

A novel averaging principle provides insights in the impact of intratumoral heterogeneity on tumor progression

Marta Leocata

University of Pisa, Italy

leocata@mail.dm.unipi.it

J. C. L. Alfonso

Braunschweig Integrated Centre of Systems Biology,

Helmholtz Centre for Infection Research, Braunschweig, Germany

JuanCarlos.LopezAlfonso@helmholtz-hzi.de

Nikos I. Kavallaris

Department of Mathematics, University of Chester,

Thornton Science Park Pool Lane, Ince, Chester CH2 4NU, UK

n.kavallaris@chester.ac.uk

Haralampos Hatzikirou

Braunschweig Integrated Centre of Systems Biology,

Helmholtz Centre for Infection Research, Braunschweig, Germany

Haralampos.Hatzikrou@helmholtz-hzi.de

Typically stochastic differential equations (SDEs) involve an additive or multiplicative noise term. Here, we are interested in stochastic differential equations for which the white noise is non-linearly integrated in the corresponding evolution term, typically termed as random ordinary differential equations (RODEs). The classical averaging methods fail to treat such RODEs. Therefore, we introduce a novel averaging method appropriate to be applied on RODEs. To exemplify the importance of our method, we apply it in an important biomedical problem, i.e. the assessment of intratumoral heterogeneity impact on tumor dynamics. In particular, we model gliomas according to a well-known Go or Grow (GoG) model and tumor heterogeneity is modelled as a stochastic process. It has been shown that this GoG model exhibits an emerging Allee effect (bistability). We analytically and computationally show that the introduction of white noise, as a model of intratumoral heterogeneity, leads to a monostable tumor growth. This monostability behaviour is also derived even when spatial cell diffusion is taking into account.

Keywords: Averaging; white noise; intrinsic heterogeneity; phenotypic switching; glioblastoma multiforme

AMS Subject Classification: 70K65, 70K20, 34F05, 35R60, 92B99

1. Introduction

In the current section we demonstrate the key ideas of the novel averaging principle is introduced in the present work. For that purpose let us consider the following simple ordinary differential equation (ODE)

$$\begin{cases} \frac{\partial \rho}{\partial t} = F(\rho, k), & t \in (0, T], \\ \rho(0) = \rho_0, \end{cases} \quad (1.1)$$

where $F : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ is a globally Lipschitz function in both variables, $\rho, k \in \mathbb{R}$. Here the variable k represents a generic parameter of the model represented by the system (1.1). Assume now that the parameter k is not constant but varies randomly

2 *M. Leocata, J.C.L. Alfonso, N.I. Kavallaris & H. Hatzikirou*

in time, and thus some noise is introduced to system (1.1). One can assume that the introduced noise resembles a Gaussian Process, or in more singular case, when it varies very strongly and at very short time scale, that has the form of a white noise. In the latter case and when the dependence on the parameter $k = k(t)$ is linear, namely

$$F(\rho, k) = \mu(\rho) + \sigma(\rho)k,$$

where then system (1.1) is reduced to the following initial value problem of a stochastic differential equation (SDE)

$$\begin{cases} \frac{d\rho}{dt} = \mu(\rho) dt + \sigma(\rho) dB_t, & t \in (0, T], \\ \rho(0) = \rho_0. \end{cases}$$

Next in the case the dependence on $k(t)$ is non linear and $k(t)$ is a real valued process, hence not white noise, then system (1.1) is well defined and its mathematical study can be delivered through a well established theory, see.¹² Otherwise, if $k(t)$ is thought as a white noise, namely as a uniform distribution will be denoted by ξ_t to make more clear the reference to a stochastic process, then we have to deal with the following system

$$\begin{cases} \frac{d\rho}{dt} = F(\rho, \xi_t), & t \in (0, T], \\ \rho(0) = \rho_0, \end{cases} \quad (1.2)$$

which is meaningless in the context of stochastic differential equations (SDES), since the drift of the ODE (1.2) is given by a non linear function of a distribution, namely of the white noise

To tackle system (1.2), the idea is to work with an approximation of the white noise ξ_t , which is defined in a complete probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, T]}, \mathbb{P})$ with filtration $(\mathcal{F}_t)_{t \in [0, T]}$. We will call one of the possible approximations of ξ_t as ξ_t^N . Then considering ξ_t^N , instead of ξ_t , we obtain the following well posed system,

$$\begin{cases} \frac{d\rho^N}{dt} = F(\rho^N, \xi_t^N(\omega)) := F_\omega(t, \rho), & t \in [0, T], \omega \in \Omega. \\ \rho(0) = \rho_0, \end{cases} \quad (1.3)$$

where the equation of (1.3) is a random ordinary differential equation (RODE), that is a non-autonomous ODE for almost every realization $\omega \in \Omega$.

The challenge is to investigate whether system (1.3) has a limit as $N \rightarrow \infty$. To do so we will focus on a specific case, in particular we assume that

$$\lim_{k \rightarrow \pm\infty} F(\rho, k) = \beta_\pm(\rho) \quad \text{for each } \rho \in \mathbb{R}.$$

Intuitively speaking ξ_t^N are independent centered Gaussian variables with variance that tends to infinite. Our main aim is to figure out how the nonlinear term F behaves when it is perturbed by an approximation of the white noise. Let us think the variable ρ being fixed, then the function F will be evaluated in extremely big positive values (on which $F \approx \beta_+$) or extremely big negative values (on which $F \approx \beta_-$).

Thus one would expect that in the limit $N \rightarrow \infty$, F becomes a sort of Bernoullian process, denoted by α . Namely on a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, T]}, \mathbb{P})$, the random variable α is defined by

$$\alpha : [0, T] \times \Omega \rightarrow \mathbb{R},$$

with

$$\alpha_t = \begin{cases} \beta_+(\rho) & p = \frac{1}{2}, \\ \beta_-(\rho) & p = \frac{1}{2}. \end{cases}$$

Furthermore for each $t_1, t_2 \in [0, T]$, we expect $(\alpha_{t_1}, \alpha_{t_2})$ are independent random variables. Consequently, in the limit $N \rightarrow +\infty$ one would expect as a limit equation of (1.3) a RODE with a drift term which maintains a random oscillation between the two values $\beta_+(\rho)$ and $\beta_-(\rho)$. Surprisingly this is not the case and as we will see in section 4 the limit equation is a deterministic ODE, whose drift is given by the mean of the two extremes, i.e. $\frac{\beta_+(\rho) + \beta_-(\rho)}{2}$.

The outline of the current manuscript is as follows. In section 2 for reader's convenience we present some preliminary material. Section 3 introduces some auxiliary results, related with the averaging approach. The previous is followed for the demonstration and the proof of our main mathematical result, the novel averaging principle presented in section 4. In section 5 an application of the introduced averaging principle on the impact of intratumoral heterogeneity in glioma progression is provided. It is actually observed that the introduction of randomness in the intratumoral heterogeneity works towards the disappearance of the Allee effect. Finally we close the paper with a discussion of our main results in section 6.

2. Preliminaries

In the current section some preliminaries are presented for reader's convenience. Throughout the manuscript we consider random approximations of a white noise $(\xi_t)_{t \in [0, T]}$. White Noise can have different equivalent definitions, however we will focus on a particular one: the one that allow us to figure out the white noise as the “derivative” of Brownian motion. Before proceeding with the required definitions we just point out that we will work in the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, T]}, \mathbb{P})$.

First, we recall the definition of Gaussian stochastic process generalized as can be found in.⁷

Definition 2.1. Let us consider a generalized stochastic process Φ , namely

$$\Phi : \Omega \rightarrow D'(U),$$

where U is an open set in \mathbb{R} and $D'(U)$ is the space of distribution on U . We say that Φ is a gaussian process if for each $\phi_1, \dots, \phi_k \in C_c^\infty(U)$ the random vector $\Phi B(\phi_1), \dots, \Phi(\phi_k)$ is normally distributed.

4 *M. Leocata, J.C.L. Alfonso, N.I. Kavallaris & H. Hatzikirou*

Notice that a Gaussian process is characterized by the following quantities: its mean

$$m_{\Phi}(\phi) := \mathbb{E}[\Phi(\phi)],$$

and its covariance

$$c_{\Phi}(\phi, \psi) := \mathbb{E}[\Phi(\phi) - m_{\Phi}(\phi)][\Phi(\psi) - m_{\Phi}(\psi)].$$

for each $\phi, \psi \in C_c^{\infty}(U)$.

Definition 2.2. Let $(B_t)_{t \in \mathbb{R}^+}$ a Brownian motion on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Associated to B we can consider the generalized stochastic process,

$$\mathcal{B} : \Omega \rightarrow D'(U),$$

$$\mathcal{B}(\phi) := \int_0^T \phi_s B_s ds, \quad \text{for any } \phi \in C_c^{\infty}([0, T]).$$

Then we define the white noise as the generalized derivative $\mathcal{B}' : \Omega \rightarrow D'([0, T])$ of the generalized Brownian motion \mathcal{B} , that is

$$\mathcal{B}'(\phi) := - \int_0^T \phi'_s B_s ds, \quad \text{for any } \phi \in C_c^{\infty}([0, T]).$$

Henceforth we will denote \mathcal{B}' by ξ and thus

$$\xi(\phi) = \int_0^T \phi_s \xi_s ds := - \int_0^T \phi'_s B_s ds, \quad \text{for any } \phi \in C_c^{\infty}([0, T]). \quad (2.1)$$

White noise is a generalized Gaussian Stochastic Process, since

$$\xi(\phi) = -\mathcal{B}(\phi')$$

and \mathcal{B} is a Gaussian process. Recalling that $\text{cov}(B_t, B_s) = \min\{t, s\}$, we can evaluate the mean and covariance of the white noise as

$$m_{\xi}(\phi) = -m_{\mathcal{B}}(\phi') = 0$$

and

$$c_{\xi}(\phi, \psi) = \int_0^T \int_0^T \min\{t, s\} \phi'_s \psi'_t ds dt = \int_0^T \int_0^T \delta(t-s) \phi_s \psi_t ds dt.$$

So we can conclude that the white noise is a δ -correlated stationary, Gaussian process with mean zero and covariance $\mathbb{E}[\xi_t \xi_s] = \delta(t-s)$.

As it has been already mentioned at the beginning of the section our main strategy is to work with an approximation of $(\xi_t)_{t \in [0, T]}$, since only the associated ODEs are well defined.

Let us now consider the stochastic process \mathcal{B}^{Δ} , defined by the difference quotient of Brownian motion, $\frac{B_{s+\Delta} - B_s}{\Delta}$, namely

$$\mathcal{B}^{\Delta}(\phi) := \int_0^T \frac{B_{s+\Delta} - B_s}{\Delta} \phi_s ds \quad \text{for any } \phi \in C_c^{\infty}([0, T]).$$

We justify our choice with the following lemma.

Lemma 2.1. *The process \mathcal{B}^Δ converges weakly to ξ , almost surely, that is*

$$\mathcal{B}^\Delta(\phi) \xrightarrow{\Delta \rightarrow 0} \xi(\phi) \quad \text{for every } \phi \in C_c^\infty([0, T]), \text{ a.s..}$$

Proof. Due to the non-differentiability of Brownian motion we first prove the lemma for a mollified approximation of Brownian motion $B_t^\varepsilon = B_t * \psi^\varepsilon$, where ψ^ε is a mollifier function, see,¹⁵ and then taking the limit as $\varepsilon \rightarrow 0$ we will deduce the desired result. It is easily seen that for any fixed ε we derive

$$\begin{aligned} \int_0^T \frac{B_{s+\Delta}^\varepsilon - B_s^\varepsilon}{\Delta} \phi_s ds &= \int_0^T \frac{B_{s+\Delta}^\varepsilon \phi_{s+\Delta} - B_s^\varepsilon \phi_s}{\Delta} ds \\ &+ \int_0^T B_{s+\Delta}^\varepsilon \frac{\phi_s - \phi_{s+\Delta}}{\Delta} ds. \end{aligned} \quad (2.2)$$

Taking now the limit of (2.1) as $\Delta \rightarrow 0$ we derive

$$\mathcal{B}^{\Delta, \varepsilon}(\phi) \xrightarrow{\Delta \rightarrow 0} \underbrace{\int_0^T (B_s^\varepsilon \phi_s)' ds}_{=0} - \int_0^T B_s^\varepsilon \phi_s' ds = - \int_0^T B_s^\varepsilon \phi_s' ds, \quad (2.3)$$

since $\phi \in C_c^\infty([0, T])$.

Finally considering the limit of (2.3) as $\varepsilon \rightarrow 0$ we derive the desired result. \square

Now we consider a discrete sampling of the process $\frac{B_{s+\Delta} - B_s}{\Delta}$ with $\Delta = \Delta_N = 1/N$. Thus we divide the interval $[0, T]$ in small sub-intervals such that $[0, T] = \bigcup_{i=1}^N A_i^N$ where $A_i^N = \left[\frac{(i-1)T}{N}, \frac{iT}{N} \right]$ with $i = 1, \dots, N$. Note that in the interval A_i^N the sample is normally distributed

$$\frac{B_{\frac{i-1}{N}T + \Delta_N} - B_{\frac{i-1}{N}T}}{\Delta_N} \sim \mathcal{N}\left(0, \frac{1}{\Delta_N}\right).$$

Consequently, inspired by Lemma 2.1, we consider the following random step function as an approximation of the white noise

$$\xi_t^N = \sum_{i=1}^N \alpha_i^N \mathbb{1}_{A_i^N}(t) \quad \text{where} \quad \alpha_i^N = \frac{B_{\frac{i-1}{N}T + \Delta_N} - B_{\frac{i-1}{N}T}}{\Delta_N} \sim \mathcal{N}(0, N), \quad (2.4)$$

see Fig. 1.

3. Some auxiliary results

Our mathematical approach to study system (1.2) is inspired by the averaging theory. Indeed, this theory treats dynamical systems in which two variables (X_t, ξ_t) coexist and interact, but in two different time scales. In particular, the time scaling

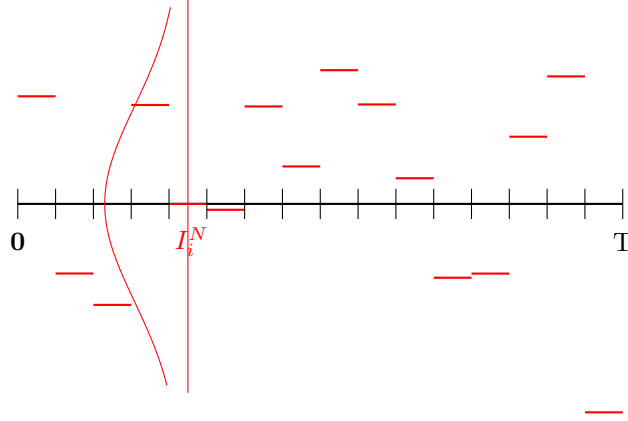


Fig. 1: Plot of the approximation ξ_t^N of white noise ξ_t .

of one of the two variables, ξ , is subjected to acceleration by a factor ϵ , which implies that its actual time scale is t/ϵ and so it contributes to the dynamics of the system in the form $\xi_{t/\epsilon}$. Via averaging theory we can investigate what is the impact to the system for small ϵ , namely when the acceleration is big, and we can conclude that the contribution of the faster process $\xi_{t/\epsilon}$ is evaluated in its time average.

Throughout the current manuscript we will try to implement the mentioned ideas into system (1.2). Since the latter system is not well defined when the random perturbation is a white noise we will deal with the approximating system (1.3), where the white noise is approximated by ξ_t^N , defined by (2.4), and the parameter of approximation N , can be reinterpreted as the new scaling time parameter⁴. In the current section we present some auxiliary results will be used for the proof of our novel averaging principle demonstrated and proven in section 4.

We now present the main hypotheses regarding the drift F term. Please note that C denotes a positive constant independent of t which might change its value from line to line.

Hypothesis 3.1.

- (1) $F : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ is Lipschitz with respect the first variable, uniformly on the second, namely there holds

$$|F(\rho_1, k) - F(\rho_2, k)| \leq C|\rho_1 - \rho_2|, \quad (3.1)$$

for each $k \in \mathbb{R}$, where the positive constant C , does not depends on k .

- (2) F is also bounded, i.e.

$$\|F(\rho, k)\|_\infty \leq C < +\infty. \quad (3.2)$$

- (3) F is quasimonote, i.e.

$$F(0, k) \geq 0. \quad (3.3)$$

(4) Finally for each $\rho \in \mathbb{R}$ the asymptotic profile of the drift term is defined as

$$|F(\rho, k) - \beta_{\pm}(\rho)| = O\left(\frac{1}{|k|^2}\right) \quad \text{for } k \rightarrow \pm\infty. \quad (3.4)$$

The way we interpret the solutions of RODE (1.3) as well as some of its features are provided below:

Definition 3.1. By a solution of the RODE (1.3) we mean a stochastic process such that for each $\omega \in \Omega$, $\rho^N(\omega) \in C([0, T])$ and satisfies:

$$\rho_t^N(\omega) = \rho_0 + \int_0^t F(\rho_s^N(\omega), \xi_s^N(\omega)) ds,$$

with a deterministic initial condition $\rho_0 \in \mathbb{R}^+$.

Remark 3.1. Under the assumption that F locally Lipschitz in its first variable then problem (1.3) has a unique pathwise solution $\rho^N(t, \omega)$ for every $\omega \in \Omega$, in the finite interval $[0, T]$,¹² Moreover the assumption (3.3) guarantee the positivity of the solutions of (1.3) initiated from positive initial condition $\rho_0 \geq 0$, see also Theorem 2.4 in,¹² which is a desired property in biological models like the one that will be considered in section 5.

In the next section, see in particular Theorem 4.1, we prove that the limit towards $N \rightarrow +\infty$ of problem (1.3) is the following deterministic problem

$$\begin{cases} \frac{d\rho}{dt} = \bar{F}(\rho) \\ \rho(0) = \rho_0, \end{cases} \quad (3.5)$$

where

$$\bar{F}(\rho) = \frac{\beta_+(\rho) + \beta_-(\rho)}{2}. \quad (3.6)$$

Definition 3.2. By a solution of the ODE (3.5) we mean a function $\rho \in C([0, T])$ satisfying:

$$\rho_t = \rho_0 + \int_0^t \bar{F}(\rho_s) ds,$$

with a deterministic initial condition $\rho_0 \in \mathbb{R}^+$.

Again if the drift term $\bar{F}(\rho_s)$ is locally Lipschitz continuous and positive, then problem (3.6) has a unique positive solution.

The rest of this section is devoted to the study of the average in time contribution of the process ξ_t^N on the drift F (asymptotically in N), more precisely see Proposition 3.1. Before proceeding with the proof of this result we need a key tool, i.e. a law of large numbers for random variables (r.v.) $F(x, \alpha_N^i), i = 1, 2, \dots, N$, which is stated and proven in the following lines.

8 *M. Leocata, J.C.L. Alfonso, N.I. Kavallaris & H. Hatzikirou*

Lemma 3.1. *Assume that F satisfies Hypothesis 3.1. Fixed $N \in \mathbb{N}$, we consider $\alpha_i^N \sim \mathcal{N}(0, N)$ as independent random variables, then for each $x \in \mathbb{R}^+$ there holds*

$$\mathbb{E} \left| \frac{1}{N} \sum_{i=1}^N F(x, \alpha_i^N) - \bar{F}(x) \right|^4 \leq \frac{C}{N^\alpha}, \quad \text{for some } \alpha > \frac{1}{4}, \quad (3.7)$$

where $\bar{F}(x) := \frac{\beta_+(x) + \beta_-(x)}{2}$ and provided that N is sufficiently large.

Proof. Let us call $S^N(x) := \frac{1}{N} \sum_{i=1}^N F(x, \alpha_i^N)$, then for $\mu^N(x) := \mathbb{E} [F(x, \alpha_i^N)]$, $i = 1, 2, \dots, N$, we have

$$\begin{aligned} \mathbb{E} |S^N(x) - \bar{F}(x)|^4 &= \mathbb{E} \left| \frac{1}{N} \sum_{i=1}^N F(x, \alpha_i^N) - \bar{F}(x) \right|^4 \\ &= \mathbb{E} \left| \left(\frac{1}{N} \sum_{i=1}^N (F(x, \alpha_i^N) - \mu^N(x)) \right) - (\mu^N(x) - \bar{F}(x)) \right|^4 \\ &\leq 8 \cdot \mathbb{E} \left| \frac{1}{N} \sum_{i=1}^N (F(x, \alpha_i^N) - \mu^N(x)) \right|^4 + 8 \cdot |\mu^N(x) - \bar{F}(x)|^4 \\ &= \frac{8}{N^4} \cdot \mathbb{E} \left| \sum_{i=1}^N (F(x, \alpha_i^N) - \mu^N(x)) \right|^4 + 8 \cdot |\mu^N(x) - \bar{F}(x)|^4. \end{aligned} \quad (3.8)$$

Now, setting $X_i := (F(x, \alpha_i^N) - \mu^N(x))$ then the first term in the r.h.s of (3.8) can be expanded as

$$\begin{aligned} &\mathbb{E} \left[\left| \sum_{i=1}^N X_i \right|^4 \right] \\ &= \mathbb{E} \left[\sum_{i=1}^N X_i^4 + \sum_{i \neq j} X_i^3 X_j + \sum_{i \neq j} X_i^2 X_j^2 + \sum_{i_1 \neq i_2 \neq i_3} X_{i_1} X_{i_2} X_{i_3}^2 + \sum_{i_1 \neq i_2 \neq i_3 \neq i_4} X_{i_1} X_{i_2} X_{i_3} X_{i_4} \right] \\ &= \sum_{i=1}^N \mathbb{E} [X_i^4] + \sum_{i \neq j} \mathbb{E} [X_i^3 X_j] + \sum_{i \neq j} \mathbb{E} [X_i^2 X_j^2] + \sum_{i_1 \neq i_2 \neq i_3} \mathbb{E} [X_{i_1} X_{i_2} X_{i_3}^2] \\ &\quad + \sum_{i_1 \neq i_2 \neq i_3 \neq i_4} \mathbb{E} [X_{i_1} X_{i_2} X_{i_3} X_{i_4}]. \end{aligned}$$

Due to the independence of the involved r.v. we derive

$$\begin{aligned} \mathbb{E} \left[\left| \sum_{i=1}^N X_i \right|^4 \right] &= \sum_{i=1}^N \mathbb{E} [X_i^4] + \sum_{i \neq j} \mathbb{E} [X_i^3] \mathbb{E} [X_j] + \sum_{i \neq j} \mathbb{E} [X_i^2] \mathbb{E} [X_j^2] \\ &\quad + \sum_{i_1 \neq i_2 \neq i_3} \mathbb{E} [X_{i_1}] \mathbb{E} [X_{i_2}] \mathbb{E} [X_{i_3}^2] + \sum_{i_1 \neq i_2 \neq i_3 \neq i_4} \mathbb{E} [X_{i_1}] \mathbb{E} [X_{i_2}] \mathbb{E} [X_{i_3}] \mathbb{E} [X_{i_4}]. \end{aligned}$$

Since $\mathbb{E}[X_i] = 0$, the second, the fourth sum and the last sum are equal to zero. So the previous quantity is reduced to

$$\mathbb{E} \left[\left| \sum_{i=1}^N X_i \right|^4 \right] = \sum_{i=1}^N \mathbb{E}[X_i^4] + \sum_{i \neq j} \mathbb{E}[X_i^2] \mathbb{E}[X_j^2].$$

Thanks to the previous computation and to the fact that the r.v. X_i are independent and identically distributed (i.i.d.) and uniformly bounded with respect to N , we obtain:

$$\begin{aligned} \frac{1}{N^4} \mathbb{E} \left| \sum_{i=1}^N X_i \right|^4 &\leq \frac{1}{N^4} \sum_{i=1}^N \mathbb{E}[X_i^4] + \frac{1}{N^4} \sum_{i \neq j} \mathbb{E}[X_i^2] \mathbb{E}[X_j^2] \\ &\leq \frac{N}{N^4} \mathbb{E}[X_1^4] + \frac{3 \cdot N \cdot (N-1)}{N^4} \mathbb{E}[X_1^2]^2 \\ &\leq \frac{C}{N^3} + \frac{C}{N^2}. \end{aligned} \quad (3.9)$$

Next we try to estimate the second term in the right hand side (r.h.s.) of (3.8). It remains to prove that

$$|\mu^N(x) - \bar{F}(x)| \leq \frac{1}{N^\alpha}, \quad \text{for some } \alpha > \frac{1}{4}.$$

In the following it will be more convenient to rewrite α_i^N as $\sqrt{N}Z_i$, where $Z_i \sim \mathcal{N}(0, 1)$ for $i = 1, 2, \dots, N$. Then considering the sequence $\epsilon_N = N^{-\beta}$ with $\beta \in (\frac{1}{3}, \frac{3}{8})$ we have

$$\begin{aligned} |\mu^N(x) - \bar{F}(x)| &= \left| \mathbb{E} \left[F(x, \sqrt{N}Z_i) \right] - \bar{F}(x) \pm \beta_+(x) \mathbb{P}[Z_i > \epsilon_N] \pm \beta_-(x) \mathbb{P}[Z_i > \epsilon_N] \right| \\ &\leq \underbrace{\left| \mathbb{E} \left[F(x, \sqrt{N}Z_i) \right] - (\beta_+(x) + \beta_-(x)) \mathbb{P}[Z_i > \epsilon_N] \right|}_{(I)} \\ &\quad + \underbrace{|\beta_+(x) + \beta_-(x) \mathbb{P}[Z_i > \epsilon_N] - \bar{F}(x)|}_{(II)}. \end{aligned}$$

The first term (I) is estimated as follows

$$\begin{aligned} \left| \mathbb{E} \left[F(x, \sqrt{N}Z_i) \right] - (\beta_+(x) + \beta_-(x)) \mathbb{P}[Z_i > \epsilon_N] \right| &\leq \\ &\leq \underbrace{\left| \mathbb{E} \left[F(x, \sqrt{N}Z_i) \mathbb{1}_{Z_i > \epsilon_N} \right] - \beta_+(x) \mathbb{P}[Z_i > \epsilon_N] \right|}_{(I)_1} \\ &\quad + \underbrace{\left| \mathbb{E} \left[F(x, \sqrt{N}Z_i) \mathbb{1}_{|Z_i| < \epsilon_N} \right] \right|}_{(I)_2} \\ &\quad + \underbrace{\left| \mathbb{E} \left[F(x, \sqrt{N}Z_i) \mathbb{1}_{Z_i < -\epsilon_N} \right] - \beta_-(x) \mathbb{P}[Z_i > \epsilon_N] \right|}_{(I)_3}, \end{aligned}$$

10 *M. Leocata, J.C.L. Alfonso, N.I. Kavallaris & H. Hatzikirou*

where $(I)_1$ can be expressed as

$$\left| \mathbb{E} \left[F(x, \sqrt{N}Z_i) \mathbf{1}_{Z_i > \epsilon_N} \right] - \beta_+(x) \mathbb{P}[Z_i > \epsilon_N] \right| = \mathbb{E} \left[\left| F(x, \sqrt{N}Z) - \beta_+(x) \right| \mathbf{1}_{Z_i > \epsilon_N} \right].$$

Due to condition (3.4), there exists $K_0^+ > 0$ such that for each $\sqrt{K} > K_0^+$

$$|F(x, \sqrt{K}) - \beta_+(x)| \leq \frac{c}{K},$$

thus choosing N such that $\sqrt{N}\epsilon_N > K_0^+$, we have

$$\mathbb{E} \left[\left| F(x, \sqrt{N}Z_i) - \beta_+(x) \right| \mathbf{1}_{Z_i > \epsilon_N} \right] \leq \frac{c}{N\epsilon_N^2} \mathbb{P}[Z_i > \epsilon_N] \leq \frac{c}{N\epsilon_N^2}.$$

Recalling that $\epsilon_N = N^{-\beta}$ we have that $N\epsilon_N^2 = N^{-2\beta+1}$ and thus

$$(I)_1 \leq \frac{C}{N^{-2\beta+1}}. \quad (3.10)$$

We proceed in a very similar way to estimate $(I)_3$, which again can be written as

$$\left| \mathbb{E} \left[F(x, \sqrt{N}Z_i) \mathbf{1}_{Z_i < -\epsilon_N} \right] - \beta_-(x) \mathbb{P}[Z_i < -\epsilon_N] \right| = \mathbb{E} \left[\left| F(x, \sqrt{N}Z) - \beta_-(x) \right| \mathbf{1}_{Z < -\epsilon_N} \right].$$

Again condition (3.4) entails the existence of some $K_0^- < 0$ such that for each $-\sqrt{K} < K_0^-$ there holds

$$|F(x, -\sqrt{K}) - \beta_-(x)| \leq \frac{c}{K},$$

and thus by choosing N such that $-\sqrt{N}\epsilon_N < K_0^-$, we obtain

$$\begin{aligned} (I)_3 &= \mathbb{E} \left[\left| F(x, \sqrt{N}Z_i) - \beta_-(x) \right| \mathbf{1}_{Z_i < -\epsilon_N} \right] \\ &= \mathbb{E} \left[\left| F(x, -\sqrt{N}|Z_i|) - \beta_-(x) \right| \mathbf{1}_{Z_i < -\epsilon_N} \right] \\ &\leq \frac{c}{N\epsilon_N^2} \mathbb{P}[Z_i < -\epsilon_N] \leq \frac{c}{N^{-2\beta+1}}. \end{aligned} \quad (3.11)$$

Regarding the estimation of term $(I)_2$, since $\epsilon_N = \frac{1}{N^\beta} \rightarrow 0$ as $N \rightarrow \infty$ and $\mathbb{P}[Z_i = 0] = 0$ then we can find N_1 such that for $N > N_1$

$$(I)_2 = \left| \mathbb{E} \left[F(x, Z_i) \mathbf{1}_{|Z_i| < \epsilon_N} \right] \right| \leq \|F\|_\infty \mathbb{P}[|Z_i| < \epsilon_N] \leq \frac{C}{N^\beta}, \quad (3.12)$$

taking also into account (3.2). Regarding the last estimation, about term (II) :

$$\begin{aligned} |(\beta_+(x) + \beta_-(x))\mathbb{P}[Z_i > \epsilon_N] - \bar{F}(x)| &\leq |(\beta_+(x) + \beta_-(x))| \left| \mathbb{P}[Z_i > \epsilon_N] - \frac{1}{2} \right| \\ &\leq |(\beta_+(x) + \beta_-(x))| \left| \operatorname{erf} \left(\frac{\epsilon_N}{\sqrt{2}} \right) \right| \lesssim \frac{1}{N^\beta} \end{aligned} \quad (3.13)$$

In summary, by virtue of (3.10), (3.11), (3.12), (3.13) and for $N > \max\{N_1, K_0^+, -K_0^-\}$ we deduce

$$|\mu^N(x) - \bar{F}(x)| \leq \frac{C}{N^\alpha},$$

where $\alpha = \min\{\beta, 1 - 2\beta\} = 1 - 2\beta > 1/4$ for $\beta \in (\frac{1}{3}, \frac{3}{8})$. Thus in conjunction with (3.9) we finally obtain the desired relation (3.7). \square

Now we are ready to prove that the sequence of solutions ρ^N of system (1.3), converges to the solution ρ of system (3.5) whose drift is given by Definition 3.2.

Proposition 3.1. *Assume that F satisfies Hypothesis 3.1. Let ξ_t^N be the approximation of white noise, defined in (2.4) then for each $t, T > 0$ and $x \in \mathbb{R}$ there holds*

$$\lim_{N \rightarrow \infty} \int_t^{t+T} F(x, \xi_s^N) ds = T\bar{F}(x) \quad a.s. \quad (3.14)$$

with $\bar{F}(x) := \frac{\beta_+(x) - \beta_-(x)}{2}$.

Proof. Note that by the definition (2.4) of the process ξ_t^N we immediately have

$$\begin{aligned} \frac{1}{T} \int_t^{t+T} F(x, \xi_s^N) ds &= \frac{1}{T} \int_t^{t+T} F\left(x, \sum_{i=1}^N \alpha_i^N \mathbb{1}_{A_i^N}(t)\right) ds \\ &= \frac{1}{T} \sum_{i=1}^N \int_{A_i^N} F(x, \alpha_i^N) ds = \frac{1}{N} \sum_{i=1}^N F(x, \alpha_i^N). \end{aligned}$$

Next we prove that the sum $\frac{1}{N} \sum_{i=1}^N F(x, \alpha_i^N)$ converges almost surely to $\bar{F}(x) := \frac{\beta_+(x) - \beta_-(x)}{2}$. To this end we can not directly apply the classical result of law of large numbers, since we are treating a triangular sequence of random variables; namely notice that r.v. α_i^N depend on the double index (i, N) . In the following, for simplicity we will denote with \tilde{S}^N the quantity, whose limit is under investigation,

$$\tilde{S}^N(x) := \frac{1}{N} \sum_{i=1}^N F(x, \alpha_i^N).$$

In order to prove that $\tilde{S}^N(x)$ converges almost surely to $\bar{F}(x)$ we equivalently demonstrate that there exists an infinitesimal sequence ϵ_N , such that

$$\mathbb{P}\left(\limsup_N A_N\right) = 0, \quad (3.15)$$

with $A_N := \{|\tilde{S}^N(x) - \mu(x)| > \epsilon_N\}$. To this end we will use Borel-Cantelli Lemma and hence it is sufficient to show the convergence of the series $\sum_{N=1}^{\infty} \mathbb{P}(A_N)$. Next by virtue of Markov inequality we derive

$$\begin{aligned} \sum_{N=1}^{\infty} \mathbb{P}(A_N) &= \sum_{N=1}^{\infty} \mathbb{P}\left(\left\{|\tilde{S}^N(x) - \mu(x)| > \epsilon_N\right\}\right) \\ &= \sum_{N=1}^{\infty} \mathbb{P}\left(\left\{|\tilde{S}^N(x) - \mu(x)|^4 > \epsilon_N^4\right\}\right) \\ &\leq \sum_{N=1}^{\infty} \frac{\mathbb{E}\left[|\tilde{S}^N(x) - \mu(x)|^4\right]}{\epsilon_N^4}, \end{aligned}$$

12 *M. Leocata, J.C.L. Alfonso, N.I. Kavallaris & H. Hatzikirou*

and thus by Lemma 3.1 we obtain

$$\sum_{N=1}^{\infty} \frac{\mathbb{E} \left[|\tilde{S}^N(x) - \mu(x)|^4 \right]}{\epsilon_N^4} \leq \sum_{N=1}^{\infty} \frac{1}{N^\alpha \epsilon_N^4} < \infty,$$

for some $\alpha > \frac{1}{4}$. Then under the choice $\epsilon_N = \frac{1}{N^\beta}$ and imposing $4\beta < \alpha - 1$ we guarantee the convergence of the series. As a consequence we derive the validity of (3.15) and thus the proof is complete. \square

4. A novel averaging principle

In the current section we present and prove our main mathematical result. Indeed, using the auxiliary results provided in section 3, we demonstrate a novel averaging principle can then be applied to the approximation (1.3) of system (1.2). As it has been already explained in the introduction the investigation of the dynamics of nonlinear system (1.2), under a white noise perturbation, passes through the study of the dynamics of its approximation (1.3) where the white noise ξ_t is substitute by the sequence ξ_t^N defined by (2.4). Then this idea can be applied, see section 5, to study the dynamics of some nonlinear models describing the tumor growth of a brain tumor (glioblastoma).

Our main mathematical result is stated as follows.

Theorem 4.1. *Let ξ_t^N be the approximation of white noise, given by (2.4). Then the solution trajectories ρ_t^N of the RODE (1.3) converge uniformly in time and almost surely to the solution trajectories of the ODE (3.5), i.e.*

$$\lim_{N \rightarrow \infty} \sup_{t \in [0, T]} |\rho_t^N - \rho_t| = 0 \quad a.s. \quad (4.1)$$

Proof. Let us take $\omega \in \Omega$. By the integral formulation of (1.3) and (3.5) we have

$$\begin{aligned} \rho_t^N(\omega) - \rho_t &= \int_0^t [F(\rho_s^N, \xi_s^N)(\omega) - \bar{F}(\rho_s)] ds \\ &= \int_0^t [F(\rho_s^N, \xi_s^N)(\omega) - F(\rho_s, \xi_s^N)(\omega)] ds \\ &\quad + \int_0^t [F(\rho_s, \xi_s^N)(\omega) - \bar{F}(\rho_s)] ds. \end{aligned}$$

By considering the Lipschitz property (3.1) of the drift term F with respect to the variable ρ we obtain

$$|\rho_t^N(\omega) - \rho_t| \leq C \int_0^t |\rho_s^N(\omega) - \rho_s| ds + \int_0^t |F(\rho_s, \xi_s^N)(\omega) - \bar{F}(\rho_s)| ds.$$

Applying now Gronwall's lemma, see,¹⁴ on the quantity $m(t) = |\rho_t^N - \rho_t|(\omega)$ we derive

$$|\rho_t^N - \rho_t|(\omega) \leq \exp(CT) \left| \int_0^t [F(\rho_s, \xi_s^N)(\omega) - \bar{F}(\rho_s)] ds \right|,$$

and thus

$$\sup_{t \in [0, T]} |\rho_t^N - \rho_t|(\omega) \leq \exp(CT) \left| \int_0^T [F(\rho_s, \xi_s^N)(\omega) - \bar{F}(\rho_s)] ds \right|.$$

Hence in order to prove (4.1) it is sufficient to show that

$$\lim_{N \rightarrow \infty} \left| \int_0^T [F(\rho_s, \xi_s^N)(\omega) - \bar{F}(\rho_s)] ds \right| = 0 \quad \text{a.s.}, \quad (4.2)$$

or equivalently to demonstrate that for any fixed $\omega \in \Omega$ and for each $\epsilon > 0$ there exists $\bar{N}(\omega) > 0$, such that for any $N > \bar{N}(\omega)$ there holds

$$\left| \int_0^T [F(\rho_s, \xi_s^N)(\omega) - \bar{F}(\rho_s)] ds \right| \leq \epsilon. \quad (4.3)$$

Now we fix ϵ and we discretize the interval $[0, T]$ in n subintervals of amplitude $\frac{T}{n}$, so we have

$$[0, T] = \bigcup_{k=0}^{n-1} \left[\frac{kT}{n}, \frac{(k+1)T}{n} \right].$$

We underline the fact that the choice of n depends on ϵ , that is as small is ϵ then as big we must choose n . Using now the introduced discretization we have

$$\begin{aligned} \int_0^T |F(\rho_s, \xi_s^N)(\omega) - \bar{F}(\rho_s)| ds &= \sum_{k=0}^{n-1} \int_{kT/n}^{(k+1)T/n} |F(\rho_s, \xi_s^N)(\omega) - \bar{F}(\rho_s)| ds \\ &= \underbrace{\sum_{k=0}^{n-1} \int_{kT/n}^{(k+1)T/n} |F(\rho_s, \xi_s^N)(\omega) - F(\rho_{kT/n}, \xi_s^N)| ds}_{(\Sigma_1)} \\ &\quad + \underbrace{\sum_{k=0}^{n-1} \int_{kT/n}^{(k+1)T/n} |\bar{F}(\rho_{kT/n}) - \bar{F}(\rho_s)| ds}_{(\Sigma_2)} \\ &\quad + \underbrace{\sum_{k=0}^{n-1} \int_{kT/n}^{(k+1)T/n} |F(\rho_{kT/n}, \xi_s^N) - \bar{F}(\rho_{kT/n})| ds}_{(\Sigma_3)}. \end{aligned} \quad (4.4)$$

We first deal with terms (Σ_1) and (Σ_2) into relation (4.4). Due to the Lipschitz

14 *M. Leocata, J.C.L. Alfonso, N.I. Kavallaris & H. Hatzikirou*

property of F, \bar{F} and ρ we obtain

$$\begin{aligned}
 (\Sigma_1) + (\Sigma_2) &= \sum_{k=0}^{n-1} \int_{kT/n}^{(k+1)T/n} |F(\rho_s, \xi_s^N)(\omega) - F(\rho_{kT/n}, \xi_s^N)| ds \\
 &+ \sum_{k=0}^{n-1} \int_{kT/n}^{(k+1)T/n} |\bar{F}(\rho_{kT/n}) - \bar{F}(\rho_s)| ds \\
 &\leq 2C \cdot \sum_{k=0}^{n-1} \int_{kT/n}^{(k+1)T/n} |\rho_s - \rho_{kT/n}| ds \\
 &\leq 2C \cdot \sum_{k=0}^{n-1} \int_{kT/n}^{(k+1)T/n} |s - kT/n| ds \\
 &\leq 2C \cdot \sum_{k=0}^{n-1} \int_{kT/n}^{(k+1)T/n} \frac{T}{n} ds \\
 &= 2C \cdot \sum_{k=0}^{n-1} \frac{T^2}{n^2} \\
 &= 2C \cdot \frac{T^2}{n} \leq \frac{\epsilon}{2}, \tag{4.5}
 \end{aligned}$$

provided that $n > \frac{4CT^2}{\epsilon}$, where C is a positive constant associated with the Lipschitz constants of F, \bar{F} and ρ .

Regarding now the estimation of term (Σ_3) , we plan to apply Proposition 3.1 on each

$$\int_{kT/n}^{kT/n+T/n} |F(\rho_{kT/n}, \xi_s^N)(\omega) - \bar{F}(\rho_{kT/n})| ds,$$

for $k = 0, \dots, n-1$.

We choose to work on a subspace of measure 1, indeed we consider $\Omega_1 := \bigcap_{n=1}^{\infty} \Omega^n$ with $\Omega^n = \bigcap_{k=1}^n \Omega_{kT/n, T/n, \rho_{kT/n}}$ and where $\Omega_{kT/n, T/n, \rho_{kT/n}}$ is the subset of Ω on which the convergence result (3.14) holds for $x = \rho_{kT/n}$, $t = kT/n$ and $T = T/n$.

Now (3.14) entails that for any fixed $\omega \in \Omega_1$ and for each $N > \bar{N}^k(\omega)$

$$\int_{kT/n}^{kT/n+T/n} |F(\rho_{kT/n}, \xi_s^N)(\omega) - \bar{F}(\rho_{kT/n})| ds \leq \frac{\epsilon}{2n}.$$

Consequently, by choosing $N > \max_{k=1, \dots, n-1} \bar{N}^k(\omega) := \bar{N}^\epsilon(\omega)$ we deduce

$$(\Sigma_3) = \sum_{k=0}^{n-1} \int_{kT/n}^{(k+1)T/n} [F(\rho_{kT/n}, \xi_s^N)(\omega) - \bar{F}(\rho_{kT/n})] ds \leq \frac{\epsilon}{2}. \tag{4.6}$$

In summary, combining (4.5) and (4.6) then for any fixed $\omega \in \Omega_1$ and for each $\epsilon > 0$ there exists $\bar{N}(\omega)$ such that (4.3) holds for any $N > \bar{N}(\omega)$

$$\int_0^T |F(\rho_s, \xi_s^N)(\omega) - \bar{F}(\rho_s)| ds \leq \epsilon,$$

and thus the proof is complete. \square

Next we focus on the infinite dimensional case, i.e. we consider the following partial differential equation (PDE) problem

$$\frac{\partial \rho^N}{\partial t} = \mathcal{A}\rho^N + F(\rho, \xi_t^N), \quad (x, t) \in D_T, \quad (4.7)$$

$$\mathcal{B}\rho^N(x, t) = 0, \quad (x, t) \in \Gamma_T, \quad (4.8)$$

$$\rho^N(x, 0) = \rho_0(x) \geq 0, \quad x \in D, \quad (4.9)$$

where $D_T := D \times (0, T)$, $\Gamma_T := \partial D \times (0, T)$ is a bounded and smooth $D \subset \mathbb{R}^\ell$, $\ell \geq 1$ and $T > 0$ stands for the maximum existence time of solution ρ^N , which is positive in D_T as long as it exists. Here \mathcal{A} represents a bounded operator who generates a strongly continuous semigroup $S_{\mathcal{A}}(t)$ whilst \mathcal{B} is a bounded boundary operator. Besides the nonlinearity $F_\omega^N(\rho, t) := F(\rho, \xi_t^N)$ is assumed to be almost periodic with respect to t uniformly for ρ in compact subsets in Banach space X .

Consider now the averaged (PDE) problem

$$\frac{\partial \rho}{\partial t} = \mathcal{A}\rho^N + \bar{F}(\rho), \quad (x, t) \in D_T, \quad (4.10)$$

$$\mathcal{B}\rho(x, t) = 0, \quad (x, t) \in \Gamma_T, \quad (4.11)$$

$$\rho(x, 0) = \rho_0(x) \geq 0, \quad x \in D, \quad (4.12)$$

where $\bar{F}(\rho)$ is given by (3.6). By virtue of a semigroup approach we can easily obtain a unique, positive solution for (4.10)-(4.12), see.¹⁴

Now by using Theorem 4.1 in conjunction with the semigroup approach introduced in¹⁰ we deduce the following infinite dimensional averaging principle

Theorem 4.2. *The solution trajectories ρ_t^N of problem (4.7)-(4.9) converge uniformly in time and almost surely to the unique solution of problem (4.10)-(4.12).*

A special version of Theorem 4.2 is proven, by using an compactness approach, in section 5 see Theorem 5.1.

5. Application: impact of intratumoral heterogeneity in glioma progression

In this section, the main purpose is to use the averaging principle, as introduced in the previous section by Theorems 4.1 and 4.2, to address the problem of intrinsic heterogeneity in glioma progression. Gliomas originate from glial cells in the brain and constitute the most common type of malignant brain tumor in adults.^{9,23} The WHO (World Health Organization) classifies gliomas into different classes, ranging from low-grade to high-grade malignancies¹⁶ depending on their proliferative potential and invasiveness. In particular, high grade gliomas are commonly characterized by diffusive infiltration into the brain tissue, behavior associated with poor prognosis and limited treatment outcomes. Glioma treatments usually involve resection

accompanied by radiotherapy and/or chemotherapy. Herein, we use a mathematical model to understand the reasons of treatment failure and optimize the combination of current treatment modalities.

Two of the most important reasons for glioma treatment failure is due to intratumoral heterogeneity comprised by intrinsic heterogeneity and phenotypic plasticity.^{17,18} The former describes the typically irreversible genetic or epigenetic diversity of tumor cells (e.g. due to mutations or clonal selection), and the latter is related to the reversible phenotypic responses towards microenvironmental cues. In gliomas, the most predominant phenotypic plasticity phenomenon is related to the “Go or Grow” (GoG) mechanism or migration/proliferation plasticity.¹³ This mechanism implies a mutually exclusive switching between migratory and proliferative phenotypes. The key question is how this tumor cell decision-making mechanism is regulated, and what is its impact on glioma growth and invasion. Concerning regulation, it was discovered a dependence of this cell decision mechanism on the local cell density without concluding on the exact functional form, by analysing images of *in vitro* experiments.²⁵ Analysing further the potential local cell density dependencies, it was also found out low-grade tumor micro-ecology potentially exhibits an emergent Allee effect, i.e. a critical tumor cell density implying both tumor growth and control.² The precise quantification of this critical tumor cell density could be a relevant prognostic criterion for the tumor fate, since it can be measured in biopsies samples. Moreover, it has been shown that this GoG mechanism explains the fast tumor recurrence time of high-grade brain tumors after resection.¹³ Finally based on our theoretical understanding of the cellular mechanism, it has been shown how personalized vasomodulatory cancer therapies can be optimized.^{6,11} In particular, it has been revealed that one-size-fits-all vaso-modulatory interventions should be expected to fail, because control of glioma invasion characteristics, such as tumor front speed and infiltration width, can be very variable and may require more personalized therapeutic interventions.⁸

All the aforementioned models assume that all glioma cells have an identical GoG mechanism. However, in reality, each cell may have an idiosyncratic migration and proliferation regulation, following the GoG mechanism, due to intrinsic heterogeneity. The question is how we can model and analyze the impact of intrinsic heterogeneity of a tumor cell population, when migration and proliferation are regulated by a non-uniform GoG mechanism. In particular, we will focus on the existence of the Allee effect in the presence of intratumoral heterogeneity.

5.1. *The Go or Grow model of glioma*

For the reader’s convenience, we briefly present the derivation of the Go or Grow model, as introduced in.¹³ Due to the migration/proliferation dichotomy, we can distinguish the total population of glioma cells ρ in two groups, ρ^p proliferating cells and ρ^m migratory cells. For each type of population we can then write down an equation describing the corresponding dynamics, so we have the following system,

which is described as a GoG model,

$$\frac{\partial \rho^m}{\partial t} = \Delta \rho^m + r_{p \rightarrow m}(\rho) \rho^m - r_{m \rightarrow p}(\rho) \rho^m - r_d \rho^m, \quad (5.1)$$

$$\frac{\partial \rho^p}{\partial t} = r_b \rho^p (1 - \rho^p) + r_{m \rightarrow p}(\rho) \rho^p - r_{p \rightarrow m}(\rho) \rho^p - r_d \rho^p, \quad (5.2)$$

where r_b is the birth rate, r_d is the death rate, $r_{p \rightarrow m}(\rho)$ is the switch rate from proliferating to motile phenotype and $r_{m \rightarrow p}(\rho)$ is the switch rate from motile to proliferating phenotype of the tumor cells.

To obtain a unique equation for $\rho_t = \rho_t^m + \rho_t^p$, we refer to detailed balance condition

$$\rho^p r_{p \rightarrow m}(\rho) = \rho^m r_{m \rightarrow p}(\rho).$$

Through this condition we can rewrite the preceding system as a single PDE

$$\frac{\partial \rho}{\partial t} = \Delta (r_{p \rightarrow m}(\rho) \rho) + r_b r_{m \rightarrow p}(\rho) \rho (1 - \rho) - r_d \rho. \quad (5.3)$$

The phenotypic switching is regulated by local microenvironmental cues lumped into local density dependence. For the sake of simplicity, we will assume that both rates depend monotonously on cell density, approximated by a sigmoidal function with slope given by k . Intuitively, the slope can be viewed as the way that single tumor cell interprets its microenvironment and decides over its phenotype. Following,² we consider that the two rates are complementary, namely if cell motility increases with cell density then cell proliferation decreases with density and vice versa: namely if k denotes the slope of the switch from motile type to proliferating type, then $-k$ denotes the slope of the switch from proliferating type to motile type. A possible choice is given by the following, see Fig. 2:

$$r_{m \rightarrow p}(\rho) = \frac{1}{2} (1 + \tanh(k(\rho - \theta))) := r_s(\rho), \quad (5.4)$$

$$r_{p \rightarrow m}(\rho) = 1 - r_s(\rho), \quad (5.5)$$

and thus (5.1) takes the form

$$\frac{\partial \rho}{\partial t} = \Delta ((1 - r_s(\rho, k)) \rho) + r_b r_s(\rho, k) \rho (1 - \rho) - r_d \rho. \quad (5.6)$$

For positive slope $k > 0$ the phenotypic switch presents an attractive behavior, while for $k < 0$ a repulsive one (see Fig. 2).² To the best of our knowledge the GoG model (5.6) has been only investigated for the case when k is a constant, see,² i.e. the tumor cell population decides in a homogeneous way over proliferation or migration. Here we assume that tumor is heterogeneous in the way cells regulate their migration/proliferation phenotype controlled by a stochastic k .

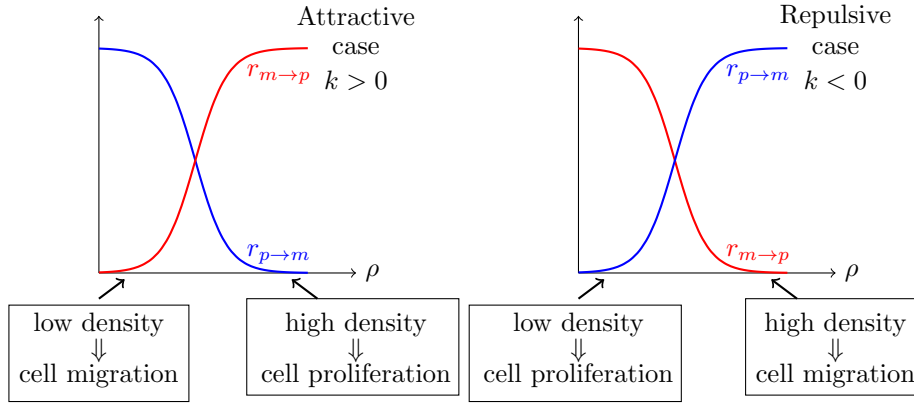


Fig. 2: Sketch of cellular mechanism for different kind of phenotypic plasticity. In the left aggregative configuration is represented, while in the right figure, the repulsive configuration is represented.

5.2. *Intrinsic intratumoral heterogeneity of glioma cells as a white noise*

We recall that the sign of k indicates the regime where we are, it identifies if we are in an attractive or repulsive regime, whilst the absolute value of k measures the intensity of the phenotypic switching. In the following, we introduce the desirable heterogeneous regulation of the switch by assuming that k follows a probability distribution, i.e. we heuristically take

$$k \mapsto k + \text{Noise}. \quad (5.7)$$

It is anticipated that the introduced intrinsic heterogeneity facilitates the tumor growth and persistence.^{17,18} Therefore it is plausible to consider the “worst” case heterogeneity scenario, hence k is considered as a white noise. As first step towards the investigation of the dynamics of the GoG model (5.6) under random perturbation (5.7), we choose to neglect the contribution of the diffusion, and thus we initially consider the following ODE

$$\begin{cases} \frac{\partial \rho}{\partial t} = F(\rho, k) := r_b r_s(\rho, k) \rho (1 - \rho) - r_d \rho, \\ 0 \leq \rho(0) = \rho_0 \leq 1, \end{cases} \quad (5.8)$$

where

$r_s(\rho, k)$ given by (5.4)

for $r_b > r_d$, $\theta \in [0, 1]$ and $\rho_0 \in \mathbb{R}^+$.

If one wants to approximate k as a white noise, clearly relapses in the case exposed in the introduction, since the occurring system

$$\begin{cases} \frac{\partial \rho^\xi}{\partial t} = F(\rho, \xi), \\ 0 \leq \rho(0) = \rho_0 \leq 1, \end{cases} \quad (5.9)$$

is not well defined.

In order to tackle system (5.9) we appeal to the averaging principle demonstrated by Theorem 4.1. To this end we just need to verify all the involved conditions of Hypothesis 3.1. We first note that condition (3.1) is trivially verified, since the drift term $F(\rho, k) = r_b r_s(\rho, k)\rho(1 - \rho) - r_d \rho$ for $\rho \in [0, 1]$ is differentiable. Besides F is bounded on both variables (ρ, k) , since k appears as an argument of the hyperbolic tangent, and the variable ρ varies in a compact set. Thus condition (3.2) also holds. Obviously we have $F(0, k) = 0$ and thus condition (3.3) is also fulfilled. It remains to check the validity of condition (3.4). It can be easily seen that

$$\lim_{k \rightarrow \pm\infty} (r_b r_s(\rho, k)\rho(1 - \rho) - r_d \rho) = r_b \beta_\pm \rho(1 - \rho) - r_d \rho,$$

with $\beta_+ = 1$ and $\beta_- = 0$ where the order of convergence is exponential, and hence condition (3.4) is also fulfilled. Consequently we have the following result which is a straightforward consequence of Theorem 4.1

Proposition 5.1. *Let ξ_t^N be the process described in (2.4). Then the solution trajectories ρ_t^N of*

$$\begin{cases} \frac{\partial \rho^N}{\partial t} = r_b r_s(\rho^N, \xi^N)\rho^N(1 - \rho^N) - r_d \rho^N, \\ 0 \leq \rho(0) = \rho_0 \leq 1, \end{cases}$$

converge uniformly in time and almost surely to the solution ρ of the following ODE problem

$$\begin{cases} \frac{\partial \rho}{\partial t} = \frac{1}{2} r_b \rho(1 - \rho) - r_d \rho \\ 0 \leq \rho(0) = \rho_0 \leq 1. \end{cases} \quad (5.10)$$

Namely,

$$\lim_{N \rightarrow \infty} \sup_{t \in [0, T]} |\rho_t^N - \rho_t| = 0 \quad a.s. .$$

Thus in order to investigate the model (5.8), Proposition 5.1 provide us with the appropriate approximating ODE (5.10) which will be analyzed with respect to tumor steady states.

Let us here recall that in the absence of intrinsic heterogeneity, the trivial solution $\rho_{nc} = 0$ is a stable point for a certain range of parameters, as shown in Fig. 3 and.² However, under the assumption of white noise intrinsic heterogeneity, the stability analysis of (5.10) leads to a monostable configuration. Specifically, the point $\rho_c = 1 - 2\frac{r_d}{r_b}$ is the only stable point, whilst $\rho_{nc} = 0$ is still a fixed point but

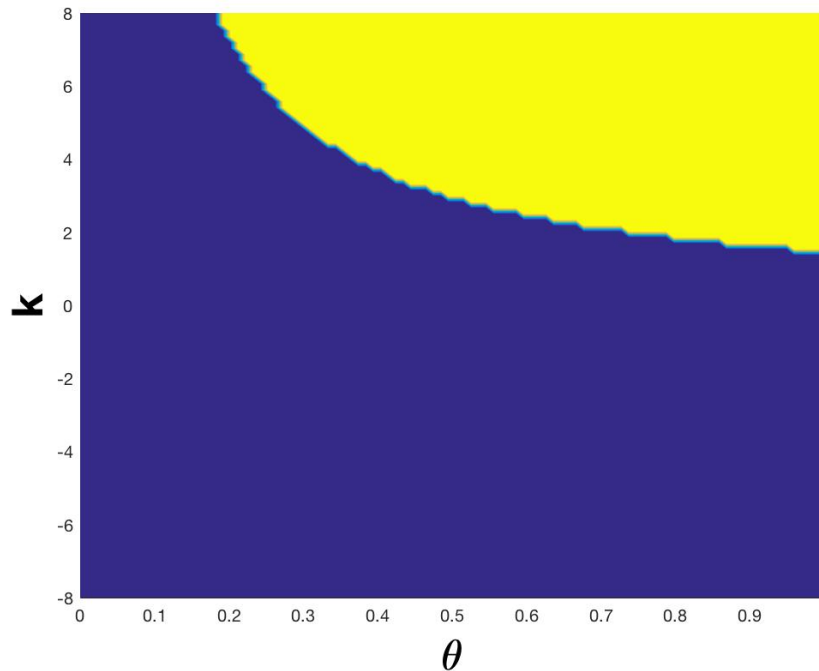


Fig. 3: Allee effect in the deterministic system: The yellow area represents the area where 0 is stable for the above system, whereas the blue one depicts the area where 0 becomes unstable.

unstable (see Fig. 4). In this case the parameters (k, θ) do not influence the model's steady states as in the deterministic one (see Fig. 3). The strength of the white noise is so large, that annihilates the existence of multiple steady states. Finally, the Allee effect that was observed in the deterministic case² now disappears and the only remaining effect is the survival of a steady cancerous population.

5.3. Impact of heterogeneity on a spatio-temporal GoG model

In the previous section, and as a first step towards the investigation of the GoG model (5.6), spatial dependence was ignored and in consequence we investigated an ODE model, where only the dynamics of resting and switching between the two species were taken into account. Nonetheless, the main goal of the current section is to investigate what is the impact, if any, of the diffusion component on the stability analysis of the GoG model. Namely, in the current section we consider the full model (5.6) where now a randomization parameter is introduced both in the diffusion and the reaction terms. To do this, we follow again the averaging approach introduced in section 4. To guarantee the well posedness of the system, we again consider the

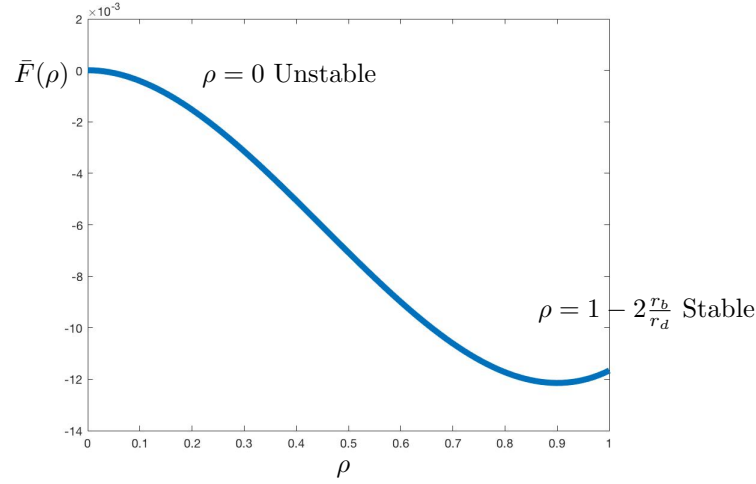


Fig. 4: Plot of the potential of equation (5.10), namely $\bar{F}(\rho) := \frac{1}{2}r_b\rho(1 - \rho) - r_d\rho$ with $r_b = 0.1$ and $r_d = 0.02$.

functional approximation of the white noise introduced in (2.4).

Thus we focus on the investigation of the following

$$\frac{\partial \rho^N}{\partial t} = \Delta((1 - r_s(\rho^N, \xi_t^N))\rho^N) + F(\rho^N, \xi_t^N), \quad (x, t) \in D_T, \quad (5.11)$$

$$\rho^N(x, t) = 0, \quad (x, t) \in \Gamma_T, \quad (5.12)$$

$$0 \leq \rho^N(x, 0) = \rho_0(x) \leq 1, \quad x \in D, \quad (5.13)$$

for a bounded and smooth $D \subset \mathbb{R}^3$, where $F(\rho^N, \xi_t^N) := r_b r_s(\rho^N, \xi_t^N) \rho^N (1 - \rho^N) - r_d \rho^N$. Here $T > 0$ again stands for the maximum existence time of solution ρ^N . Note that the Dirichlet type boundary condition (5.12) means that there are no cancerous cells on the boundary of the domain under investigation D , which is a plausible assumption. Using [5, Theorem 9 in Chapter 7] we immediately obtain existence, uniqueness and positivity of a global-in-time solution, since also the drift term $F(\rho^N, \xi_t^N)$ is bounded, for the random partial differential equation (RPDE) problem (5.11)-(5.13).

By virtue of an energy approach or via the maximum principle, we derive the positivity and boundeness of the solution of the limiting problem, while its uniqueness is derived using Gronwall's lemma, see also,²¹

$$\frac{\partial \rho}{\partial t} = \Delta((1 - \bar{r}_s(\rho))\rho) + r_b \bar{r}_s(\rho) \rho (1 - \rho) - r_d \rho, \quad (x, t) \in D_T, \quad (5.14)$$

$$\rho(x, t) = 0, \quad (x, t) \in \Gamma_T, \quad (5.15)$$

$$0 \leq \rho(x, 0) = \rho_0(x) \leq 1, \quad x \in D, \quad (5.16)$$

where $\bar{r}_s(\rho) = \frac{1}{2}$.

Denote $H = L^2(D)$ and $\mathcal{H}_T = C(0, T; H)$ then making use of the novel averaging principle developed in section 4 we can prove the following averaging result, which is inspired by,³

Theorem 5.1. *For every $T > 0$ the unique solution ρ^N of problem (5.11)-(5.13) converges as $N \rightarrow \infty$ into \mathcal{H}_T to the unique solution ρ of the average problem (5.14)-(5.16).*

Proof. First note that the term $1 - r_s(\rho^N, \xi_t^N)$ is bounded due the definition of $r_s(\rho^N, \xi_t^N)$ given by (5.4) and hence both the diffusion coefficient and the reaction term in problem (5.11)-(5.12) are also bounded. Thus via parabolic regularity we derive the a priori estimate

$$\|\rho^N\|_{\mathcal{H}_T} \leq C_T < +\infty \quad \text{for any } T > 0,$$

therefore we can find a subsequence denoted again by (ρ^N) without any confusion and a function $\psi \in \mathcal{H}_T$ such that

$$\|\rho^N - \psi\|_{\mathcal{H}_T} \rightarrow 0 \quad \text{as } N \rightarrow +\infty. \quad (5.17)$$

Considering the weak formulation of (5.11)-(5.13) we have

$$\begin{aligned} \langle \rho^N, \phi \rangle &= \langle \rho_0, \phi \rangle + \left\langle \int_0^t (1 - r_s(\rho^N, \xi_\sigma^N)) \rho^N d\sigma, \Delta \phi \right\rangle \\ &\quad + \left\langle \int_0^t r_b r_s(\rho^N, \xi_\sigma^N) \rho^N (1 - \rho^N) - r_d \rho^N d\sigma, \phi \right\rangle, \end{aligned} \quad (5.18)$$

for any test function $\phi \in C_0^\infty(D)$, where $\langle \cdot, \cdot \rangle$ stands for the inner product in $H = L^2(D)$.

Using some algebraic manipulations then (5.18) infers

$$\begin{aligned} \langle \rho^N, \phi \rangle &= \langle \rho_0, \phi \rangle + \left\langle \int_0^t (1 - r_s(\rho^N, \xi_\sigma^N)) \rho^N d\sigma, \Delta \phi \right\rangle \\ &\quad + \left\langle \int_0^t r_b r_s(\rho^N, \xi_\sigma^N) \rho^N (1 - \rho^N) - r_d \rho^N d\sigma, \phi \right\rangle + \left\langle \int_0^t (1 - \bar{r}_s(\rho^N)) \rho^N d\sigma, \Delta \phi \right\rangle \\ &\quad - \left\langle \int_0^t (1 - \bar{r}_s(\rho^N)) \rho^N d\sigma, \Delta \phi \right\rangle + \int_0^t \langle r_b \bar{r}_s(\rho^N) \rho^N (1 - \rho^N) - r_d \rho^N, \phi \rangle d\sigma \\ &\quad - \left\langle \int_0^t r_b \bar{r}(\rho^N) \rho^N (1 - \rho^N) - r_d \rho^N d\sigma, \phi \right\rangle, \end{aligned}$$

and after rearranging the preceding relation is reduced to

$$\begin{aligned} \langle \rho^N, \phi \rangle &= \langle \rho_0, \phi \rangle + \left\langle \int_0^t (1 - \bar{r}_s(\rho^N)) \rho^N d\sigma, \Delta \phi \right\rangle \\ &\quad + \left\langle \int_0^t [r_b \bar{r}_s(\rho^N) \rho^N (1 - \rho^N) - r_d \rho^N] d\sigma, \phi \right\rangle \\ &\quad + R_1^N(t) + R_2^N(t), \end{aligned} \quad (5.19)$$

where

$$\begin{aligned} R_1^N(t) &:= \left\langle \int_0^t [(1 - r_s(\rho^N, \xi_\sigma^N)) - (1 - \bar{r}_s(\rho^N))] \rho^N d\sigma, \Delta\phi \right\rangle \\ &= \left\langle \int_0^t [\bar{r}_s(\rho^N) - r_s(\rho^N, \xi_\sigma^N)] \rho^N d\sigma, \Delta\phi \right\rangle, \end{aligned}$$

and

$$R_2^N(t) := \left\langle \int_0^t r_b(r_s(\rho^N, \xi_\sigma^N) - \bar{r}(\rho^N)) \rho^N (1 - \rho^N) d\sigma, \phi \right\rangle.$$

Considering the term $R_1^N(t)$ we derive the following estimate:

$$R_1^N(t) \leq \tilde{C}_T \cdot \|\Delta\phi\|_\infty \delta_1(N),$$

where $\tilde{C}_T = T \cdot C_T$ and

$$\delta_1(N) := \sup_{\|\rho\|_{\mathcal{H}_T} \leq C_T} \sup_{0 \leq t < T} \left\| \frac{1}{t} \int_0^t (r_s(\rho, \xi_\sigma^N) - \bar{r}_s(\rho)) d\sigma \right\|_\infty \rightarrow 0 \quad \text{as } N \rightarrow +\infty,$$

for any $T > 0$ by virtue of Lemma 5.1. Therefore for given $T > 0$

$$R_1^N(t) \rightarrow 0 \quad \text{as } N \rightarrow +\infty \quad \text{for any } 0 < t < T. \quad (5.20)$$

In a similar manner we have

$$R_2^N(t) \leq \hat{C}_T \sup_{\|\rho\|_{\mathcal{H}_T} \leq C_T} \sup_{0 \leq t < T} \left\| \frac{1}{t} \int_0^t (r_s(\rho, \xi_\sigma^N) - \bar{r}_s(\rho)) d\sigma \right\|_\infty,$$

for any $T > 0$ where $\hat{C}_T := r_b T C_T (1 + C_T) \|\phi\|_\infty$ and using again Lemma 5.1 we infer

$$R_2^N(t) \rightarrow 0 \quad \text{as } N \rightarrow +\infty \quad \text{for any } 0 < t < T. \quad (5.21)$$

Letting now $N \rightarrow +\infty$ into (5.19) and using (5.17), (5.20) and (5.21) we derive

$$\langle \psi, \phi \rangle = \langle \rho_0, \phi \rangle + \left\langle \int_0^t (1 - \bar{r}_s(\psi)) \psi d\sigma, \Delta\phi \right\rangle + \left\langle \int_0^t [r_b \bar{r}_s(\psi) \psi (1 - \psi) - r_d \psi] d\sigma, \phi \right\rangle,$$

i.e. ψ is a weak solution of (5.14)-(5.16), and since this problem has a unique solution we finally infer that $\psi = \rho$. This completes the proof of the Theorem 5.1. \square

By the linear stability of problem (5.14)-(5.16), since the diffusion coefficient is small (c.f.¹⁹), we again obtain that $\rho_c = 1 - 2\frac{r_d}{r_b}$ is stable whilst the trivial steady-state $\rho_{nc} = 0$ is unstable. Therefore, the diffusion has no effect on the stability of the spatial homogeneous steady-states and again the cancerous population survives.

24 *M. Leocata, J.C.L. Alfonso, N.I. Kavallaris & H. Hatzikirou*

6. Discussion

In the current paper we have introduced a novel average method to deal with the study of the dynamics of a class of RODEs and RPDEs since classical averaging methods fail to treat this kind of problems. To elucidate the relevance of our theoretical results, we apply this method into an important biomedical problem, i.e. the assessment of intratumoral heterogeneity impact on tumor dynamics. In particular, we consider the development of gliomas according to a well-known Go or Grow (GoG) model, where intratumoral heterogeneity is modelled as a stochastic process. The deterministic version of the considered GoG model has been shown that there exhibits an emerging Allee effect (bistability). On the other hand, for the novel stochastic version of GoG model we demonstrate that the introduction of white noise, as a model of intratumoral heterogeneity, leads to a monostable tumor growth. The latter entails the disappearance of the Allee effect, and thus we conclude that the extinction is impossible under parametric variations. Consequently, our results suggest that the existence of heterogeneity worsens the prognosis of tumor growth, which is actually in accordance with the clinical experience and literature. However, the assumption of white noise is the worst possible scenario related to tumor heterogeneity, and therefore other noise distributions should be analyzed such as Gaussian noise. Moreover, there have been a plethora of studies trying to quantify intratumoral heterogeneity, see,^{1,20,22,24} nevertheless our method is able to include the existing literature and analyze the impact of data-driven heterogeneity distribution in more realistic tumor models. Furthermore, our method can be also implemented to investigate the long-time dynamics of the full GoG system (5.1)-(5.2), which will be the subject of a forthcoming work.

Acknowledgment

Part of the current work was inspired and initiated when NK was visiting the Center of Interdisciplinary Research (ZIF) and Helmholtz Centre for Infection Research. He would like to express his gratitude for the warm hospitality of both institutes and the financial support from ZIF. HH is supported by Mic2Mode-I2T (01ZX1710B) and HH is supported by SYSIMIT (01ZX1308D) and MulticellML (01ZX1707C) by the Federal Ministry of Education and Research (BMBF) and by the SYSMIFTA (031L0085B) of the ERACOSYSMED initiative. JCLA and HH gratefully acknowledges the funding support of the Helmholtz Association of German Research Centers - Initiative and Networking Fund for the project on Reduced Complexity Models (ZT-I-0010)

References

1. L. Alic, W. J. Niessen and J. F. Veenland, Quantification of heterogeneity as a biomarker in tumor imaging: A systematic review, *PLoS ONE* **9** (2014) 1–15.
2. K. Böttger, H. Hatzikirou, A. Voss-Böhme, E. A. Cavalcanti-Adam, M. A. Herrero and

- A. Deutsch, An Emerging Allee Effect Is Critical for Tumor Initiation and Persistence, *PLoS Computational Biology* **11** (2015) e1004366.
3. J. Duan and W. Wang, *Effective Dynamics of Stochastic Partial Differential Equations* (Elsevier, 2014).
 4. M. I. Freidlin and A. D. Wentzell, Random perturbations, in *Random perturbations of dynamical systems* (Springer, 1998), pp. 15–43.
 5. A. Friedman, *Partial Differential Equations of Parabolic Type* (Prentice-Hall, Inc., Englewood Cliffs, N.J., 1964).
 6. J. C. L. Alfonso, A. Köhn-Luque, T. Stylianopoulos, F. Feuerhake, A. Deutsch and H. Hatzikirou, Why one-size-fits-all vaso-modulatory interventions fail to control glioma invasion: in silico insights, *Scientific Reports* **6** (2016) 37283.
 7. I. M. Gelfand and G. E. Shilov, *Generalized functions, Vol. 4: applications of harmonic analysis* (Academic Press, 1964).
 8. J. C. L. Alfonso, K. Talkenberger, M. Seifert, B. Klink, A. Hawkins-Daarud, K. R. Swanson, H. Hatzikirou, A. Deutsch The biology and mathematical modelling of glioma invasion: a review, *J. R. Soc. Interface* **14**(136) (2017) 1–20.
 9. M. L. Goodenberger and R. B. Jenkins, Genetics of adult glioma, *Cancer genetics* **205** (2012) 613–621.
 10. J. Hale and S. Verduyn Lunel, Averaging in infinite dimensions, *Journal of Integral Equations and Applications* **2** (1990) 463–494.
 11. H. Hatzikirou, J. C. L. Alfonso, S. Mühle, C. Stern, S. Weiss and M. Meyer-Hermann, Cancer therapeutic potential of combinatorial immuno- and vaso-modulatory interventions, *J. R. Soc. Interface* **12** (2015) 20150439.
 12. X. Han and P. E. Kloeden, *Random Ordinary Differential Equations and Their Numerical Solution*, Probability Theory and Stochastic Modelling 85 (Springer, 2017).
 13. H. Hatzikirou, D. Basanta, M. Simon, K. Schaller and A. Deutsch, go or grow: the key to the emergence of invasion in tumour progression?, *Mathematical medicine and biology: a journal of the IMA* **29** (2012) 49–65.
 14. D. Henry, *Geometric Theory of Semilinear Parabolic Equations*, Lecture Notes in Mathematics (Springer Berlin Heidelberg, 2006).
 15. W. S. Ikeda N., *Stochastic Differential Equations and Diffusion Processes* (North-Holland Publishing Co., Amsterdam, 1981).
 16. D. N. Louis, H. Ohgaki, O. D. Wiestler, W. K. Cavenee, P. C. Burger, A. Jouvett, B. W. Scheithauer and P. Kleihues, The 2007 WHO classification of tumours of the central nervous system, *Acta neuropathologica* **114** (2007) 97–109.
 17. N. McGranahan and C. Swanton, Biological and therapeutic impact of intratumor heterogeneity in cancer evolution, *Cancer Cell* **27** (2015) 15–26.
 18. C. E. Meacham and S. J. Morrison, Tumour heterogeneity and cancer cell plasticity, *Nature* **501** (2013) 328–337.
 19. J. Murray, *Mathematical Biology Vol II: Spatial Models and Biomedical Applications* (Springer, 2003).
 20. L. Oesper, G. Satas and B. J. Raphael, Quantifying tumor heterogeneity in whole-genome and whole-exome sequencing data, *Bioinformatics* **30** (2014) 3532–3540.
 21. P. Quittner and S. Ph., *Superlinear parabolic problems. Blow-up, global existence & steady states*, Birkhäuser Adv. Texts Basler Lehrbücher (Birkhäuser, 2007).
 22. E. M. Rutter, H. T. Banks and K. B. Flores, Estimating intratumoral heterogeneity from spatiotemporal data, *Journal of Mathematical Biology* .
 23. N. Sanai, A. Alvarez-Buylla and M. S. Berger, Neural stem cells and the origin of gliomas, *New England Journal of Medicine* **353** (2005) 811–822.
 24. A. Sottoriva, I. Spiteri, S. G. M. Piccirillo, A. Touloumis, V. P. Collins, J. C. Marioni,

26 *M. Leocata, J.C.L. Alfonso, N.I. Kavallaris & H. Hatzikirou*

C. Curtis, C. Watts and S. Tavaré, Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics., *Proceedings of the National Academy of Sciences of the United States of America* **110**.

25. M. Tektonidis, H. Hatzikirou, A. Chauvière, M. Simon, K. Schaller and A. Deutsch, Identification of intrinsic in vitro cellular mechanisms for glioma invasion., *Journal of theoretical biology* **287** (2011) 131–47.