

Chapter 4

Stochastic Modeling and Simulation Methods for Biological Processes: Overview

Annelies Lejon and Giovanni Samaey

Abstract The use of stochastic modeling and simulation techniques is widespread in computational biology when fluctuations become important. In this chapter, we give a high-level overview of stochastic modeling techniques for biological problems, focussing on some common individual-based modeling and simulation methods. We pay particular attention to the equivalence between the stochastic process that governs the evolution of individual agents and the deterministic behaviour of the involved probability distributions, and we discuss numerical methods that exploit this relation for variance reduction purposes. The discussion will be illustrated using examples involving intracellular chemical reactions, bacterial chemotaxis and tumor growth, showing the effects of stochasticity at different scales and different levels of description.

Keywords Stochasticity · Stochastic differential equations · Velocity-jump processes · Asymptotic variance reduction

4.1 Introduction

Stochastic effects are ubiquitous in biological systems, at multiple scales. At an intracellular level, for instance, gene regulatory networks often exhibit different metastable states. Since the number of molecules in a cell is not very high, significant fluctuations in concentrations can occur, triggering transitions between these metastable states [20, 28]. At a cellular level, individual cells can be modeled as agents that interact with each other and with their environment.

Stochasticity can then be introduced to account for differences between individual cells or to incorporate the coarse-grained effect of phenomena that occur at more microscopic scales directly at the cellular level. (An example of the latter would be the use of a Brownian motion to model the net effect of a large number of collisions of

A. Lejon · G. Samaey (✉)

Department of Computer Science, KU Leuven, Celestijnenlaan 200A, 3001 Leuven, Belgium
e-mail: Giovanni.Samaey@cs.kuleuven.be

a large molecule with surrounding—but not explicitly modeled—solvent molecules.) Such an approach has been followed in many settings, including the applications that will be considered as illustrative examples in this chapter: bacterial chemotaxis (see, e.g., [5, 15] and references therein) and tumor growth (see, e.g., [43] for a recent review and references). At a population level, one usually models the evolution of cell densities via partial differential equations (PDEs) of reaction-advection-diffusion type. At this level, one can introduce, for instance, stochastic parameters and geometries to account for differences between individuals, resulting in PDEs with stochastic coefficients [22].

Despite the stochastic nature of the time evolution, one is usually interested in deterministic quantities, such as the mean switching time between metastable states or the expected behaviour of a large population of cells or individuals. Additionally, one may also be interested in deviations with respect to this mean behaviour, requiring information on higher order statistics or on the complete probability distribution of possible states of the system. While, in principle, the time evolution of these probability distributions can be modeled using *deterministic* evolution laws, the associated computational cost is usually prohibitive due to the high dimensionality of the resulting equations. One therefore needs to resort to some form of Monte Carlo simulation of the stochastic process [7].

In this chapter, we discuss stochastic individual-based modeling techniques and show the equivalence with deterministic techniques for modeling the involved probability distributions. In Sect. 4.2, we introduce stochastic models for chemically reacting systems with low number of molecules (as would occur in modeling intracellular dynamics). We discuss both the time-discrete and time-continuous case, and show how these models relate to classical mean-field equations for the evolution of concentrations. In Sect. 4.3, we consider advection-diffusion processes, as they occur, for instance, in agent-based models for bacterial chemotaxis and tumor growth. We briefly describe cellular automata and Markov jump processes, before giving a more detailed discussion of stochastic differential equations (SDEs). We show the relation between an SDE for an individual particle and an advection-diffusion equation (the Fokker–Planck equation) for the population density. In Sect. 4.4, we turn to more realistic microscopic processes, and relate these to kinetic theory and Boltzmann-type equations. For each of the described modeling techniques, we discuss the mathematical formulation of the stochastic process as well as practical simulation algorithms. In Sect. 4.5, we discuss Monte Carlo simulation using these stochastic processes. We show how to compute confidence intervals for the obtained results and introduce numerical algorithms that can yield results with significantly reduced variance.

We illustrate the introduced concepts for two specific applications: bacterial chemotaxis (Sect. 4.6) and tumor growth (Sect. 4.7). For each of these applications, we discuss in detail the modeled processes and introduce a dedicated simulation technique.

We conclude this introduction with a few remarks on topics that will not be treated in this chapter. First, we will not discuss uncertainty propagation resulting from uncertainty in parameters or geometries. For a mathematical and algorithmic introduction to this topic, we refer to [22]. Second, in many applications, one may

have, besides the mathematical model, some observation data available. Two questions can then be posed: (i) does the data provide sufficient support to validate the model; and (ii) can one use the data to estimate unknown parameters in the model? These topics form the subject of intense current research, and we refer to [51] for an introduction.

4.2 Stochastic Modeling of Chemical Reactions

In this section, we provide an individual-based description of chemically reacting systems. For concreteness, we start from the following system of chemical reactions, as introduced by Schlögl [46],



in which the rate constants k_i ($1 \leq i \leq 4$) indicate that the probability for two randomly chosen molecules of type A to react according to reaction i in any time interval $[t, t + dt)$ is given by $k_i dt$.

In Sect. 4.2.1, we discuss discrete-time simulation. It will appear that the corresponding algorithm introduces a time discretization error, and at the same time is also very inefficient. We therefore turn to continuous-time simulation algorithms in Sect. 4.2.2. Next, in Sect. 4.2.3, we introduce the chemical master equation, a deterministic system of equations for the probability of finding a given number of molecules at any given time. From this master equation, we can obtain information on the mean behaviour of the system and on fluctuations. We conclude in Sect. 4.2.4 with a numerical illustration on the Schlögl model. The exposition in this section is based on [17].

4.2.1 Discrete-Time Simulation

Let us denote by $A(t)$ the number of molecules of type A in the system at time $t \geq 0$, and assume as initial condition $A(0) = A_0$. We are interested in an approximate solution $A^k \approx A(t^k)$, with $t^k = k\Delta t$ and $\Delta t > 0$ a small time step. (The symbol k that indicates the discrete time instance is added as a superscript for notational consistency throughout the chapter.)

At time $t = t^k$, the probability of reaction 1 (with rate k_1) taking place in $(t, t + \Delta t)$ is approximately given by

$$\alpha_1^k \Delta t = A^k (A^k - 1) k_1 \Delta t, \quad (4.2)$$

with α_1^k the *propensity function*, since $k_1 \Delta t$ is the probability of two randomly chosen molecules of type A to react according to reaction 1, and $A^k(A^k - 1)$ is the number of pairs of molecules that can be randomly chosen. (Note that this is an approximation due to the fact that we have replaced the infinitesimal interval of length dt by a finite interval of length Δt .) Similarly, the propensity functions for the other reactions can be seen to be

$$\alpha_2^k = A^k(A^k - 1)(A^k - 2)k_2, \quad \alpha_3^k = k_3, \quad \alpha_4^k = A^k k_4. \quad (4.3)$$

To take a time step from time t^k to t^{k+1} , one needs to decide for each reaction whether or not it has occurred during the time step. This can be done using the following algorithm:

Algorithm 1 (*Time-discrete simulation of Schlögl's model*) Given the concentration A^k at time t^k , we compute A^{k+1} at time t^{k+1} as follows:

1. For each reaction i :
 - Compute the propensity α_i^k using (4.2) or (4.3);
 - Generate an independent random number r_i^k from the uniform distribution on the interval $(0, 1)$;
 - Decide that reaction i has occurred during the time interval if $r_i^k \leq \alpha_i^k \Delta t$ (else, the reaction has not occurred);
2. Compute A^{k+1} by applying all reactions that occurred during the time step (for instance, if only reaction 1 occurred, we have $A^{k+1} = A^k + 1$).

A simulation is then performed by repeating the above time step over the time interval of interest. The algorithm can easily be generalized to systems with multiple species and any number I of possible reactions.

There are two problems when using the above algorithm, which will both be dealt with when switching to continuous-time simulation in Sect. 4.2.2. First, Algorithm 1 introduces a time discretization error since it replaces an infinitesimal time interval of length dt by a finite interval of length Δt . This error manifests itself in two ways. First, we neglect the (non-zero) probability that two reactions of the same type occur within the time step of size Δt . Moreover, every reaction event influences the propensity function of other reactions, since the propensity functions depend on the state of the system. As a consequence, there will be an error in the used reaction probabilities when multiple reactions are performed during a single time step.

To limit the impact of these time discretization errors, Δt should be chosen fairly small. A common guideline is to choose Δt such that the probability of having a reaction is less than 1 % per time step. However, this implies that 99 % of the time steps are taken just to conclude that nothing happened! This is detrimental to computational efficiency, which is the second problem with the above algorithm.

4.2.2 Continuous-Time Simulation

Instead of taking discrete time steps of fixed size Δt from t to $t + \Delta t$, and calculating the probability that a reaction takes place in that time interval $[t, t + \Delta t)$, a continuous-time simulation computes the (random) time increment τ until the next reaction takes place by sampling from the corresponding probability distribution. Afterwards, a second random number is generated to decide *which* reaction occurred.

Let us first characterize the relevant probability distribution for τ . Consider a system with one reaction with propensity $\alpha_1(t)$. (Recall that the propensity is time-dependent due to the time-dependence of the concentrations of the present species.) Further, denote as $f_1(t, s)ds$ the probability, at time t , that the next reaction occurs in the time interval $[t + s, t + s + ds]$, and denote the probability that no reaction occurs in the interval $[t, t + s]$ as $g_1(t, s)$. We then have

$$f_1(t, s)ds = g_1(t, s)\alpha_1(t + s)ds = g_1(t, s)\alpha_1(t)ds, \quad (4.4)$$

in which we used the fact that $\alpha_1(t + s) = \alpha_1(t)$ in the absence of reactions, and independence of individual reaction events. (Then, the probability that the *first* reaction occurs in the time interval $[t + s, t + s + ds]$ is given by the product of the probability $\alpha_1(t + s)ds$ that there is a reaction in that time interval and the probability $g_1(t, s)$ that no reaction occurred earlier.)

The two quantities $f_1(t, s)$ and $g_1(t, s)$ are, as for now, unknown. We derive a differential equation for $g_1(t, s)$. We start by observing that the probability of not having a reaction in the time interval $[t, t + s + ds]$ can be written as the product of the probability of not having a reaction in the time interval $[t, t + s]$ and the probability of not having a reaction in the time interval $[t + s, t + s + ds]$, i.e.,

$$g_1(t, s + ds) = g_1(t, s)(1 - \alpha_1(t)ds), \quad (4.5)$$

from which we obtain

$$\frac{g_1(t, s + ds) - g_1(t, s)}{ds} = -\alpha_1(t)g_1(t, s). \quad (4.6)$$

Given that $g(t, 0) = 1$ (zero probability of having the reaction exactly at time t), we obtain

$$g_1(t, s) = \exp(-\alpha_1(t)s), \quad (4.7)$$

and hence, using (4.4), the probability that the first reaction occurs in the infinitesimal interval $t + s$ is given by

$$f_1(t, s)ds = \alpha_1(t) \exp(-\alpha_1(t)s)ds, \quad s \geq 0. \quad (4.8)$$

The corresponding cumulative distribution is given as

$$F_1(t, \tau) = \int_0^\tau f_1(t, s) ds = 1 - \exp(-\alpha_1(t)\tau), \quad \tau \geq 0. \quad (4.9)$$

(Note that we have $F_1(t, 0) = 0$ and $\lim_{\tau \rightarrow \infty} F_1(t, \tau) = 1$.)

Now that we know the probability distribution $f_1(t, s)$, we are ready to generate a random time increment $\tau_1(t)$, sampled from $f_1(t, s)$, which we denote as $\tau_1(t) \sim f_1(t, s)$. Using the transformation method for the generation of random numbers [7], the increment $\tau_1(t)$ until the next event of reaction 1 can be computed from a uniformly distributed random number $\tilde{\theta}_1$ in $(0, 1)$ via

$$\tau_1 = F^{-1}(t, \tilde{\theta}_1) = -\frac{1}{\alpha_1(t)} \log(1 - \tilde{\theta}_1), \quad \text{or} \quad \tau_1 = -\frac{1}{\alpha_1(t)} \log(\theta_1), \quad (4.10)$$

with θ_1 also a uniformly distributed random number in $(0, 1)$ and $F^{-1}(t, \theta)$ the inverse of the cumulative density $F(t, s)$ with respect to s , treating t as a parameter.

Remark 1 (Markov property) The probability distribution $f_1(t, s)$ is called the exponential distribution with rate $\alpha_1(t)$. Its main property is that it is memoryless, which implies that the evolution is completely determined by the current state (and no information from the past is required). In particular, this implies that, to determine when the next reaction will occur, it is irrelevant how long the system is already in its current state. Mathematically, this can be seen by checking that, for any positive T and τ , we have

$$\Pr(\tau_1(t) > T) = \Pr(\tau_1(t) > \tau + T | \tau_1(t) > \tau), \quad (4.11)$$

i.e., the probability that the next reaction will not occur within a time interval of length T from the current time, does not depend on the amount of time τ that the system is already in the current state.

If multiple reactions can occur in the system (as is the case for the Schlögl model (4.1)), a naive way of proceeding is to generate the next reaction time for each of the reactions and only select the reaction that occurs first, after which the system clock is updated and the procedure is repeated. With I possible reactions, this requires the generation of I exponentially distributed random numbers (independent of the number of species in the system) to choose a single reaction. This algorithm can be made much more efficient by making use of the following theorem:

Theorem 2 (Exponential distributions) *Consider I exponential distributions $f_i(t, s)$ with rates $\alpha_i(t)$, and consider an independent set of random numbers $\tau_i(t)$, each sampled from the corresponding distribution $f_i(t, s)$. Let $\tau(t)$ be the minimum of these, $\tau(t) = \min_i \tau_i(t)$. Then, the probability distribution for $\tau(t)$ is an exponential distribution with rate $\alpha(t) = \sum_i \alpha_i(t)$. Moreover, the probability that $\tau(t) = \tau_i(t)$ (the probability that the i -th reaction occurs first) is given by $\alpha_i(t)/\alpha(t)$.*

Using this theorem, only two random numbers need to be generated per time step: one to determine the time increment until the next reaction, and one to choose the next reaction. This gives rise to the following classical algorithm, due to Gillespie [23], which is immediately written for a system consisting of I reactions:

Algorithm 2 (*Stochastic simulation algorithm (SSA) for chemically reacting systems*) Given the concentration A^k at time t^k , we compute A^{k+1} at time t^{k+1} as follows:

1. For each reaction i , compute the propensity $\alpha_i^k = \alpha_i(t^k)$, and compute the total propensity $\alpha^k = \sum_i \alpha_i^k$;
2. Generate a uniformly distributed random number θ^k in $(0, 1)$, and compute the time increment until the next reaction occurs as $\tau^k = -\frac{1}{\alpha^k} \log(\theta^k)$,
3. Generate a uniformly distributed random number y^k in $(0, 1)$ and select the reaction i for which

$$\sum_{j \leq i-1} \alpha_j^k / \alpha^k \leq y < \sum_{j \leq i} \alpha_j^k / \alpha^k;$$

4. Compute A^{k+1} by applying the selected reaction i (for instance, if reaction 1 was selected, we have $A^{k+1} = A^k + 1$).

Again, the algorithm can easily be generalized to systems with multiple species. We refer to [26] for more details and variants. Here, we only state the main convergence result: the stochastic simulation algorithm (SSA) is *exact*, in the sense that it does not contain any time discretization error.

Remark 3 (Acceleration of SSA) While this method is significantly faster than the time-discrete Algorithm 1 without introducing a time discretization error, the resulting algorithm can still be computationally prohibitive, especially in situations with many chemical reactions with disparate time scales. Consider for instance a system with a fast, reversible reaction and an irreversible but very slow reaction. In such a system, most of the time steps will select the fast reversible reaction, resulting in very small time steps that go back and forth along the fast reaction. More sophisticated algorithms, tailored to these situations, have been developed. These include, for instance, the τ -leaping method [9, 25], the slow-scale stochastic simulation algorithm [8], and R-leaping [3]. Note, however, that such methods accelerate simulation at the expense of re-introducing a (small) time-discretization error.

4.2.3 Population-Level Dynamics and Mean-Field Approximation

In general, repeated simulation of a stochastic process for a chemically reacting system yields different results for each stochastic realization. The precise results depend on the generated sequence of random numbers. Usually, one is not interested in the

detailed behaviour of such an individual realization, but in quantitative statements on the mean behaviour and on fluctuations.

In this section, we first derive the *chemical master equation*, which gives a complete description of the (time-dependent) probability distribution of possible states for the system. Afterwards, we discuss the potential and limitations of using this equation to derive information on the statistics of the process.

In the previous sections, we denoted by the random variable $A(t)$ the number of molecules of type A at time t . Here, we define the *deterministic* quantity $p_n(t)$, which represents the *probability* that the system contains exactly n molecules of type A at time t , i.e.,

$$p_n(t) = \Pr(A(t) = n).$$

In the incremental time interval $[t, t + dt)$, the state can only change by ± 1 , since all reactions either create or destroy one molecule of type A . We can use the definition of the reactions (4.1) to compute $p_n(t + dt)$,

$$\begin{aligned} p_n(t + dt) &= p_n(t) + \underbrace{\{[k_1(n-1)(n-2) + k_3] p_{n-1}(t)dt + [k_2(n+1)n(n-1) + k_4(n+1)] p_{n+1}(t)dt\}}_{\text{gain}} \\ &\quad - \underbrace{\{[k_1n(n-1) + k_3] + [k_2n(n-1)(n-2) + k_4n]\} p_n(t)dt}_{\text{loss}}, \end{aligned} \quad (4.12)$$

in which we recognize

- a gain term that expresses the sum of (i) the probability that the system contains $n - 1$ molecules at time t and a molecule is produced during the time increment; and (ii) the probability that the system contains $n + 1$ molecules at time t and a molecule is destroyed during the time increment;
- a loss term that expresses the probability that the system contained n molecules at time t and a molecule is either created or destroyed during the time increment.

Reordering the terms, we get an (infinite-dimensional) system of ordinary differential equations, the chemical master equation:

$$\begin{aligned} \dot{p}_n(t) &= [k_1(n-1)(n-2) + k_3] p_{n-1}(t) + [k_2(n+1)n(n-1) + k_4(n+1)] p_{n+1}(t) \\ &\quad - [k_1n(n-1) + k_3 + k_2n(n-1)(n-2) + k_4n] p_n(t), \quad n \geq 0, \end{aligned} \quad (4.13)$$

in which, formally, $p_{-1}(t) = 0$.

The chemical master equation is equivalent to the stochastic description, but of limited practical use due to its infinite-dimensional nature. However, it can be used as a starting point to obtain information on the statistics of the stochastic process. In general, we denote the expectation of a function of the number of molecules, $f(n)$, by

$$F(t) := \mathbb{E}[f(A(t))] = \sum_{n=0}^{\infty} f(n)p_n(t), \quad (4.14)$$

which amounts to a weighted average of f , weighted by the probability density for n .

Let us consider the mean behaviour. The expected number of molecules of type A in the system at any given time is given by the (deterministic) quantity

$$M(t) := \mathbb{E}[A(t)] = \sum_{n=0}^{\infty} np_n(t). \quad (4.15)$$

To obtain an ordinary differential equation for the evolution of $M(t)$, we start from (4.13), and write

$$\dot{M}(t) = \sum_{n=0}^{\infty} n \dot{p}_n(t). \quad (4.16)$$

Using (4.13), we get

$$\begin{aligned} \dot{M}(t) &= \sum_{n=0}^{\infty} n \dot{p}_n(t) \\ &= \sum_{n=0}^{\infty} n \{ [k_1(n-1)(n-2) + k_3] p_{n-1}(t) + [k_2(n+1)n(n-1) + k_4(n+1)] p_{n+1}(t) \\ &\quad - [k_1n(n-1) + k_3 + k_2n(n-1)(n-2) + k_4n] p_n(t) \}. \end{aligned}$$

We regroup the terms per reaction. We first consider reaction 3, for which we have

$$\begin{aligned} \sum_{n=0}^{\infty} k_3n (p_{n-1}(t) - p_n(t)) &= k_3 \sum_{n=0}^{\infty} ((n+1)p_n(t) - np_n(t)) \\ &= k_3 \sum_{n=0}^{\infty} p_n(t) = k_3. \end{aligned} \quad (4.17)$$

Next, consider reaction 4. Here, we have

$$\begin{aligned} \sum_{n=0}^{\infty} k_4n ((n+1)p_{n+1}(t) - np_n(t)) &= k_4 \sum_{n=0}^{\infty} ((n-1)np_n(t) - n^2p_n(t)) \\ &= -k_4 \sum_{n=0}^{\infty} np_n(t) = -k_4M(t). \end{aligned} \quad (4.18)$$

Similarly, we obtain for the reactions 1 and 2,

$$\sum_{n=0}^{\infty} k_1 n ((n-1)(n-2)p_{n-1}(t) - n(n-1)p_n(t)) = k_1 \sum_{n=0}^{\infty} n(n-1)p_n(t), \quad (4.19)$$

$$\sum_{n=0}^{\infty} k_2 n ((n+1)n(n-1)p_{n+1}(t) - n(n-1)(n-2)p_n(t)) = -k_2 \sum_{n=0}^{\infty} n(n-1)(n-2)p_n(t), \quad (4.20)$$

resulting in the equation

$$\dot{M}(t) = k_3 - k_4 M(t) + k_1 \sum_{n=0}^{\infty} n(n-1)p_n(t) - k_2 \sum_{n=0}^{\infty} n(n-1)(n-2)p_n(t). \quad (4.21)$$

We notice immediately that the evolution of the mean $M(t)$ does not only depend on the mean itself, but also on higher-order statistics of the distribution, such as the second moment

$$\mathbb{E} \left[A(t)^2 \right] = \sum_{n=0}^{\infty} n^2 p_n(t).$$

We can proceed similarly to derive an evolution equation for the variance

$$V(t) = \mathbb{E} \left[(A(t) - M(t))^2 \right]. \quad (4.22)$$

However, we should expect even higher order moments to appear in the corresponding righthand side, leading to an infinite cascade. If we want to obtain an evolution law in terms of only the mean number of molecules $M(t)$, we will therefore be obliged to resort to an approximation.

Let us now take a more detailed look into the fluctuations around the mean, as measured by the variance $V(t)$. We are specifically interested in systems with large numbers of molecules, for which we assume a mean-field approximation to hold. We therefore introduce a characteristic number of molecules per unit volume N , and look at the *concentrations* $a(t) = A(t)/N$ and $\rho(t) = M(t)/N$. Then, the variance can be written as

$$V(t) = \frac{1}{N^2} \mathbb{E} \left[(a(t) - \rho(t))^2 \right]. \quad (4.23)$$

We conclude that fluctuations around the mean concentration become negligible as the number of molecules per unit volume N tends to infinity. In that limit, the quantized concentration n/N approaches a continuous variable a , and the probability distribution $(p_n(t))_{n=0}^{\infty}$ approaches a continuous probability distribution $p(a, t)$, $a \in [0, \infty)$. The fact that the fluctuations vanish in that limit implies that $p(a, t) = \delta_{M(t)}(a)$, i.e., the concentration $a(t) = M(t)$ almost surely. (We note that the above,

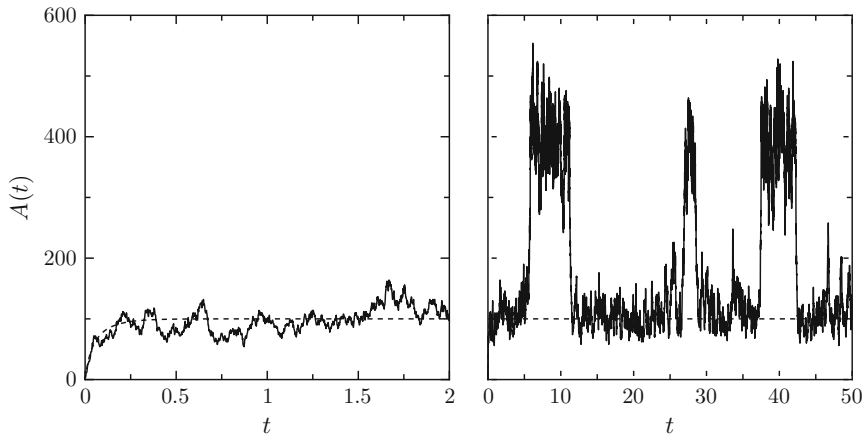


Fig. 4.1 Simulation of the Schlögl model (4.1): number of molecules as a function of time for a single realization of the stochastic simulation, using Algorithm 2 (*solid*) and the corresponding mean-field approximation (*dashed*). *Left* short time-scale; *right* long time-scale. Parameter values are in the text

rather heuristic, reasoning can be turned into a rigorous mathematical theory, see, e.g., [24].)

Using the above reasoning in the limit when N tends to infinity, one can derive from Eq. (4.21) the mean-field approximation for $\rho(t)$ as

$$\dot{\rho} = k_3 - k_4\rho + k_1\rho^2 - k_2\rho^3. \quad (4.24)$$

From the above derivation, one concludes that stochastic modeling of chemical reactions is mainly useful when the number of molecules present is not too large. In that case, results can deviate from the mean-field approximation for two reasons: (i) stochastic fluctuations around the mean can become important; and (ii) due to the present nonlinearities, the ensemble average of a large number of systems with a low number of molecules does not necessarily follow the mean-field behaviour (Fig. 4.1).

4.2.4 Numerical Simulations for the Schlögl Model

Let us now consider system (4.1) with (non-dimensionalized) reaction rates $k_1 = 0.18$, $k_2 = 2.5 \times 10^{-4}$, $k_3 = 2200$ and $k_4 = 37.5$. We simulate one realization of a stochastic simulation using Algorithm 2, as well as a forward Euler simulation of the mean field Eq. (4.24) with time step $\Delta t = 5 \times 10^{-3}$. As an initial condition, we take $A(0) = 0$. For the chosen parameter values, Eq. (4.24) has two stable steady state, $A_1 = 100$ and $A_2 = 400$, and an unstable steady state $A_u = 220$. Thus, the Schlögl model represents a bistable system. Figure 4.1 shows the results. The mean-field

equation converges to one of the two stable equilibria, depending solely on the initial condition. When $A(0) \in [0, A_u]$, the solution converges to A_1 ; when $A(0) > A_u$ the solution converges to A_2 . On short time-scales (left figure), the stochastic simulation fluctuates around the stable equilibrium of the mean-field equation. However, over long time scales (right figure), we notice that the fluctuations cause the system to occasionally switch between the two steady states. Such occasional switches are called *rare events*, and they occur when large fluctuations can occur; for instance, in gene regulatory networks [20, 28]. If one is interested in quantities such as mean switching times, one cannot use a mean-field approximation and needs to resort to stochastic simulation. We refer to [52] for theoretical and computational work related to stochastic simulation of rare events.

4.3 Stochastic Modeling of Advection-Diffusion Processes

In the previous section, all systems were assumed to be well mixed, such that only the temporal evolution of concentrations needed to be considered. In many processes, however, interesting dynamics arises from spatial heterogeneity. In a biological context, one can, for instance, think of bacterial chemotaxis, tumor growth, or bone tissue engineering. In this section, we give an overview of individual-based modeling techniques for biological systems consisting of moving individuals that are able to reproduce and die.

In Sect. 4.3.1, we consider the positions of the individuals to be discrete (on a lattice). We briefly discuss cellular automata and Markov jump processes, and give references to the corresponding literature. In Sect. 4.3.2, we introduce Brownian motion and stochastic differential equations (SDEs), which are used for space/time continuous modeling of random motion. Subsequently, in Sect. 4.3.3, we discuss the equivalence between the SDE for an individual particle and the (deterministic) Fokker–Planck equation that describes the evolution of the particle density.

4.3.1 Discrete-Space Modeling

Several techniques exist for the discrete stochastic modeling of biological particles. We briefly discuss cellular automata and Markov jump processes.

Cellular automata In cellular automata, one considers space to be discretized as a grid, say $\Pi(x) = \{x_n\}_{n=0}^N$ in one space dimension. The state is then given as the number of particles A_n^k at each grid location x_n at each discrete moment in time t^k . (Clearly, one can incorporate the presence of particles of multiple types.) The cellular automaton then defines an evolution law that determines the state $A^{k+1} = (A_n^{k+1})_{n=0}^N$ at time t^{k+1} from the state A^k . This evolution law can contain reactions with associated rates, as in the time-discrete schemes in Sect. 4.2.1. Movement on the grid can be modeled using *hops*, which can be seen as a reaction event (with an

associated reaction rate) in which a particle moves from one lattice site to another. Then, the algorithmic structure that resulted in Algorithm 1 can be reused to model advection-diffusion processes [17].

One particularly appealing feature of cellular automata is their modeling flexibility: in defining the evolution laws, any set of rules can be allowed. One can, for instance, change the rules depending on the number of neighbors, or let the evolution of individual particles depend on some internal state variable (see also Sect. 4.4). We refer to [21, 33, 39] for a number of cellular automata models in the context of tumor growth, and to [12, 31] for examples in bone tissue engineering.

Markov jump processes When keeping a discrete state space, but allowing time to be continuous, one ends up with a Markov jump process. Consider a particle with position $X(t)$ that is allowed to reside on any position in the lattice $\Pi(x)$. Given that the current position $X^k = x_n$ at time t^k , we can introduce a propensity $\alpha_{n,m}^k$, $1 \leq m \leq N$, such that $\alpha_{n,m}^k dt$ represents the probability that the particle moves from x_n to x_m in the infinitesimal time interval $[t, t + dt)$. Then, such a movement can be added to the table of reactions in Algorithm 2, and the same algorithm can be used.

4.3.2 Stochastic Differential Equations (SDEs)

When space and time are allowed to be continuous, the corresponding model becomes an SDE. In this section, we start from a definition of Brownian motion (Sect. 4.3.2.1). We then proceed to the construction of general SDEs in the Itô sense (Sect. 4.3.2.2) and discuss numerical methods (Sect. 4.3.2.3). We conclude in Sect. 4.3.2.4 with a numerical example that illustrates the results. The exposition is partly based on [29].

4.3.2.1 Brownian Motion

A scalar standard Brownian motion essentially describes an unbiased random walk in one space dimension. (Generalizations to multiple space dimensions are, of course, straightforward.) The description is phenomenological. We define a Brownian motion as a random variable $W(t)$, continuous in time $t \in [0, T]$, that satisfies the following conditions:

1. $W(0) = 0$ (with probability 1);
2. For $0 \leq s \leq t \leq T$, the *increment* $W(t) - W(s)$ is a normally distributed random variable with mean 0 and variance $t - s$, i.e., $W(t) - W(s) \sim \sqrt{t - s}N(0, 1)$, where $N(0, 1)$ denotes a standard normally distributed random variable (with mean 0 and variance 1);
3. For $0 \leq s \leq t \leq u \leq v \leq T$ are independent.

There are several ways of justifying this definition. One way is to start by defining a grid $\Pi(x) = \{-N\Delta x, \dots, 0, \dots, N\Delta x\}$ and letting a particle move one grid cell to the left or to the right (each with probability 1/2) in each time step of size Δt ,

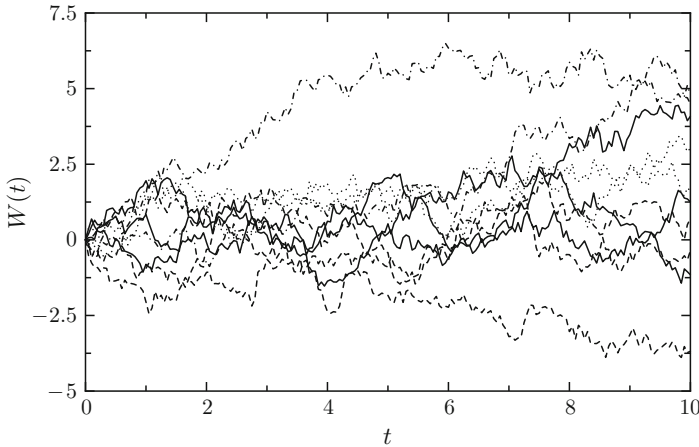


Fig. 4.2 Realizations of Brownian motion: $W(t)$ as a function of time, using (4.25) with $\Delta t = 5 \times 10^{-2}$

starting at position $X(t_0) = 0$ at time $t_0 = 0$. When choosing $\Delta x^2 = \Delta t$ and taking the limit of $\Delta t \rightarrow 0$, we obtain the standard Brownian motion.

To visualize realizations of Brownian motion, we consider time-discretized Brownian paths, generated as $W^k \approx W(t^k)$, with $t^k = k\Delta t$ via time-stepping,

$$W^{k+1} = W^k + \sqrt{\Delta t}\xi^k, \quad \xi^k \sim N(0, 1), \quad (4.25)$$

where the random numbers ξ^k are independent and identically distributed (i.i.d.). In Fig. 4.2, we show 10 realizations of a Brownian motion with $\Delta t = 5 \times 10^{-2}$.

The figure illustrates some properties of Brownian motion. A proper definition of the probability spaces generated by Brownian motion is out of the scope of the present chapter. Let us just suffice by stating that, when writing $\mathbb{E}[\cdot]$, we imply the mean with respect to all possible Brownian paths $W(t)$. It can be proved (see, e.g., [11]) that the expected value of a Brownian motion $\mathbb{E}[W(t)] = 0$ for any $t \in [0, T]$. This is easily seen intuitively, as the Brownian increments do not have a preferred direction. Moreover, we observe that $\mathbb{E}[W(t)^2] = t$. Both properties can easily be proved in the time-discrete setting of Eq. (4.25) by using the basic rules of probability on the normal random variables ξ^k . A final property is that Brownian motion is nowhere differentiable with probability 1. This can be understood by realizing that

$$\text{Var} \left[\frac{W(t + \Delta t) - W(t)}{\Delta t} \right] = \frac{1}{\Delta t^2} \text{Var} [W(t + \Delta t) - W(t)] = \frac{1}{\Delta t}. \quad (4.26)$$

Note that, while $\mathbb{E}[W(t + \Delta t)] = \mathbb{E}[W(t)] = 0$, we have $\mathbb{E}[|W(t + \Delta t) - W(t)|] = \mathcal{O}(\Delta t^{1/2})$, which is proportional to the standard deviation.

4.3.2.2 Itô Stochastic Differential Equations

Often, the evolution of the position $X(t)$ of a particle is composed of a deterministic (mean) component, supplemented with a stochastic (fluctuating) part, modeled using a stochastic differential equation of the form

$$dX(t) = a(X(t))dt + b(X(t))dW(t), \quad X(0) = X_0, \quad (4.27)$$

in which a is called the *drift coefficient*, b is the diffusion coefficient, $W(t)$ is a Brownian motion and X_0 is the initial condition. For instance, one can consider the particle to model the position of a bacterium that is biasing its random motion to favor directions that are in line with the gradient of a chemoattractant. Then, the drift coefficient $a(X(t))$ models a preferred direction, whereas the second term takes into account the randomness of the motion. The diffusion coefficient $b(X(t))$ then defines, at the position $X(t)$, how strongly the evolution is affected by the Brownian motion $W(t)$.

Under mild assumptions on a and b , the stochastic differential equation (SDE) (4.27) has exactly one solution per Brownian path $W(t)$. To obtain this solution, one first needs to make sense of Eq. (4.27), something that will turn out to be nontrivial. Consider a classical solution,

$$X(t) = X_0 + \int_0^t a(X(s))ds + \int_0^t b(X(s))dW(s). \quad (4.28)$$

The first integral is a well-defined integral with respect to time; the second integral, however, is not well-defined and we need to be specific about its meaning.

Consider the integral to be defined via a Riemann sum using subintervals $[t^k, t^{k+1}]$, with $t^k = k\Delta t$ and $0 \leq k \leq K$, $K\Delta t = t$,

$$\int_0^t b(X(s))dW(s) = \lim_{\Delta \rightarrow 0} \sum_{k=1}^K b(X(s^k))\Delta W^k, \quad (4.29)$$

in which $\Delta W^k = W(t^{k+1}) - W(t^k)$ and $s^k \in [t^k, t^{k+1}]$. It is now easy to see, using only standard rules of probability, that the choice of s^k has a significant influence on the value of the integral. (This is due to the fact that $W(s)$ does not have bounded variations, which is related to (4.26).) The most common interpretation (the Itô interpretation) is obtained by choosing $s^k = t^k$, i.e., a left-point rule. In that case, we have

$$\begin{aligned} \mathbb{E} \left[\int_0^t b(X(s))dW(s) \right] &= \lim_{\Delta \rightarrow 0} \sum_{k=0}^{K-1} \mathbb{E} \left[b(X(t^k))\Delta W^k \right] \\ &= \lim_{\Delta \rightarrow 0} \sum_{k=0}^{K-1} \mathbb{E} \left[b(X(t^k)) \right] \mathbb{E} \left[\Delta W^k \right] = 0, \end{aligned}$$

where the first equality is due to the definition of the Itô integral and the linearity of the expectation, the second equality is due to independence of $b(X(t^k))$ and ΔW^k , and the last equality is due to the definition of Brownian motion.

A second popular interpretation (the Stratonovich interpretation) of (4.29) is obtained when choosing $s^k = (t^k + t^{k+1})/2$. It is clear that the above reasoning then can no longer be used, as $b(X(s^k))$ cannot be independent of ΔW^k . The resulting stochastic integral will then, in general, take a different value. We denote the Stratonovic integral as

$$\int_0^t b(X(s)) \circ dW(s)$$

and compute, as an example,

$$\mathbb{E} \left[\int_0^t W(s) \circ dW(s) \right] = \lim_{\Delta \rightarrow 0} \sum_{k=0}^{K-1} \mathbb{E} \left[W((t^k + t^{k+1})/2) \Delta W^k \right].$$

To compute this integral, we need to evaluate $W((t^k + t^{k+1})/2)$. It can be shown that this quantity is statistically equivalent to the quantity

$$\frac{W(t^k) + W(t^{k+1})}{2} + \Delta Z^k,$$

in which $\Delta Z^k \sim N(0, \Delta t/4)$ and independent of $W(t^k)$ and $W(t^{k+1})$. We can thus replace $W((t^k + t^{k+1})/2)$ by this alternative in computing the expectation, and we obtain

$$\mathbb{E} \left[\int_0^t W(s) \circ dW(s) \right] = \lim_{\Delta \rightarrow 0} \sum_{k=0}^{K-1} \left(\mathbb{E} \left[\left(\frac{W(t^k) + W(t^{k+1})}{2} \right) \Delta W_k \right] + \mathbb{E} \left[\Delta Z^k \Delta W^k \right] \right).$$

Since ΔZ^k and ΔW^k are independent, the second term is zero, and we continue with only the first term:

$$\begin{aligned} \mathbb{E} \left[\int_0^t W(s) \circ dW(s) \right] &= \lim_{\Delta \rightarrow 0} \sum_{k=0}^{K-1} \mathbb{E} \left[\left(\frac{W(t^k) + W(t^{k+1})}{2} \right) (W(t^{k+1}) - W(t^k)) \right] \\ &= \lim_{\Delta \rightarrow 0} \frac{1}{2} \sum_{k=0}^{K-1} \mathbb{E} \left[(W(t^{k+1}))^2 - (W(t^k))^2 \right] = \frac{1}{2} (W(t))^2, \end{aligned}$$

which is clearly different from the corresponding Itô integral (which is zero). For more details, we refer to [29] and references therein. In this chapter, we will always work with the Itô interpretation.

If $W(t)$ does not have bounded variations, then neither does $X(t)$. Consequently, also $X(t)$ will be nowhere differentiable (with probability 1). In particular, we have

$$\mathbb{E} [|X(t + \Delta t) - X(t)|] = \mathcal{O}(\Delta t^{1/2}).$$

Remark 4 (Further reading) We have deliberately kept the introduction on SDEs very brief. For more information, we refer the interested reader to the excellent books [19, 34].

Remark 5 (Reproduction and death) To model individuals that are also able to reproduce and die, one needs to add a (potentially stochastic) process that determines for each individual the moment at which it reproduces (and thus generates an additional individual) or dies (and is therefore removed from the system). To this end, one can, for instance, use the Markov processes that were discussed in Sect. 4.2 in the context of chemical reactions. Section 4.7 will contain additional modeling techniques that are of a more mechanistic nature.

4.3.2.3 Euler-Maruyama Method

Once an SDE model is obtained for a specific problem, a (numerical) solution needs to be computed. The most straightforward way to discretize an SDE of the type (4.27) is by using the stochastic extension of the forward Euler method, called Euler-Maruyama,

$$X^{k+1} = X^k + a(X^k)\Delta t + b(X^k)\Delta W^k, \quad (4.30)$$

in which Δt is the time step, and ΔW^k is sampled from a normal distribution with zero mean and variance Δt , i.e., $\Delta W^k \sim N(0, \Delta t)$. The scheme can equivalently be written as

$$X^{k+1} = X^k + a(X^k)\Delta t + b(X^k)\sqrt{\Delta t}\xi^k, \quad (4.31)$$

with $\xi^k \sim N(0, 1)$.

For a deterministic system, the convergence behaviour of the forward Euler method can be derived in a straightforward manner: given the numerical solution $X^K \approx X(t^K)$, the error $X^K - X(t^K)$ can be bounded as

$$\left| X^K - X(t^K) \right| \leq C \Delta t,$$

where K and Δt are varied simultaneously such that $K \Delta t = t^*$.

In the SDE case, this is no longer true. In fact, since both the numerical solution X^K and the exact solution $X(t^K)$ are random, only statistical statements can be made about the numerical error. One can immediately come up with two different definitions. We can define the *strong error* at time t^K as

$$e_{\Delta t}^K = \mathbb{E} \left[\left| X^K - X(t^K) \right| \right], \quad (4.32)$$

i.e., the expectation of the absolute value of the error on individual trajectories. Since it measures the “mean of the error”, averaged over all possible Brownian motions, the strong error gives an indication on the size of the error on an individual trajectory (defined by the Brownian path that generated it). Alternatively, one can define the *weak error* as

$$E_{\Delta t}^K = \left| \mathbb{E} \left[X^K \right] - \mathbb{E} \left[X(t^K) \right] \right|, \quad (4.33)$$

or, more generally, for an arbitrary function f in an appropriate function class,

$$E_{\Delta t}^K[f] = \left| \mathbb{E} \left[f(X^K) \right] - \mathbb{E} \left[f(X(t^K)) \right] \right|, \quad (4.34)$$

i.e., the error in the expectation of the function f when computed using time-discretized trajectories.

In general, these types of error are not the same, and also the order of convergence (as a function of Δt) differs. For the Euler-Maruyama method, we have

$$e_{\Delta t}^K \leq C \Delta t^{1/2}, \quad E_{\Delta t}^K \leq C \Delta t, \quad (4.35)$$

i.e., the Euler-Maruyama method has a strong order 1/2 and a weak order 1. Proving these orders would lead us too far. In this chapter, we simply illustrate this result numerically (see Sect. 4.3.2.4).

Remark 6 (Stability) The order of convergence of the Euler-Maruyama method only gives information on the asymptotic behavior of the error as Δt tends to zero. In practice, one will always take a finite time step. In that case, one needs the time step to be such that the Euler-Maruyama method is stable, i.e., loosely speaking, one needs to ensure that the numerical solution does not blow up for the chosen value of Δt when the exact solution remains bounded. While stability of time integration is a relatively straightforward concept for deterministic ODEs [27], this is no longer true for SDEs. As for convergence, multiple definitions of stability exist, see, e.g., [29, 32] for more details.

Remark 7 (Higher-order methods) Due to the presence of stochastic integrals, the definition of higher-order methods for SDEs is far more complicated than for ODEs. We refer to [32] for details.

4.3.2.4 Numerical Example: Geometric Brownian Motion

To illustrate the most important concepts in the previous sections, we perform some numerical experiments on a simple linear SDE, namely a geometric Brownian motion,

$$dX(t) = \lambda X(t)dt + \mu X(t)dW(t), \quad (4.36)$$

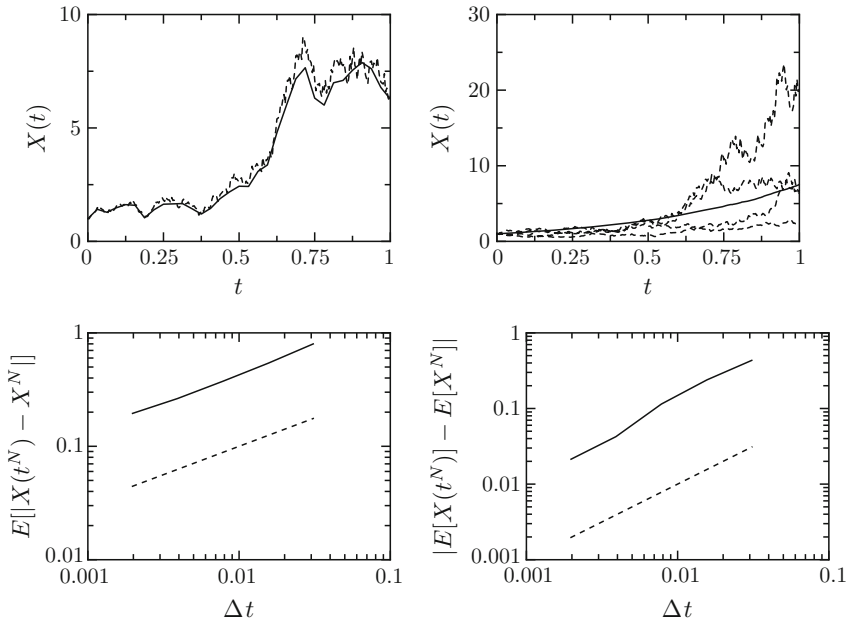


Fig. 4.3 Simulations of a geometric Brownian motion (4.36). *Top left* a single trajectory using the Euler-Maruyama method (4.31) with time step $\Delta t = 1/2^8$ (solid), compared to the exact solution for the same Brownian path (dashed); *Top right* four individual trajectories (dashed) and the empirical average of $M = 1000$ trajectories; *Bottom left* strong error of Euler-Maruyama as a function of Δt (solid), compared to the predicted strong order $1/2$ (dashed); *Bottom right* weak error of Euler-Maruyama as a function of Δt (solid), compared to the predicted weak order 1 (dashed). Remaining parameters are in the text

which we discretize with the Euler-Maruyama method (4.31) with time step Δt . For this SDE, an exact solution is known analytically (given the Brownian path $W(t)$), and given by

$$X(t) = X_0 \exp \left(\left(\lambda - \mu^2/2 \right) t + \mu W(t) \right). \tag{4.37}$$

In our experiments, we choose $\lambda = 2$, $\mu = 1$ and $X(0) = X_0 = 1$ with probability 1. The example is based on [29], in which also Matlab code can be found.

We first compare, for a single realization of $W(t)$, the exact solution $X(t)$ with the numerical solution $(X^k)_{k=0}^K$, with X^k obtained via Euler-Maruyama (4.31) with time step $\Delta t = 1/2^5$ on the time interval $t \in [0, 1]$ (hence $K \Delta t = 1$). The results are shown in Fig. 4.3, top left. We clearly see a discretization error.

To quantify this discretization error, we repeat the simulation for $M = 1000$ realizations of the Brownian path, $(W_m(t))_{m=1}^M$, resulting in M trajectories $(X_m(t))_{m=1}^M$, and compute an approximation to the strong and weak errors (4.32) and (4.33) as

$$\hat{e}_{\Delta t}^K = \hat{\mathbb{E}}_M \left[\left| X^K - X(t^K) \right| \right] = \frac{1}{M} \sum_{m=1}^M \left| X_m^K - X_m(t^K) \right|, \quad (4.38)$$

$$\hat{E}_{\Delta t}^K = \hat{\mathbb{E}}_M \left[X^K \right] - \hat{\mathbb{E}} \left[X(t^K) \right] = \frac{1}{M} \left| \sum_{m=1}^M \left(X_m^K - X_m(t^K) \right) \right|. \quad (4.39)$$

To achieve this, we first generate Brownian paths with time step $\Delta t = 1/2^9$, and subsequently use these Brownian paths to perform Euler-Maruyama simulations with time step $R\Delta t$, $R = 1, 2, 4, 8, 16$. The results are shown in Fig. 4.3, bottom. We clearly observe the predicted theoretical strong order 1/2 and weak order 1.

Remark 8 (Statistical error) The estimates (4.38) and (4.39) contain statistical error due to the finite number M of realizations. The problem and method parameters have been chosen such that this statistical error is negligible with respect to the time discretization error.

Finally, we look at the evolution of $F(t) := \mathbb{E}[X(t)]$ as a function of time. Being the expectation of the (time-dependent) random variable $X(t)$, $F(t)$ is a deterministic function of time. Figure 4.3 shows that $F(t)$ (unlike $X(t)$) is a smooth, differentiable function of time. This observation will be elaborated in the next section.

4.3.3 Population-Level Dynamics and Fokker-Planck Equation

Usually, one is not interested in the detailed stochastic behaviour of a single individual (cell, bacterium), but rather in the evolution of a large *population* of such cells. One then has given an initial *density* $\rho_0(x)$ of individuals as a function of space $x \in D \subset \mathbb{R}$, with D the domain. We interpret $\rho_0(x)$ as a probability density (this can always be done with a proper normalization), i.e., for an individual particle with position X_0 at time $t = 0$, we have

$$\Pr(x \leq X_0 < x + dx) = \rho_0(x)dx.$$

Note that lowercase x represents a possible position in the domain D , whereas uppercase X denotes the (random) position of an individual cell.

Each of the individuals behaves according to (4.27), generating a path $X(t)$. The question then becomes: defining the time-dependent density $\rho(x, t)$ as

$$\Pr(x \leq X(t) < x + dx) = \rho(x, t)dx,$$

can one obtain a corresponding evolution equation for $\rho(x, t)$? In this section, we will show that $\rho(x, t)$ satisfies a advection-diffusion partial differential equation (PDE), the Fokker-Planck equation. We perform the derivation only for a pure Brownian motion, after which we simply state the result for the general SDE (4.27). Our exposition closely follows [11].

In the pure diffusion case, we start from a (scaled) Brownian motion,

$$dX(t) = b dW(t), \quad (4.40)$$

with a constant scaling parameter $b > 0$.

Given the probability density $\rho(x, t)$ at time t , we can write the density $\rho(x, t + \Delta t)$ at time $t + \Delta t$ as

$$\rho(x, t + \Delta t) = \int_{-\infty}^{\infty} \rho(x + y, t) \cdot \psi(x, y, \Delta t) dy, \quad (4.41)$$

where the *transition probability kernel* $\psi(x, y, \Delta t)$ is the probability of ending up at position x at time $t + \Delta t$, given that one started at position $x + y$ at time t , i.e.,

$$\psi(x, y, \Delta t) = \Pr(X(t + \Delta t) \in [x, x + dx] | X(t) \in [x + y, x + y + dx]) \quad (4.42)$$

Eq. (4.41) states that the probability of finding a particle at position x at time $t + \Delta t$ is equal to the probability of finding the particle at position $x + y$ at time t , multiplied by the probability of moving from $x + y$ to x during the time step of size Δt , and this integrated over every possible value of y (i.e., integrated over every possible position $x + y$ at time t).

Remark 9 (Notation) It may seem odd to introduce the auxiliary variable y , instead of simply integrating over all possible original positions $z = x + y$. This is done because we intend to make a Taylor expansion of $\rho(x + y, t)$ around $\rho(x, t)$.

Let us now obtain an expression for $\psi(x, y, \Delta t)$. We know that $-y$ is the increment that was generated to move from $x + y$ to x . This increment is normally distributed with mean 0 and variance $b^2 \Delta t$, and therefore,

$$\psi(x, y, \Delta t) = \frac{1}{b\sqrt{2\pi\Delta t}} \exp\left(-\frac{y^2}{2b^2\Delta t}\right). \quad (4.43)$$

Now, we are ready to expand Eq. (4.41) by performing a Taylor expansion of $\rho(x + y, t)$ around $\rho(x, t)$,

$$\rho(x + y, t) = \rho(x, t) + y\partial_x\rho(x, t) + \frac{y^2}{2}\partial_{xx}\rho(x, t) + \text{h.o.t.} \quad (4.44)$$

in which h.o.t. stands for “higher order terms”. This leads to

$$\begin{aligned}
 \rho(x, t + \Delta t) &= \int_{-\infty}^{\infty} \rho(x, t) \psi(x, y, \Delta t) dy + \int_{-\infty}^{\infty} y \partial_x \rho(x, t) \psi(x, y, \Delta t) dy \\
 &\quad + \int_{-\infty}^{\infty} \frac{y^2}{2} \partial_{xx} \rho(x, t) \psi(x, y, \Delta t) dy + \text{h.o.t.} \\
 &= \rho(x, t) \underbrace{\int_{-\infty}^{\infty} \psi(x, y, \Delta t) dy}_{I_1} + \partial_x \rho(x, t) \underbrace{\int_{-\infty}^{\infty} y \psi(x, y, \Delta t) dy}_{I_2} \\
 &\quad + \frac{1}{2} \partial_{xx} \rho(x, t) \underbrace{\int_{-\infty}^{\infty} y^2 \psi(x, y, \Delta t) dy}_{I_3} + \text{h.o.t.} \tag{4.45}
 \end{aligned}$$

We will now separately look into each of the terms $I_{1,2,3}$. First, we have

$$I_1 = 1, \tag{4.46}$$

since $\psi(x, y, \Delta t)$ is independent of x and a probability density for y . Next, we have

$$I_2 = 0, \tag{4.47}$$

since y is an odd function and $\psi(x, y, \Delta t)$ is even. Finally, we obtain

$$I_3 = \text{Var} [y] = b^2 \Delta t, \tag{4.48}$$

in which $\text{Var} [y]$ is to be interpreted as the variance of y with respect to the probability density ψ . This leads to

$$\rho(x, t + \Delta t) = \rho(x, t) + \frac{b^2}{2} \Delta t \partial_{xx} \rho(x, t) + \text{h.o.t.}, \tag{4.49}$$

which is a time-discretized version of the diffusion equation. Taking the limit of $\Delta t \rightarrow 0$, we obtain the desired result:

$$\partial_t \rho(x, t) = \frac{b^2}{2} \partial_{xx} \rho(x, t). \tag{4.50}$$

Thus, the density evolves according to the diffusion equation. This is to be expected, as the particles have no preferred direction and will therefore spread evenly. This derivation explains why the Brownian motion is also called a diffusion process.

In the more general case of the SDE (4.27), the drift term will introduce a systematic bias in the motion of individual particles, resulting in an advective behaviour of the density $\rho(x, t)$. It can be shown that, in that case, the density $\rho(x, t)$ satisfies the

advection-diffusion equation

$$\partial_t \rho(x, t) + \partial_x (a(x)\rho(x, t)) = \frac{1}{2} \partial_{xx} \left((b(x))^2 \rho(x, t) \right). \quad (4.51)$$

We refer to [19, 34] for details.

4.4 More Realistic Microscopic Processes

Whereas stochastic differential equations are useful to describe a wide range of stochastic processes in biological applications, they remain in essence *phenomenological*, and thus descriptive. Even if these models have predictive power, it is often impossible to use such models for a detailed understanding of the *mechanism* that generate the dynamics. In this section, we therefore discuss more mechanistic velocity-jump processes and their relation to bacterial chemotaxis. The section follows the same structure as that of the previous sections: we first discuss the stochastic individual-based model (Sect. 4.4.1), after which we continue with an equivalent continuum description (Sect. 4.4.2). In Sect. 4.4.3, we relate the resulting stochastic processes with the SDEs of Sect. 4.3.2, and we conclude with some bibliographical remarks on generalizations in Sect. 4.4.4.

4.4.1 Velocity-Jump Processes for Bacterial Chemotaxis

Generally, the motion of flagellated bacteria consists of a sequence of *run* phases, during which a bacterium moves in a straight line at constant speed. The bacterium changes direction in a *tumble* phase, which is typically much shorter than the run phase and acts as a reorientation. Hence, the motion of an individual bacterium can be modeled as a velocity-jump process. To bias movement towards regions with high concentration of chemoattractant, the bacterium adjusts its turning rate to increase, resp. decrease, the chance of tumbling when moving in an unfavorable, resp. favorable, direction [2, 50]. The velocity-jump models described here are based on [15] and [44, 45].

We consider bacteria that are sensitive to the concentration of a chemoattractant $S(x) \geq 0$ for $x \in \mathbb{R}^d$, where x is the present position of the bacterium. While we do not consider time dependence of chemoattractant via production or consumption by the bacteria, a generalization to this situation is straightforward, at least for the definition of the models and the numerical method. Bacteria move with a constant speed v (run), and change direction at random instances in time (tumble), in an attempt to move towards regions with high chemoattractant concentrations.

The position of an individual bacterium is given by $X(t)$, its velocity is

$$\frac{dX_c(t)}{dt} = \epsilon V_c(t), \quad V_c(t) \in \mathbb{V} = \mathbb{S}^{d-1},$$

with \mathbb{S}^{d-1} the unit sphere in \mathbb{R}^d . Hence, $V_c(t)$ represents the direction and the scaling parameter $\epsilon > 0$ represents the size of the velocity. (The reason for the introduction of the subscript c will become clear in Sect. 4.6.) The velocity of each bacterium is switched at random jump times $(T_c^k)_{k \geq 1}$ that are generated via a Poisson process with a time dependent turning rate $\lambda_c^\epsilon(x, v)$ that depends on the bacterium's current position and velocity. The new velocity at time T_c^k is generated at random according to a centered probability distribution $\mathcal{M}(dv)$ with $\int v \mathcal{M}(dv) = 0$, typically

$$\mathcal{M}(dv) = \sigma_{\mathbb{S}^{d-1}}(dv),$$

where $\sigma_{\mathbb{S}^{d-1}}$ is the uniform distribution on the unit sphere.

The turning rate is assumed to satisfy

$$0 < \lambda_{\min} \leq \lambda_c^\epsilon(x, v) \leq \lambda_{\max}, \tag{4.52}$$

as well as, for small values of ϵ ,

$$\lambda_c^\epsilon(x, v) := \lambda^0 - \epsilon A_\epsilon^T(x)v + O(\epsilon^2). \tag{4.53}$$

Typically, $\lambda_c^\epsilon(x, v)$ is a function of $\nabla S(x)$, so that the model (4.53) may describe a large bacterium that is able to directly sense chemoattractant gradients. When the turning rate (4.53) is proportional to $\nabla S(x)v$, it can be interpreted as follows: the rate at which a bacterium will change its velocity direction depends on the alignment of the velocity with the gradient of the chemoattractant concentration $\nabla S(x)$, resulting in a transport towards areas with higher chemoattractant concentrations.

The resulting stochastic process can be written as

$$\begin{cases} \frac{dX_c(t)}{dt} = \epsilon V_c(t), \\ \int_{T_c^k}^{T_c^{k+1}} \lambda_c^\epsilon(X_c(t), V_c(t)) dt = \theta^{k+1}, \\ V_c(t) = \mathcal{V}^k \quad \text{for } t \in [T_c^k, T_c^{k+1}], \end{cases} \tag{4.54}$$

with initial condition $X(0), \mathcal{V}(0) \in \mathbb{R}^d$. In (4.54), $(\theta^k)_{k \geq 1}$ denote i.i.d. random variables with normalized exponential distribution, and $(\mathcal{V}^k)_{k \geq 1}$ denote i.i.d. random variables with distribution $\mathcal{M}(dv)$.

4.4.2 Population-Level Dynamics and Kinetic Equations

At the population level, there are two main differences with respect to the Fokker-Planck Eq. (4.51) for SDEs. First, the stochastic process is written in terms of positions x and velocities v , implying that the density of interest will be a density $p_c(x, v, t)$ in position-velocity phase space. Second, the stochastic process is discontinuous in the velocity component, which will result in a “collision operator” that models the discontinuous velocity changes probabilistically.

The resulting evolution equation for the density $p_c(x, v, t)$ turns out to be a Boltzmann equation with BGK-type collision operator,

$$\partial_t p_c + \epsilon v \cdot \nabla_x p_c = (R(\lambda_c^\epsilon p) - \lambda_c^\epsilon p_c), \quad (4.55)$$

where

$$R(p_c) := \int_{\mathbb{V}} p_c(\cdot, v, \cdot) \mathcal{M}(dv)$$

is the operator integrating velocities with respect to \mathcal{M} , and λ_c^ϵ is defined as in (4.53). We will not derive this equation here, but instead refer the interested reader to [18] for the derivation of master equations associated to Markov jump processes. Here, we suffice by pointing out that the advection term models the effect of the velocity on positions, and the righthand side models the effect of the random velocity changes.

4.4.3 Coarse-Graining and Approximate Macroscopic Descriptions

The explicit modeling of these individual velocity changes is necessary to have a biologically relevant mechanistic description of bacterial motion. In general, however, one is not really interested in the detailed phase-space distribution $p_c(x, v, t)$, but rather in the the position density $\rho_c(x, t) = R(p_c(x, v, t))$, and this for (at least) two reasons: (i) it is usually impossible to obtain experimental data on the velocity distribution of the bacteria; and (ii) the bacteria typically travel only a microscopic distance between velocity changes, such that the observed macroscopic motion is the averaged effect on long time scales of a large number of velocity changes. One can expect the position density $\rho_c(x, t)$ to satisfy a partial differential equation advection-diffusion type, such as (4.51). The velocity-jump process (4.54) and the advection-diffusion SDE (4.27) are therefore related.

To consider the behaviour of Eq. (4.55) on long time scales and for small bacterial velocities, we let ϵ tend to 0 and introduce the rescaled time $\bar{t} = t\epsilon^2$. (Then, when \bar{t} is $O(1)$, this corresponds to a physical time t that is $O(1/\epsilon^2)$.) In that case, the position density $\rho_c(x, t)$ satisfies the advection-diffusion PDE

$$\partial_{\bar{t}}\rho_c = \frac{1}{\lambda_0} \operatorname{div}_x (D\nabla_x \rho_c - DA_0(x)\rho_c), \quad (4.56)$$

in which the diffusion matrix is given by the covariance of the Maxwellian distribution,

$$D = \int_{\mathbb{S}^{d-1}} v \otimes v \mathcal{M}(dv) \in \mathbb{R}^{d \times d}. \quad (4.57)$$

We refer to [35] for justifications of this result based on Hilbert expansions. The result implies that the position density obtained via (4.56) and via (4.55) are the same in the limit when ϵ tends to 0.

Additionally, it is shown in [45] that the position $X^c(\bar{t})$ of an individual trajectory generated by the stochastic process (4.54) converges to a trajectory of the SDE

$$dX_c(\bar{t}) = \frac{DA_0(X_c(\bar{t}))}{\lambda^0} d\bar{t} + \left(\frac{2D}{\lambda_0}\right)^{1/2} dW(\bar{t}), \quad (4.58)$$

where $\bar{t} \mapsto W_{\bar{t}}$ is a standard Brownian motion, as ϵ tends to zero. (Note that this second result implies the convergence of the position densities to a solution of (4.56), but not vice versa.)

4.4.4 Further Comments and Remarks

The main modeling limitation of the models discussed so far is that they deal with *non-interacting* particles, i.e., every individual follows its own stochastic path, independently of all other individuals present. Many generalizations exist to introduce interactions between particles, either as two-particle collision operators, via long-range interactions or via interactions of individual particles with the position density. Giving an overview of all these generalizations would lead too far. We refer to the two excellent books [49] and [40] and references therein.

4.5 Monte Carlo Simulation and Variance Reduction

4.5.1 The Need for Monte Carlo Simulation

In most situations, we are interested in the evolution of a large population of individuals (cells, bacteria). To simulate this evolution, two courses of action are possible:

- a (stochastic) Monte Carlo simulation of an ensemble of M realizations of the stochastic process (such as (4.27) or (4.54)), from which information on the population density can be obtained using histograms or kernel density estimation [47, 48];

- A deterministic (grid-based) simulation of the corresponding PDE (such as (4.51) or (4.55)).

Both options have advantages and drawbacks. The clear drawback of a Monte Carlo simulation is the appearance of statistical error on the obtained population density, which is absent in a deterministic simulation—the variance on the obtained result being of the order of $O(1/\sqrt{M})$ [7]. The drawback of a grid-based simulation is that the computational cost of mesh refinement depends crucially on the number of dimensions of the PDE. Considering a system of particles in 3 spatial dimensions, the kinetic equation (4.55) is a 6-dimensional PDE. Doubling the number of mesh points in each spatial dimension already increases the total number of unknowns by a factor of 2^6 , even if one can still take the same time step. In contrast, the computational cost of refining a Monte Carlo simulation is independent of the dimension of the problem: one can simply augment the number M of simulated particles.

In more realistic applications, the computational complexity of the PDE-based description can be even higher, due to several reasons. When particles also have internal state, the dimension of the kinetic equation (4.55) increases even further (see Sect. 4.6). Moreover, if the particles are interacting, the collision operator becomes non-local, requiring the evaluation of an integral over velocity space at each spatial mesh point. The situation becomes even more difficult when particles experience long-range interactions.

4.5.2 Variance Reduction Techniques

Because of the problems associated with simulating high-dimensional PDEs, Monte Carlo simulation is a viable alternative, provided one can control the variance of the simulation. As a consequence, there exists a large literature on variance reduction techniques. The most popular techniques can roughly be categorized in two classes: importance sampling and the use of a control variate. These techniques are well-established in the computation of integrals with respect to a known probability distribution, see [7] and references therein.

In importance sampling, the key idea is to adaptively sample the density of interest using *weighted particles*, such that more particles are placed (with correspondingly lower weights) in regions in which the variance is expected to be higher. Goal is to obtain a variance that is evenly distributed over the computational domain. With a control variate, the key idea is to compute an approximation to the quantity of interest *deterministically* based on the solution of a related but simpler problem, for instance analytically or by numerically solving a PDE of lower dimension. One then uses the Monte Carlo only to sample the *correction* with respect to the deterministically computed quantity.

Several research groups are currently working along these lines to develop hybrid Monte Carlo/PDE methods. We refer to [13, 14] and related papers and to [1, 42] and related papers in the context of the Boltzmann equation, and to [4] for an example in

the context of radiation transport. We have developed a strategy based on a control variate [44] that will be detailed for bacterial chemotaxis in Sect. 4.6. A related method that we are currently developing for tumor growth will be discussed in Sect. 4.7.

4.6 Application 1: Bacterial Chemotaxis

As a first application, we return to bacterial chemotaxis, which was also used in Sect. 4.4.1. Since many species are unable to sense chemoattractant gradients reliably due to their small size, adjustment of their turning rate to bias motion in favorable directions is often done via an intracellular mechanism that allows the bacterium to retain information on the history of the chemoattractant concentrations along its path [6]. The resulting model, which will be called the “internal state” or “fine-scale” model in this text, can be formulated as a velocity-jump process, combined with an ordinary differential equation (ODE) that describes the evolution of an internal state that incorporates this memory effect [16]. The probability density distribution of the velocity-jump process evolves according to a kinetic equation, in which the internal variables appear as additional dimensions. A direct deterministic simulation of this equation is therefore prohibitively expensive, and one needs to resort to a stochastic particle method.

Unfortunately, a direct fine-scale simulation using stochastic particles presents a large statistical variance, even in the diffusive asymptotic regime when ϵ is small. In that regime, the bacterial density of the fine-scale model is known explicitly to satisfy a Keller-Segel advection-diffusion equation. Consequently, it is difficult to assess accurately how the solutions of the fine-scale model differ from their advection-diffusion limit in intermediate regimes.

In this section, we discuss a numerical method to simulate individual-based models for chemotaxis of bacteria with internal dynamics with reduced variance, introduced in [44]. The variance reduction is based on a coupling technique (control variate): the main idea is to simultaneously simulate, using the same random numbers, a simpler, “coarse” process where the internal dynamics is replaced by a direct “gradient sensing” mechanism (see [2, 38, 41] for references on such gradient sensing models). The probability density of the latter satisfies a kinetic equation without the additional dimensions of the internal state, and converges to a similar advection-diffusion limit as the model with internal state, see e.g. [10, 35, 37, 45]. The precise coarse model will be (4.54) with a suitable choice for $A_\epsilon(x)$ in (4.53) (see later), such that the coarse and fine-scale model have *exactly* the same advection-diffusion limit.

We first discuss the fine-scale model with internal dynamics in Sect. 4.6.1. The model is a simplification (for expository purposes) of the more general model in [44, 45]. We describe the variance reduction technique in detail in Sect. 4.6.3. Some numerical results are given in Sect. 4.6.4. A detailed analysis of the method can be found in [44].

4.6.1 Bacterial Chemotaxis with Internal State

We again consider bacteria that are sensitive to the concentration of a chemoattractant $S(x) \geq 0$ for $x \in \mathbb{R}^d$. As in Sect. 4.4.1, the bacteria follow a velocity-jump process in which $X(t)$ is the position of an individual bacterium, and the normalized velocity is given by

$$\frac{dX(t)}{dt} = \epsilon V(t), \quad V(t) \in \mathbb{V} = \mathbb{S}^{d-1},$$

with \mathbb{S}^{d-1} the unit sphere in \mathbb{R}^d . Hence, $V(t)$ represents the direction and the parameter ϵ represents the size of the velocity. The difference with respect to the process (4.54) with direct gradient sensing is in the definition of the turning rate. As in [15], the turning rate is made to depend upon an internal state $y \in \mathbb{Y} \subset \mathbb{R}$ of each individual bacterium, which models the memory of the bacterium and is subject to an evolution mechanism attracted by the chemoattractant concentration $S(x)$. (The model in [44, 45] is more general and can take into account multiple chemoattractants and higher-dimensional internal states.)

The internal state adapts to the local chemoattractant concentration through an ODE,

$$\frac{dY(t)}{dt} = F_\epsilon(Y(t), S(X(t))), \quad (4.59)$$

which is required to have a unique fixed point $y^* = S(x^*)$ for every fixed value $x^* \in \mathbb{R}^d$. We also introduce the deviations from equilibrium $Z(t) = S(X(t)) - Y(t)$.

The velocity of each bacterium is switched at random jump times $(T^k)_{k \geq 1}$ that are generated via a Poisson process with a time dependent rate given by $\lambda(Z(t))$, where $z \mapsto \lambda(z)$ is a smooth function satisfying

$$0 < \lambda_{\min} \leq \lambda(z) \leq \lambda_{\max}, \quad (4.60)$$

as well as (for small values of z),

$$\lambda(z) = \lambda^0 - bz + c_\lambda \mathcal{O}(|z|^\gamma), \quad (4.61)$$

with $b \in \mathbb{R}$, $\gamma \geq 2$. As before, the new velocity at time T^k is generated at random according to a centered probability distribution $\mathcal{M}(dv)$ with $\int v \mathcal{M}(dv) = 0$, typically

$$\mathcal{M}(dv) = \sigma_{\mathbb{S}^{d-1}}(dv),$$

where $\sigma_{\mathbb{S}^{d-1}}$ is the uniform distribution on the unit sphere.

The resulting fine-scale stochastic evolution of a bacterium is then described by the following differential velocity-jump equation,

$$\begin{cases} \frac{dX(t)}{dt} = \epsilon V(t), \\ \frac{dY(t)}{dt} = F_\epsilon(Y(t), S(X(t))), \\ \int_{T^k}^{T^{k+1}} \lambda(Z(t)) dt = \theta^{k+1}, \quad \text{with } Z(t) := S(X(t)) - Y(t), \\ V(t) = \mathcal{V}^k \quad \text{for } t \in [T^k, T^{k+1}), \end{cases} \quad (4.62)$$

with initial condition $X(0) \in \mathbb{R}^d$, $Y(0) \in \mathbb{R}$ and $T^0 = 0$. In (4.62), $(\theta^k)_{k \geq 1}$ denote i.i.d. random variables with normalized exponential distribution, and $(\mathcal{V}^k)_{k \geq 1}$ denote i.i.d. random variables with distribution $\mathcal{M}(dv)$.

In the numerical experiments, we will use a specific example, adapted from [15]. For the internal dynamics (4.59), we choose a linear equation

$$\frac{dy}{dt} = \frac{S(x) - y}{\tau} = \frac{z}{\tau}. \quad (4.63)$$

For the turning rate $z \mapsto \lambda(z)$, we choose the following nonlinear strictly decreasing smooth function

$$\lambda(z) = 2\lambda_0 \left(\frac{1}{2} - \frac{1}{\pi} \arctan \left(\frac{\pi}{2\lambda_0} z \right) \right). \quad (4.64)$$

The probability distribution density of the fine-scale process with internal state at time t with respect to the measure $dx \mathcal{M}(dv) dy$ is denoted as $p(x, v, y, t)$, suppressing the dependence on ϵ for notational convenience, and evolves according to the Kolmogorov forward evolution equation (or master equation). In the present context, the latter is the following kinetic equation

$$\partial_t p + \epsilon v \cdot \nabla_x p + \operatorname{div}_v (F_\epsilon(x, y) p) = \lambda(S(x) - y) (R(p) - p), \quad (4.65)$$

where

$$R(p) := \int_{\mathbb{V}} p(\cdot, v, \cdot) \mathcal{M}(dv)$$

is again the operator integrating velocities with respect to \mathcal{M} .

4.6.2 Relation Between Fine-Scale and Coarse Process

In [45], it is shown, using probabilistic arguments, that, in the limit $\epsilon \rightarrow 0$, both the equation for the coarse process (4.55) and the equation for the process with internal state (4.65) converge to an advection-diffusion limit on diffusive time scales. Convergence is to be understood *pathwise*, i.e., in the sense of individual trajectories.

For the coarse process, this result has already been stated, see Sect. 4.4.3.

In the same way, a standard probabilistic diffusion approximation argument can be used to derive the pathwise diffusive limit of the process with internal state (4.62), see [45]. For $\epsilon \rightarrow 0$, the process $\bar{t} \mapsto X^\epsilon(\bar{t})$, solution of (4.62), converges towards an advection-diffusion process, satisfying the stochastic differential equation (SDE) (4.58), where A_0 originates from

$$A_0(x) = b \lim_{\epsilon \rightarrow 0} \frac{\tau}{\lambda^0 \tau + 1} \nabla S(x), \quad (4.66)$$

in which b , τ , and λ^0 were introduced in (4.61)–(4.63) as parameters of the process with internal state. Again, the diffusion matrix D is given by the covariance of the Maxwellian distribution (4.57).

Introducing the bacterial density of the process with internal state as

$$\rho(x, t) = \int_{\mathbb{Y}} \int_{\mathbb{V}} p(x, v, y, t) \mathcal{M}(dv) dy, \quad (4.67)$$

this implies that the evolution of ρ converges to (4.56) on diffusive time scales in the limit of $\epsilon \rightarrow 0$.

4.6.3 Asymptotic Variance Reduction

As discussed in Sect. 4.5, obtaining the position density of bacteria by solving the kinetic equation (4.65) over diffusive time scales can be cumbersome, due to the additional dimensions associated with the internal state. The alternative is to use stochastic particles. However, a particle-based simulation of Eq. (4.65) is subject to a large statistical variance of the order $\mathcal{O}(M^{-1/2})$, where M is the number of simulated particles. Additionally, the asymptotic analysis shows that the position bacterial density approaches an advection-diffusion limit (4.56) when $\epsilon \rightarrow 0$; moreover, this advection-diffusion limit is shared with a simpler, coarse model without internal dynamics. Consequently, to accurately assess the deviations of the process with internal state (4.62) as compared to its advection-diffusion limit (for small but non-vanishing (intermediate) values of ϵ), the required number of particles needs to increase substantially with decreasing ϵ , which may become prohibitive from a computational point of view.

The idea that will be discussed here is to construct a hybrid method, based on the principle of control variates, that couples the process with internal dynamics with the coarse process, which is simulated simultaneously using a grid-based method. The stochastic particles are then only used to perform a Monte Carlo simulation of the deviations of the model with internal state with respect to the coarse model.

To explain the method, let us first assume that we are able to compute the exact solution of the kinetic equation for the coarse process, (4.55), with infinite precision in space and time. From now on, we will refer to (4.55) as the *control process* as it is used as a control variate. (This explains the addition of the subscript c .)

The algorithm of asymptotic variance reduction is based on a coupling between an ensemble of realizations evolving according to the process with internal state (4.62), denoted as

$$\{X_m(t), V_m(t), Y_m(t)\}_{m=1}^M,$$

and an ensemble of realization of the control process (4.54), denoted as

$$\{X_{m,c}(t), V_{m,c}(t)\}_{m=1}^M.$$

We denote the empirical measure of the particles with internal state in position-velocity space as

$$\mu_{\bar{t}}^M(x, v) = \frac{1}{M} \sum_{m=1}^M \delta_{X_m(\bar{t}/\epsilon^2), V_m(\bar{t}/\epsilon^2)},$$

and, correspondingly, the empirical measure of the control particles as

$$\mu_{\bar{t}}^{M,c} = \frac{1}{M} \sum_{m=1}^M \delta_{X_{m,c}(\bar{t}/\epsilon^2), V_{m,c}(\bar{t}/\epsilon^2)},$$

with $\delta_{x,v}$ the Dirac delta centered at (x, v) . (These empirical measures can be interpreted as a discrete particle approximation to the phase-space density of the bacteria.)

A coupling between the two ensembles is obtained by ensuring that both simulations use *the same random numbers* $(\theta^k)_{k \geq 1}$ and $(\mathcal{V}^k)_{k \geq 0}$, which results in a strong correlation between $(X_m(t), V_m(t))$ and $(X_{m,c}(t), V_{m,c}(t))$ for each realization. Simultaneously, the kinetic equation for the control process (4.55) is also solved using a deterministic method (which, for now, is assumed to be exact). We formally denote the corresponding semi-group evolution (a formal description of the exact solution) as

$$e^{\bar{t}L_c}, \quad \text{with } L_c(p_c) = -\epsilon v \cdot \nabla_x p_c + (R(\lambda_c^\epsilon p_c) - \lambda_c^\epsilon p_c).$$

Besides the two particle measures $\mu_{\bar{t}}^M$ and $\mu_{\bar{t}}^{M,c}$, we denote by $\bar{\mu}_{\bar{t}}^M$ the variance reduced measure, which will be defined by the algorithm below. Since, with increasing diffusive time, the variance of the algorithm increases due to a loss of coupling between the particles with internal state and the control particles, the variance reduced algorithm will also make use of a reinitialization time step $\bar{\delta}t_{ri}$, which is defined on the diffusive time scale. The corresponding time instances are denoted as $\bar{t}^\ell = \ell \bar{\delta}t_{ri}$ on the diffusive time scale, or equivalently, on the original time scale as $t^\ell = \ell \bar{\delta}t_{ri} / \epsilon^2$.

Starting from an initial probability measure μ_0 at time $t = 0$, we sample μ_0 to obtain the ensemble $\{X_m(t), V_m(t), Y_m(t)\}_{m=1}^M$, corresponding to μ_0^M , and then set $\mu_0^{M,c} := \mu_0^M$, i.e., $X_{m,c}(0) = X_m(0)$ and $V_{m,c}(0) = V_m(0)$ for all $m = 1, \dots, M$. Furthermore, we set the variance reduced estimator as $\bar{\mu}_0^M := \mu_0 = \mathbb{E}(\mu_0^M)$. We then use the following algorithm to advance from \bar{t}^ℓ to $\bar{t}^{\ell+1}$, (see also Fig. 4.4):

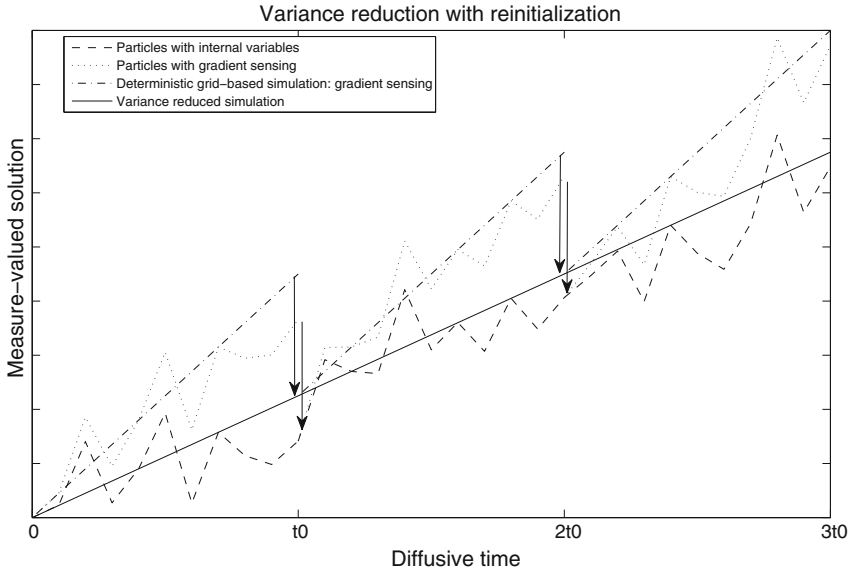


Fig. 4.4 A schematic description of Algorithm 3. The *dashed line* represent the evolution of M bacteria with internal state. The *dotted line* represent the coupled evolution of M bacteria with gradient sensing, subject to regular reinitializations. The *dashed-dotted line* is computed according to a deterministic method simulating the density of the model with gradient sensing, and subject to regular reinitializations. The *solid line* is the variance reduced simulation of the internal state dynamics, and is computed by adding the difference between the particle computation with internal state, and the particle simulation with gradient sensing to the deterministic gradient sensing simulation. At each reinitialization step, the two simulations (deterministic and particles) of the gradient sensing dynamics are reinitialized to the values of their internal state simulation counterpart (as represented by the *arrows*)

Algorithm 3 At time t^ℓ , we have that the particle measure $\mu_{\tilde{t}^\ell}^{M,c} = \mu_{\tilde{t}^\ell}^M$, and the variance reduced measure is given by $\bar{\mu}_{\tilde{t}^\ell}^M$. To advance from time \tilde{t}^ℓ to $\tilde{t}^{\ell+1}$, we perform the following steps :

- Evolve the particles $\{X_m(t), V_m(t), Y_m(t)\}_{m=1}^M$ from t^ℓ to $t^{\ell+1}$, according to (4.62);
- Evolve the particles $\{X_{m,c}(t), V_{m,c}(t)\}_{m=1}^M$ according to (4.54), using the same random numbers as for the process with internal state;
- Compute the variance reduced evolution

$$\bar{\mu}_{\tilde{t}^{\ell+1}}^M = \bar{\mu}_{\tilde{t}^\ell}^M e^{\delta \tilde{t}_{ri} / \epsilon^2 L_c} + \left(\mu_{\tilde{t}^{\ell+1}}^M - \mu_{\tilde{t}^{\ell+1}}^{M,c} \right). \tag{4.68}$$

(Note that this implies that we start the deterministic simulation for this time step from $\bar{\mu}_{\tilde{t}^\ell}^M$.)

- Reinitialize the control particles by setting

$$X_{m,c}(t^{\ell+1}) = X_m(t^{\ell+1}), \quad V_{m,c}(t^{\ell+1}) = V_m(t^{\ell+1}), \quad m = 1, \dots, M,$$

i.e., we set the state of the control particles to be identical to the state of the particles with internal state.

In (4.68), we use the symbol $\bar{t}^{\ell+1}$ to emphasize that the involved particle positions and velocities are those obtained *before* the reinitialization. An easy computation shows that the algorithm is unbiased in the sense that for any $\ell \geq 0$,

$$\mathbb{E} \left[\bar{\mu}_{\bar{t}^\ell}^M \right] = \mathbb{E} \left[\mu_{\bar{t}^\ell}^M \right],$$

since the particles with internal dynamics are unaffected by the reinitialization, and, additionally,

$$\mathbb{E} \left[\mu_{\bar{t}^{\ell+1}}^{M,c} \right] = \mathbb{E} \left[\bar{\mu}_{\bar{t}^\ell}^M e^{\bar{\delta} t_{ri} / \epsilon^2 L_c} \right].$$

Moreover, the variance is controlled by the coupling between the two processes. Indeed, using the independence of the random numbers between two steps of Algorithm 3, and introducing φ as a position and velocity dependent test function, we get (see [44]),

$$\text{stdev}(\bar{\mu}_{\bar{t}^\ell}^M(\varphi)) \leq C \frac{\epsilon}{M}, \quad (4.69)$$

where in the last line, C is independent of ℓ , ϵ , and M .

Remark 10 (Sharpness of the variance estimate) In some generic situations, we can argue that the statistical error in Algorithm 3 coming from the coupling is “sharp” with respect to the order in ϵ . This means that the difference between the probability distribution of the model with internal state and the probability distribution of the model with gradient sensing is of the same order. This would imply that, with the asymptotic variance reduction technique, one is able to reliably assess the true deviation of the process with internal variables from the control process using a number of particles M that is independent of ϵ .

Remark 11 (Effect of time discretization) The analysis in [44] reveals that the variance reduction is *asymptotic*, in the sense that the variance vanishes in the diffusion limit. To ensure this asymptotic variance reduction during actual simulations, one needs to ensure that the time discretization preserves the diffusion limits of the time-continuous process. An appropriate time discretization is highly non-trivial, and is discussed in [44].

4.6.4 Numerical Results

To illustrate the algorithm, we consider a simulation of the density of an ensemble of particles, with and without variance reduction. We restrict ourselves to one space

dimension, with domain $x \in [0, 20]$ and periodic boundary conditions. In this case, the kinetic equation corresponding to the control process reduces to the system

$$\begin{cases} \partial_t p_c^+ + \epsilon \partial_x p_c^+ = -\frac{\lambda_c(x, +1)}{2} p_c^+ + \frac{\lambda_c(x, -1)}{2} p_c^- \\ \partial_t p_c^- - \epsilon \partial_x p_c^- = \frac{\lambda_c(x, +1)}{2} p_c^+ - \frac{\lambda_c(x, -1)}{2} p_c^- \end{cases}. \quad (4.70)$$

of two PDEs, which is straightforward to simulate using finite differences.

We fix the chemoattractant concentration field as

$$S(x) = \alpha \left(\exp \left(-\beta (x - \xi)^2 \right) + \exp \left(-\beta (x - \eta)^2 \right) \right), \quad (4.71)$$

with parameters $\alpha = 2$, $\beta = 1$, $\xi = 7.5$ and $\eta = 12.5$. For the internal dynamics, the model ((4.63–4.64), (4.53–4.66)) is used. The parameters are $\epsilon = 0.5$, $\lambda_0 = 1$, $\tau = 1$, $\delta t = 0.1$.

All simulations are performed with $M = 5000$ particles. The initial positions are uniformly distributed in the interval $x \in [13, 15]$; the initial velocities are chosen uniformly, i.e., each particle has an equal probability of having an initial velocity of $\pm\epsilon$. The initial condition for the internal variable is chosen to be in local equilibrium, i.e., $Y_m(0) = S(X_m(0))$. The initial positions and velocities of the control particles are chosen to be identical.

We discretize the continuum description (4.70) on a mesh with $\Delta x = 0.1$ using a third-order upwind-biased scheme, and perform time integration using the standard fourth order Runge–Kutta method with time step $\delta t_{pde} = 10^{-1}$. The initial position density is given as

$$\rho(x, 0) = \begin{cases} 0.25, & x \in [13, 15], \\ 0, & \text{otherwise.} \end{cases} \quad (4.72)$$

Simulation without variance reduction First, we simulate both stochastic processes up to time $\bar{t} = 50$ ($t = 50/\epsilon^2$) and estimate the density of each of these processes $\hat{\rho}_c^M(x, \bar{t})$, resp. $\hat{\rho}_c^M(x, \bar{t})$, without variance reduction. The density is obtained via binning in a histogram, in which the grid points of the deterministic simulation are the centers of the bins. Figure 4.5 (left) shows the results for a single realization. We see that, given the fluctuations on the obtained density, it is impossible to conclude on differences between the two models. This observation is confirmed by computing the average density of both processes over 100 realizations. The mean densities are shown in Fig. 4.5 (right), which also reveals that the mean density of the control process is within the 95% confidence interval of the process with internal state. Both figures also show the density that is computed using the continuum description, which coincides with the mean of the density of the control particles.

Simulation with variance reduction Next, we compare the variance reduced estimation (4.68) with the density of the control PDE. We reinitialize the control par-

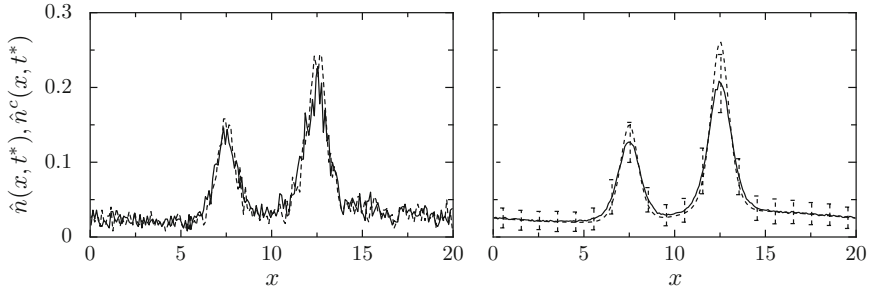


Fig. 4.5 Bacterial density as a function of space at $t = 50/\epsilon^2$ without variance reduction. *Left* one realization. *Right* mean over 100 realizations and 95% confidence interval. The *solid line* is the estimated density from a particle simulation using the process with internal state; the *dashed line* is estimated from a particle simulation using the control process. Both used $M = 5000$ particles. The *dotted line* is the density obtained from the deterministic PDE (4.70)

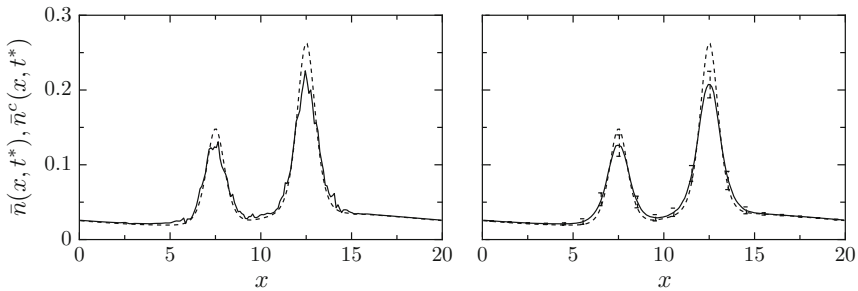


Fig. 4.6 Bacterial density as a function of space at $t = 50/\epsilon^2$ with variance reduction and reinitialization. *Left* variance reduced density estimation of one realization with $M = 5000$ particles (*solid*) and density obtained from a deterministic solution for the control process (4.70) (*dashed*). *Right* mean over 100 realization and 95% confidence interval (*solid*) and density obtained from a deterministic solution for the control process (4.70) (*dashed*)

ticles after each coarse-scale step, i.e., each k steps of the particle scheme, where $k\delta t = \delta t_{pde}$, (here $k = 1$). The results are shown in Fig. 4.6. We see that, using this reinitialization, the difference between the behaviour of the two processes is visually clear from one realization (left figure). Also, the resulting variance is such that the density of the control PDE is no longer within the 95% confidence interval of the variance reduced density estimation (right figure). We see that there is a significant difference between both models: the density corresponding to the control process is more peaked, indicating that bacteria that follow the control process are more sensitive to sudden changes in chemoattractant gradient. This difference can be interpreted from the fact that the bacteria with internal state do not adjust themselves instantaneously to their environment, but instead with a time constant τ .

4.7 Application 2: Tumor Growth

As a second application, we turn to tumor growth, which is a complex biological phenomenon consisting of processes on different scales. On the cellular level, one has to track the random motion of the cells, as well as cell division and cell death (apoptosis). As before, we deal with random motion by means of a velocity-jump process, in which we additionally ensure that the concentration of cells in a certain volume remains restricted. To achieve this, the spatial evolution of the individuals will be coupled through the local cell density. Cell division and apoptosis are modeled by means of two extra (intracellular) variables (a cell cycle variable ϕ and an apoptosis variable z), which in turn depend on intracellular and environmental concentrations of a number of chemical compounds (described via reaction diffusion equations). The resulting *fine-scale* model therefore consists of a velocity-jump process, supplemented with a set of ODEs describing the (sub)-cellular state of the individual cells and a set of reaction-diffusion PDEs describing the environment.

As for bacterial chemotaxis, it is possible to equivalently write the fine-scale model as a kinetic equation for each cell type p that then models the phase space density $p_p(x, v, \phi, z, t)$. However, a direct simulation of this fine-scale kinetic model is again not feasible because of the high-dimensional character of the resulting system, while a stochastic particle discretization is significantly influenced by Monte Carlo noise. Therefore, we propose a tailored variance reduction technique. The key point is to simulate the kinetic description of a simpler *control model* that only contains the motion of the cells and to couple this deterministic simulation with a stochastic agent-based simulation to obtain information on cell divisions and apoptosis.

We first give a detailed overview of the different layers of the model in Sect. 4.7.1. This model is similar to the model (4.62) used to describe bacterial chemotaxis, and reproduces the features of the cellular automaton model proposed by Owen et al. [39]. We describe the variance reduction algorithm in Sect. 4.7.2. Finally, we illustrate the technique with some numerical experiments in Sect. 4.7.3.

4.7.1 Model

The model consists of two main components: an agent-based model, describing the individual cellular motion and internal processes attached to each cell (cell cycle and apoptosis) and the environment, modeled by a set of reaction diffusion equations. We start by giving an overview of the model structure and notations that will be used throughout the section (Sect. 4.7.1.1), after which we describe the agent-based model (Sect. 4.7.1.2) and the evolution laws for the environment (Sect. 4.7.1.3).

4.7.1.1 Overview and Notation

We consider three types of cells, indexed by $1 \leq p \leq P = 3$: normal cells ($p = 1$), cancer cells ($p = 2$), and endothelial cells (that build up blood vessels, $p = 3$). For

each of these cell types, we consider an ensemble of $M_p(t)$ cells, characterized by their positions $X_{m,p}(t)$, velocities $V_{m,p}(t)$, cell cycle phase $\Phi_{m,p}(t)$, and remaining internal state variables $Z_{m,p}(t)$, for $1 \leq m \leq M_p(t)$ and $1 \leq p \leq P$. (The time dependence of the number of particles is due to the fact that cells divide and die.) In the numerical experiments in this chapter, blood vessel growth is not taken into consideration. We thus only consider normal cells and cancer cells.

To keep a consistent notation throughout the section, we introduce the following convention. If, at a moment $t = t^*$, the cell with index m^* in population p divides, we set

$$M_p(t^*) = M_p(t^*_-) + 1, \quad (4.73)$$

in which the symbol t^*_- is used to emphasize that the involved number of cells is meant to be taken *just before* the division. Simultaneously, we introduce a new cell as specified below (see Eq. 4.79). When a cell undergoes apoptosis, it is removed from the simulation. To avoid cumbersome renumbering of the cells in the text, we associate a weight $w_{m,p}(t)$ to each of the cells. If the cell is alive, the corresponding weight is one; upon apoptosis, it becomes zero. The active number of cells is therefore

$$\bar{M}_p(t) = \sum_{m=1}^{M_p(t)} w_{m,p}(t). \quad (4.74)$$

Given the positions of these cells, the empirical cell number density is then obtained as

$$\rho_p(x, t) = \sum_{m=1}^{M_p(t)} w_{m,p}(t) \delta_{X_{m,p}(t)}. \quad (4.75)$$

The agent-based cellular model is coupled with the environment, consisting of oxygen and vascular endothelial growth factor (VEGF). We denote by $C(x, t)$ the concentration of oxygen and by $G(x, t)$ the concentration of VEGF; both fields evolve according to PDEs in which the cell number density (4.75) appears. The different behaviour for different cell types originates both from different intracellular mechanisms and from the cell-type dependency of the coefficients in the agent-based model, see below. A table containing the parameter values for all cell types is given in Table 4.1.

4.7.1.2 Agent-Based Model

Motion Cellular motion is composed of two components: random motion, which will be modeled by means of a velocity jump process, and deterministic chemotactic motion towards high VEGF concentrations. The deterministic term also contains a volume factor, restricting motion towards regions where the number density $\rho_p(X_{m,p}(t), t)$ is higher than a threshold value $\rho_{\max,p}$

$$\frac{dX_p(t)}{dt} = \epsilon V_p(t) + \epsilon^2 \chi_p \nabla G(X_p(t), t) \left(1 - \frac{\rho_p(X(t), t)}{\rho_{\max, p}} \right), \tag{4.76}$$

$$V_p(t) = \mathcal{V}_p^k \text{ for } t \in [T_p^k, T_p^{k+1}], \quad \mathcal{V}_p^k \in \mathbb{V} = V_p^* \mathbb{S}^{d-1}, \tag{4.77}$$

in which $0 < \epsilon \ll 1$ and χ_p is the cell-type dependent chemotactic sensitivity. In this chapter, we choose the same velocity V_p^* for all cell types. Similar to the bacterial chemotaxis case, the velocities \mathcal{V}_p^k are sampled from the uniform distribution

$$\mathcal{M}_p(dv) = \sigma_{V_p^* \mathbb{S}^{d-1}}(dv),$$

on the sphere $V_p^* \mathbb{S}^{d-1}$, and the jump times T_p^k are generated from a Poisson process with constant rate λ_p ,

$$\int_{T_p^k}^{T_p^{k+1}} \lambda_p dt = \lambda_p (T_p^{k+1} - T_p^k) = \theta_p^{k+1},$$

where θ_p^{k+1} are i.i.d. random numbers, sampled from a normalized exponential distribution. During the numerical experiments, we choose $\lambda_p = 1$, independent of the cell type. In the bacterial chemotaxis case, bacteria possessed internal dynamics

Table 4.1 Parameter values related to the populations

Parameter	Normal population	Cancer population	Endothelial cells
χ	0.0	0.0	0.33333
ρ_{\max}	0.1	0.2	0.1
ϵ	3.536×10^{-4}	3.536×10^{-4}	–
C_ϕ	399.96	186.64	–
T_{\min}	1.8×10^5	9.6×10^4	–
z_{high}	0.8	–	–
z_{low}	0.08	–	–
n_{thr}	0.75	–	–
c_1	3.3333×10^{-5}	–	–
c_2	1.6667×10^{-4}	–	–
c_3	3.3333×10^{-5}	–	–
c_4	3.3333×10^{-5}	–	–
c_5	1.6667×10^{-4}	–	–
J_5	0.04	–	–
C_{VEGF}	0.01	–	–
C_{p53}	0.01	–	–
A	–	1	–
B	–	1	–

to bias their chemotactic motion. In the tumor model, internal dynamics governs cell division and cell death (apoptosis). (Hence, the time-dependent number of cells $M_p(t)$.) Let us now look at these intracellular mechanisms (Table 4.1)

Cell Cycle Dynamics Cells evolve according to a cell cycle, and division occurs as soon as this cycle is completed. In our model, this is represented by a cell phase cycle variable $\Phi_p(t)$ that is zero at cell birth. The cell cycle is completed when $\Phi_p(t) = 1$. The cell cycle speed depends on the local oxygen concentration $C(X_p(t), t)$ that is observed by the cell as it is simultaneously moving through space and evolving through the cycle. The higher the oxygen concentration, the faster the cycle proceeds, while the cell cycle is put on hold once the oxygen concentration is approaching zero. This behaviour is modeled by means of the following ODE,

$$\frac{d\Phi_p(t)}{dt} = \frac{C(X_p(t), t)}{\tau_{\min,p}(C_{\phi,p} + C(X_p(t), t))}, \quad (4.78)$$

in which we introduce the cell-type dependent parameters $C_{\phi,p}$ and $\tau_{\min,p}$, the minimal time needed for a cell to complete one cell cycle. From Table 4.1, we see that cancer cells are able to proceed twice as fast as normal cells during the cell cycle in a given environment.

If, for the cell with index m^* in population p , at time $t = t^*$, we obtain $\Phi_p(t^*) \geq 1$, we introduce a new cell in the simulation. We adjust $M_p(t)$ according to (4.73) and set $\Phi_{m^*,p}(t) = 0$. The new cell inherits the complete state from the cell that divides:

$$\begin{aligned} X_{M_p(t),p}(t^*) &= X_{m^*,p}(t^*), & V_{M_p(t),p}(t^*) &= V_{m^*,p}(t^*), \\ \Phi_{M_p(t),p}(t^*) &= \Phi_{m^*,p}(t^*), & Z_{M_p(t),p}(t^*) &= Z_{m^*,p}(t^*). \end{aligned} \quad (4.79)$$

More details on this cell cycle model can be found in [39] and its supplementary material.

Remaining Internal State Variables To account for apoptosis, we introduce a second sub-cellular model, consistent with [39],

$$\frac{dZ_p(t)}{dt} = F_p(X_p(t), Z_p(t)),$$

for the internal state $Z_p(t) \in \mathbb{R}^q$, in which q may depend on the cell type. This internal dynamics follows a different mechanism, depending on the cell type. For the normal tissue ($p = 1$), the variable $Z_1(t)$ contains two components, namely the p53 concentration $Z_{1,1}(t)$ and the intracellular VEGF concentration $Z_{1,2}(t)$. The former can be seen as an estimator for the number of mutations that a cell has undergone during its lifetime. The latter models the process that allows cells to store VEGF during hypoxic conditions and release it once this intracellular concentration has reached a certain threshold level.

$$\frac{dZ_{1,1}(t)}{dt} = c_1 - c_2 \frac{C(X_1(t), t)}{C_{p53} + C(X_1(t), t)} Z_{1,1}(t), \quad (4.80)$$

$$\frac{dZ_{1,2}(t)}{dt} = c_3 - c_4 \frac{Z_{1,1}(t)Z_{1,2}(t)}{J_5 + Z_{1,2}(t)} + c_5 \frac{C(X_1(t), t)}{C_{VEGF} + C(X_1, t)} Z_{1,2}(t), \quad (4.81)$$

where c_i , $1 \leq i \leq 5$, C_{p53} , J_5 , and C_{VEGF} are parameters of which the value can be found in Table 4.1. As soon as oxygen is available, the second term in (4.80) ensures exponential decay of $Z_{1,1}(t)$; the first term in (4.80) models linear growth of $Z_{1,1}(t)$, an effect that is only dominant in the absence of oxygen. A cell undergoes apoptosis if $Z_{1,1}(t)$ reaches a threshold value $\gamma_{\text{apt}}(\rho_1(X_1(t), t))$ that depends on the local density of normal cells. The threshold value is lower in case of a harsh environment (with low cell number density), i.e.,

$$\gamma_{\text{apt}}(\rho) = \begin{cases} z_{\text{high}} & \text{if } \rho \leq \rho_{\text{thr}} \\ z_{\text{low}} & \text{else} \end{cases}.$$

See Table 4.1 for parameter values.

The internal dynamics of tumor cells does not depend on the p53 concentration, since this mechanism to regulate the normal cell cycle does not function properly anymore in a tumor. Cancer cells are able to go into a quiescent state when expressed to hypoxic circumstances, meaning that they don't consume any nutrients. However, the maximal duration of this quiescent state is limited, which implies that cancer cells will also undergo apoptosis when the hypoxia holds too long. On the other hand, cancer cells have the ability to recover quickly once oxygen becomes available again. This mechanism can be modeled by the following equation:

$$\frac{dZ_2(t)}{dt} = \underbrace{A \text{H}(C_{\text{threshold}} - C(X_2(t), t))}_{\text{Linear increase during hypoxia}} - \underbrace{B Z_2(t) \text{H}(C(X_2(t), t) - C_{\text{threshold}})}_{\text{Exponential decay if } C(X_2(t), t) > C_{\text{threshold}}}, \quad (4.82)$$

where A , B are constants (see Table 4.1) and H is the Heaviside function. The first term models the reaction to a hypoxic state, i.e., when the local oxygen concentration $C(X_2(t), t)$ drops below the threshold level $C_{\text{threshold}}$. During this hypoxic period, the internal variable $Z_2(t)$ increases linearly as a function of time. On the other hand, the second term describes the recovery of the cancer cells if the environment is not hypoxic anymore, which is captured by the exponential decay term of $Z_2(t)$. Cancer cells die as soon as $Z_2(t) \geq 1$, corresponding to $\gamma_{\text{apt}} = 1$.

Complete Agent-Based Model Combining all components described above, we end up with the following set of differential equations governing the behaviour of an individual cell:

$$\left\{ \begin{array}{l} \frac{dX_p(t)}{dt} = \epsilon V_p(t) + \epsilon^2 \chi_p \nabla G(X_p(t), t) \left(1 - \frac{\rho_p(X_p(t), t)}{\rho_{\max}} \right), \\ (T^{k+1} - T^k) \lambda_p = \theta_p^{k+1}, \\ V_p(t) = \mathcal{V}_p^k \quad \text{for } t \in [T_p^k, T_p^{k+1}), \\ \frac{d\Phi_p(t)}{dt} = \frac{C(X_p(t), t)}{\tau_{\min, p}(C_{\phi, p} + C(X_p(t), t))}, \\ \frac{dZ_p(t)}{dt} = F_p(Z_p(t), t), \end{array} \right. \quad (4.83)$$

combined with the division rule (4.79) and the rule that adjusts the weights $w_{m,p}(t)$ upon apoptosis.

Coarse (population-Level) Description As in Sect. 4.4.2, we again have a kinetic equation of the phase space density $p_p(x, v, z, \phi, t)$, which becomes quite complicated due to the cell cycle and apoptosis that govern cell division and cell death. Ideally, a coarse-grained model would be written in terms of the cell number density $\rho_p(x, t)$. We expect to obtain a reaction-advection-diffusion equation. However, due to the modeling detail for the cell cycle and apoptosis, it is unrealistic to expect one can write the reaction term as a closed-form function of $\rho_p(x, t)$. When ignoring the intracellular dynamics (and therefore considering a system with constant number of cells) and using a diffusive scaling $\bar{t} = t\epsilon^2$ [36], one can obtain an advection-diffusion equation where no reactions (cell divisions, cell deaths) are taken into account [30]:

$$\partial_{\bar{t}} \rho_p(x, \bar{t}) = D_p \nabla^2 \rho_p(x, \bar{t}) - \chi_p \nabla \cdot \left[\rho_p(x, \bar{t}) \left(1 - \frac{\rho_p(x, \bar{t})}{\rho_{\max}} \right) \nabla G(x, t) \right], \quad (4.84)$$

in which $D_p = \int_{\mathbb{V}} v \otimes v \mathcal{M}(dv)$.

In the variance reduction algorithm, we will also consider the *kinetic number density*,

$$N_p(x, v, t) = \int \int p_p(x, v, z, \phi, t) dz d\phi = \sum_{m=1}^{M_p(t)} w_{m,p}(t) \delta_{X_{m,p}(t), V_{m,p}(t)}, \quad (4.85)$$

which counts the number of particles with a position x and velocity v at time t , regardless of their internal state.

4.7.1.3 Environment

The cellular environment consists of two diffusible components regulating the behavior of the cells in various ways: oxygen $C(x, t)$ and VEGF concentration $G(x, t)$. Oxygen is evidently important for the cells to proceed through the cell cycle and survive, see Eqs. (4.78) and (4.80). The local oxygen concentration is determined from the following advection-diffusion equation:

Table 4.2 Parameter values reaction diffusion equations

Parameter	Oxygen	VEGF
D	2.4167×10^{-9}	1.6667×10^{-11}
ψ	6	1.6667×10^{-9}
δ	0	$1.667e - 4$
k_{normal}	-0.2167	0.01
k_{cancer}	-0.2167	0.01

$$\partial_t C(x, t) = \underbrace{D_C \nabla^2 C(x, t)}_{\text{diffusion}} + \underbrace{\psi_C \rho_v(x, t) (C_{\text{blood}}(x) - C(x, t))}_{\text{exchange with blood}} - \underbrace{C(x, t) \sum_{p=1}^P k_{C,p} \rho_p(x, t)}_{\text{consumption by cells}}, \quad (4.86)$$

where D_C is the diffusion coefficient, ψ_C denotes the permeability of the oxygen through the vessels, $\rho_v(x, t)$ describes the surface area occupied by blood vessels at position x , and $C_{\text{blood}}(x)$ defines the oxygen concentration in a blood vessel located at position x . In general the blood vessel concentration $\rho_v(x, t)$ is computed from the agent-based evolution of endothelial cells. In this paper, we present numerical tests for a case where $\rho_v(x, t) \equiv \rho_v(x)$ is fixed during the simulation, and no vessel growth is incorporated. We choose $\rho_v(x) = 1$ in grid cells where blood vessels are present. During the experiments we have taken $C_{\text{blood}}(x, t) = 400$ at vessel locations and zero elsewhere. The last term in Eq. (4.86) reflects the fact that all cell types consume oxygen with a cell specific rate $k_{C,p}$. Recall that the cell number density $\rho_p(x, t)$ is defined via Eq. (4.75).

A similar approach is used to describe the local concentration of VEGF, which is responsible for the growth of new blood vessels. This is especially important for larger tumors, since endothelial cells will grow towards regions with higher VEGF concentrations. Initially, the tumor can benefit from the existing vasculature, but when the tumor occupies a larger volume, the oxygen supply does not suffice and the cells are obliged to use their ability to ask for new vessels by secreting VEGF. Endothelial cells – the building blocks of blood vessels– can then react and move chemotactically towards the hypoxic regions. The corresponding reaction diffusion equation for VEGF reads:

$$\partial_t G(x, t) = \underbrace{D_G \nabla^2 G(x, t)}_{\text{diffusion}} - \underbrace{\psi_G \rho_v(x, t) G(x, t)}_{\text{exchange with blood}} + \underbrace{\sum_{p=1}^P k_{G,p} \rho_p(x, t)}_{\text{production}} - \underbrace{\delta_G G(x, t)}_{\text{decay}}. \quad (4.87)$$

In contrast to oxygen, the VEGF concentration is assumed to be zero in the blood and the growth factors contained in the tissue is degrading with rate δ_G when time evolves (Table 4.2).

Because the diffusible components equilibrate much faster than the individual cells, we adopt a steady state approximation, which implies that $\partial_t G(x, t)$ and $\partial_t C(x, t)$ are set to zero. Then, a time step of the agent-based model is performed first, after which the steady state equations for $C(x, t)$ and $G(x, t)$ are solved.

4.7.2 Variance Reduction

In this section, we propose an variance reduction algorithm similar to the technique used for bacterial chemotaxis. The main differences are due to the fact that (i) the tumor model is not conservative; and (ii) the internal dynamics only relates to cell division and apoptosis and not to advection-diffusion behaviour.

Again, the algorithm conceptually relies on the combination of three simulations: a stochastic simulation with the full fine-scale model, as well as with a coarse, approximation, combined with a deterministic, grid-based simulation of the coarse model. The full fine-scale model uses $M_p(t)$ particles with state variables

$$\{X_{m,p}(t), V_{m,p}(t), \Phi_{m,p}(t), Z_{m,p}(t)\}_{m=1}^{M_p(t)}. \quad (4.88)$$

As the coarse agent-based model, we conceptually consider an agent-based model in which the internal state has been suppressed and only position and velocity remain:

$$\{X_{m,p}^c(t), V_{m,p}^c(t)\}_{m=1}^{M_p(t)}.$$

Since no internal dynamics is present, cells cannot divide or die. (In practice, we will use the results obtained from the full fine-scale model, in which we neglect apoptosis and cell division, see later.) The only dynamics is motion, which can be modeled with a kinetic equation for the phase space density $N_p^c(x, v, t)$,

$$\partial_t N_p^c + \epsilon v \cdot \nabla_x N_p^c = \lambda \left(R(N_p^c) - \lambda N_p^c \right), \quad (4.89)$$

see also (4.55). We again call this coarse approximation the *control process*. (Recall that the kinetic number density for the original process is denoted as $N_p(x, v, t)$, see (4.85).) We also introduce the formal semigroup notation

$$e^{tL_p^c}, \quad \text{with } L_p^c(N_p^c) = -(\epsilon v + \epsilon^2 \chi \nabla G(x, t)) \cdot \nabla_x N_p^c + \left(\lambda(R(N_p^c) - N_p^c) \right) \quad (4.90)$$

that represents the exact solution of the kinetic equation (4.89). In practice, this exact solution will be approximated by a deterministic simulation on a grid.

It should be clear that the advection-diffusion behaviour in both agent-based models is identical. Thus, the only difference between the two models occurs when cells divide or die. Assuming no reactions take place, the three processes thus have the

same expectation. This observation leads to the following variance reduction algorithm. As an initial condition, we start from $M_p(0)$ particles sampled from the kinetic probability densities $\mu_{p,0}^{M_p(0)}(x, v)$, resulting in the number density $N_p(x, v, 0)$. For each particle, we choose a given internal state, for instance $\Phi_{m,p}(0) = Z_{m,p}(0) = 0$, $1 \leq m \leq M_p(0)$, $1 \leq p \leq P$. (These internal states could also be sampled from an appropriate probability distribution.) Additionally, we introduce the variance reduced measure $\bar{N}(x, v, t)$, which we initialize as $\bar{N}_p(x, v, 0) = N_p(x, v, 0)$. We denote the time step δt and the discrete time instances $t^\ell = \ell \delta t$, $\ell = 0, 1, \dots$

Algorithm 4 (*Variance reduction for tumor growth*) We advance the variance reduced kinetic number density $\bar{N}(x, v, t)$ from time t^ℓ to $t^{\ell+1}$ as follows:

- Evolve the particle states (4.88) from t^ℓ to $t^{\ell+1}$ using the agent-based model (4.83).
- Compute the kinetic number density for the stochastic fine-scale model using (4.85), as well as the kinetic number density for the coarse process as

$$N_p^c(x, v, t^{\ell+1}) = \sum_{m=1}^{M_p(t^\ell)} w_{m,p}(t^\ell) \delta_{X_{m,p}(t^{\ell+1}), V_{m,p}(t^{\ell+1})}, \quad (4.91)$$

i.e., we compute the kinetic number density for the control process based on particle positions and velocities at time $t^{\ell+1}$, taking into account only the particles that were present in the simulation at time t^ℓ .

- Evolve the control kinetic number density $N_p^c(x, v, t)$ using a grid-based method based on (4.90) and add the reactions (the difference in kinetic number density due to cell division and apoptosis)

$$\bar{N}_p(x, v, t^{\ell+1}) := \bar{N}_p(x, v, t^\ell) e^{\delta t / \epsilon^2 L^c} + N_p(x, v, t^{\ell+1}) - N_p^c(x, v, t^{\ell+1}), \quad (4.92)$$

Again, in the absence of discretization errors in the grid-based method, we have

$$\mathbb{E} \left[\bar{N}_p(x, v, t^{\ell+1}) \right] = \mathbb{E} \left[N_p(x, v, t^{\ell+1}) \right],$$

since $\bar{N}_p(x, v, t^\ell) = N_p^c(x, v, t^\ell)$ and $\mathbb{E} \left[N_p^c(x, v, t^\ell) e^{\delta t / \epsilon^2 L^c} \right] = \mathbb{E} \left[N_p^c(x, v, t^{\ell+1}) \right]$. Moreover, we expect the variance of \bar{N}_p to be significantly lower than that of N_p , since all randomness due to random motion has been removed and only the location of the reactions remains random.

4.7.3 Numerical Experiments

In this section, we illustrate both the model and the variance reduction algorithm by means of some numerical experiments. First, we initialize a normal tissue consisting

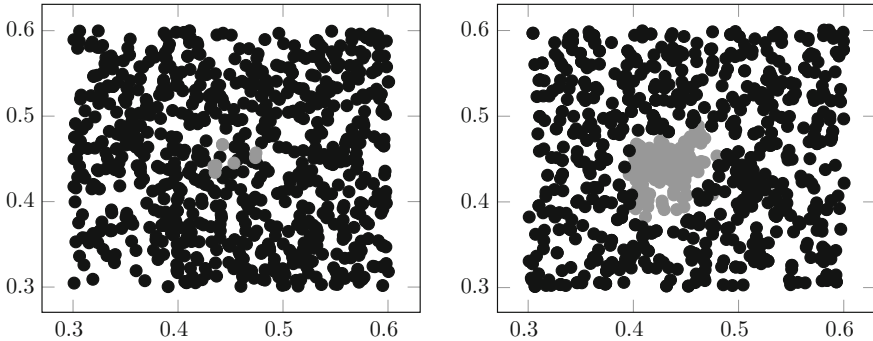


Fig. 4.7 Evolution of normal cell population (*black*) with small tumor (*gray*). *Left* initial state (1600 normal cells and 20 cancer cells). *Right* result after 490 steps

of 1600 cells, with initial positions sampled from the uniform distribution on the domain $[0.3, 0.6] \times [0.3, 0.6]$. The initial velocities are sampled from the uniform distribution on the sphere, and the internal state variables $\Phi(0) = 0.0$, $Z(0) = 0$ for every particle. We initialize a small tumor, containing 20 cells, with initial positions samples from a uniform distribution in $[0.445, 0.455] \times [0.445, 0.455]$, also with initial velocities sampled from the uniform distribution on the sphere with radius $V_p^* = 3.5 \times 10^{-6}$, and the internal state variables $\Phi(0) = Z(0) = 0$ for every particle. In the following numerical experiments, we adopt a static vasculature, consisting of two straight vertical vessels at $x = 0.4$ and at $x = 0.8$ to be specific. The first agent-based experiment was performed in a non-scaled way with discretization parameters $\delta t = 1.8 \times 10^3$ s and $\Delta x = 4 \times 10^{-5}$ m for the agent-based model. We use reflective boundary conditions. In Fig. 4.7, we show the evolution of both cell populations. On the left, one can observe the initial configuration, while on the right hand side, we see the resulting configuration after 490 steps. Apart from the fact that cells performed a random walk, a significant amount of the normal tissue died because of the presence of the tumor.

In a second experiment, we illustrate the performance of the variance reduction algorithm as it was explained in Sect. 4.7.2 in a similar setting as the previous experiment but on a diffusive scale, i.e. the normal tissue containing 2500 cells is uniformly distributed on the square $[0.2, 0.8] \times [0.2, 0.8]$ and a small tumor originally consisting of 20 cells is also uniformly distributed within the area $[0.39, 0.41] \times [0.39, 0.41]$. Furthermore, the scaled cellular velocity was chosen $\bar{V}^* = \sqrt{2}/2$ and we modified the minimal cell cycle durations ($T_{\min, \text{cancer}} = 9.6 \times 10^3$ s, $T_{\min, \text{normal}} = 1.8 \times 10^4$ s) to demonstrate the performance of the algorithm in an extreme (not biologically realistic) setting where cells are able to divide very quickly. As in the bacterial chemotaxis application, the coarse equation is modeled on a diffusive timescale, where we choose $\delta \bar{t} = 0.8\epsilon^2 = 1 \times 10^{-7}$, $\Delta \bar{x} = 2 \times 10^{-2}$ as mesh parameters to simulate the deterministic kinetic equation (4.89) needed to apply the variance reduction algorithm. To simulate this, a second order central finite volume scheme was adopted to discretize

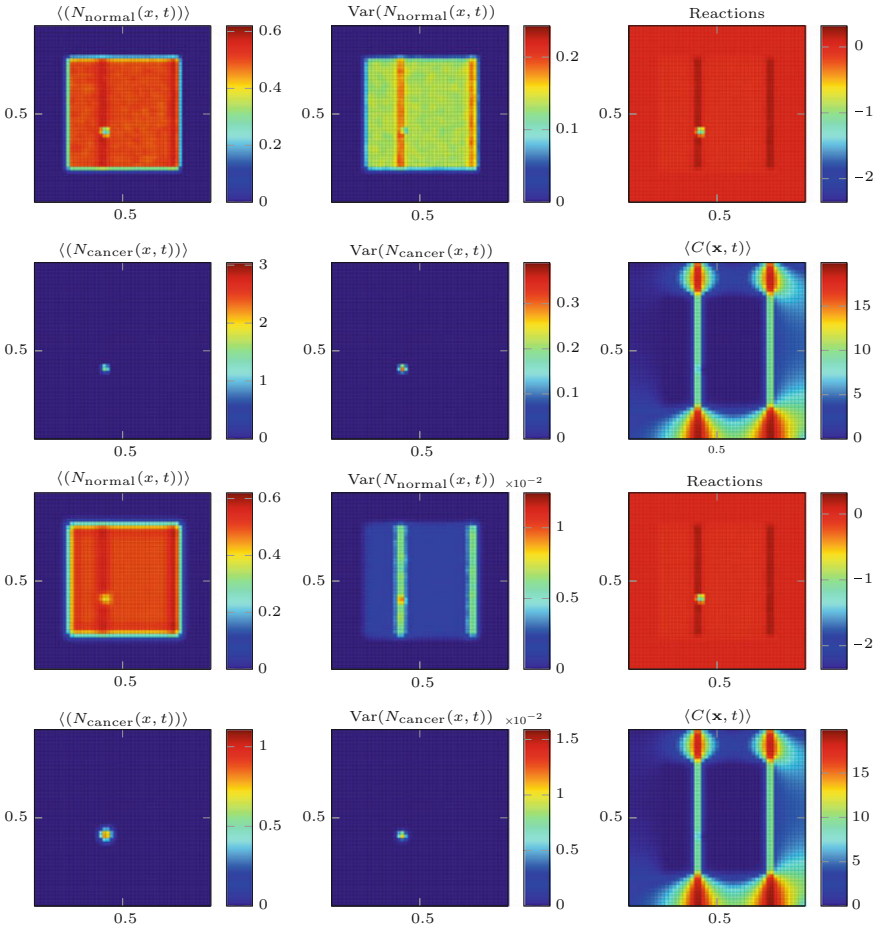


Fig. 4.8 Illustration variance reduction algorithm for tumor growth: Mean cell distribution (*first column*) and variance (*2nd column*) for normal and cancer population after 5000 timesteps. The *first two rows* show the results without applying the variance reduction algorithm, which can be compared with the results displayed in the *last two rows* where the variance reduction algorithm was applied

the spatial derivative in the kinetic equation (4.89) and a first order forward Euler scheme for the time derivative.

The mean normal and cancer cell distribution and there variance with and without applying the variance reduction algorithm are displayed. Furthermore the mean oxygen distribution and the corresponding mean reaction field are also shown next to the variance. First, we observe that variance on the normal cell distributed has been reduced significantly for both cell populations by applying Algorithm 4. Additionally, one can observe a clear correlation with the oxygen field. The latter is a logical consequence from the fact that the progress of the cell cycle is closely related to the local oxygen concentration along the track, meaning that a cell is more likely to

divide in oxygen rich environments. Moreover, the variance is almost eliminated in the other regions (Fig. 4.8).

In this section, the parameters used in the numerical experiments are listed. They are based on the parameters used in [39]. One can find them in the supplementary material corresponding to [39].

4.8 Conclusions

We gave a broad overview of the use of stochastic modeling and simulation techniques in computational biology, focussing on some common individual-based modeling and simulation methods. We payed particular attention to the equivalence between the stochastic process that governs the evolution of individual agents and the deterministic behaviour of the involved probability distributions, and we discussed numerical methods that exploit this relation for variance reduction purposes. Using examples involving intracellular chemical reactions, bacterial chemotaxis and tumor growth, we showed the effects of stochasticity at different scales and different levels of description.

A main focus of the chapter was the design of dedicated simulation algorithms. Two main computational bottlenecks arise. The first is related to the time-scale separation between the fast processes (that determine the maximal time step that is allowed) and slow processes (that determine the time scale over which the simulation needs to be performed). The second is related to noise that appears in the simulation and that requires dedicated variance reduction techniques. These problems are not completely solved, and stochastic simulation therefore remains an active research topic.

Acknowledgments This work was partially supported by the Research Council of the K.U. Leuven through grant OT/13/66 and by the Interuniversity Attraction Poles Programme of the Belgian Science Policy Office through grant IUAP/V/22 and by the agency for Innovation by Science and Technology in Flanders (IWT). We thank Bert Mortier, Delphine Vandamme and the editors for their careful proofreading and comments that helped us improve the text.

References

1. Al-Mohssen, H.A., Hadjiconstantinou, N.G.: Low-variance direct Monte Carlo simulations using importance weights. *ESAIM. Math. Model. Numer. Anal.* **44**(05), 1069–1083 (2010)
2. Alt, W.: Biased random walk models for chemotaxis and related diffusion approximations. *J. Math. Biol.* **9**(2), 147–177 (1980)
3. Auger, A., Chatelain, Ph. Koumoutsakos, P.: R-leaping: accelerating the stochastic simulation algorithm by reaction leaps. *J. Chem. Phys.* **125**(8), 084103 (2006)
4. Bal, G., Davis, A.B., Langmore, I.: A hybrid (Monte Carlo/deterministic) approach for multi-dimensional radiation transport. *J. Comput. Phys.* **230**(20), 7723–7735 (2011)
5. Berg, H.C.: *Random Walks in Biology*. Princeton University Press, Princeton (1993)

6. Bren, A., Eisenbach, M.: How signals are heard during bacterial chemotaxis: protein-protein interactions in sensory signal propagation. *J. Bacteriol.* **182**(24), 6865–6873 (2000)
7. Caffisch, R.E.: Monte Carlo and quasi-Monte Carlo methods. *Acta Numer.* **7**, 1–49 (1998)
8. Cao, Y., Gillespie, D.T., Petzold, L.R.: The slow-scale stochastic simulation algorithm. *J. Chem. Phys.* **122**(1), 014116 (2004)
9. Cao, Y., Gillespie, D.T., Petzold, L.R.: Efficient step size selection for the tau-leaping simulation method. *J. Chem. Phys.* **124**(4), 044–109 (2006)
10. Chalub, F.A.C.C., Markowich, P.A., Perthame, B., Schmeiser, C.: Kinetic models for chemotaxis and their drift-diffusion limits. *Monatsh. für Math.* **142**(1–2), 123–141 (2004)
11. Chorin, A.J., Hald, O.H.: *Stochastic Tools in Mathematics and Science*. Springer, Texts in applied mathematics (2013)
12. Chung, C.A., Lin, T.-H., Chen, S.-D., Huang, H.-I.: Hybrid cellular automaton modeling of nutrient modulated cell growth in tissue engineering constructs. *J. Theor. Biol.* **262**(2), 78–267 (2010)
13. Degond, P., Dimarco, G., Pareschi, L.: The moment-guided Monte Carlo method. *Int. J. Numer. Meth. Fluids* **67**(2), 1–24 (2009)
14. Dimarco, G., Pareschi, L.: Hybrid multiscale methods II. kinetic equations. *Multiscale Model. Simul.* **6**(4), 1169–1197 (2008)
15. Erban, R., Othmer, H.G.: From individual to collective behavior in bacterial chemotaxis. *SIAM J. Appl. Math.* **65**(2), 361–391 (2004)
16. Erban, R., Othmer, H.G.: From signal transduction to spatial pattern formation in *E. coli*: a paradigm for multiscale modeling in biology. *Multiscale Model. Simul.* **3**(2), 362–394 (2005)
17. Erban, R., Chapman, J., Maini, Ph., A practical guide to stochastic simulations of reaction-diffusion processes. Technical report, 2007
18. Ethier, S.N., Kurtz, T.G.: *Markov processes: characterization and convergence*. Wiley, New York (2009)
19. Gardiner, C.: *Stochastic methods: a handbook for the natural and social sciences*, vol. 13. Springer Series in Synergetics. Springer, Berlin (2009)
20. Gardner, T.S., Cantor, C.R., Collins, J.J.: Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403**(6767), 339–342 (2000)
21. Gerlee, P., Anderson, A.R.A.: An evolutionary hybrid cellular automaton model of solid tumour growth. *J. Theor. Biol.* **246**(4), 583–603 (2007)
22. Ghanem, R.G., Spanos, P.D.: *Stochastic Finite Elements: A Spectral Approach*. Springer-Verlag New York Inc, New York (1991)
23. Gillespie, D.T.: Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* **81**(25), 2340–2361 (1977)
24. Gillespie, D.T.: The chemical langevin equation. *J. Chem. Phys.* **113**(1), 297–306 (2000)
25. Gillespie, D.T.: Approximate accelerated stochastic simulation of chemically reacting systems. *J. Chem. Phys.* **115**(4), 1716 (2001)
26. Gillespie, D.T.: Stochastic simulation of chemical kinetics. *Annu. Rev. Phys. Chem.* **58**, 35–55 (2007)
27. Hairer, E., Wanner, G.: *Solving Ordinary Differential Equations II*. Springer Series in Computational Mathematics, Springer, Berlin (1996)
28. Hasty, J., McMillen, D., Collins, J.J.: Engineered gene circuits. *Nature* **420**(6912), 224–30 (2002)
29. Higham, D.J.: An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM Rev.* **43**(3), 525–546 (2001)
30. Hillen, T., Painter, K.J.: A user’s guide to PDE models for chemotaxis. *J. Math. Biol.* **58**(1–2), 183–217 (2009)
31. Kiskowski, M.A., Alber, M.S., Thomas, G.L., Glazier, J.A., Bronstein, N.B., Pu, J., Newman, S.A.: Interplay between activator-inhibitor coupling and cell-matrix adhesion in a cellular automaton model for chondrogenic patterning. *Dev. Biol.* **271**(2), 87–372 (2004)
32. Kloeden, P.E., Platen, E.: *Numerical Solution of Stochastic Differential Equations*. Springer, Stochastic modeling and applied probability (1992)

33. Mallet, D.G., De Pillis, L.G.: A cellular automata model of tumor-immune system interactions. *J. Theor. Biol.* **239**(3), 50–334 (2006)
34. Øksendal, B.: *Stochastic Differential Equations*. Springer (2003)
35. Othmer, H.G., Hillen, T.: The diffusion limit of transport equations derived from velocity-jump processes. *SIAM J. Appl. Math.* **61**(3), 751 (2000)
36. Othmer, H., Hillen, T.: The diffusion limit of transport equations derived from velocity-jump processes. *SIAM J. Appl. Math.* **61**(3), 751–775 (2000)
37. Othmer, H.G., Hillen, T.: The diffusion limit of transport equations II: chemotaxis equations. *SIAM J. Appl. Math.* **62**(4), 1222 (2002)
38. Othmer, H.G., Dunbar, S.R., Alt, W.: Models of dispersal in biological systems. *J. Math. Biol.* **26**(3), 263–298 (1988)
39. Owen, M.R., Stamper, I.J., Muthana, M., Richardson, G.W., Dobson, J., Lewis, C.E., Byrne, H.M.: Mathematical modeling predicts synergistic antitumor effects of combining a macrophage-based, hypoxia-targeted gene therapy with chemotherapy. *Cancer Res.* **71**(8), 37–2826 (2011)
40. Pareschi, L., Toscani, G.: *Interacting Multiagent Systems: Kinetic Equations and Monte Carlo Methods*. Oxford University Press (2013)
41. Patlak, C.S.: Random walk with persistence and external bias. *Bull. Math. Biophys.* **15**(3), 311–338 (1953)
42. Radtke, G., Hadjiconstantinou, N.: Variance-reduced particle simulation of the Boltzmann transport equation in the relaxation-time approximation. *Phys. Rev. E* **79**(5), 1–9 (2009)
43. Roose, T., Chapman, S.J., Maini, Ph K: Mathematical models of avascular tumor growth. *SIAM Rev.* **49**(2), 179–208 (2007)
44. Rousset, M., Samaey, G.: Simulating individual-based models of bacterial chemotaxis with asymptotic variance reduction. *Math. Models Methods Appl. Sci.* **23**(12):2155–2191 (2013)
45. Rousset, M., Samaey, G.: Individual-based models for bacterial chemotaxis in the diffusion asymptotics. *Math. Models Methods Appl. Sci.* **23**(11), 2005–2037 (2013)
46. Schlögl, F.: Chemical reaction models for non-equilibrium phase transitions. *Zeitschrift für Physik* **253**(2), 147–161 (1972)
47. Scott, D.W.: *Multivariate density estimation: theory, practice, and visualization*. Wiley Series in Probability and Statistics. Wiley, New York (2009)
48. Silverman, B.W.: *Density Estimation for Statistics and Data Analysis*. Chapman and Hall/CRC, Boca Raton (1986)
49. Spohn, H.: *Large scale dynamics of interacting particles*, Springer, vol. 825 (1991)
50. Stock, A.M.: A nonlinear stimulus-response relation in bacterial chemotaxis. *Proc. Natl. Acad. Sci.* **96**(20), 10945–10947 (1999)
51. Stuart, A.M.: Inverse problems: a Bayesian perspective. *Acta Numer.* **19**, 451–559 (2010)
52. Weinan, E., Vanden-Eijnden, E.: Transition-path theory and path-finding algorithms for the study of rare events. *Annu. Rev. Phys. Chem.* **61**:391–420 (2010)