data D has been of great interest for many years. In some respects, the subject goes back some 650 years to Ockham's razor, advanced by the 14th century friar, William of Ockham, and generally interpreted to be the principle: *use the simplest model to make predictions among all models that are sufficient to explain observed phenomena.* The debate as to whether this is indeed a principle to guide scientific discovery has been waged for over a century, and has involved such combatants as Popper,[33] Jaynes,[23] and many others. Generally, the popular view has been that Ockham's razor is not a principle that can guide scientific discovery. However, Jaynes[23] (p. 601), claims that "only through modern Bayesian analysis has it (Ockham's razor) been well understood". In more recent times, frameworks for models selection have been proposed in the Bayesian statistics literature by Berger and Pericchi[7] and others; in engineering applications by Beck *et al.*;[3,29] a full chapter in the recent book of Robert[39] is devoted to model choice; and Jaynes[23] presents a lengthy discussion of model choice. In this section, we describe a general framework for model selection with some slight generalizations that pull in notions of model validation discussed earlier.

5.1. *Model evidence and plausibility*

Besides the manifolds V_Q and $V_{\boldsymbol{\theta}}$ mentioned in Sec. 4, we also assume we are given a data manifold V_D, equipped with a measure μ_D, as well as any arbitrarily fixed subset $D \subset V_D$. We then define a "parametric model class" (PMC) as any collection of assumptions and entities that completely determine *a posterior* PDF through the Bayesian formula (3.4). A PMC is composed of:

(i) a set of mathematical equations and functions used to compute values Y comparable to (points in) D;
(ii) a prior PDF $\pi_{\text{prior}} : V_{\boldsymbol{\theta}} \to \mathbb{R}_+$ for such input parameters; and
(iii) a conditional PDF $\pi_{\text{like}}(D|\boldsymbol{\theta}) : V_D \to \mathbb{R}_+$ for any given $\boldsymbol{\theta} \in V_{\boldsymbol{\theta}}$.

The set of equations and functions of the PMC usually reflect physical principles and assumptions about the phenomenon being modeled, and is sometimes referred to as "model structure". Each point $\boldsymbol{\theta} \in V_{\boldsymbol{\theta}}$ specifies one (stochastic) model. The likelihood PDF $\pi_{\text{like}}(D|\boldsymbol{\theta})$, a PDF of D, encapsulates assumptions about the discrepancy between the values D that are observed in experiments and the values Y that are computed with the particular (stochastic) model specified by $\boldsymbol{\theta}$. As a PDF, $\pi_{\text{like}}(D|\boldsymbol{\theta})$ satisfies

$$\int_{V_D} \pi_{\text{like}}(D|\boldsymbol{\theta}) \mu_D(dD) = 1 \quad \forall\, \boldsymbol{\theta} \in V_{\boldsymbol{\theta}}. \tag{5.1}$$

The likelihood PDF can also be seen as a function of $\boldsymbol{\theta}$ (called likelihood function) for any arbitrarily fixed point D in V_D:

$$L : V_{\boldsymbol{\theta}} \to \mathbb{R}_+,$$
$$\boldsymbol{\theta} \mapsto L(\boldsymbol{\theta}) = \pi_{\text{like}}(D|\boldsymbol{\theta}).$$

In other words, all possible (stochastic) models specified by points $\boldsymbol{\theta} \in V_{\boldsymbol{\theta}}$, together with all the assumptions implicitly and explicitly made: (i) by the mathematical modeling structure, (ii) by $\pi_{\text{prior}}(\boldsymbol{\theta})$, and (iii) by $\pi_{\text{like}}(D|\boldsymbol{\theta})$, determine one (stochastic) PMC. We denote a PMC by the capital letter M, and rewrite the Bayesian formula (3.4) as

$$\pi_{\text{post}}(\boldsymbol{\theta}|D, M) = \frac{\pi_{\text{like}}(D|\boldsymbol{\theta}, M) \cdot \pi_{\text{prior}}(\boldsymbol{\theta}|M)}{\pi_{\text{data}}(D|M)}, \tag{5.2}$$

where

$$\pi_{\text{data}}(D|M) = \int_{V_{\boldsymbol{\theta}}} \pi_{\text{like}}(D|\boldsymbol{\theta}, M) \cdot \pi_{\text{prior}}(\boldsymbol{\theta}|M)\mu_{\boldsymbol{\theta}}(d\boldsymbol{\theta}) \tag{5.3}$$

is the normalization value (for a given D) that makes (5.2) a PDF. Since $\pi_{\text{like}}(D|\boldsymbol{\theta}, M)$ is a PDF, as stated by (5.1), it immediately follows that (5.3) is also a PDF. Indeed,

$$\begin{aligned}\int_{V_D} \pi_{\text{data}}(D|M)\mu_D(dD) &= \int_{V_D}\int_{V_{\boldsymbol{\theta}}} \pi_{\text{like}}(D|\boldsymbol{\theta}, M) \cdot \pi_{\text{prior}}(\boldsymbol{\theta}|M)\mu_{\boldsymbol{\theta}}(d\boldsymbol{\theta})\mu_D(dD)\\ &= \int_{V_{\boldsymbol{\theta}}}\int_{V_D} \pi_{\text{like}}(D|\boldsymbol{\theta}, M) \cdot \pi_{\text{prior}}(\boldsymbol{\theta}|M)\mu_D(dD)\mu_{\boldsymbol{\theta}}(d\boldsymbol{\theta})\\ &= \int_{V_{\boldsymbol{\theta}}} 1 \cdot \pi_{\text{prior}}(\boldsymbol{\theta}|M)\mu_{\boldsymbol{\theta}}(d\boldsymbol{\theta})\\ &= 1.\end{aligned} \tag{5.4}$$

The notation $\pi_{\text{post}}(\boldsymbol{\theta}|D, M)$ indicates the conditional probability of parameters $\boldsymbol{\theta}$ given both data D and PMC M. Jaynes[23] refers to (5.2) as a "higher form" of Bayes' theorem in which we acknowledge the PMC M as the underlying model for the data D. The quantity (5.3) is called the *model evidence* for M. It defines the probability of obtaining the reference data D given by the PMC M and its underlying assumptions.

Let us now suppose that we have a set

$$\mathcal{M} = \{M_1, M_2, \ldots, M_m\}$$

of m different PMCs put forward to model a specific physical event and to allow us to make predictions. Equation (5.2) can be used *intra-PMC* for each PMC M_j, in $\mathcal{M}$, as follows:

$$\pi_{\text{post}}(\boldsymbol{\theta_j}|D, M_j) = \frac{\pi_{\text{like}}(D|\boldsymbol{\theta_j}, M_j) \cdot \pi_{\text{prior}}(\boldsymbol{\theta_j}|M_j)}{\pi_{\text{data}}(D|M_j)}. \tag{5.5}$$

However, one can apply Bayes' theorem not only to compare (the infinitely many) models $\boldsymbol{\theta_j}$ inside M_j, but also to compare (the finitely many) PMCs M_j inside $\mathcal{M}$. That is, Bayes' theorem can be used in an *inter-PMC* way as follows (note the use of "p" instead of "π", since we are dealing with probability mass functions instead of PDFs):

$$p_{\text{post}}(M_j|D,\mathcal{M}) = \frac{\pi_{\text{data}}(D|M_j)\cdot p_{\text{prior}}(M_j|\mathcal{M})}{\pi_{\text{data}}(D|\mathcal{M})}, \tag{5.6}$$

where (i) $p_{\text{prior}}(M_j|\mathcal{M})$ is the *prior plausibility* of M_j, i.e. the prior probability of M_j in the set $\mathcal{M}$, (ii) $p_{\text{post}}(M_j|D,\mathcal{M})$ is the *posterior plausibility* of M_j, (iii) $\pi_{\text{data}}(D|M_j)$ is the model evidence appearing in (5.5), i.e. a PDF describing how probable D is according to computations performed with M_j, and (iv) the constant

$$\pi_{\text{data}}(D|\mathcal{M}) = \sum_{j=1}^{m} \pi_{\text{data}}(D|M_j)\cdot p_{\text{prior}}(M_j|\mathcal{M}) \tag{5.7}$$

is the normalization value (for a given D) that makes (5.6) a probability mass function, that is,

$$\sum_{j=1}^{m} p_{\text{post}}(M_j|D,\mathcal{M}) = 1.$$

Since $\pi_{\text{data}}(D|M_j)$ is a PDF, as stated by (5.4), it immediately follows that (5.7) is also a PDF.

Equation (5.6) provides a mean to compare PMCs for the given set D of reference data and to pick the most plausible one. A PMC M_{j_1} is deemed "better" than PMC M_{j_2} if $p_{\text{post}}(M_{j_1}|D,\mathcal{M}) > p_{\text{post}}(M_{j_2}|D,\mathcal{M})$. The superiority of one PMC over another clearly depends upon the particular D, as indicated by the conditional notation. It is also clear that the evidence (5.3) is a fundamental quantity for the determination of the status of PMC M relative to other PMCs.[23]

5.2. *A quantitative interpretation of Ockham's razor*

We now discuss an interpretation of Ockham's razor, following Ref. 29. From (5.5) it is immediate that

$$\ln(\pi_{\text{data}}(D|M_j)) = \ln(\pi_{\text{like}}(D|\boldsymbol{\theta_j},M_j))\ln\left(\frac{\pi_{\text{post}}(\boldsymbol{\theta_j}|D,M_j)}{\pi_{\text{prior}}(\boldsymbol{\theta_j}|M_j)}\right).$$

Now, multiplying all terms by $\pi_{\text{post}}(\boldsymbol{\theta_j}|D,M_j)$ and integrating over $V_{\boldsymbol{\theta_j}}$, we obtain

$$\ln(\pi_{\text{data}}(D|M_j)) = E[\ln(\pi_{\text{like}}(D|\boldsymbol{\theta_j},M_j))] - E\left[\ln\left(\frac{\pi_{\text{post}}(\boldsymbol{\theta_j}|D,M_j)}{\pi_{\text{prior}}(\boldsymbol{\theta_j}|M_j)}\right)\right], \tag{5.8}$$

where the expectation is with respect to the posterior $\pi_{\text{post}}(\boldsymbol{\theta_j}|D,M_j)$.

In Eq. (5.8), the term

$$E[\ln(\pi_{\text{like}}(D|\boldsymbol{\theta_j}, M_j))] \tag{5.9}$$

can be interpreted as a measure of how well the PMC M_j fits the data, since it involves an average of the likelihood of each model $\boldsymbol{\theta_j}$ in M_j. The term

$$E\left[\ln\left(\frac{\pi_{\text{post}}(\boldsymbol{\theta_j}|D, M_j)}{\pi_{\text{prior}}(\boldsymbol{\theta_j}|M_j)}\right)\right], \tag{5.10}$$

on the other hand, is precisely the Kullback–Leibler divergence introduced in (4.4). It is a measure of how much was learned from the prior knowledge and the posterior knowledge, by incorporating D. Informally speaking, if one makes gradually more complex PMCs (e.g. by gradually adding model parameters or equations), in order to gradually learn all possible details from a given D, the PMC fitting capability, indicated by (5.9), will probably gradually increase, but at the expense of too much complexity, which is indicated by (5.10). That is, at some point the log-evidence $\ln(\pi_{\text{data}}(D|M_j))$ will stop increasing and then stall or decrease with extra complexity in the PMC. In order words, mathematical models should not be made too complex for data D, which could be interpreted as a statement of Ockham's razor.

5.3. *Averaging of model predictions*

Given a candidate PMC M_j, a QoI Q, and a conditional PDF $\pi_{\text{pred}}(Q|\boldsymbol{\theta_j}, D, M_j)$, the predictive PDF of Q is

$$\pi_{\text{pred}}(Q|D, M_j) = \int_{V_{\boldsymbol{\theta_j}}} \pi_{\text{pred}}(Q|\boldsymbol{\theta_j}, D, M_j) \cdot \pi_{\text{post}}(\boldsymbol{\theta_j}|D, M_j)\mu_{\boldsymbol{\theta_j}}(d\boldsymbol{\theta_j}). \tag{5.11}$$

Since $\pi_{\text{pred}}(Q|\boldsymbol{\theta_j}, D, M_j)$ is a PDF, it immediately follows that (5.11) is also a PDF. Given the set $\mathcal{M}$ of candidate PMCs, the probabilistic information for Q is[22,37]

$$\pi_{\text{avg}}(Q|D, \mathcal{M}) = \sum_{j=1}^{m} \pi_{\text{pred}}(Q|D, M_j) \cdot p_{\text{post}}(M_j|D, \mathcal{M}), \tag{5.12}$$

an equation referred to as *posterior model averaging* in the Bayesian literature. Since $\pi_{\text{pred}}(Q|D, M_j)$ is a PDF, it immediately follows that (5.12) is also a PDF.

5.4. *Calculation of integrals*

One of the most formidable obstacles in implementing Bayesian methodologies has historically been the evaluation of integrals such as (5.3), (5.9)–(5.11), and this remains a challenging area of research. This has gradually yielded to modern algorithms and high-performance computing (HPC) systems. One possible approach is to use Markov chain Monte Carlo (MCMC) algorithms for sampling posterior PDFs (and computing integrals), and then Monte Carlo algorithms for sampling

QoI PDFs (as well as computing integrals), usually in combination with HPC, as e.g. in Refs. 35 and 36.

A posterior PDF might have multiple modes, and unprepared MCMC algorithms might inappropriately explore just some of the modes while missing the others, with the collateral damage of wrongly computing estimates for the integrals. One simple idea to improve the chances of exploring all existing modes, and therefore computing good estimates for the integrals, is to sample increasingly difficult intermediate distributions, accumulating information from one intermediate distribution to the next, until the target posterior distribution is better sampled. Assuming one is able to sample the prior (either because it is a simple PDF or because it comes from a previously sampled posterior), possible intermediate distributions are given by

$$\pi_{\text{int}}^{(\ell)}(\boldsymbol{\theta_j}|D, M_j) = [\pi_{\text{like}}(D|\boldsymbol{\theta_j}, M_j)]^{\alpha_\ell} \cdot \pi_{\text{prior}}(\boldsymbol{\theta_j}|M_j), \quad \ell = 0, 1, \ldots, L, \tag{5.13}$$

for a given $L > 0$ and a given sequence $0 = \alpha_0 < \alpha_1 < \cdots < \alpha_L = 1$ of exponents. When $\alpha_\ell = 0$, the distribution is the prior, and when $\alpha_\ell = 1$, the distribution is the posterior. As ℓ increases from $\ell = 0$ until to $\ell = L$, the distribution transitions from the initial prior to the (eventually multimodal) posterior.

Sampling algorithms following (5.13) can greatly benefit from parallel computing. At each (level) ℓ, many compute nodes can be used to sample the parameter space collectively. Beginning with $\ell = 0$, the compute nodes: (a) sample $\pi_{\text{int}}^{(\ell)}(\boldsymbol{\theta_j}|D, M_j)$, (b) select some of the generated samples to serve as initial positions of Markov chains for the next distribution $\pi_{\text{int}}^{(\ell+1)}(\boldsymbol{\theta_j}|D, M_j)$, and (c) generate the Markov chains for $\pi_{\text{int}}^{(\ell+1)}(\boldsymbol{\theta_j}|D, M_j)$. The process (a)–(c) continues until the final posterior distribution is sampled. The selection process in step (b), as ℓ increases, tends to value samples that are located in the modes, and such modes gradually get more distinguishable as α_ℓ increases. So, as ℓ increases, if the samples selected from the ℓth level to the $(\ell + 1)$th level are not redistributed among compute nodes before step (c), the "lucky" compute nodes (that is, the ones that had samples of initial levels already in the final posterior modes) will tend to accumulate increasingly more samples. Therefore, as the exponent α_ℓ increases, care needs to be taken in order to maintain a load balance among all compute nodes. A more careful analysis of these and related issues is available in Ref. 35.

5.5. *Remarks on Bayes' factors and Ockham's factors*

The idea of using Bayesian approaches for model selection and so-called hypothesis testing goes back to the early works of Jeffreys in 1935 (see Ref. 25) and 1961 (see Ref. 26), and is addressed in the work of Jefferys and Berger,[24] as well as in modern books on Bayesian methods, such as Robert[39] and Jaynes.[23] In order to make a connection between these ideas and those developed in Secs. 5.1 and 5.2, we first

define the posterior *odds ratio* of any two PMCs $M_j, M_k \in \mathcal{M}$ as

$$\mathcal{O}_{jk} \equiv \frac{p_{\text{post}}(M_j|D)}{p_{\text{post}}(M_k|D)} = \frac{p_{\text{prior}}(M_j)}{p_{\text{prior}}(M_k)} \cdot \frac{\pi_{\text{data}}(D|M_j)}{\pi_{\text{data}}(D|M_k)}. \tag{5.14}$$

The ratio $p_{\text{prior}}(M_j)/p_{\text{prior}}(M_k)$ is the prior odds of PMC M_j to M_k and represents the prior knowledge of the relative subjective superiority of one PMC over another. If we denote the maximum likelihood $L^{(j)}_{\max}$ by

$$L^{(j)}_{\max} = \sup_{\boldsymbol{\theta_j} \in V_{\boldsymbol{\theta_j}}} \pi_{\text{data}}(D|\boldsymbol{\theta_j}, M_j),$$

then

$$\pi_{\text{data}}(D|M_j) = L^{(j)}_{\max} \cdot W_j,$$

where W_j is the *Ockham's factor*,[23] (p. 610),

$$W_j = \int_{V_{\boldsymbol{\theta_j}}} \pi_{\text{prior}}(\boldsymbol{\theta_j}|M_j) \cdot \frac{L^{(j)}(\boldsymbol{\theta_j})}{L^{(j)}_{\max}} \, \mu_{\boldsymbol{\theta_j}}(\boldsymbol{\theta_j}). \tag{5.15}$$

Then the model comparison (5.14) can be written

$$\mathcal{O}_{jk} = \frac{p_{\text{prior}}(M_j)}{p_{\text{prior}}(M_k)} \cdot \frac{L^{(j)}_{\max}}{L^{k}_{\max}} \cdot \frac{W_j}{W_k}, \tag{5.16}$$

which can also be written

$$\mathcal{O}_{jk} = B_{jk} \cdot \frac{p_{\text{prior}}(M_j)}{p_{\text{prior}}(M_k)}, \tag{5.17}$$

where B_{jk} are the so-called *Bayes' factor* (see Refs. 7 and 39 for an exhaustive discussion),

$$B_{jk} = \frac{L^{(j)}_{\max}}{L^{k}_{\max}} \cdot \frac{W_j}{W_k} = \frac{\pi_{\text{data}}(D|M_j)}{\pi_{\text{data}}(D|M_k)}. \tag{5.18}$$

Thus, the Bayes factor B_{jk} is the ratio between the model evidences for M_j and M_k. The posterior odds ratio is then the prior odds ratio multiplied by the ratio of model evidences.

6. Computational Results

In this section, we apply the Bayesian methodology described above to the determination of the most plausible PMC between two candidate PMCs M_1 and M_2, which compete to better simulate a given data set D. The PMC ranking example of this section can be easily generalized for any number and variety of PMCs.

6.1. *General considerations*

Throughout Sec. 6, the physical domain corresponding to a tissue is the square $\Omega_{\text{actual}} = (0, 25.6) \times (0, 25.6)$. The state variable is the concentration $u(\mathbf{x}, t)$ of cancer cells in the tissue for any $\mathbf{x} \in \Omega_{\text{actual}}$ and any $t \geq 0$. Given the physical meaning of u, we demand that

$$u(\mathbf{x}, t) \geq 0 \quad \forall\, (\mathbf{x}, t) \in \Omega_{\text{actual}} \times (0, +\infty). \tag{6.1}$$

Due to the symmetry of the problem with respect to the domain center, we solve only for $\Omega = (12.8, 25.6) \times (12.8, 25.6)$, thus considerably diminishing the computational cost. An initial state

$$u_0(\mathbf{x}) \equiv u(\mathbf{x}, 0) \tag{6.2}$$

is given at $t = 0$. We discretize the equations in space with a finite element triangular mesh Ω_h, and in time with an explicit time stepping scheme with a fixed time step $\Delta t > 0$. We solve the corresponding discretized equations with Newton's method and a direct linear solver. Mesh adaptation is performed at some time steps during the simulations, and we denote by $\Omega_h(t)$ the mesh at instant $t \geq 0$. We use parallel computing in all computational steps. More specifically, we use the Lonestar computational platform at the Texas Advanced Computing Center (TACC[41]), where each computational node contains 24 GB of memory and 12 processing cores of 2 GHz each.

In order to have a PMC, one has to:

(1) decide which equations to use;
(2) decide which parameters to treat randomly (instead of deterministically);
(3) decide which prior PDFs to use for the random parameters; and
(4) decide which likelihood to use, that is, decide on how to model the discrepancy between the reference data and the quantities computed with the equations.

All these four decisions will be exemplified below for the two PMCs M_1 and M_2. It is clear that, even keeping the same equations, one can propose a rich variety of PMCs, all sharing such same equations.

For the purpose of likelihood computations, as well as plotting figures, we proceed as follows. We first divide Ω into a uniform grid of 70×70. Let us denote such squares by $\Omega_{i,j}$, and their centers by $\mathbf{x}_{i,j}$, $1 \leq i, j \leq 70$. Each square can be thought as a pixel in a tissue image (see Fig. 2). Then, for any given $u_{\text{given}}(\cdot, t)$ (necessary for some plotting or for some likelihood calculation), we define

$$u_{\text{avg}}(\mathbf{x}, t) \equiv \frac{\int_{\Omega_{i,j}} u_{\text{given}}(\mathbf{x}, t) d\mathbf{x}}{\int_{\Omega_{i,j}} d\mathbf{x}} \quad \forall\, (\mathbf{x}, t) \in \Omega_{i,j} \times [0, +\infty). \tag{6.3}$$

Then, for a given threshold value

$$u_{\text{threshold}} > 0, \tag{6.4}$$

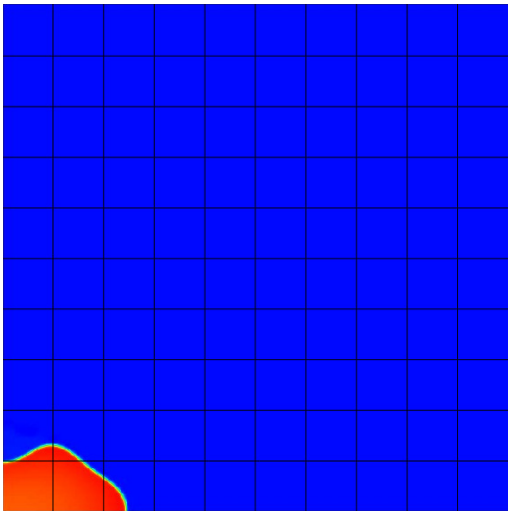

Fig. 2. The synthetic images of the tissue under consideration are thought to be broken down into a union of square pixels, as illustrated here. In the computations, the images were broken down into 70×70 pixels.

we define the vector

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{4900} \end{bmatrix}_{4900\times 1} \in (0, +\infty)^{4900} \tag{6.5}$$

by

$$u_{i\times 70+j} = \begin{cases} u_{\text{avg}}(\mathbf{x}_{i,j}, t) & \text{if } u_{\text{avg}}(\mathbf{x}_{i,j}, t) \geq u_{\text{threshold}}, \\ u_{\text{threshold}} & \text{if } u_{\text{avg}}(\mathbf{x}_{i,j}, t) < u_{\text{threshold}}, \end{cases} \tag{6.6}$$

where $1 \leq i, j \leq 70$. The value $u_{\text{threshold}}$ is necessary because we will work with logarithms in Secs. 6.2 and 6.3 below: see Eqs. (6.12) and (6.21).

For the sake of notation to be used below, let us summarize the whole procedure (6.3)–(6.6) just described by the operator "F", as follows:

$$\mathbf{u} = F(u_{\text{given}}). \tag{6.7}$$

Corresponding to the initial state (6.2), we then define

$$\mathbf{u}_0 = F(u_0), \tag{6.8}$$

which is plotted in Fig. 3.

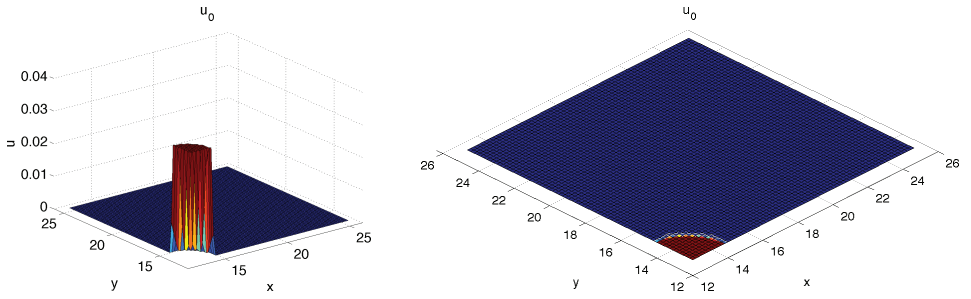

Fig. 3. Initial state used in all our numerical experiments. The tumor is located in the left-lower corner of the domain.

We also define

$$\ln(\mathbf{u}) \equiv \begin{bmatrix} \ln(u_1) \\ \ln(u_2) \\ \vdots \\ \ln(u_{4900}) \end{bmatrix}_{4900\times 1} \qquad \forall\, \mathbf{u} \in (0,+\infty)^{4900}, \tag{6.9}$$

and, regarding PDF notation, we denote by $\mathcal{U}(B)$ a uniform distribution over a given set B, and by $\mathcal{N}(\mathbf{m}, \mathbf{C})$ a (multivariate) Gaussian distribution of mean $\mathbf{m}$ and (co)variance (matrix) $\mathbf{C}$.

6.2. *The reference data D*

The reference data D used in our experiments is fabricated with a PMC M_{ref} that is different, and presumably more sophisticated, than the PMCs M_1 and M_2 below. All parameters $\boldsymbol{\theta}_{\text{ref}}$ of M_{ref} are treated deterministically (that is, their PDFs are delta functions), and M_{ref} is simulated from $t_0 = 0$ until $t = T > 0$ with time step $\Delta t < T$, with the state variable evolving from $u_0(\cdot)$ until

$$u_{\text{ref}}(\cdot) \equiv u(\cdot, T). \tag{6.10}$$

Corresponding to the reference solution (6.10), we define

$$\mathbf{u}_{\text{ref}} = F(u_{\text{ref}}) \tag{6.11}$$

and

$$D = \ln(\mathbf{u}_{\text{ref}}). \tag{6.12}$$

Figure 4 shows the reference solution with $T = 1$. Throughout the rest of Sec. 6 we will use a D computed with $T = 0.3$.

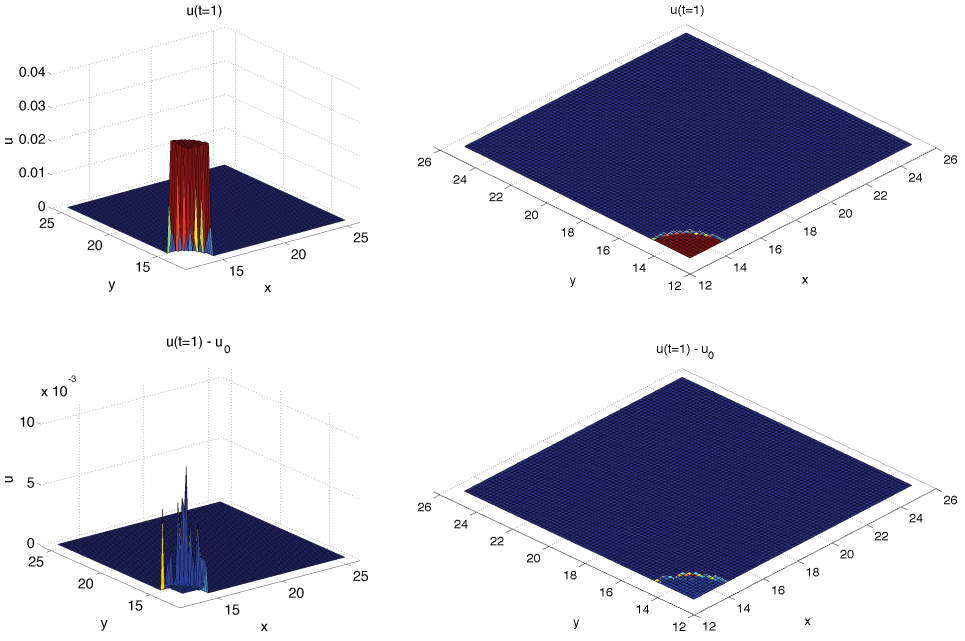

Fig. 4. Reference solution computed with $T = 1$. The tumor region expands with respect to the region of Fig. 3. The concentration of tumor cells grows (less) not only in the region of Fig. 3, but also (and more) in the expanded region.

6.3. *The likelihood*

For any given candidate PMC M (either M_1 or M_2 in Sec. 6), we first compute the solution $u_M(\cdot, T)$, and then compute the corresponding vector

$$\mathbf{u}_M = \begin{bmatrix} u_{M,1} \\ u_{M,2} \\ \vdots \\ u_{M,4900} \end{bmatrix}_{4900\times 1} = F(u_M), \tag{6.13}$$

following the procedure F explained above in (6.3)–(6.7).

Now we come up with three important hypotheses about the pixels corresponding to the actual (and unknown) concentration field $u_{\text{actual}}(\cdot, T)$ of tumor cells. Let us refer to this vector of 4900 unknown non-negative values as $\mathbf{u}_{\text{actual}}$. The three hypotheses are:

(i) We *believe* that each component of $\mathbf{u}_{\text{actual}}$ is a random variable satisfying

$$[\mathbf{u}_{\text{actual}}]_k = u_{\text{ref},k} \cdot e^{\nu_{\text{data}}} \quad \forall\, 1 \le k \le 4900, \tag{6.14}$$

where ν_{data} is a Gaussian random variable of mean 0 and (unknown) variance σ^2_{data}, that is,

$$\nu_{\text{data}} \sim \mathcal{N}(0, \sigma^2_{\text{data}}). \tag{6.15}$$

In other words, we assume that the ratio $\frac{\mathbf{u}_{\text{actual}_k}}{u_{\text{ref},k}}$ follows a log-normal distribution for each kth square in Ω, $1 \leq k \leq 4900$.

(ii) We also *believe* that each component of $\mathbf{u}_{\text{actual}}$ is a random variable satisfying

$$[\mathbf{u}_{\text{actual}}]_k = u_{M,k} \cdot e^{\nu_{\text{model}}} \quad \forall\, 1 \leq k \leq 4900, \tag{6.16}$$

where ν_{model} is a Gaussian random variable of mean 0 and (unknown) variance σ^2_{model}, that is,

$$\nu_{\text{model}} \sim \mathcal{N}(0, \sigma^2_{\text{model}}). \tag{6.17}$$

(iii) We assume that

$$\text{all components of } \mathbf{u}_{\text{actual}} \text{ are independent of each other.} \tag{6.18}$$

It is important to note that both assumptions (6.14) and (6.16) guarantee that only non-negative values are possible for all the components of $\mathbf{u}_{\text{actual}}$. Of course, one could assume many other forms of discrepancy, such as a stationary Gaussian field with zero mean and an unknown correlation length.

Now, equalizing (6.14) and (6.16), as well as using (6.12) and (6.18), we conclude that the random variable D satisfies

$$D = \ln(\mathbf{u}_M) + \boldsymbol{\nu}_{4900\times 1},$$

where

$$\boldsymbol{\nu}_{4900\times 1} \sim \mathcal{N}(\mathbf{0}_{4900\times 1}, \sigma^2 \cdot \mathbf{I}_{4900\times 4900}), \tag{6.19}$$

and

$$\sigma^2 = \sigma^2_{\text{model}} + \sigma^2_{\text{data}}. \tag{6.20}$$

Therefore, we can write, defining $N = 4900$ and $\boldsymbol{\Delta}(\boldsymbol{\theta}) = \ln(\mathbf{u}_{\text{ref}}) - \ln(\mathbf{u}_M(\boldsymbol{\theta}))$,

$$\begin{aligned}
\pi_{\text{like}}(D|\boldsymbol{\theta}, M) &= \frac{1}{\sqrt{2\pi(\sigma^2)^N}} \cdot e^{\left\{-\frac{1}{2}\cdot[\boldsymbol{\Delta}(\boldsymbol{\theta})]^T\cdot[\sigma^2]^{-1}\cdot[\boldsymbol{\Delta}(\boldsymbol{\theta})]\right\}} \\
&= \frac{1}{\sqrt{2\pi(\sigma^2)^N}} \cdot e^{\left\{-\frac{1}{2}\cdot\frac{\|\ln(\mathbf{u}_{\text{ref}})-\ln(\mathbf{u}_M(\boldsymbol{\theta}))\|_2^2}{\sigma^2}\right\}} \\
&= \frac{1}{\sqrt{2\pi(\sigma^2)^N}} \cdot e^{\left\{-\frac{1}{2}\cdot\frac{\sum_{k=1}^N[\ln(u_{\text{ref},k})-\ln(u_{M,k}(\boldsymbol{\theta}))]^2}{\sigma^2}\right\}}.
\end{aligned} \tag{6.21}$$

Formula (6.21) helps explain the motivation for the threshold value (6.4). Indeed, if both $u_{\text{ref},k}$ and $u_{M,k}$ are small enough to be smaller than $u_{\text{threshold}}$, then we consider both to be equal to a common "noise" value. The cancellation of both terms $u_{\text{ref},k}$ and $u_{M,k}$ in such cases helps avoid inappropriate contributions to the final likelihood value (6.21) from pixels that show tiny concentrations of tumor cells.

We can also rewrite (6.21) as

$$\ln[\pi_{\text{like}}(D|\boldsymbol{\theta}, M)] = -\frac{1}{2} \cdot \ln(2\pi) - N \cdot \ln(\sigma) - \frac{1}{2} \cdot \frac{\| \ln(\mathbf{u}_{\text{ref}}) - \ln(\mathbf{u}_M(\boldsymbol{\theta}))\|_2^2}{\sigma^2}. \tag{6.22}$$

It is easy to see that, for any fixed $\boldsymbol{\theta}^*$ the maximum likelihood value occurs for

$$\sigma^* = \frac{\|\ln(\mathbf{u}_{\text{ref}}) - \ln(\mathbf{u}_M(\boldsymbol{\theta}^*))\|_2}{\sqrt{N}}. \tag{6.23}$$

Indeed, as $\sigma \to 0$ or $\sigma \to +\infty$, one has $\ln[\pi_{\text{like}}(D|\boldsymbol{\theta}, M)] \to -\infty$.

6.4. *The PMC M_1*

PMC M_1 uses the model equations A_1 of Sec. 2, with only the parameter P being treated randomly besides σ from (6.20), that is,

$$\boldsymbol{\theta_1} = (P, \sigma).$$

We assume a prior PDF $\mathcal{U}((0.001, 0.201) \times (0.001, 0.301))$ for (P, σ). The other parameter values are shown in Table 1. Following (6.22), the likelihood $\pi_{\text{like}}(D|\boldsymbol{\theta_1}, M_1)$ for M_1 is such that

$$\ln[\pi_{\text{like}}(D|\boldsymbol{\theta_1}, M_1)] = -\frac{1}{2} \cdot \ln(2\pi) - N \cdot \ln(\sigma) - \frac{1}{2} \cdot \frac{\|\ln(\mathbf{u}_{\text{ref}}) - \ln(\mathbf{u}_{M_1}(\boldsymbol{\theta_1}))\|_2^2}{\sigma^2}. \tag{6.24}$$

6.5. *The PMC M_2*

PMC M_2 uses the model equations A_2 of Sec. 2, with only the parameter P being treated randomly besides σ from (6.20), that is,

$$\boldsymbol{\theta_2} = (P, \sigma).$$

We assume a prior PDF $\mathcal{U}((0.001, 0.201) \times (0.001, 0.301))$ for (P, σ). The other parameter values are shown in Table 1. Following (6.22), the likelihood

Table 1. Parameters involved in both PMCs.

Parameter	PMC M_1	PMC M_2
ζ	200	200
P	$\mathcal{U}((0.001, 0.201))$	$\mathcal{U}((0.001, 0.201))$
A	Not applicable	0
ϵ	0.005	0.005
τ	Not applicable	7
ξ	Not applicable	1
γ	0.045	0.045
σ	$\mathcal{U}((0.001, 0.301))$	$\mathcal{U}((0.001, 0.301))$

$\pi_{\text{like}}(D|\boldsymbol{\theta}_2, M_2)$ for M_2 is such that

$$\ln[\pi_{\text{like}}(D|\boldsymbol{\theta}_2, M_2)] = -\frac{1}{2}\cdot\ln(2\pi) - N\cdot\ln(\sigma) - \frac{1}{2}\cdot\frac{\|\ln(\mathbf{u}_{\text{ref}}) - \ln(\mathbf{u}_{M_2}(\boldsymbol{\theta}_2))\|_2^2}{\sigma^2}. \tag{6.25}$$

6.6. *Some implementation aspects and checks*

Although our work relies on many existing libraries, such as QUESO,[36] libMesh,[27] PETSc, GSL, BOOST, and STL, we also use two new codes to perform the experiments of Sec. 6. One code, referred to as Model Library, deals with the tumor modeling equations.[19] The other code, referred to as Top Application, deals with the definition of statistical inverse problems (parameter spaces, prior PDFs, likelihood functions, reference data), as well as with the proper use of QUESO C++ classes in order to solve such statistical inverse problems through Bayesian formula and MCMC algorithms, and in order to calculate evidences and other integrals. Figure 5 gives an overview of the software modules we use in our work, while Table 2 gives the values of some key parameters involved in the solution of each sampled tumor growth problem.

Before performing the long MCMC runs for computing the evidences (to be reported in Sec. 6.7), we vary only one parameter, fixing all remaining parameters, and plot the values of (6.24) and (6.25). Figure 6 shows the plots. These tests

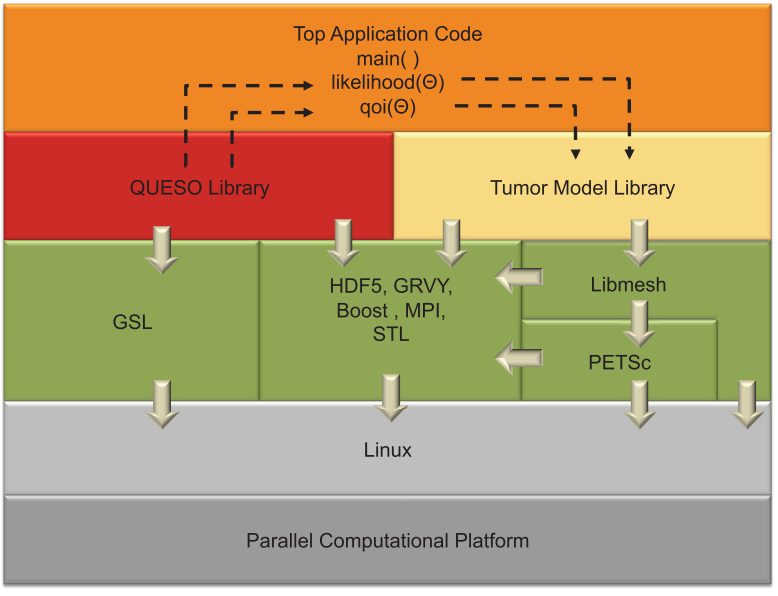

Fig. 5. The software stack related to our experiments.

Table 2. Numerical parameters used in our time-dependent experiments, with time $0 \leq t \leq T = 0.3 = \Delta t \times n_{\Delta t}$.

Numerical parameter	PMC M_1	PMC M_2
Time step Δt	0.075	0.075
Number $n_{\Delta t}$ of time steps	40	40
Time integration scheme	FullLinSplitLinMob	FullLinSplitNonMob
Refinement period in time steps	35	30
$u_{\text{threshold}}$	10^{-5}	10^{-5}
Number of Markov chains	16	16
Number of processors/chain	6	6
Total number of processors	96	96
Total number of samples/MCMC level	1024	1024
(for last level)	(2048)	(2048)
Number of samples/MCMC level/chain	64	64
(for last level)	(128)	(128)

Both M_1 and M_2 each used a total of eight MCMC levels (see explanation on Sec. 5.4).

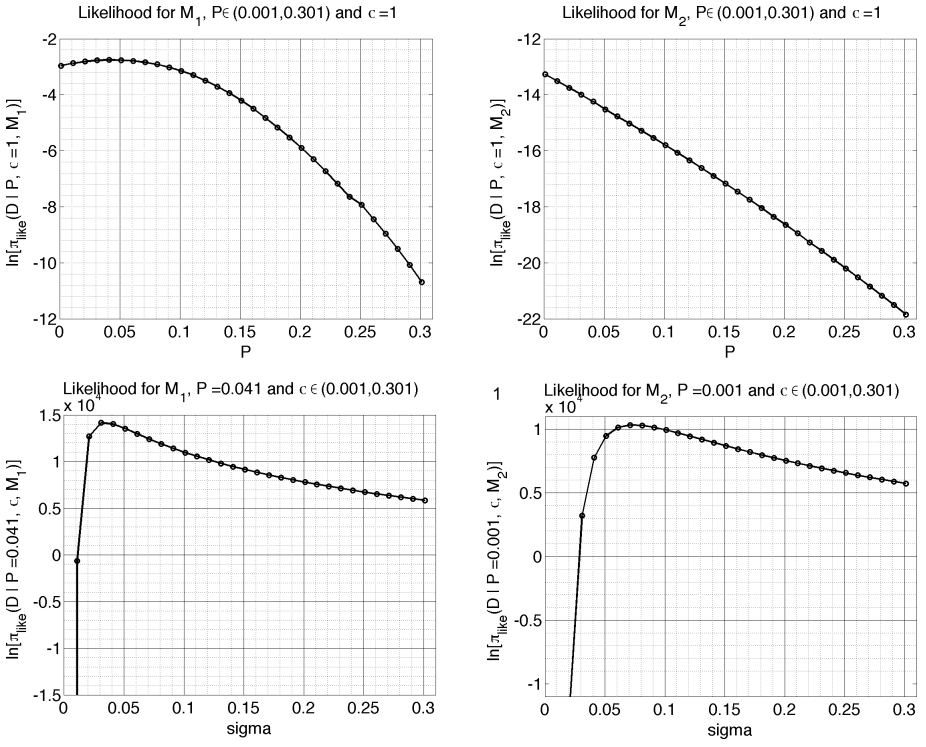

Fig. 6. Some checks with M_1 (left column) and M_2 (right column). One should note that these plots are *not* marginal PDFs, but rather just slices of the joint PDF of (P, σ) at fixed values of P or σ. That is, these plots give us some information, but should *not* be directly compared to the marginal plots of Fig. 7. Also, here we plot the natural logarithmic values, while in Fig. 7 we plot the marginal PDFs themselves.

help us to check the stability of the Model Library, judge the correctness of the Top Application code and of the likelihood routines, and gain confidence that the suggested prior PDFs for (P, σ) make sense.

6.7. *Numerical results*

While the prior $p_{\text{prior}}(M_j|\mathcal{M})$ should be chosen to reflect previous experience, for simplicity we choose equal plausibilities of $\frac{1}{2}$ for each PMC. Figure 7 shows the

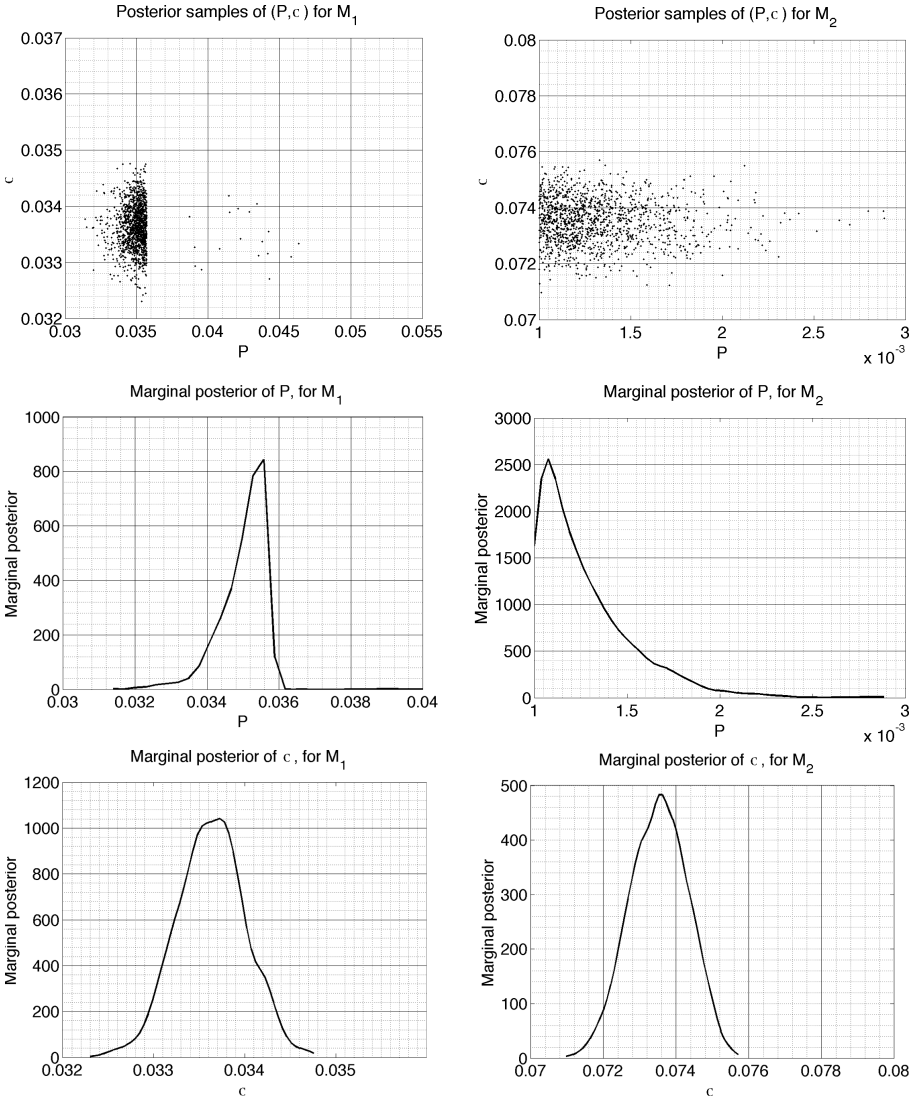

Fig. 7. Posterior samples and posterior marginal PDFs for M_1 (left column) and M_2 (right column). A total of 2048 samples is used for each plot.

Table 3. Numerical results on log evidence, model fitness and model complexity (see Eq. (5.8)).

PMC	Log evidence $\ln(\pi(D\|M_j))$	Model fitness $E[\ln(\pi_{\text{like}}(\cdot))]$	Model complexity $E\left[\ln \frac{\pi_{\text{post}}(\cdot)}{\pi_{\text{prior}}(\cdot)}\right]$	Prior plausibility $p(M_j\|\mathcal{M})$	Posterior plausibility $p(M_j\|D,\mathcal{M})$
M_1	14171.9	14182.8	10.9	1/2	$\approx$1
M_2	10329.8	10340.6	10.8	1/2	$\approx$0

computed posterior samples and posterior marginal PDFs, while Table 3 shows the computed evidences, model fitness and model complexity. Some comments:

- As T becomes smaller (e.g. $T = 0.3$, as in our reported experiments), the optimum values (6.23) will tend to become very small, causing the values (6.22) to become very big, as shown in Table 3.
- The σ term can be interpreted as a measure of the overall discrepancy between PMC and D, and so it is expected that the PMC with smaller σ has higher evidence, again as shown by Fig. 7 and Table 3.
- The (much) worse evidence of M_2 should *not* be interpreted as a sign of the inappropriateness of the equations in M_2, but rather of our choices: which parameters to treat deterministically in M_2, as well as of which values to give them.

7. Conclusion

Perhaps the central issues in the development of predictive phenomenological models of tumor growth in living tissue are: (i) the selection of models best informed by available data and (ii) the development of methods to quantify uncertainty in key quantities of interest. In this work, we described a Bayesian framework for uncertainty quantification, calibration, validation, and selection of models, as well as demonstrated applications on model problems. Of particular importance is the study of models able to cope with the possible change, over time, of the type of cancer in a living tissue. Since the evolution of the type of cancer depends on the underlying dynamics at the microscale, it does not seem appropriate to postulate a model structure *a priori*. Instead, the structure of a candidate model should be allowed to change over time, and Bayesian filtering could be applied so that the candidate model is calibrated every time some data is collected from the evolving tissue. In Bayesian filtering approaches, a model update does not happen only once, in a batch mode, e.g. long after data is collected, but instead in a real-time fashion. That is, models are given the chance to relearn from newer collected data, and they can readapt themselves to the instantaneous information on reality. Given candidate models competing to describe an unsteady cancer-type situation, some models will rank better than others, as it also happens in steady-state cases. Even the relative ranking among competing models might change over time. It should be noted that the Bayesian framework for model calibration, comparison, and averaging, laid out

in Secs. 4 and 5 can be applied to either fixed or adaptable models, and to either steady-state or time-evolving data.

The development of successful methodologies thus draws on a diverse range of technologies, including tumor morphology, Bayesian inference, information theory, parallel sampling algorithms, imaging, and high performance computing. The foundations for a class of approaches to all these problems are laid down in this work. To advance this work further will require deeper interactions with data-retrieving technologies, imaging, *in vitro* experiments, and data-intensive computing algorithms. We hope to take up these extensions in future work.

Acknowledgments

The support of this work under DOE contract DE-FC52-08NA28615 in connection with the Predictive Science Academic Alliance Program, the DOE Multiscale Mathematics Program through contract DE-FG02-0SER25701, and NSF Grant DMS-1115865, is gratefully acknowledged. The second author is thankful to Dr. Sai-Hung Cheung for numerous technical discussions on Bayesian inference methodologies.

References

1. R. P. Araujo and D. L. S. McElwain, A mixture theory for the genesis of residual stresses in growing tissues I: A general formulation, *SIAM J. Appl. Math.* **65** (2005) 1261–1284.
2. I. Babuška, F. Nobile and R. Tempone, A systematic approach to model validation based on Bayesian updates and prediction related rejection criteria, *Comput. Methods Appl. Mech. Engrg.* **197** (2008) 2517–2539.
3. J. L. Beck and K.-V. Yuen, Model selection using response measurements: Bayesian probabilistic approach, *J. Engrg. Mech.* **130** (2004) 192–203.
4. N. Bellomo, A. Bellouquid, J. Nieto and J. Soler, On the asymptotic theory from microscopic to macroscopic growing tissue models: An overview with perspectives, *Math. Models Methods Appl. Sci.* **22** (2012) 1130001, 37 pp.
5. N. Bellomo, M. Chaplain and E. De Angelis (eds.), *Selected Topics in Cancer Modeling*: *Genesis*, *Evolution*, *Immune Competition*, *and Therapy* (Springer, 2008).
6. N. Bellomo, N. K. Li and P. K. Maini, On the foundations of cancer modelling: Selected topics, speculations, and perspectives, *Math. Models Methods Appl. Sci.* **18** (2008) 593–646.
7. J. O. Berger, L. R. Pericchi, J. K. Ghosh, T. Samanta and F. De Santis, Objective Bayesian methods for model selection: Introduction and comparison, *IMS Lecture Notes Monogr. Ser.* **38** (2001) 135–207.
8. R. M. Bowen, Theory of mixtures, in *Continuum Physics*, ed. A. D. Eringen, Vol. 3 (Academic Press, 1976).
9. F. Cavallo, C. De Giovanni, P. Nanni, G. Forni and P.-L. Lollini, 2011: The immune hallmarks of cancer, *Cancer Immunol. Immunother.* **60** (2011) 319–326.
10. V. Cristini, X. Li, J. S. Lowengrub and S. Wise, Nonlinear simulations of solid tumor growth using a mixture model: Invasion and branching, *J. Math. Biol.* **58** (2009) 723–763.
11. V. Cristini and J. S. Lowengrub, *Multiscale Modeling of Cancer*: *An Integrated Experimental and Mathematical Approach* (Cambridge Univ. Press, 2010).

12. H. Darcy, *Les Fontaines Publiques de la Ville de Dijon* (Dalmont, 1856).
13. A. Fick, Poggendorff's Annel, *Physik* **94** (1855) 59.
14. H. B. Frieboes, M. A. Chaplain, A. Thompson, E. L. Bearer, J. S. Lowengrub and V. Cristini, Physical oncology: A bench-to-bedside quantitative and predictive approach, *Cancer Res.* **71** (2011) 298–302.
15. H. B. Frieboes, J. S. Lowengrub, S. Wise, X. Zheng, P. Macklin, E. L. Bearer and V. Cristini, Computer simulation of glioma growth and morphology, *NeuroImage* **37** (2007) S59–S70.
16. D. Hanahan and R. A. Weinburg, The hallmarks of cancer, *Cell* **100** (2000) 57–70.
17. D. Hanahan and R. A. Weinberg, Hallmarks of cancer: The next generation, *Cell* **144** (2011).
18. L. H. Hartwell, J. J. Hopfield, S. Leibler and A. W. Murray, From molecular to modular cell biology, *Nature* **402** (1999) c47–c52.
19. A. Hawkins-Daarud, Toward a predictive model of tumor growth, Ph.D. thesis, The University of Texas at Austin (2011).
20. A. Hawkins-Daarud, S. Prudhomme, K. G. van der Zee and J. T. Oden, Bayesian calibration, validation, and uncertainty quantification of diffuse interface models of tumor growth, ICES Report 10-44, The University of Texas at Austin (2010).
21. A. Hawkins-Daarud, K. G. van der Zee and J. T. Oden, Numerical simulation of a thermodynamically-consistent four-species tumor growth model, *Int. J. Num. Meth. Biomed. Engrg.* **28** (2012) 3–24.
22. J. A. Hoeting, D. Madigan, A. E. Raftery and C. T. Volinsky, Bayesian model averaging: A tutorial (with discussion), *Statist. Sci.* **14** (1999) 382–417.
23. E. T. Jaynes, Probability theory: The logic of science (Cambridge Univ. Press, 2003).
24. W. Jefferys and J. O. Berger, Ockham's razor and Bayesian analysis, *Amer. Sci.* January–February (1992) 64–77.
25. H. Jeffreys, Some tests of significance treated by the theory of probability, *Proc. Cambridge Philos. Soc.* **31** (1935) 203–222.
26. H. Jeffreys, *Theory of Probability*, 3rd edn. (Oxford Univ. Press, 1961).
27. B. S. Kirk, J. W. Peterson, R. H. Stogner and G. F. Carey, libMesh: A c++ library for parallel adaptive mesh refinement/coarsening simulations, *Engrg. Comput.* **22** (2006) 237–254.
28. S. Kullback and R. A. Leibler, On information and sufficiency, *Ann. Math. Statist.* **22** (1951) 79–86.
29. M. Muto and J. L. Beck, Bayesian updating and model class selection for hysteretic structural models using stochastic simulation, *J. Vib. Cont.* **14** (2008) 7–34.
30. J. T. Oden, A. Hawkins and S. Prudhomme, General diffuse-interface theories and an approach to predictive tumor growth modeling, *Math. Models Methods Appl. Sci.* **20** (2010) 477.
31. J. T. Oden, R. Moser and O. Ghattas, Computer predictions with quantified uncertainty, part I, *SIAM News* **43**(9) (2010) 1–3.
32. J. T. Oden, R. Moser and O. Ghattas, Computer predictions with quantified uncertainty, part II, *SIAM News* **43**(10) (2010) 1–4.
33. K. Popper, *The Logic of Scientific Discovery* (Hutchinson and Co., 1958).
34. L. Preziosi (ed.), *Cancer Modelling and Simulation* (Chapman and Hall/CRC, 2003).
35. E. E. Prudencio and S. H. Cheung, Parallel adaptive multilevel sampling algorithms for the Bayesian analysis of mathematical models, *Int. J. Uncertain. Quantification* **2** (2012) 215–237.
36. E. E. Prudencio and K. W. Schulz, The parallel C++ statistical library QUESO: Quantification of uncertainty for estimation, simulation and optimization, in *Euro-Par*

2011 Workshops, Part I, eds. M. Alexander *et al.*, Lecture Notes in Computer Science, Vol. 7155 (Springer-Verlag, 2012), pp. 398–407.
37. A. E. Raftery, D. Madigan and J. A. Hoeting, Bayesian model averaging for linear regression models, *J. Amer. Statist. Assoc.* **92** (1997) 179–191.
38. K. R. Rajagopal, On a hierarchy of approximate models for flows of incompressible fluids through porous solids, *Math. Models Methods Appl. Sci.* **17** (2007) 215–252.
39. C. P. Robert, *The Bayesian Choice*: *From Decision-Theoretic Foundations to Computational Implementation* (Springer Science+Business Media, 2007).
40. K. R. Swanson, R. C. Rockne, J. Claridge, M. A. Chaplain, E. C. Alvord Jr. and A. R. Anderson, Quantifying the role of angiogenesis in malignant progression of gliomas: *In silico* modeling integrates imaging and histology, *Cancer Res.* **71** (2011) 7366–7375.
41. TACC, Texas Advanced Computing Center, http://www.tacc.utexas.edu/ (2008–2012).
42. C. Truesdell, Sulle basi della thermomeccanica, *Rend. Lincei* **22** (1957) 33–38, 158–166.
43. C. Truesdell and R. Toupin, The classical field theories, in *Handbuch der Physik*, ed. S. Flugge, Vol. III/I (Springer-Verlag, 1960).
44. C. H. Wang, J. K. Rockhill, M. Mrugala, D. L. Peacock, A. Lai, K. Jusenius, J. M. Wardlaw, T. Cloughesy, A. M. Spence, R. Rockne, E. C. Alvord Jr. and K. R. Swanson, Prognostic significance of growth kinetics in newly diagnosed glioblastomas revealed by combining serial imaging with a novel biomathematical model, *Cancer Res.* **69** (2009) 9133–9140.
45. S. M. Wise, J. S. Lowengrub, H. B. Frieboes and V. Cristini, Three-dimensional diffuse-interface simulation of multispecies tumor growth-I. Model and numerical method, *J. Theor. Biol.* **253** (2008) 523–543.