MRI Based Bayesian Personalization of a Tumor Growth Model - Review and analysis

Charlotte Desanges, Philippine Cordelle and Julien Bergerot

I. Introduction

This paper focuses on the mathematical modeling of brain tumor growth. If most of the models use the reaction-diffusion equation, the difficulty lies in the personalization, i.e. identifying the parameters involved in the diffusion and the proliferation for a particular patient. The paper presents two previous personalization works, and compares them with a new method based on a Bayesian personalization. From a clinical point of view, they focus on Glioblastomas multiformes (GBM) as they are the most common and aggressive sub-type of primary brain tumors (10.000 new cases each year in the US).

II. Model

A. Previous work

The goal is to model the spatio-temporal evolution of the tumor cells density in the considered domain which is mainly impacted by two effects : the proliferation of the tumor cells, and the diffusion in neighbouring tissues. The Reaction-Diffusion Equation (RDE) reflects those two effects through its two terms: one depending on the diffusion tensor D and another reflecting the reaction.

$$\frac{\partial u}{\partial t} = \nabla(D.\nabla u) + R(u) \tag{1}$$

Since 1989, when the RDE was first used, the model has evolved and anatomical constraints have been added. For instance the modeling of mass effect, by coupling a mechanical model (2005), modeling of higher diffusion in the white matter (2005), division of tumor cells into sub-categories (2011, 2014), or modeling of the edema apparition.

B. Reaction-Diffusion Model

The Reaction-diffusion equation used in this paper is the following:

$$\frac{\partial u}{\partial t} = \nabla(D.\nabla u) + \rho u(1-u) \tag{2}$$

$$D\nabla u.\vec{n}_{\partial\Omega} = 0 \tag{3}$$

$$\theta = (D, \rho) \tag{4}$$

where ρ is the proliferation rate.
Assuming that the solution is a wave travelling at speed $v = 2\sqrt{\rho D}$, we can look for solution $u(x,t) = u(x-vt) = u(\xi)$. By linearizing for $u << 1$ we have an equation of the form:

$$n' D n \frac{\partial^2 u}{\partial^2 \xi} + v\frac{\partial u}{\partial \xi} + \rho u = 0 \tag{5}$$

which admits solutions of the form $u(\xi) = (A\xi + B)exp(-x/\lambda)$ where A and B are constant and $\lambda = \sqrt{D/\rho}$. λ is known as the **invisibility index**.

In order to take into account the different motility in the white and gray matter, the diffusion tensor is adapted :
$D = d_\omega \mathbb{1}$ for WM, and $D = (d_\omega/10)\mathbb{1}$ for GM.

C. Model and MRI

The data used in the models are acquired on MRI, using T1Gd and T2-FLAIR to reveal different parts of the tumor: T1Gd reveals the active part, and T2-FLAIR the edema. The data are collected on two consecutive time points, resulting in 4 abnormalities segmented by the clinician. In order to relate the tumor cell density u to the MRIs, the frontier of the abnormalities is assumed to correspond to a threshold value of the cell density.

III. Simulation

A. Data Preprocessing

The preprocessing step is very crucial for this project. This step can be separated into two.
The first step relates the segmentation of the abnormality in the different images. For each case, we will be working with two MRIs T1Gd (time t1 and t2), two MRIs T2-FLAIR (time t1 and t2) and the fractional anisotropy (FA) that is extracted from the Diffusion Tensor Image (DTI). For the first four images, a clinician segments the abnormalities. Then we need a way to map the segmentation on time t2 to the one on t1. This is done in three steps. Firstly, the T2-FLAIR and FA MRIs are registered linearly to the T1Gd at each time (the RR arrows on Figure 1). In a second time, the T1Gd at time t2 is non linearly registered to the T1Gd at time t1 (NLR arrow on Figure 1). This is done using the FNRIT function of the FSL software. During this part, the segmentation at t2 (because it is bigger) is used to exclude the abnormality from the similarity criterion. Therefore, the tumor is ignored and simply considered as its surroundings. In a third time, these transformations are used to transport the abnormalities (T2-FLAIR at t1 and t2 and T1Gd) on the T1Gd MRI at t1.
The second step aims at extracting white and gray matter and cerebrospinal fluid (CSF) from the MR1 at time t1 (as this separation is not supposed to change over time). To do so, the brain is firstly extracted from the skull. Then, the white matter, gray matter and CSF are segmented. We proceed to the separation of the right and left hemisphere. Furthermore, the separation between the left and right is labeled as CSF to prevent the tumor from invading the contra-lateral hemisphere

through the falx cerebri. Finally, voxels with FA ≥ 0.45 and labeled as white matter (initially) are not affected by the aforementioned process so the corpus callosum remains white matter.

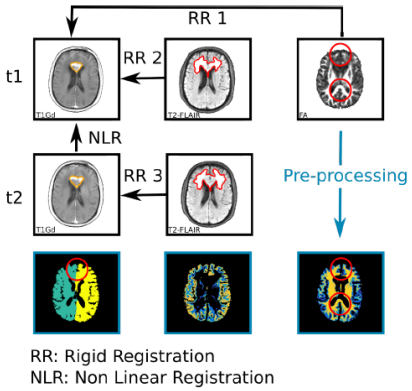

Figure 1: Preprocessing steps of the different MRI images. The arrows show the different transformations.

B. Latice Boltzmann Method

In this paper, the method used to process the reaction-diffusion equation relies on a set of fictitious particles living in a cartesian grid. These particles will collide into each other and stream on the grid. To do so, they use the D3Q7 (resemble to the 6 connectivity in 3D with the addition of the 0 vector) is order to discretize the solution.

$$[e_0, e_1, e_2, e_3, e_4, e_5, e_6] = \begin{bmatrix} 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}$$

Each of these velocities are weighted according to the position (center is more important). The algorithm consists in 4 steps :

- Initialisation : $u_\alpha = w_\alpha * u$
- Collision : $u^*_\alpha = u_\alpha - \Sigma A_{\alpha,i} u_i + \Delta t w_\alpha \rho u(1-u)$ where $A = M^{-1}SM$ is the collision matrix. M projects the vector on the moment space.

$$M = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 6 & -1 & -1 & -1 & -1 & -1 & -1 \\ 0 & 2 & 2 & -1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 1 & 1 & -1 & -1 \end{bmatrix}$$

S is the relaxation time matrix.

$$S^{-1} = \begin{bmatrix} \tau_0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \tau_{xx} & \tau_{xy} & \tau_{xz} & 0 & 0 & 0 \\ 0 & \tau_{xy} & \tau_{yy} & \tau_{yz} & 0 & 0 & 0 \\ 0 & \tau_{xz} & \tau_{yz} & \tau_{zz} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \tau_4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \tau_5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \tau_6 \end{bmatrix}$$

where $\tau_{ij} = 1/2\delta_{ij} + 4\Delta t/(\Delta x^2) D_{ij}$ and $\tau_k = 1.33$. τ_k only affects the stability of the method and according to the paper, the value of 1.33 showed reasonable results.

- Streaming : $u_\alpha(x + \delta x e_\alpha, t + \delta t) = u^*_\alpha(x,t)$. The particle are going forward in time.
- Neumann boundary condition : There are two equations, use one depending the distance with the boundary noted Δ ((6) if $\Delta > 1/2$ (7) else).

$$u_{\alpha^{-1}}(x, t+\delta t) = \frac{1}{2\Delta} * u_\alpha(x,t) + \frac{2\Delta - 1}{2\Delta} u_{\alpha^{-1}}(x,t) \quad (6)$$

$$u_{\alpha^{-1}}(x, t+\delta t) = 2\Delta * u_\alpha(x,t) + (1-2\Delta) u_{\alpha^{-1}}(x - \Delta x e_\alpha, t) \quad (7)$$

C. Initialization

The initialization of $u(x, t = t1)$ is very important as the rest of the simulation will somehow entirely rely on it for its evolution. To do so, the tumor tail extrapolation algorithm is used. This algorithm is an approximation of the static equation (1). This is computed on the T1Gd MRI abnormality border. Only the invisibility index λ is used during the process. By construction, the found density at time t1 is exactly the same as the T1Gd abnormality at time t1.

IV. Personalization

To perform the initialization of the density and also to compute its evolution via the LBM, we need to have access to the diffusion factor D and the proliferation rate ρ. This can be obtained by a simple analysis of the evolution of the tumor. Also, BOBYQA, a free-derivative optimization method can be used to infer these two parameters. Finally, this paper suggests a third method that relies on the estimation of posterior probability of D and ρ, using the Gaussian Process Hamiltonian Monte Carlo.

A. Spherical Asymptotic Analysis

We can relate D and ρ to the velocity $v = 2\sqrt{D\rho}$ and the invisibility index $\lambda = \sqrt{D/\rho}$. Finding these two new values is equivalent to finding D and ρ.

This paper makes the hypothesis that a simple relation links those values and the radius of the abnormalities at different time. Therefore, we can measure the velocity growth as the temporal variation of the radius. This can be calculated on the T1Gd and T1-FLAIR MRIs and may give some different results. In order to merge these results, the geometric mean is computed to estimate the final v. The final formulas can be found on Figure 8.

In order to measure the invisibility index, several simulations were done with different values of D and ρ. For each simulation, the difference of the radius (T2-FLAIR – T1Gd) was computed and this value was plotted according to the index. With different threshold (80% and 16% to estimate the radius in T1 and T2 MRI), we can infer the second relation from Figure **??**.

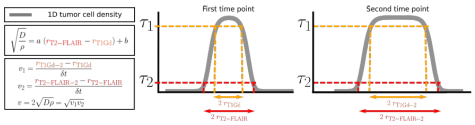

Figure 2: Left - Formulas to infer the velocity and invisibility index from the radius. Center and Left - Measure of the radius at different times with two different threshold according to the MRI.

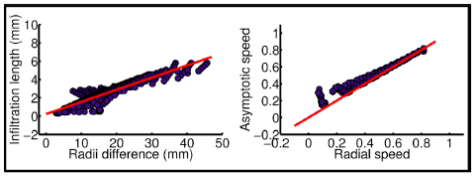

Figure 3: Left - Invisibility index according to the difference of the radius. Right - Velocity according to the growth of the radius of the abnormality.

B. BOBYQA Optimisation

This method is an improvement of the previous as it aims at considering the heterogeneity and anisotropy of the brain (white matter and gray matter will not give the same evolution) and account for the anatomical barriers (CSF for instance).
To do so, θ is initialized and then the simulation is initialized too, using the first T1Gd abnormality. Then the simulation is processed with the LBM. The abnormalities are extracted of the results at time t1 and t2 with the correct threshold (τ_1 and τ_2). The 95^{th} percentile of the symmetric Hausdorff distance between the results and the ground truth is computed for the other 3 MRI. The mean of these errors is used as the error for this simulation. They use this error as it is non related to the size of the sample (unlike DICE score). This error is minimized via the BOBYQA (Bound Optimization BY Quadratic Approximation) optimization, which is a derivative-free algorithm. This algorithm relies on quadratic approximation, where the values are bounded.

This algorithm is run with different initializations to explore different local minima, and the best is kept at the end of the day.

C. Bayesian Personalization

The Bayesian Personalization focuses on posterior probability $P(\theta|S) \propto P(\theta)P(S|\theta)$ where S is the target set of clinicians segmentations.
We know how to compute the likelihood : it is proportional to $exp(-H^2/\sigma^2)$ where H is the mean of the 95th percentile symmetric Hausdorff distance between the segmentations S and the T1Gd and the T2-FLAIR lines simulated with θ.The noise level σ is a normalization term.
To estimate the posterior distribution $P(\theta|S)$, the paper uses a version of the Hamiltonian Monte Carlo (HMC), which is a Markov Chain Monte Carlo algorithm based on Hamilton Dynamics [2]. At each iteration, the algorithm randomly samples a momentum p from a Gaussian distribution. p and θ are then updated by simulating Hamiltonian dynamics :

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = \frac{\partial H}{\partial p}$$
$$\frac{\mathrm{d}p}{\mathrm{d}t} = -\frac{\partial H}{\partial \theta}$$

where :

$$\begin{aligned} \mathrm{H}(\theta; p) &= -logP(p, \theta|S) \\ &= \frac{p^2}{2} - logP(\theta|S) \\ &= E_{kin} + E_{pot} \end{aligned}$$

This movement is simulated with "leapfrogs" : the time is discretized in small stepsizes, then it computes a half-step update of the momenta p, then a full-step of the position θ using the new p, then the last half-step of p using θ.

The new proposal state $(\theta*; p*)$ is then accepted with probability $A = min[1; exp(-H(\theta*; p*) + H(\theta; p))]$, which should be close to 1 : if we could have computed exactly the Hamiltonian movement, the energy H would have been conserved and A=1. If the move is accepted, only the state $\theta*$ is stored as the next state of the Markov Chain.
Finally, the boundary conditions on θ (biological constraints depending namely on the type of tumor) are respected by "bouncing back" the particle (the moment p is reversed) if θ crosses a boundary when proposing a new state.

The idea behind Hamilton Dynamics is to imagine a small particle in movement ; instead of doing the equivalent of a random walk (and then accepting or rejecting it) like Metropolis-Hasting sampling, the walk is attracted to higher-probability zones : the potential energy $E_{pot} = -logP(\theta|S)$ attracts the proposal state towards its local minimas –the density's local maximas.
To reduce the number of model evaluations, the E_{pot} can be approximated with a Gaussian process: that is the global idea of Gaussian Process Hamiltonian Monte Carlo.

In particular, for these experiences, the prior $P(\theta)$ is assumed log-uniformed inside its boundaries.

V. Results

A. Synthetic case

A simulation has been run on the synthetic case, (with parameters $D = 1mm^2.day^{-1}$, and $\rho = 0.18day^{-1}$ for 30 days), underlying the results of the three different personalization methods and the influence of parameters.
Firstly, if the BOBYQA's result is close to the mode of the Bayesian model, the spherical asymptotic analysis personalization is largely under-estimating the diffusion parameter D. Moreover, this simulation shows the impact of different parameters: for instance, using a uniform prior results in a more concentrated sample, but with similar shape and location of the mode. Reducing the noise level leads to a posterior more peaked around the mode. The distance choice has also an

impact: choosing the Hausdorff distance instead of the 95th percentile results in a more spread sample.

B. Glioblastoma patients

The method has been applied to 7 patients under therapy. In the Figure 4, we can see the log-likelihood of the posterior $P(\theta|S)$ distribution, and the result of the three personalization methods. As in the synthetic case, the asymptotic analysis provides underestimated results of ρ and D. The best BOBYQA solution always falls close to a mode, and sometimes the second best falls close to a second mode.

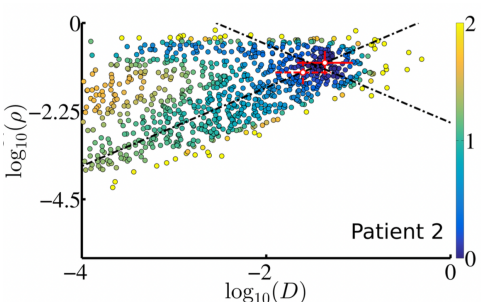

Figure 4: The result of the asymptotic personalization (dashed red lines), the result of the BOBYQA optimization (solid red lines). The color scale indicates the negative log-likelihood of the posterior P(—S), which is equal to the potential energy Epot.

This method provides several advantages. First, the posterior distribution provides further information: the colors allow to visualize the region of low potential energy (of high probability), and have an idea of the distribution, and reflects the correlation between the two parameters. It also reveal the presence of several modes if there are any. Moreover, this method doesn't seem sensitive to the presence of boundaries whereas the two other methods are, leading sometimes to nonsense results. And finally, a last key advantage, is to visualize the uncertainty of the results. For each sample results, it is possible to compute the segmentation for this result and thus compute the probability to lie in one of those segmentation, giving further information on the reliability of the model and data.

The simulation of this method has been done for 2 representative patients, and the results of the segmentation prediction is visible on Figure 5: once θ, the Maximum a Posteriori (MAP) has been determined, this parameter is used in the simulation described in this paper for the two patients. The isolines of probability 10% and 90% are also calculated, allowing to visualize the uncertainty of the MAP segmentation.

VI. Our LBM implementation

We wanted to test the Lattice Boltzmann Method depicted in the paper, here. We followed this procedure and adapted it to work on a 2D image. We only kept the 5 first components of the e_i vectors and deleted the appropriate rows and columns of S and M to only have a 5x5 matrix. We used a value of $dt = dx = 1$ and took $D = 10^(-1)$ or $D = 10^(-2)$,

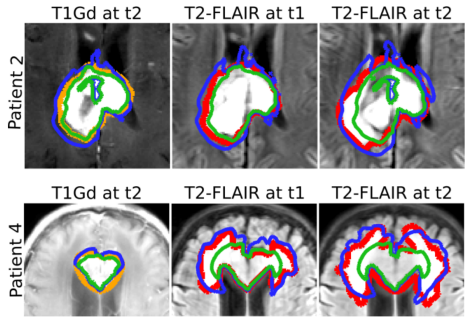

Figure 5: Clinician segmentation (orange for the T1Gd and red for the T2-FLAIR).The blue outline (resp. green) encloses the voxels which were present in at least 10% (resp. 90%) of the segmentations deduced from the samples.

depending on the position in the image, to see the difference, for instance between white and gray matter. The ρ value we selected was $10^(-1)$. We ran the algorithm for a few step and the results are presented in Figure 6.

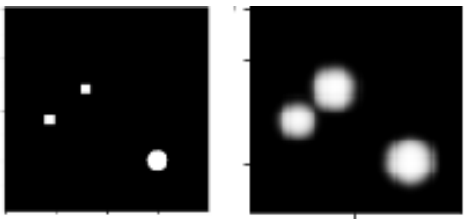

Figure 6: Left : Starting point. There are three different spots. The upper spot is in a location with a higher dilation factor. Right : Results obtained after 100 iterations.

VII. Discussion

This paper presents an implementation of the reaction-diffusion equation, based on the Lattice Boltzmann Model, and 3 methods of different complexities for the personalization. The asymptotic method is the simplest but it under-estimates the parameters ρ and D, in particular in presence of boundaries. The second method, the BOBYQA, is more accurate but can also fail when in presence of boundaries. Moreover, some differently initialized solutions lead to very poor results: the method requires on average 20 simulations for each initialization. The third method is based on the GPHMC: it provides additional information, such as the correlation between the parameters, the shape of the posterior or the presence of several modes, which can be very valuable. However, it requires a total of 1100 model evaluations, around 6 times what is necessary for BOBYQA.

VIII. Critics and other ideas

A. The simple reaction-diffusion model

A first criticism that could be made on this article is the very simple equation used to model the reaction-diffusion. The spatio-temporal evolution of the tumor cell density is

indeed only described by an equation with two additional variables : the diffusion tensor D and the net proliferation rate ρ (and another equation enforcing the Neumann boundary conditions).
In some others papers, the equations sometimes take into account many other elements :

1) The influence of oxygen concentration and vasculature: In [5], the evolution of the tumor cell density is tightly linked with the concentration of oxygen $n(x,t)$ (normalized between 0 and 1) and the density of functional vasculature $v(x,t)$. In one dimension :

$$\frac{\partial c}{\partial t} = D\frac{\partial^2}{\partial x^2}(\alpha c) + \rho\beta c(1-c)$$
$$\frac{\partial n}{\partial t} = D_n\frac{\partial^2 n}{\partial x^2} + h_1 v\,(n(t=0) - n) - h_2 cn$$
$$\frac{\partial v}{\partial t} = D_v\frac{\partial^2 v}{\partial x^2} + g_1\mathcal{H}\,(n - n(t=0))\,v(1-v) - g_2 v c^{\delta}$$

with $\mathcal{H}(\cdot)$ a sigmoidal function, h1 the oxygen consumption by tumor cells, g1 the vascular formation due to the tumor, and g2 the occlusion rate by compression from tumor cells. The functions $\alpha = \alpha(n)$ and $\beta(n)$ account for the dependence of D and ρ on the oxygen level.
Actually, the tumor cell denisty, the oxygen concentration and the vasculature are tightly linked and evolve together, so it's important to write these tangled equations.
Note that if we fix n=1 (the maximal oxygen rate), we get back the original paper's equation.

2) The tumor's heterogeneity: Another biological parameter to take into account is the tumor heterogeneity: a tumor is composed of Tumor Cells (TC) and Cancer Stem Cells (CSC) which possess specific characteristics such as the ability to generate tumors, causing relapse and metastasis, driving treatment resistance, and tumor progression. [6] The study of the evolution of cells density should differentiate the TCs and CSCs density, as they play different roles: increasing cell death could even accelerate tumor growth, as TCs compete for space and resources with CSCs, preventing division and driving tumors into dormancy. The death of TCs could result in a liberation of space for CSCs and then a renewed proliferation.

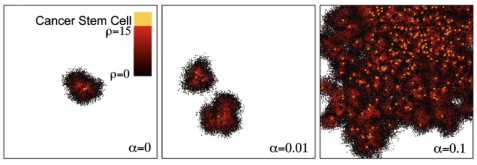

Figure 7: Representative simulations of tumors developing with spontaneous cell death rates $p\alpha = 0, 0.01, 0.1$ after t = 8 months (ρ = 15, $\mu = 15$)

The mathematical model modelling this differentiated evolution is composed of two coupled equations describing CSC et TC dynamics, here using simplifying assumptions presented in the article:

$$\frac{\partial u}{\partial t}(x,t) = D_u\Delta u + \delta\gamma k(p(x,t))\bar{u}(t)$$
$$\frac{\partial v}{\partial t}(x,t) = D_v\Delta v + (1-\delta)\gamma k(p(x,t))\bar{u}(t) - \alpha v(x,t) + \rho k(p(x,t))\bar{v}(t)$$

with:
- $k(x,y,p)$ the spatial distribution kernel describing the rate of progeny contribution to location x from a cell at location y,
$\bar{u}(t)$ and $\bar{v}(t)$ the mean density of CSCs, and TCs
- γ and ρ, the number of cell cycle times per unit time of CSCs and TCs respectively,
- δ the fraction of CSC divisions that give two CSCs,
- α the TC death rate (α=0 for CSCs).
In the end, the growth rate doesn't only depend on the tumoral cell concentration, so a distinction between CSC and TC is necessary to better model the growth.

3) Drug Resistance: Drug resistance is one of the greatest obstacles to the treatments' success in oncology: as it was stated that all patients were under chemotherapy, it would have been very interesting to add a parameter modeling and personalizing a patient's resistance to anticancer treatment, in order to apply afterwards a precision medicine accordingly.
A common assumption is to differentiate two types of tumor cells : the drug-sensitive and the drug-resistant ones [1]. A new parameter coding for the mutation rate from sensitive to resistant cells is introduced, and the additional treatment term only affects sensitive cells:

$$\frac{\partial c}{\partial t} = F(c) - K_D(t)Exposure(t)c$$

with $K_D(t) = K_{D,0}e^{-\lambda t}$ with $K_D(t)$ is the cell kill rate of the drug : it decreases exponentially with time to account for the development of resistance. Finally, $Exposure(t)$ codes for the drug exposure at time t. K_D and λ vary between patients, like D and ρ in the original model.

Of course a lot of others biological parameters could be added in the modelization to better predict the tumor growth, but simplifying assumptions are necessary to compute a mathematical model. For instance, the assumption of a constant proliferation and diffusion for all time points limit the model's ability to describe the nonlinear character of the real tumor progression.

B. Another model for parameter estimation

It's also possible to estimate the parameters by other ways. In 2020, another model - based on a mixture-density network[3] - has been proposed to estimate the parameters, and compare the results to those obtained with a Bayesian method with MCMC sampling. The technique allows an evaluation of the posterior distribution of the parameters by training a network. The tumor growth model is the same, using the same

reaction-diffusion equation, different degree of infiltration in the white and gray matters, and introducing thresholds to simulate isolines of the tumor cell density matching outlines of the tumor. At the end, there is a correlation between the parameters obtained with a neural posterior estimator (NPE) and the explicit Bayesian inference with MCMC sampling. However, the NPE is more computationally efficient, giving a relevant estimation after 4 sampling, against 20 for the MCMC method.

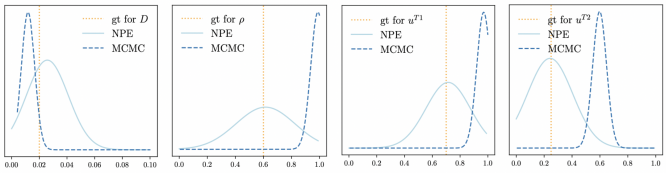

Figure 8: Comparison for NPE and MCMC methods. Depicted by red stars and orange vertical lines are ground truth data

Hence this method allows for more efficient parametric estimation, and presents another advantage: this network can serve as a universal function approximation, and exempts from the necessity to introduce an explicit form for the likelihood.

C. Adding automatic initialization

Finally, a major drawback of this method is the need of having clinicians to draw the segmentation at two time dates in order to do the personalization step.
This automatic segmentation step can be done with mere edge detection with the study of pixel-wise gradient : edge detection algorithms, image thresholding, ...

Today, Deep Learning is more and more used [4] : tumors in the brain can have many kind of shapes, sizes and contrasts, so Deep Learning is quite suitable for this complex task.

We could also think of solving the whole problem with Deep Learning : Giving as input of the neural network two tumor images at two different time dates, and applying a loss function on the segmentation of the predicted tumor months after at the output.
A first obstacle is to maintain the same time gap between the two images : this parameter would be hard to communicate to the network.
Moreover, it is well known that neural networks are suitable to deal with the underlying complexity of biomedical problems, but they often do so without taking into account the knowledge we have of the physical and biological interactions : for example, the diffusion rate that is taken several times bigger in specific areas since tumors progress faster in white matter is an important prior : the network would probably deduce it by itself, but the overall model would just be a black box, so it might not be desirable.

D. To go further

An exciting side of this article is that through Bayesian personalization, we get a rather accurate measure of the characteristics of the growth of a patient's tumor. If this kind of metric could be extended to a large amount of patients (which would require a large number of segmentation by a clinician), it would probably be interesting to link these parameters ρ and D to other variables : patient's age, tumor's age, location of the tumor, prior history of cancer, alimentation habits, carcinogenic factors, ... : The gap of the personalized ρ and D to the default model ones might have a detectable physical and biological explanation.

References

[1] Johan G.C. van Hasselt Jesse J. Swen Anyue Yin, Dirk Jan A.R. Moes and Henk-Jan Guchelaar. A review of mathematical models for tumor dynamics and treatment resistance evolution of solid tumors. 2019.

[2] Gregory Gundersen. *Hamiltonian Monte Carlo*. 2020. https://gregorygundersen.com/blog/2020/07/05/hmc/.

[3] Suprosanna Shit Florian Kofler Nore Collomb Benjamin Lemasson Emmanuel Barbier Ivan Ezhov, Jana Lipkova and Bjoern Menze. *Neural parameters estimation for brain tumor growth modeling*. 2020.

[4] David Warde-Farley Antoine Biard Aaron Courville Yoshua Bengio Chris Pal Pierre-Marc Jodoin Hugo Larochelle Mohammad Havaei, Axel Davy. *Brain Tumor Segmentation with Deep Neural Networks*. 2016.

[5] Juan Carlos López Alfonso1 Michael Meyer-Hermann Haralampos Hatzikirou Pietro Mascheroni, Symeon Savvopoulos. *Improving personalized tumor growth predictions using a Bayesian combination of mechanistic modeling and machine learning*. https://doi.org/10.1038/s43856-021-00020-4.

[6] Philip Hahnfeldt Thomas Hillen, Heiko Enderling. *The Tumor Growth Paradox and Immune System-Mediated Selection for Cancer Stem Cells*. 2012.