

METHODS FOR STOCHASTIC SYSTEMS
LECTURE NOTES
MA3/50257



DR CHRISTIAN A. YATES (based on a lecture course and book: “Stochastic Modelling of Reaction-Diffusion Processes” by Professor Radek Erban)

SEMESTER 1 2018/2019

These lecture notes must not be reproduced, republished or shared without
the explicit permission of the lecturer

Contents

1 Lecture 1 - Stochastic simulation of chemical reactions	5
1.1 Stochastic simulation of degradation	5
2 Lecture 2 - Stochastic simulation of chemical reactions II	11
2.1 Stochastic simulation of production and degradation	11
3 Lecture 3 - Higher-order chemical reactions	17
3.1 Implementation of Figure 2.1 in Matlab	17
3.2 Higher-order chemical reactions	20
3.3 The Gillespie algorithm	23
4 Lecture 4 - Approximate stochastic simulation algorithms	29
4.1 Poisson processes	29
4.2 τ -leap algorithm	30
4.3 A fixed choice of τ	31
4.4 Meeting the leap condition	35
4.5 Negative populations	35
5 Lecture 5 - Deterministic versus stochastic modelling I	36
5.1 Stochastic simulation of dimerisation	36
6 Lecture 6 - Deterministic versus stochastic modelling II	43
6.1 Stochastic focusing	43
7 Lecture 7 - Deterministic versus stochastic modelling III	50
7.1 Systems with multiple favourable states	50
7.2 A research example - Modelling the collective behaviour of locusts	53
8 Lecture 8 - Deterministic versus stochastic modelling IV	57
8.1 Self-induced stochastic resonance	57
9 Lecture 9 - Moment closure approximations I	61
9.1 Cumulants	61
9.2 Reversible Dimerisation Model	62
9.3 Mean-field model	63
9.4 Pairwise approximation	64
9.5 Cumulant neglect	65
9.6 Kirkwood superposition approximation	66
10 Lecture 10 - Moment closure approximations II	68
10.1 SIR example	68
10.2 Mean-field model	69

10.3 Pairwise approximation	70
10.4 Cumulant neglect	71
10.5 Kirkwood superposition approximation	74
11 Lecture 11 - A realistic model of cell proliferation I	75
11.1 Multi-stage model of the cell cycle	75
12 Lecture 12 - A realistic model of cell proliferation II	80
12.1 Mean behaviour	80
12.2 Limiting behaviour	81
13 Lecture 13 - Stochastic differential equations I	84
13.1 A computational definition of an SDE	84
13.2 Examples of SDEs	86
14 Lecture 14 - Stochastic differential equations II - Fokker-Planck equation	90
15 Lecture 15 - Stochastic differential equations III	95
15.1 The backward Kolmogorov equation	95
15.2 SDEs with multiple favourable states	96
16 Lecture 16 - Chemical Fokker-Planck equation I	100
16.1 Derivation of the Chemical Fokker-Planck equation	100
17 Lecture 17 - Chemical Fokker-Planck equation II	104
17.1 Analysis of example from Section 7.1	104
17.2 Analysis of example from Section 8.1	107
18 Lecture 18 - Diffusion and position jump processes	110
18.1 Diffusion modelled by SDEs	110
18.2 Compartment-based approach to diffusion	114
19 Lecture 19 - Diffusion and velocity jump processes	120
19.1 Introducing velocity-jump models	120
19.2 Einstein-Smoluchowski relation	127
20 Lecture 20 - Boundary conditions for diffusion	130
20.1 Diffusion to adsorbing surfaces	130
20.1.1 Adsorbing boundaries in SDE-based models of diffusion	130
20.1.2 Adsorbing boundaries in compartment-based models of diffusion	134
20.2 Reactive boundary conditions	136
21 Lecture 21 - Compartment-based reaction-diffusion	141
21.1 Stochastic simulation of diffusion, production and degradation	141
21.2 Compartment-based SSA for higher-order reactions	144
21.3 Choice of compartment size h	147
21.4 Pattern formation	150
22 Molecular-based approaches to reaction-diffusion modelling	153
22.1 Molecular-based stochastic model of diffusion, production and degradation	153
22.2 Molecular-based approaches for second-order reactions	154
22.3 Reaction radius and reaction probability	156

22.4 Modelling reversible reactions	161
23 Lecture 23 - Collective behaviour of cells and animals	165
23.1 Stochastic modelling of diffusion-advection processes	165
23.2 Bacterial chemotaxis	168
23.3 Collective behaviour of locusts	173
23.4 Reaction-diffusion-advection SSAs	177
24 Lecture 24 - Metropolis-Hastings algorithm	180
25 Lecture 25 - Multiscale modelling	185
25.1 A simple multiscale problem	185
25.2 Multiscale SSA with partial equilibrium assumption	188
25.2.1 Distributions of fast variables	191
25.2.2 Discussion	192
25.3 Other methods for efficient simulations of well-mixed chemical systems	194
25.4 Multiscale and hybrid modelling for spatio-temporal processes	195
25.4.1 Coupling compartment-based and molecular-based models of diffusion	196
25.4.2 Other methods for efficient simulations of spatio-temporal models	201
26 Lecture 26 - Examination and beyond	202

Lecture 1 - Stochastic simulation of chemical reactions

We introduce basic concepts using a simple degradation (1.1). We will build on this material in Lecture 3 when we discuss stochastic methods for modelling spatially homogeneous systems of chemical reactions.

1.1 Stochastic simulation of degradation

Consider the single chemical reaction



where A is the chemical species of interest and k is the rate constant of the reaction. The symbol \emptyset denotes chemical species which are of no further interest. The rate constant k in (1.1) is defined so that kdt gives the probability that a randomly chosen particle of the chemical species A reacts (is degraded) during the time interval $[t, t+dt]$ where t is time and dt an (infinitesimally) small time-step.

Let us denote the number of particles of chemical species A at time t by $A(t)$ (a convention which will be used throughout the course). Then, in the time interval $[t, t + dt]$, a number of things may happen: none of the particles may react, exactly one may react, or more than one may react. Assuming that each particle acts independently, we may combine the individual probabilities of reaction to deduce that:

- no reactions occur with probability $1 - A(t)kdt + O(dt^2)$,
- exactly one reaction occurs with probability $A(t)kdt + O(dt^2)$,
- two or more reactions occur with probability $O(dt^2)$,

where $O(dt^2)$ signifies terms proportional to dt^2 , which we neglect as dt is small.

Let us assume that we have n_0 particles of A in the system at time $t = 0$, i.e. $A(0) = n_0$. Our first goal is to compute the number of particles $A(t)$ for times $t > 0$. To do that, we need a computer routine generating random numbers uniformly distributed in the interval $(0, 1)$. Such a program is included in many modern programming languages (e.g. the function `rand` in Matlab - we will discuss Matlab in Lecture 3). The routine will generate a number $r \in (0, 1)$, such that the probability that r is in a subinterval $(a, b) \subset (0, 1)$ is equal to $b - a$ for any $a, b \in (0, 1)$, $a < b$.

The mathematical definition of the chemical reaction (1.1) can be directly used to design a “naive” stochastic simulation algorithm (SSA) for simulating it. We choose a small time-step Δt , and compute the number of particles $A(t)$ at times $t = i\Delta t$, $i = 1, 2, 3, \dots$, as follows. Starting with $t = 0$ and $A(0) = n_0$, we perform two steps at time t :

Algorithm 1

- (1a) Generate a random number, r , uniformly distributed in the interval $(0, 1)$.
- (1b) If $r < A(t)k\Delta t$, then put $A(t + \Delta t) = A(t) - 1$; otherwise, put $A(t + \Delta t) = A(t)$. Then continue with step (1a) for time $t + \Delta t$.

Since r is a random number uniformly distributed in the interval $(0, 1)$, the probability that $r < A(t)k\Delta t$ is equal to $A(t)k\Delta t$. Consequently, step (1b) says that the probability that the chemical reaction (1.1) occurs in the time interval $[t, t + \Delta t]$ is equal to $A(t)k\Delta t$. Thus step (1b) correctly implements the definition of (1.1) provided that Δt is small. The time evolution of A obtained by the naive SSA (1a)-(1b) is given in Figure 1.1 (a) for $k = 0.1 \text{ sec}^{-1}$, $A(0) = 20$ and $\Delta t = 0.005 \text{ sec}$. We repeated the stochastic simulation twice and we plotted two realisations of the SSA (1a)-(1b). We see in Figure 1.1 (a) that these two realisations of the SSA (1a)-(1b) give two different evolutions. Each time we run the algorithm, we obtain different results. This is generally true for any SSA. Therefore, one might reasonably ask what useful and reproducible information can be obtained from stochastic simulations? We will come back to this question later in this section.

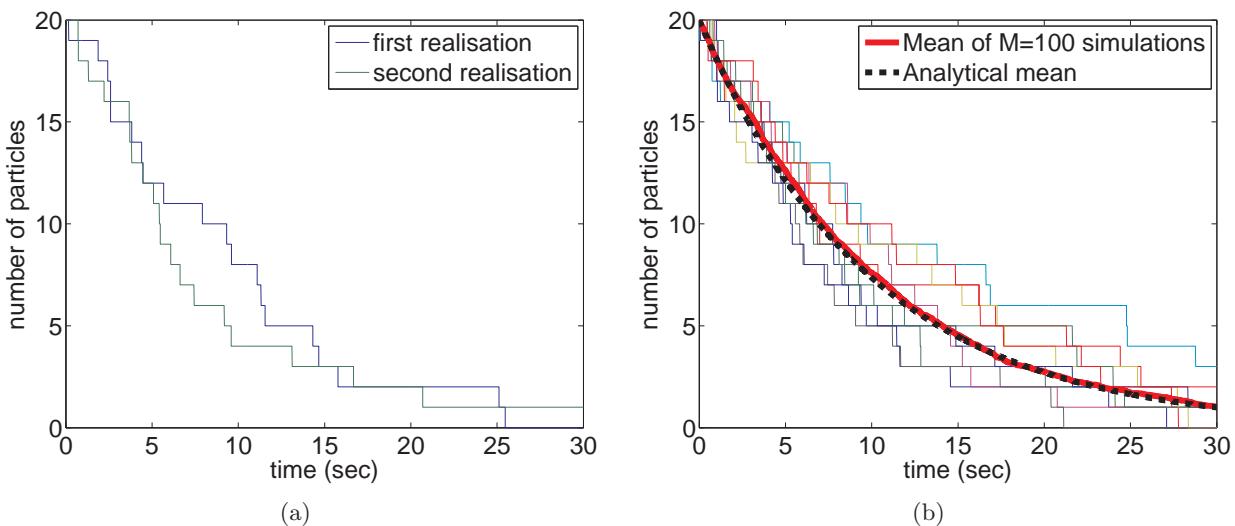


Figure 1.1: Stochastic simulation of chemical reaction (1.1) for $k = 0.1 \text{ sec}^{-1}$ and $A(0) = 20$. (a) Number of particles of A as a function of time for two realisations of the “naive” SSA (1a)-(1b) for $\Delta t = 0.005 \text{ sec}$; (b) results of ten realisations of the SSA (2a)-(2c) (thin solid lines; different colours show different realisations), the mean averaged over 100 repeat simulations (thick solid red line) and mean given by equation (1.8) (black dashed line).

The probability that exactly one reaction (1.1) occurs during the infinitesimal time interval $[t, t + dt]$ is equal to $A(t)kdt$. To design the SSA (1a)-(1b), we replaced dt by the finite time-step Δt . In order to get reasonably accurate results, we must ensure that $A(t)k\Delta t \ll 1$. For our simulations we used $k = 0.1 \text{ sec}^{-1}$ and $\Delta t = 0.005 \text{ sec}$. Since $A(t) \leq A(0) = 20$ for any $t \geq 0$, this gives $A(t)k\Delta t \in [0, 0.01]$ for any $t \geq 0$. Consequently, the condition $A(t)k\Delta t \ll 1$ is reasonably satisfied during the simulation. We might further increase the accuracy of the SSA (1a)-(1b) by decreasing Δt . However, decreasing Δt increases the computational intensity of the algorithm. The probability that the reaction (1.1) occurs during the time interval $[t, t + \Delta t]$ is less than or equal to 1% for our parameter values. Thus

during most time-steps, we generate a random number, r , in step **(1a)** only to find out that no reaction occurs in step **(1b)**: we need to generate a lot of random numbers before the reaction takes place. This naive SSA is extremely inefficient: we can do a lot better.

The key to improving the algorithm is a change in viewpoint. Instead of focussing on time we focus on events: rather than stepping forward in time and asking did a reaction take place, we ask at what time will the next reaction occur?

If the time now is t , our goal is to compute the time, $t + \tau$, when the next reaction (1.1) takes place. Of course, τ is a random variable, so that, in fact, what we need to calculate is its probability distribution function. Let us denote by $f(s; A(t))ds$ the probability that, given there are $A(t)$ particles in the system at time t , the next reaction occurs during the time interval $[t+s, t+s+ds]$ where ds is an (infinitesimally) small time-step. In order for this to happen, there must have been no reaction during the interval $[t, t+s]$, and then a reaction must have occurred during the interval $[t+s, t+s+ds]$. Thus, if we let $g(s; A(t))$ be the probability that no reaction occurs in interval $[t, t+s]$, given there are $A(t)$ particles in the system at time t , the probability $f(s; A(t))ds$ can be computed as a product of $g(s; A(t))$ and $A(t+s)kds$:

$$f(s; A(t))ds = g(s; A(t))A(t+s)kds + O(ds^2).$$

Since no reaction occurs in $[t, t+s]$, we have $A(t+s) = A(t)$, so that in fact

$$f(s; A(t))ds = g(s; A(t))A(t)kds + O(ds^2). \quad (1.2)$$

It remains for us to calculate $g(s; A(t))$. For any $\sigma > 0$, the probability that no reaction occurs in the interval $[t, t+\sigma+d\sigma]$ can be computed as the product of the probability that no reaction occurs in the interval $[t, t+\sigma]$ and the probability that no reaction occurs in the interval $[t+\sigma, t+\sigma+d\sigma]$. Hence

$$g(\sigma + d\sigma; A(t)) = g(\sigma; A(t))[1 - A(t+\sigma)kd\sigma] + O(d\sigma^2).$$

Since no reaction occurs in the interval $[t, t+\sigma]$, we have $A(t+\sigma) = A(t)$. Consequently, after some rearrangement,

$$\frac{g(\sigma + d\sigma; A(t)) - g(\sigma; A(t))}{d\sigma} = -A(t)kg(\sigma; A(t)) + O(d\sigma).$$

Passing to the limit $d\sigma \rightarrow 0$, we obtain the ordinary differential equation (in the σ variable)

$$\frac{dg(\sigma; A(t))}{d\sigma} = -A(t)kg(\sigma; A(t)).$$

Solving this equation with initial condition $g(0; A(t)) = 1$, we obtain

$$g(\sigma; A(t)) = \exp[-A(t)k\sigma].$$

Now, in the limit $ds \rightarrow 0$, equation (1.2) can be written as

$$f(s; A(t)) = A(t)k \exp[-A(t)ks]. \quad (1.3)$$

Thus we have found the probability density function for the time interval to the next reaction, τ . To use this in our simulation algorithm we need to generate random numbers τ distributed according to (1.3). The easiest way to accomplish this is to use the following auxiliary function

$$F(\tau) = \exp[-A(t)k\tau]. \quad (1.4)$$

The function $F(\cdot)$ is monotone decreasing for $A(t) > 0$. If τ is a random number from the interval

$(0, \infty)$, then $F(\tau)$ is a random number from the interval $(0, 1)$. Moreover, if τ is a random number distributed according to the probability density function (1.3), then $F(\tau)$ is a random number uniformly distributed in the interval $(0, 1)$, which can be shown as follows.

Let $a < b$, be chosen arbitrarily in the interval $(0, 1)$. The probability that $F(\tau) \in (a, b)$ is equal to the probability that $\tau \in (F^{-1}(b), F^{-1}(a))$ which is given by the integral of $f(A(t), s)$ over s in the interval $(F^{-1}(b), F^{-1}(a))$. Using (1.3) and (1.4), we obtain

$$\begin{aligned} \int_{F^{-1}(b)}^{F^{-1}(a)} f(A(t), s) ds &= \int_{F^{-1}(b)}^{F^{-1}(a)} A(t)k \exp[-A(t)ks] ds \\ &= - \int_{F^{-1}(b)}^{F^{-1}(a)} \frac{dF}{ds} ds = -F[F^{-1}(a)] + F[F^{-1}(b)] = b - a. \end{aligned}$$

Thus, if we have an algorithm which generates a random number, r , uniformly distributed on $(0, 1)$, we can generate the time of the next reaction by setting

$$r = F(\tau) = \exp[-A(t)k\tau].$$

Solving for τ , we obtain the formula

$$\tau = \frac{1}{A(t)k} \ln \left[\frac{1}{r} \right]. \quad (1.5)$$

Consequently, the improved SSA for the chemical reaction (1.1) can be written as follows. Starting with $t = 0$ and $A(0) = n_0$, we perform three steps at time t :

Algorithm 2

- (2a) Generate a random number, r , uniformly distributed in the interval $(0, 1)$.
 - (2b) Compute the time when the next reaction (1.1) occurs as $t + \tau$ where τ is given by (1.5).
 - (2c) Compute the number of particles at time $t + \tau$ by $A(t + \tau) = A(t) - 1$.
- Then continue with step (2a) for time $t + \tau$.

Steps (2a)-(2c) are repeated until we reach the time when there is no particle of A in the system, i.e. $A = 0$. The SSA (2a)-(2c) computes the time of the next reaction $t + \tau$ using formula (1.5) in step (2b) with the help of one random number only. Then the reaction is performed in step (2c) by decreasing the number of particles of chemical species A by 1. The time evolution of A obtained by the SSA (2a)-(2c) is given in Figure 1.1 (b). We plot ten realisations of the SSA (2a)-(2c) for $k = 0.1 \text{ sec}^{-1}$ and $A(0) = 20$. Since the function $A(t)$ has only integer values $\{0, 1, 2, \dots, 20\}$, it is not surprising that some of the computed curves $A(t)$ partially overlap. On the other hand, all ten realisations yield different functions $A(t)$. Even if we made millions of stochastic realisations, it would be very unlikely (with probability zero) that there would be two realisations giving exactly the same results. Therefore, the details of one realisation $A(t)$ are of no special interest (they depend on the sequence of random numbers obtained from the random number generator). However, averaging values of A at time t over many realisations (that is, computing the stochastic mean of A), we obtain a reproducible characteristic of the system - see the thick red line in Figure 1.1 (b). The mean of $A(t)$ over (infinitely) many realisations can also be computed theoretically as follows.

Let us denote by $p_n(t)$ the probability that there are n particles of A at time t in the system, i.e. $A(t) = n$. Let us consider an (infinitesimally) small time-step dt chosen such that the probability

that two particles are degraded during $[t, t + dt]$ is negligible compared to the probability that only one particle is degraded during $[t, t + dt]$. Then there are two possible ways for $A(t + dt)$ to take the value n : either $A(t) = n$ and no reaction occurred in $[t, t + dt]$, or $A(t) = n + 1$ and one particle was degraded in $[t, t + dt]$, i.e.

$$p_n(t + dt) = p_n(t) \times (1 - kn dt) + p_{n+1}(t) \times k(n + 1)dt + O(dt^2).$$

A simple algebraic manipulation yields

$$\frac{p_n(t + dt) - p_n(t)}{dt} = k(n + 1)p_{n+1}(t) - knp_n(t) + O(dt).$$

Passing to the limit $dt \rightarrow 0$, we obtain the so-called *chemical master equation* in the form

$$\frac{dp_n}{dt} = k(n + 1)p_{n+1} - knp_n. \quad (1.6)$$

Equation (1.6) looks like an infinite system of ordinary differential equations (ODEs) for $p_n, n = 0, 1, 2, 3, \dots$. However, our initial condition $A(0) = n_0$ implies that there are never more than n_0 particles in the system, so that $p_n \equiv 0$ for $n > n_0$ and the system (1.6) reduces to a system of $(n_0 + 1)$ ODEs for $p_n, n \leq n_0$. The equation for $n = n_0$ reads as follows

$$\frac{dp_{n_0}}{dt} = -kn_0p_{n_0}.$$

Solving this equation with initial condition $p_{n_0}(0) = 1$, we get $p_{n_0}(t) = \exp[-kn_0t]$. Using this formula in the chemical master equation (1.6) for $p_{n_0-1}(t)$, we obtain

$$\frac{d}{dt}p_{n_0-1}(t) = kn_0 \exp[-kn_0t] - k(n_0 - 1)p_{n_0-1}(t).$$

Solving this equation with initial condition $p_{n_0-1}(0) = 0$, we obtain $p_{n_0-1}(t) = \exp[-k(n_0 - 1)t]n_0(1 - \exp[-kt])$. Using mathematical induction, it is possible to show (see Problem Sheet 1)

$$p_n(t) = \exp[-knt] \binom{n_0}{n} \{1 - \exp[-kt]\}^{n_0-n}, \quad (1.7)$$

where

$$\binom{n_0}{n} = \frac{n_0!}{n!(n_0 - n)!},$$

is the binomial coefficient. The formula (1.7) provides all the information about the stochastic process which is defined by (1.1) with the initial condition $A(0) = n_0$. We can never say for sure that $A(t) = n$; we can only say that $A(t) = n$ with probability $p_n(t)$. In particular, formula (1.7) can be used to derive a formula for the mean value of $A(t)$ over (infinitely) many realisations, which is defined by

$$M(t) = \sum_{n=0}^{n_0} np_n(t).$$

Using (1.7), we deduce

$$\begin{aligned}
 M(t) &= \sum_{n=0}^{n_0} np_n(t) = \sum_{n=0}^{n_0} n \exp[-knt] \binom{n_0}{n} \{1 - \exp[-kt]\}^{n_0-n}, \\
 &= n_0 \exp[-kt] \sum_{n=1}^{n_0} \binom{n_0-1}{n-1} \{1 - \exp[-kt]\}^{(n_0-1)-(n-1)} \{\exp[-kt]\}^{n-1}, \\
 &= n_0 \exp[-kt].
 \end{aligned} \tag{1.8}$$

where the last step follows from the binomial theorem applied to the identity $(\{1 - \exp[-kt]\} + \exp[-kt])^{n_0-1} = 1$. The chemical master equation (1.6) and its solution (1.7) can also be used to quantify the stochastic fluctuations around the mean value (1.8), i.e. how much can an individual realisation of the SSA (2a)-(2c) differ from the mean value given by (1.8). We will present the corresponding theory and results on a more complicated illustrative example in Section 2.1.

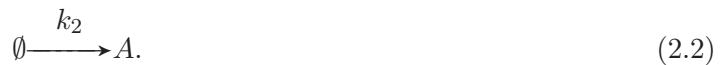
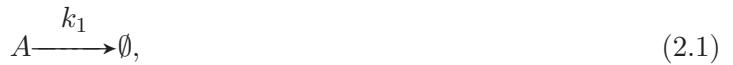
Finally, let us note that a classical deterministic description of the chemical reaction (1.1) is given by the ODE $da/dt = -ka$. Solving this equation with initial condition $a(0) = n_0$, we obtain the function (1.8), i.e. the stochastic mean is equal to the solution of the corresponding deterministic ODE. However, we should emphasise that this is not true for general systems of chemical reactions, as we will see in Lectures 5, 6, 7 and 8.

Lecture 2 - Stochastic simulation of chemical reactions II

We continue to introduce basic concepts using a simple production and degradation - reaction (2.2). We will build on this material in Lecture 3 when we discuss stochastic methods for modelling spatially homogeneous systems of chemical reactions.

2.1 Stochastic simulation of production and degradation

The reaction (1.1) eventually leads to the elimination of all particles of A. We now make it more interesting by introducing a reaction which produces A. Thus let us suppose that we have a chemical species A in a container of volume ν which is subject to the following two chemical reactions



The first reaction (2.1) describes the degradation of chemical A with the rate constant k_1 previously studied. We couple it with reaction (2.2) which represents the production of chemical A with the rate constant k_2 per unit volume. The exact meaning of reaction (2.2) is that one particle of A is created during the time interval $[t, t + dt]$ with probability $k_2\nu dt$ where ν is the system volume. As before, the symbol \emptyset denotes chemical species which are of no special interest to us. The impact of other chemical species on the rate of production of A is assumed to be time independent and is already incorporated in the rate constant k_2 .

The rate constants k_1 and k_2 have different physical units. The rate constant k_1 is expressed in the units of $[\text{sec}^{-1}]$. The units of the rate constant k_2 are $[\text{m}^{-3} \text{sec}^{-1}]$. It is the production rate per unit of volume and per unit of time, so that the probability that one particle of A is created during the time interval $[t, t + dt]$ is equal to $k_2\nu dt$. The scaling with the volume ν is natural: if we divide the container into two equal parts, the production rate in each part will be half of the production rate in the whole container. In this section, the scaling with the system volume ν is not crucial: to simulate the production of particles in a container of the volume ν , we do not need to specify k_2 and ν individually but only the product $k_2\nu$, which is the global production rate (with units $[\text{sec}^{-1}]$). The scaling of the reaction rates with the volume ν will be more important in later lectures, when we consider spatially inhomogeneous systems.

To simulate the system of chemical reactions (2.1)-(2.2), we perform the following four steps at time t (starting with $A(0) = n_0$ at time $t = 0$):

Algorithm 3

- (3a) Generate two random numbers r_1, r_2 uniformly distributed in $(0, 1)$.
 (3b) Compute $\alpha_0 = A(t)k_1 + k_2\nu$.
 (3c) Compute the time when the next chemical reaction takes place as $t + \tau$ where

$$\tau = \frac{1}{\alpha_0} \ln \left[\frac{1}{r_1} \right]. \quad (2.3)$$

- (3d) Compute the number of particles at time $t + \tau$ by

$$A(t + \tau) = \begin{cases} A(t) + 1 & \text{if } r_2 < k_2 \nu / \alpha_0; \\ A(t) - 1 & \text{if } r_2 \geq k_2 \nu / \alpha_0. \end{cases}$$

Then continue with step **(3a)** for time $t + \tau$.

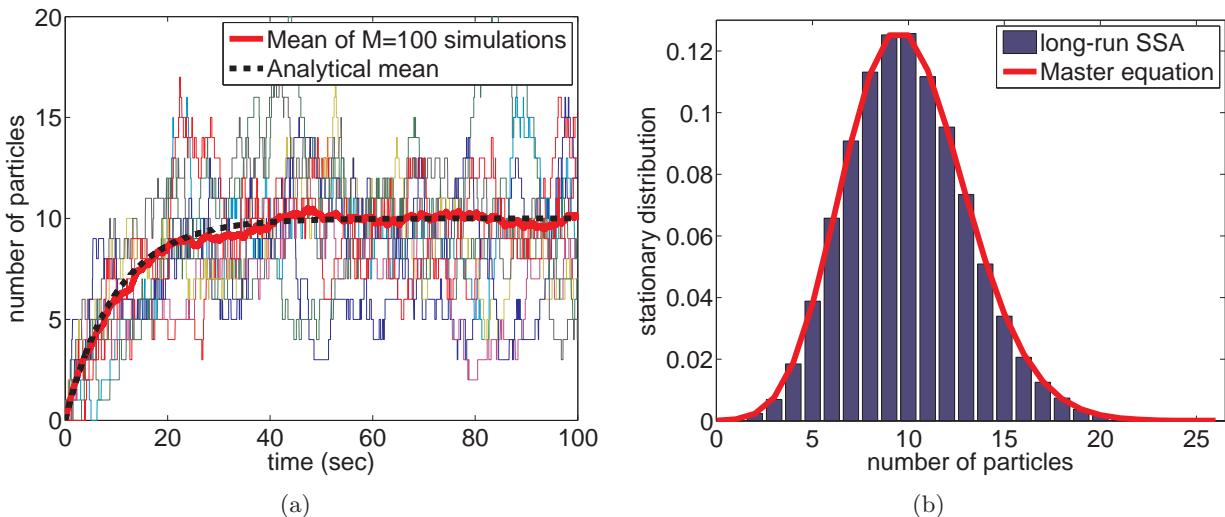


Figure 2.1: Stochastic simulation of the system of chemical reactions (2.1)-(2.2) for $A(0) = 0$, $k_1 = 0.1\text{sec}^{-1}$ and $k_2\nu = 1\text{sec}^{-1}$ (a) $A(t)$ given by five realisations of the SSA (3a)-(3d) (thin solid coloured lines), the mean averaged over 100 repeat simulations (thick red solid line) and stochastic mean (dashed black line). (b) Stationary distribution, $\phi(n)$, obtained by long-time simulation of the SSA (3a)-(3d) (blue histogram) and by formulae (2.14)-(2.15) (red solid line).

To justify that the SSA **(3a)-(3d)** correctly simulates (2.1)-(2.2), let us note that the probability that any of the reactions (2.1)-(2.2) takes place in the time interval $[t, t+dt]$ is equal to $\alpha_0 dt$, which is the sum of the probability that the first reaction occurs, $A(t)k_1 dt$, and the probability that the second reaction occurs, $k_2\nu dt$ (since the probability that both reactions occur is infinitesimal). Formula (2.3) gives the time $t+\tau$ when the next reaction takes place; it can be justified using the same arguments as for the formula (1.5). Once we know the time $t+\tau$, whether that reaction is production or degradation depends only on the relative probabilities of the two reactions: a particle is produced with probability $k_2\nu/\alpha_0$ and degraded otherwise. The decision as to which reaction takes place is given in step **(3d)** with the help of the second uniformly distributed random number r_2 .

Ten realisations of the SSA (3a)-(3d) are presented in Figure 2.1 (a) as solid lines. We plot the number of particles of A as a function of time for $A(0) = 0$, $k_1 = 0.1 \text{ sec}^{-1}$ and $k_2\nu = 1 \text{ sec}^{-1}$. We see that, after an initial transient, the number of particles $A(t)$ fluctuates around its mean value.

To compute the mean and quantify the stochastic fluctuations, we again use the chemical master equation.

As before, let $p_n(t)$ denote the probability that $A(t) = n$ for $n = 0, 1, 2, 3, \dots$. This time there are three ways we can arrive at n particles at time $t + dt$: there could have been n particles at time t and no reactions happened, or there could have been $n + 1$ particles and one was degraded, or there could have been $n - 1$ particles and one was produced. Thus

$$\begin{aligned} p_n(t + dt) &= p_n(t) \times (1 - k_1 n dt - k_2 \nu dt) \\ &\quad + p_{n+1}(t) \times k_1(n + 1)dt + p_{n-1}(t) \times k_2 \nu dt + O(dt^2). \end{aligned}$$

Rearranging and passing to the limit $dt \rightarrow 0$ gives

$$\frac{dp_n}{dt} = k_1(n + 1)p_{n+1} - k_1 n p_n + k_2 \nu p_{n-1} - k_2 \nu p_n. \quad (2.4)$$

Equation (2.4) needs to be slightly modified in the case $n = 0$: it is not possible to arrive at zero particles at time $t + dt$ by having -1 particles at time t and producing one particle! This means that the third term on the right hand side of (2.4) is missing for $n = 0$. To save ourselves the bother of writing a separate equation for the $n = 0$ case we can continue to use (2.4) if we adopt the convention that $p_{-1} \equiv 0$.

The mean, $M(t)$, and variance, $V(t)$, are defined by

$$M(t) = \sum_{n=0}^{\infty} np_n(t), \quad V(t) = \sum_{n=0}^{\infty} (n - M(t))^2 p_n(t). \quad (2.5)$$

The mean $M(t)$ gives the average number of particles of A at time t , while the variance $V(t)$ describes the fluctuations. In Section 1.1, we first solved the chemical master equation (1.6) and then we used its solution (1.7) to compute $M(t)$. An alternative approach is to try to use the chemical master equation to derive an evolution equation for $M(t)$; if we can do this we could find $M(t)$ without having to first solve the chemical master equation. Let us see how such an approach might be carried out.

Multiplying (2.4) by n and summing over n , we obtain

$$\begin{aligned} \frac{d}{dt} \sum_{n=0}^{\infty} np_n &= k_1 \sum_{n=0}^{\infty} n(n + 1)p_{n+1} - k_1 \sum_{n=0}^{\infty} n^2 p_n \\ &\quad + k_2 \nu \sum_{n=1}^{\infty} np_{n-1} - k_2 \nu \sum_{n=0}^{\infty} np_n. \end{aligned}$$

Using definition (2.5) on the left hand side and changing indices $n + 1 \rightarrow n$ (resp. $n - 1 \rightarrow n$) in the first (resp. third) sum on the right hand side, we obtain

$$\frac{dM}{dt} = k_1 \sum_{n=0}^{\infty} (n - 1)np_n - k_1 \sum_{n=0}^{\infty} n^2 p_n + k_2 \nu \sum_{n=0}^{\infty} (n + 1)p_n - k_2 \nu \sum_{n=0}^{\infty} np_n.$$

Now combining the first and the second sums, and the third and fourth sums, gives

$$\frac{dM}{dt} = -k_1 \sum_{n=0}^{\infty} np_n + k_2 \nu \sum_{n=0}^{\infty} p_n. \quad (2.6)$$

Since $p_n(t)$ is the probability that $A(t) = n$ and $A(t)$ is equal to a non-negative integer with probability

1, we have

$$\sum_{n=0}^{\infty} p_n(t) = 1. \quad (2.7)$$

Using this fact together with the definition of $M(t)$, equation (2.6) implies an evolution equation for $M(t)$ in the form

$$\frac{dM}{dt} = -k_1 M + k_2 \nu. \quad (2.8)$$

The solution of (2.8) with initial condition $M(0) = 0$ is

$$M(t) = \frac{k_2 \nu (1 - \exp(-k_1 t))}{k_1}. \quad (2.9)$$

and is plotted as a dashed line in Figure 2.1 (a).

To derive the evolution equation for the variance, $V(t)$, let us first observe that definition (2.5) implies

$$\sum_{n=0}^{\infty} n^2 p_n(t) = V(t) + M(t)^2. \quad (2.10)$$

Multiplying (2.4) by n^2 and summing over n , we obtain

$$\begin{aligned} \frac{d}{dt} \sum_{n=0}^{\infty} n^2 p_n &= k_1 \sum_{n=0}^{\infty} n^2 (n+1) p_{n+1} - k_1 \sum_{n=0}^{\infty} n^3 p_n \\ &\quad + k_2 \nu \sum_{n=1}^{\infty} n^2 p_{n-1} - k_2 \nu \sum_{n=0}^{\infty} n^2 p_n. \end{aligned}$$

Changing indices $n+1 \rightarrow n$ (resp. $n-1 \rightarrow n$) in the first (resp. third) sum on the right hand side and combining the first and the second sum (resp. the third and the fourth sum) on the right hand side gives

$$\frac{d}{dt} \sum_{n=0}^{\infty} n^2 p_n = k_1 \sum_{n=0}^{\infty} (-2n^2 + n) p_n + k_2 \nu \sum_{n=0}^{\infty} (2n+1) p_n.$$

Using (2.10), (2.7) and (2.5), we obtain

$$\frac{dV}{dt} + 2M \frac{dM}{dt} = -2k_1[V + M^2] + k_1 M + 2k_2 \nu M + k_2 \nu.$$

Substituting (2.8) for dM/dt , we derive the evolution equation for the variance $V(t)$ in the following form

$$\frac{dV}{dt} = -2k_1 V + k_1 M + k_2 \nu. \quad (2.11)$$

The time evolution of $M(t)$ and $V(t)$ are described by (2.8) and (2.11). Let us define the stationary values of $M(t)$ and $V(t)$ by

$$M_s = \lim_{t \rightarrow \infty} M(t), \quad V_s = \lim_{t \rightarrow \infty} V(t). \quad (2.12)$$

The values of M_s and V_s can be computed using the steady state equations corresponding to (2.8) and (2.11), namely by solving

$$0 = -k_1 M_s + k_2 \nu$$

and

$$0 = -2k_1 V_s + k_1 M_s + k_2 \nu.$$

Consequently,

$$M_s = V_s = \frac{k_2 \nu}{k_1}.$$

For our parameter values $k_1 = 0.1 \text{ sec}^{-1}$ and $k_2\nu = 1 \text{ sec}^{-1}$, we obtain $M_s = V_s = 10$. We see in Figure 2.1 (a) that $A(t)$ fluctuates after a sufficiently long time around the mean value $M_s = 10$. To quantify the fluctuations, one often uses the square root of V_s , the so-called standard deviation which in this case is equal to $\sqrt{10}$.

More detailed information about the fluctuations is given by the so-called *stationary distribution*, $\phi(n)$, $n = 0, 1, 2, 3, \dots$, which is defined as

$$\phi(n) = \lim_{t \rightarrow \infty} p_n(t). \quad (2.13)$$

This means that $\phi(n)$ is the probability that $A(t) = n$ after an (infinitely) long time. One way to compute $\phi(n)$ is to run the SSA (3a)-(3d) for a long time and create a histogram of values of $A(t)$ at given time intervals. Using $k_1 = 0.1 \text{ sec}^{-1}$ and $k_2\nu = 1 \text{ sec}^{-1}$, the results of such a long-time computation are presented in Figure 2.1 (b) as a blue histogram. To compute it, we ran the SSA (3a)-(3d) for 10^5 seconds, recording the value of $A(t)$ every second and then divided the whole histogram by the number of recordings, i.e. by 10^5 . An alternative way to compute $\phi(n)$ is to use the steady state version of the chemical master equation (2.4), namely

$$\begin{aligned} 0 &= k_1\phi(1) - k_2\nu\phi(0), \\ 0 &= k_1(n+1)\phi(n+1) - k_1n\phi(n) + k_2\nu\phi(n-1) - k_2\nu\phi(n), \text{ for } n \geq 1, \end{aligned}$$

which implies

$$\phi(1) = \frac{k_2\nu}{k_1}\phi(0), \quad (2.14)$$

$$\phi(n+1) = \frac{1}{k_1(n+1)}[k_1n\phi(n) + k_2\nu\phi(n) - k_2\nu\phi(n-1)], \quad (2.15)$$

for $n \geq 1$. By iterating using (2.15) we can express $\phi(n)$ for any $n \geq 1$ in terms of $\phi(0)$. This remaining constant is then fixed by applying the normalisation condition

$$\sum_{n=0}^{\infty} \phi(n) = 1, \quad (2.16)$$

which follows from (2.7) and (2.13). This enables us to compute $\phi(n)$, as follows. First set $\phi(0) = 1$ and compute $\phi(n)$, for sufficiently many n . Then normalise by dividing each $\phi(n)$ by $\sum \phi(n)$. The results obtained by (2.14)-(2.15) are plotted in Figure 2.1 (b) as a (red) solid line. As expected, the results compare well with the results obtained by the long-time stochastic simulation.

In fact, the recurrence relations (2.14)-(2.15) are sufficiently simple that we can find the formula for $\phi(n)$ directly. We leave it as an exercise (Problem Sheet 1) to verify that the solution of the (2.14)-(2.15) can be written as

$$\phi(n) = \frac{C}{n!} \left(\frac{k_2\nu}{k_1} \right)^n, \quad (2.17)$$

where C is a real constant. Substituting (2.17) into the normalisation condition (2.16), we get

$$1 = \sum_{n=0}^{\infty} \frac{C}{n!} \left(\frac{k_2\nu}{k_1} \right)^n = C \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{k_2\nu}{k_1} \right)^n = C \exp \left[\frac{k_2\nu}{k_1} \right],$$

where we used the Taylor series for the exponential function to get the last equality. Consequently,

$$C = \exp \left[-\frac{k_2\nu}{k_1} \right],$$

which, together with (2.17), implies that the stationary distribution, $\phi(n)$, is the Poisson distribution

$$\phi(n) = \frac{1}{n!} \left(\frac{k_2 \nu}{k_1} \right)^n \exp \left[-\frac{k_2 \nu}{k_1} \right]. \quad (2.18)$$

Thus the red solid line in Figure 2.1 (b) which was obtained numerically by the recurrence formula (2.14)-(2.15) can also be viewed as the stationary distribution $\phi(n)$ given by the explicit exact formula (2.18).

Lecture 3 - Higher-order chemical reactions

The best way to learn stochastic simulation methods is to attempt to implement examples from the Lecture Notes and Problem Sheets using a computer language of your choice. One option is to use **Matlab** which every Bath student can use for free (see <http://www.bath.ac.uk/bucs/tools/software/Matlab/>). The choice of programming language is not important, if you are more familiar with another language then you should feel free to use this, but **Matlab** provides a relatively straightforward environment for students who are new to programming. Hands-on experience with implementing algorithms from the Lecture Notes in the computer can improve your understanding of the underlying mathematical methods.

In the first half of this lecture, I will show how you can design and write a **Matlab** code which reproduces Figure 2.1. If you try to write your own computer code for Figure 2.1, then it might take you more time and your first result might be different than mine, but this is part of the learning process. If you do not get the answer which you expect, you have to go back to your computer code and break it into small components to find out what was not correctly implemented. If you do not get the correct answer after a few back-and-forth iterations, then it is a good idea to consult with other people on the course or the class tutor. However, please first try to identify why your code does not work. By attempting it, you might actually fix the problem and learn more during this process.

In the second half of this lecture, we will continue in our discussion of stochastic simulations of chemical reactions. We discussed one chemical reaction in Section 1.1 and a system of two chemical reactions in Section 2.1. In Section 3.3 we will generalise stochastic simulation methods to arbitrary systems of chemical reactions.

3.1 Implementation of Figure 2.1 in Matlab

We start with Figure 2.1 (a). It shows results of five realisations of the SSA (3a)-(3d) together with the mean $M(t)$ (dashed line). Let us start with the dashed line. It is the solution of equation (2.8) with the initial condition $M(0) = 0$. This equation can be solved analytically to obtain

$$M(t) = \frac{k_2\nu}{k_1}(1 - \exp(-k_1 t)). \quad (3.1)$$

Therefore we can plot the time evolution of $M(t)$ using the **Matlab** code in Scheme 3.1:

```
k1=0.1;          1
k2nu=1;          2
t=[0:0.2:100];  3
M=(k2nu/k1)*(1-\exp(-k1*t)); 4
plot (t,M);      5
```

Scheme 3.1: Matlab code for plotting the mean value of $A(t)$ according to equation (3.1).

The first two lines specify the values of parameters $k_1 = 0.1$ and $k_2\nu = 1$. The third line defines the values of time t : it is a vector with entries $[0, 0.2, 0.4, 0.8, 1, 1.2, 1.4, \dots, 99.6, 99.8, 100]$. The fourth

line is the equation (2.16) and the last line asks Matlab to plot M against time t . The above code correctly plots $M(t)$ but the resulting plot looks slightly different from Figure 2.1 (a). To get Figure 2.1 (a), I substituted the simple plotting command `plot(t, M)` with a few more lines to improve the output. The resulting code is given in scheme 3.2:

```

k1=0.1;                                1
k2nu=1;                                 2
t=[0:0.2:100];                          3
M=(k2nu/k1)*(1-\exp(-k1*t));          4
close all;                               5
figure(1);                             6
set(gca,'Fontsize',20);                7
plot(t,M,'--k','Linewidth',4);         8
axis([0 100 0 20]);                   9
xlabel('time, [sec]');                 10
ylabel('number of particles');          11

```

Scheme 3.2: Matlab code for plotting the mean value of $A(t)$ according to equation (3.1).

The output given by this code is shown in Figure 3.1 (a). The first four lines are the same as before and they do the actual computation. The fifth and sixth lines are not important to get Figure 3.1 (a), but I included them because you might find them useful later. The command `close all` on the fifth line will close all figure windows, if you have some figure windows open from the previous run of the code. The sixth line, `figure(1)`, specifies that we will plot into “Figure 1” (it is useful if the output of your code includes more than one figure). The seventh line specifies the font size of the axis labels (and other text) in the figure. The plotting command has changed from simple `plot(t,M);` to `plot(t,M,'--k','Linewidth',4);`. It includes a few options, specifying that we will use a black dashed line and its width. The ninth line says that we will plot the interval $[0, 100]$ on the time axis and the range of M will be interval $[0, 20]$. Finally, the last two lines specify labels on each axis.

Next, we implement the SSA (3a)-(3d) in Matlab. We start by initialising our main variables `A` (number of particles) and time, `t`. We also introduce vectors `Aplot` and `timeplot` which will be used for storing and plotting our results. Thus the initialisation part of our code is:

```

A=0;                                     1
time=0;                                   2
Aplot=A;                                  3
timeplot=time;                            4

```

Scheme 3.3: Matlab code initialising variable for the SSA.

We need to sample random numbers which are uniformly distributed in $(0, 1)$. In Matlab, we can use the command `rand(2, 1)` which gives us a vector of two random numbers `r(1)` and `r(2)` which are needed during each iteration of the SSA (3a)-(3d). It is good practice to initialise the random number generator. Matlab uses the following command:

```
rand('state',5);
```

Then we get the same sequence of “random numbers” whenever we run the Matlab code. This is very helpful whenever you try to identify errors in your computer code which might otherwise be difficult to find.

We add the red lines after the blue lines. In particular, the values of constants k_1 and k_2 are already specified in the blue part of the Matlab code. In each iteration, we will store the number of

particles and the corresponding time in vectors `Aplot` and `timeplot`. Typing `Aplot=[Aplot A]` or `timeplot =[timeplot time]`, Matlab adds the current value of `A` or time as the last entry of the vector where we store our results. This construction is then used in the main loop of our computer implementation of the SSA (3a)-(3d):

```

while (time<100)                                1
    r=rand(2,1);                                 2
    a0=k1*A+k2nu;                               3
    time=time+(1/a0)*log(1/r(1));               4
    if (r(2)*a0<k1*A)                         5
        A=A-1;                                    6
    else                                         7
        A=A+1;                                    8
    end;                                       9
    Aplot=[Aplot A];                            10
    timeplot=[timeplot time];                   11
end;                                         12

```

Scheme 3.4: Matlab code for running and storing the result of SSA (3a)-(3d).

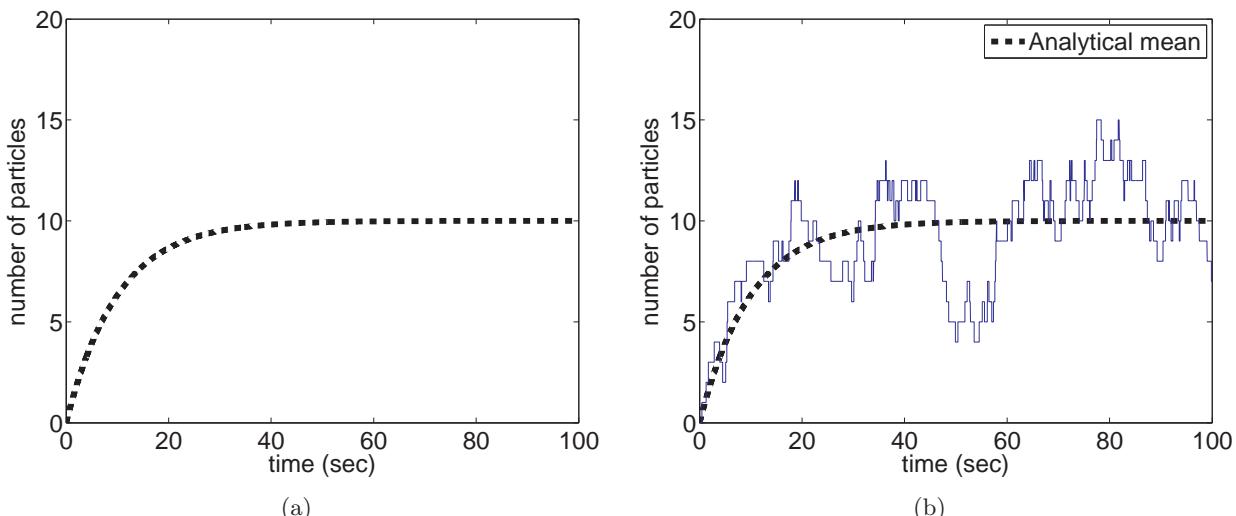


Figure 3.1: (a) Plot of equation (3.1) given by the blue Matlab code in the text. (b) Plot of equation (3.1) (black dashed line) and one realisation of the SSA (3a)-(3d) (thin solid blue line) given by the combined (blue and red) Matlab code in the text.

The results stored in vectors `Aplot` and `timeplot` are plotted using the following commands:

```

hold on;                                         1
h=stairs(timeplot,Aplot);                      2
set(h,'Color','m','Linewidth',1);              3
legend('mean',4);                           4

```

Scheme 3.5: Matlab code for plotting the results of the SSA (3a)-(3d).

where the first command, `hold on`, helps us to plot the computed realisation of the SSA (3a)-(3d) into the same figure as the result of the blue part of our code. To plot `Aplot` against `timeplot`, we use the command `stairs` instead of `plot`. This helps use to emphasise that the value of `A` stays constant

until a reaction occurs. The result computed by the combined blue and red codes is plotted in Figure 3.1 (b).

Finally, we can obtain five realisations as plotted in Figure 2.1 (a) by using the above code five times. We have to make sure that we use different random numbers to get different time evolutions of A. The stationary distribution in Figure 2.1 (b) was plotted using the `Matlab` code in Scheme 3.6.

The code in Scheme 3.6 demonstrates that the computation of the stationary distribution is based on the red computer code. As you can see, we only store what is actually plotted (histogram). We do not store the computed trajectories. Otherwise, we could end up with slower code (if we attempt to store and work with unnecessary large data files).

If “Modelling Stochastic Biological Processes” was being taught to students of Biology, then the emphasis would be on algorithms and their implementations. They would be focussing on solving complex biological problems using the algorithms from the Lecture Notes. However, this course is offered by the Department of Mathematical Sciences and it is not primarily about programming. Our aim is to understand the mathematics which can be found behind the algorithms from the Lecture Notes. Therefore, once you have a code for the corresponding figure in the Lecture Notes, your next question should be whether this code gives you the answer which you expected.

For all problems (figures) in these notes, the objective of your simulations is to get the same answer (figure) as it is presented in the Lecture Notes. However, one can also write suitable mathematical equations and get the same results without doing any simulations, if the problem is reasonably simple. For example, in Figure 2.1 (b), you can see the red line which can be obtained analytically using formula (2.18) or by solving equations (2.14)-(2.15).

Obviously, in real complex biological problems, you do not know *a priori* the answer which you should expect from your model or computer code. You can replace a complex biological model (which you do not fully understand) by a complex computer code. Such a computer code can give you interesting (and perhaps counterintuitive) graphs, but you need some confidence that your computer code is doing what it should do. One option is to break your large code into small subsystems which you fully understand, i.e. small subsystems which you can analytically solve. In this course we aim to study such simple models. In fact, two chemical reactions which form the underlying model behind Figure 2.1 can be found as a simple subsystem in larger biological models.

The figures in the Lecture Notes often show that the same information can be obtained from analysis of suitable differential equations as from stochastic simulations. In particular, there is no need to do simulations for many examples. However, teaching this course without discussing stochastic simulation algorithms could give you the wrong impression that “Modelling Stochastic Biological Processes” is a computer-free subject. In the Lecture Notes, we will focus on simple systems where analysis is possible. Complex biological models can still be analysed by writing computer codes, but their mathematical analysis is complicated or even impossible (unless we start to use some simplifying assumptions). However, even the computer-assisted analysis can be limited, because the resulting simulations can be computationally intensive. This is another area where mathematicians can usefully contribute. We will discuss in Lectures 24-25 a small selection of methods which can be used when computer codes are very slow.

3.2 Higher-order chemical reactions

The chemical reaction (1.1) is an example of a *first-order* reaction, i.e. only one particle is needed for the reaction to take place. Consequently, the rate of the first-order reaction depends linearly on the number of particles of that one reactant. Another important class of chemical reactions are the

```

close all;                                         1
rand('state',10);                                2
k1=0.1;                                         3
k2nu=1;                                         4
A=k2nu/k1;                                       5
numberofsavings=100000;                           6
p=zeros(22,1);                                    7
r=rand(2,1);                                     8
a0=k1*A+k2nu;                                   9
time=(1/a0)*log(1/r(1));                         10
for i=1:numberofsavings                         11
    while (time<i)                             12
        if (r(2)*a0<k1*A)                      13
            A=A-1;                               14
        else                                     15
            A=A+1;                               16
        end;                                    17
        r=rand(2,1);                            18
        a0=k1*A+k2nu;                          19
        time=time+(1/a0)*log(1/r(1));          20
    end;                                      21
    if ((A>0)&(A<23))                      22
        p(A)=p(A)+1;                           23
    end;                                      24
end;                                            25
p=p/numberofsavings;                           26
n=[0:25];                                       27
pCME=1./factorial(n).*(k2nu/k1).^n*exp(-k2nu/k1) 28
figure(1);                                     29
set(gca,'Fontsize',20);                        30
h1=bar(p);                                     31
colormap([0.88 0.88 0.88]);                  32
hold on;                                       33
h2=plot(n,pCME,'r','Linewidth',4);           34
xlabel('number of particles');                 35
ylabel('stationary distribution');             36
legend([h1 h2],'Gillespie SSA','master equation'); 37
axis([0 22.5 0 1.05*max(p)]);                38

```

Scheme 3.6: Matlab code for plotting the stationary distribution (3a)-(3d).

Chemical Reaction	order	propensity function, $\alpha(t)$	units of k
$\emptyset \xrightarrow{k} A$	zeroth-order	$k\nu$	$\text{m}^{-3} \text{ sec}^{-1}$
$A \xrightarrow{k} \emptyset$	first-order	$A(t)k$	sec^{-1}
$A + B \xrightarrow{k} \emptyset$	second-order	$A(t)B(t)k/\nu$	$\text{m}^3 \text{ sec}^{-1}$
$A + A \xrightarrow{k} \emptyset$	second-order	$A(t)(A(t) - 1)k/\nu$	$\text{m}^3 \text{ sec}^{-1}$

Table 3.1: The propensity functions and rates of the basic chemical reactions.

so-called *second-order* (or bimolecular) chemical reactions. A simple example can be written as follows:



Here, one particle of A and one particle of B react, with the rate constant k , to produce a particle of C.

Now we have to think carefully about how the rate of reaction should scale with the volume of the system. We would expect that a reaction between one particle of A and one particle of B is more likely in a small box than in a large box. In fact, we would expect the two particles to collide twice as often in a box which is half the size. Thus for one particle of A and one particle of B the probability of a reaction occurring in the time interval $[t, t + dt]$ is $k/\nu dt$. In particular, this means that the units of the rate constant k of the reaction (3.2) are $[\text{m}^3 \text{ sec}^{-1}]$. Now the number of different pairs of particles A and B is equal to the product $A(t)B(t)$. Consequently, the probability that a reaction (3.2) takes place in the time interval $[t, t + dt]$ between any of these pairs is equal to $A(t)B(t)k/\nu dt$.

The *propensity function*, $\alpha(t)$, of a reaction is defined to be such that the probability that the reaction occurs in the infinitesimally small time interval $[t, t + dt]$ is $\alpha(t)dt$. For the reaction (3.2) we have $\alpha(t) = A(t)B(t)k/\nu$. Other examples of second-order chemical reactions are:



In the chemical reaction 3.3, one particle of A and one particle of B react to produce a particle of C and a particle of D. The propensity function of the first reaction is the same as in the case of the reaction (3.2), i.e. $\alpha(t) = A(t)B(t)k/\nu$. The number of product particles does not influence the characterisation of the order of the chemical reaction, and does not change its propensity function. In the chemical reaction 3.4, two particles of A react with the rate constant k to produce C. To derive the propensity function for this reaction we follow the same reasoning as before. The key difference is that now the number of different pairs of A particles is equal to

$$\binom{A(t)}{2} = \frac{A(t)(A(t) - 1)}{2}.$$

It is common to absorb the factor of 2 into the rate constant, so that the propensity of this reaction is written as $\alpha(t) = A(t)(A(t) - 1)k/\nu$. Note in particular that α is zero if A is zero or one: it takes two particles of A for the reaction to happen.

The units of the rate constant of any second-order chemical reaction are equal to $[\text{m}^3 \text{ sec}^{-1}]$.

Chemical Reaction	order	propensity function, $\alpha(t)$	units of k
$A + B + C \xrightarrow{k} \emptyset$	third-order	$A(t)B(t)C(t)k/\nu^2$	$\text{m}^6 \text{ sec}^{-1}$
$2A + B \xrightarrow{k} \emptyset$	third-order	$A(t)(A(t) - 1)B(t)k/\nu^2$	$\text{m}^6 \text{ sec}^{-1}$
$3A \xrightarrow{k} \emptyset$	third-order	$A(t)(A(t) - 1)(A(t) - 2)k/\nu^2$	$\text{m}^6 \text{ sec}^{-1}$

Table 3.2: The propensity functions and rates of the third-order chemical reactions.

In Table 3.1, we summarise the order of the basic chemical reactions and their propensity functions. Since the number of product particles does not influence the propensities and the order of the chemical reaction, we denote the products by \emptyset for simplicity. We also include the production reaction (2.2) which we call a zero-order chemical reaction.

In a similar way, we can define *third-order* chemical reactions as reactions which require a collision of three particles to take place. In Table 3.2, we specify the propensity functions of three possible types of the third-order reactions. The rate constants of the third-order chemical reactions have units of $[\text{m}^6 \text{ sec}^{-1}]$.

3.3 The Gillespie algorithm

The SSAs (2a)-(2c) and (3a)-(3d) were special forms of the so-called Gillespie SSA [1]. To conclude this lecture, we formulate the Gillespie SSA in its full generality. Let us consider that we have a system of q chemical reactions and N chemical species. Let $\mathbf{X}(t) = (X_1(t), \dots, X_N)$ be the state vector which holds the number of particles of each species at time t . Let $\alpha_i(t)$ be the propensity function of the i -th reaction, $i = 1, 2, \dots, q$, at time t , that is, $\alpha_i(t)dt$ is the probability that the i -th reaction occurs in the time interval $[t, t + dt]$ (the propensity functions for basic chemical reactions are given in Tables 3.1 and 3.2). In general the propensity functions will be functions of the state vector.

If the time now is t , our goal is to compute the time, $t + \tau$, when the next reaction takes place and which reaction, j , it is that occurs. Of course, τ ($0 \leq \tau < \infty$) and j ($j = 1, 2, \dots, q$) are random variables, so that, in fact, what we need to calculate is their joint probability distribution function. Let us denote by $f(s, j; \mathbf{X})ds$ the probability that, given the system is in state $\mathbf{X}(t)$ at time t , the next reaction occurs during the time interval $[t+s, t+s+ds]$ and is of type j , where ds is an (infinitesimally) small time-step. In order for this to happen, there must have been no reaction during the interval $[t, t+s]$, and then a reaction of type j must have occurred during the interval $[t+s, t+s+ds]$. Thus, if we let $g(s; \mathbf{X})$ be the probability that no reaction occurs in interval $[t, t+s]$, given the system is in state $\mathbf{X}(t)$ at time t , the probability $f(s, j; \mathbf{X})ds$ can be computed as a product of $g(s; \mathbf{X})$ and $\alpha_j(t+s)ds$:

$$f(s, j; \mathbf{X})ds = g(s; \mathbf{X})\alpha_j(t+s)ds.$$

Since no reaction occurs in $[t, t+s]$, we have $\alpha_j(t+s) = \alpha_j(t)$ for $j = 1, 2, \dots, q$, so that in fact

$$f(s, j; \mathbf{X})ds = g(s; \mathbf{X})\alpha_j(t)ds. \quad (3.5)$$

It remains for us to calculate $g(s)$. For any $\sigma > 0$, the probability that no reaction occurs in the interval $[t, t + \sigma + d\sigma]$ can be computed as the product of the probability that no reaction occurs in the interval $[t, t + \sigma]$ and the probability that no reaction occurs in the interval $[t + \sigma, t + \sigma + d\sigma]$.

Hence

$$g(\sigma + d\sigma; \mathbf{X}) = g(\sigma; \mathbf{X}) \left[1 - \sum_{i=1}^q \alpha_i(t + \sigma) d\sigma \right].$$

Since no reaction occurs in the interval $[t, t + \sigma]$, we have $\alpha_i(t + \sigma) = \alpha_i(t)$. Consequently, after some rearrangement,

$$\frac{g(\sigma + d\sigma; \mathbf{X}) - g(\sigma; \mathbf{X})}{d\sigma} = - \sum_{i=1}^q \alpha_i(t) g(\sigma; \mathbf{X}).$$

Passing to the limit $d\sigma \rightarrow 0$, we obtain the ordinary differential equation (in the σ variable)

$$\frac{dg(\sigma; \mathbf{X})}{d\sigma} = - \sum_{i=1}^q \alpha_i(t) g(\sigma; \mathbf{X}).$$

Solving this equation with initial condition $g(0; \mathbf{X}) = 1$, we obtain

$$g(\sigma; \mathbf{X}) = \exp \left[- \sum_{i=1}^q \alpha_i(t) \sigma \right].$$

Now equation (3.5) can be written as

$$f(s, j; \mathbf{X}) = \alpha_j(t) \exp \left[- \sum_{i=1}^q \alpha_i(t) s \right] = \alpha_j(t) \exp [-\alpha_0], \quad (3.6)$$

where $\alpha_0 = \sum_{i=1}^q \alpha_i(t)s$.

Note that this probability density function is properly normalised over its domain of definition:

$$\int_0^\infty \sum_{j=1}^q f(s, j; \mathbf{X}) ds = \int_0^\infty \exp [-\alpha_0 s] ds \sum_{j=1}^q \alpha_j(t) = 1. \quad (3.7)$$

From the joint probability distribution given by equation (3.6) it is straight forward to derive equations for the marginal distributions for the time of the next reaction:

$$f_1(s) = \sum_{i=1}^q f(s, i; \mathbf{X}) = \alpha_0 \exp [-\alpha_0 s], \quad (3.8)$$

and the reaction index:

$$f_2(j) = \int_0^\infty f(s, j; \mathbf{X}) ds = \alpha_j(t) \int_0^\infty \exp [-\alpha_0 s] ds = \frac{\alpha_j(t)}{\alpha_0}. \quad (3.9)$$

Consequently the Gillespie SSA consists of the following four steps at time t .

Algorithm 4

(4a) Generate two random numbers r_1, r_2 uniformly distributed in $(0, 1)$.

(4b) Compute the propensity function $\alpha_i(t)$ of each reaction. Compute

$$\alpha_0 = \sum_{i=1}^q \alpha_i(t). \quad (3.10)$$

(4c) Compute the time when the next chemical reaction takes place as $t + \tau$ where

$$\tau = \frac{1}{\alpha_0} \ln \left[\frac{1}{r_1} \right]. \quad (3.11)$$

(4d) Compute which reaction occurs at time $t + \tau$. Find j such that

$$r_2 \geq \frac{1}{\alpha_0} \sum_{i=1}^{j-1} \alpha_i(t) \text{ and } r_2 < \frac{1}{\alpha_0} \sum_{i=1}^j \alpha_i(t).$$

Then the j -th reaction takes place, so update numbers of species according to the j -th reaction.

Continue with step (4a) for time $t + \tau$.

The Gillespie SSA (4a)-(4d) provides an exact method for the stochastic simulation of systems of chemical reactions. It was applied previously as the SSA (2a)-(2c) for the chemical reaction (1.1) and as the SSA (3a)-(3d) for the chemical system (2.1)-(2.2). At each time-step, we first ask the question when will the next reaction occur? The answer is given by formula (3.11) which can be justified using the same arguments as formulae (1.5) or (2.3). Then we ask the question which reaction takes place. The probability that the i -th chemical reaction occurs is given by α_i/α_0 for $i = 1, 2, \dots, q$. The decision on which reaction takes place is given in step (4d) with the help of the second uniformly distributed random number r_2 . Then we update the number of reactants and products accordingly.

Our simple examples can be simulated quickly in **Matlab** (in less than a second on present-day computers). If one considers systems of many chemical reactions and many chemical species, then the SSA (4a)-(4d) might be computationally intensive. In such a case, there are ways to make the Gillespie SSA more efficient [2–6]. For example, it would be a waste of time to recompute all the propensity functions at each time-step (step (4b)). We simulate one reaction per one time-step. Therefore, it makes sense to update only those propensity functions which are changed by the chemical reaction which was selected in step (4d) of the SSA (4a)-(4d).

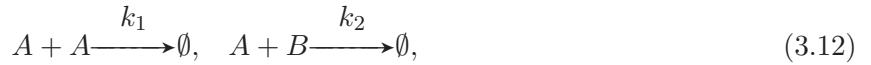
A Research Aside:

There are ways to make the Gillespie SSA more efficient [2–6].

1. Update only those propensity functions changed by the reaction.
2. Sort the list of propensity functions so largest are at the top [3, 5] - reduces linear search depth in step (4d).
3. Search in an efficient manner (e.g. binary search) [6].
4. Reverse the order of steps (4c) and (4d) and re-use the random number [4] (saves on number of random numbers to be generated).
5. Re-use random numbers in a more complex way [2].

In this section, we present the SSA (4a)-(4d) for a more complicated illustrative example which will also involve the second-order chemical reactions (3.3). We consider two chemical species A and B in a container of volume ν . We assume that A and B are subject to the following system of four

chemical reactions:



Let us note that we are not interested in chemical species C and D (which were present as products in the illustrative example (3.3)). Hence, we replaced them by \emptyset , consistent with our previous notation of unimportant chemical species. To simulate the system of chemical reactions (3.12)-(3.13), we apply the SSA (4a)-(4d). To do that, we have to find the propensities of every reaction. Using Table 3.1, we find that the propensities are: $\alpha_1(t) = A(t)(A(t) - 1)k_1/\nu$, $\alpha_2(t) = A(t)B(t)k_2/\nu$, $\alpha_3 = k_3\nu$ and $\alpha_4 = k_4\nu$. In step (4b), we compute $\alpha_0 = \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4$ which is used in the step (4c) to determine the time when the next reaction occurs (formula (3.11)). The decision on which reaction should occur is done in the step (4d) with the help of the second uniformly distributed random number r_2 . In the case of the chemical system (3.12)-(3.13), the step (4d) reads as follows:

$$A(t + \tau) = \begin{cases} A(t) - 2 & \text{if } 0 \leq r_2 < \alpha_1/\alpha_0; \\ A(t) - 1 & \text{if } \alpha_1/\alpha_0 \leq r_2 < (\alpha_1 + \alpha_2)/\alpha_0; \\ A(t) + 1 & \text{if } (\alpha_1 + \alpha_2)/\alpha_0 \leq r_2 < (\alpha_1 + \alpha_2 + \alpha_3)/\alpha_0; \\ A(t) & \text{if } (\alpha_1 + \alpha_2 + \alpha_3)/\alpha_0 \leq r_2 < 1; \\ B(t) & \text{if } 0 \leq r_2 < \alpha_1/\alpha_0; \\ B(t) - 1 & \text{if } \alpha_1/\alpha_0 \leq r_2 < (\alpha_1 + \alpha_2)/\alpha_0; \\ B(t) & \text{if } (\alpha_1 + \alpha_2)/\alpha_0 \leq r_2 < (\alpha_1 + \alpha_2 + \alpha_3)/\alpha_0; \\ B(t) + 1 & \text{if } (\alpha_1 + \alpha_2 + \alpha_3)/\alpha_0 \leq r_2 < 1. \end{cases}$$

Results of five realisations of the SSA (4a)-(4d) are plotted in Figure 3.2 as solid lines. We use $A(0) = 0$, $B(0) = 0$, $k_1/\nu = 10^{-3} \text{ sec}^{-1}$, $k_2/\nu = 10^{-2} \text{ sec}^{-1}$, $k_3\nu = 1.2 \text{ sec}^{-1}$ and $k_4\nu = 1 \text{ sec}^{-1}$. We plot the number of particles of chemical species A and B as functions of time. We see that, after initial transients, $A(t)$ and $B(t)$ fluctuate around their average values. These average values can be estimated, from long-time stochastic simulations, as 9.6 for A and 12.2 for B .

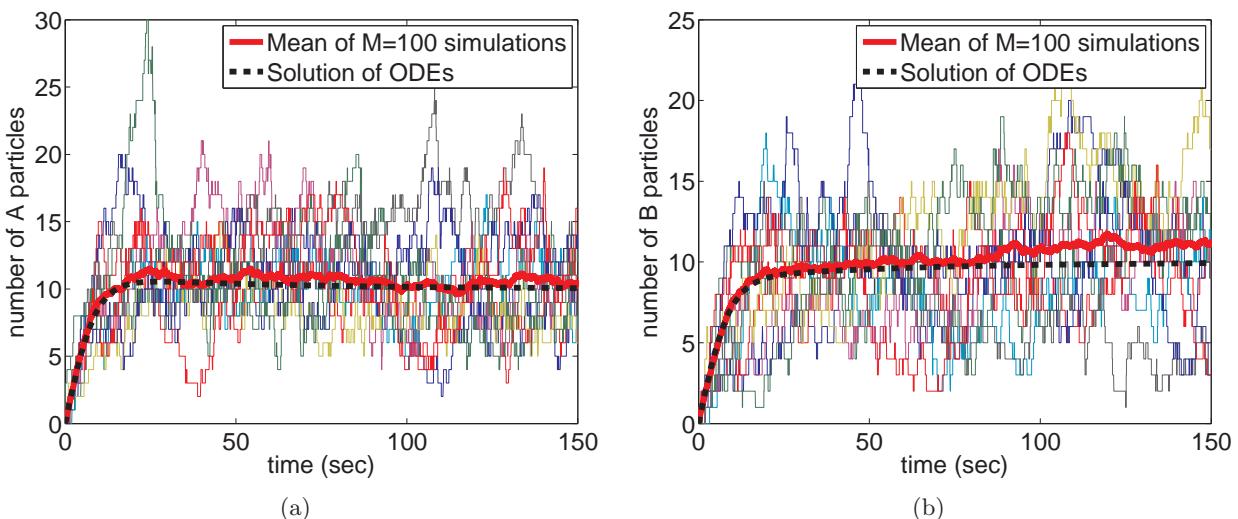


Figure 3.2: Simulations of reaction system (3.12)-(3.13) using the SSA (4a)-(4d). Ten independent repeats for the number of particles of chemical species A (a) and B (b) are plotted as functions of time (thin solid coloured lines). The mean behaviour for A and B averaged over 100 repeats are given by the thick solid red line. The solutions of equations (3.20)-(3.21) are given by the black dashed lines. We use $A(0) = 0$, $B(0) = 0$, $k_1/\nu = 10^{-3} \text{ sec}^{-1}$, $k_2/\nu = 10^{-2} \text{ sec}^{-1}$, $k_3\nu = 1.2 \text{ sec}^{-1}$ and $k_4\nu = 1 \text{ sec}^{-1}$.

Let $p_{n,m}(t)$ be the probability that $A(t) = n$ and $B(t) = m$. The chemical master equation can be written in the following form

$$\begin{aligned}\frac{dp_{n,m}}{dt} = & \frac{k_1}{\nu}(n+2)(n+1)p_{n+2,m} - \frac{k_1}{\nu}n(n-1)p_{n,m} \\ & + \frac{k_2}{\nu}(n+1)(m+1)p_{n+1,m+1} - \frac{k_2}{\nu}nmp_{n,m} \\ & + k_3\nu p_{n-1,m} - k_3\nu p_{n,m} + k_4\nu p_{n,m-1} - k_4\nu p_{n,m},\end{aligned}\quad (3.14)$$

for $n, m \geq 0$, with the convention that $p_{n,m} \equiv 0$ if $n < 0$ or $m < 0$. The difference between (3.14) and the chemical master equations from the previous sections is that equation (3.14) is parametrised by two indices n and m . The probability $p_{n,m}(t)$ is sometimes denoted by $p(n, m, t)$; such a notational convention is often used when we consider systems of many chemical species. We will use it in the following lectures to avoid long subscripts.

With two variables, in order to find the mean number of A particles, $\langle n \rangle$ we must multiply $p_{n,m}(t)$ by n and sum over both n and m :

$$M_A(t) = \langle n \rangle = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} np_{n,m}(t). \quad (3.15)$$

Similarly

$$M_B(t) = \langle m \rangle = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} mp_{n,m}(t). \quad (3.16)$$

In the same way, to calculate higher-order moments like $\langle nm \rangle$ we must multiply by nm and sum over both n and m :

$$\langle nm \rangle = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} nmp_{n,m}(t). \quad (3.17)$$

Since the system contains second-order chemical reactions, we cannot solve (3.14) analytically as we did with (1.6) and the equation (3.14) does not lead to closed evolution equations for the stochastic mean and variance (although see Lectures 9 and 10 for methods to derive approximate closed evolution equations). This means we cannot follow the same technique as in the case of equation (2.4). Such methods are applicable only to zero-order and first-order chemical reactions [7].

The classical deterministic description of the chemical system (3.12)-(3.13) is given for concentrations $a(t) = A(t)/\nu$ and $b(t) = B(t)/\nu$ by the system of ODEs

$$\frac{da}{dt} = -2k_1a^2 - k_2ab + k_3, \quad (3.18)$$

$$\frac{db}{dt} = -k_2ab + k_4. \quad (3.19)$$

The results given by (3.18)-(3.19) are also plotted in Figure 3.2 for comparison. More precisely, we plot

$$\bar{A}(t) = a(t)\nu, \quad \bar{B}(t) = b(t)\nu,$$

which give numbers of particles of chemical species A and B in the volume with concentrations $a(t)$ and $b(t)$, respectively. Multiplying (3.18)-(3.19) by ν , we obtain that $\bar{A}(t)$ and $\bar{B}(t)$ satisfy the system of ODEs

$$\frac{d\bar{A}}{dt} = -2k_1/\nu \bar{A}^2 - k_2/\nu \bar{A} \bar{B} + k_3\nu, \quad (3.20)$$

$$\frac{d\bar{B}}{dt} = -k_2/\nu \bar{A} \bar{B} + k_4\nu. \quad (3.21)$$

The solution of (3.20)-(3.21) with initial conditions $\bar{A}(0) = 0$ and $\bar{B}(0) = 0$ is plotted using dashed lines in Figure 3.2. Let us note that the equations (3.20)-(3.21) do not describe the average values of $A(t)$ and $B(t)$. For example, the steady state values of (3.20)-(3.21) are (for the parameter values of Figure 5.2) equal to $\bar{A}_s = \bar{B}_s = 10$. On the other hand, the average values estimated from long-time stochastic simulations are 9.6 particles for A and 12.2 particles for B . We will see in Lectures 5, 6, 7 and 8 that the difference between the results of stochastic simulations and the corresponding ODEs can be even more significant.

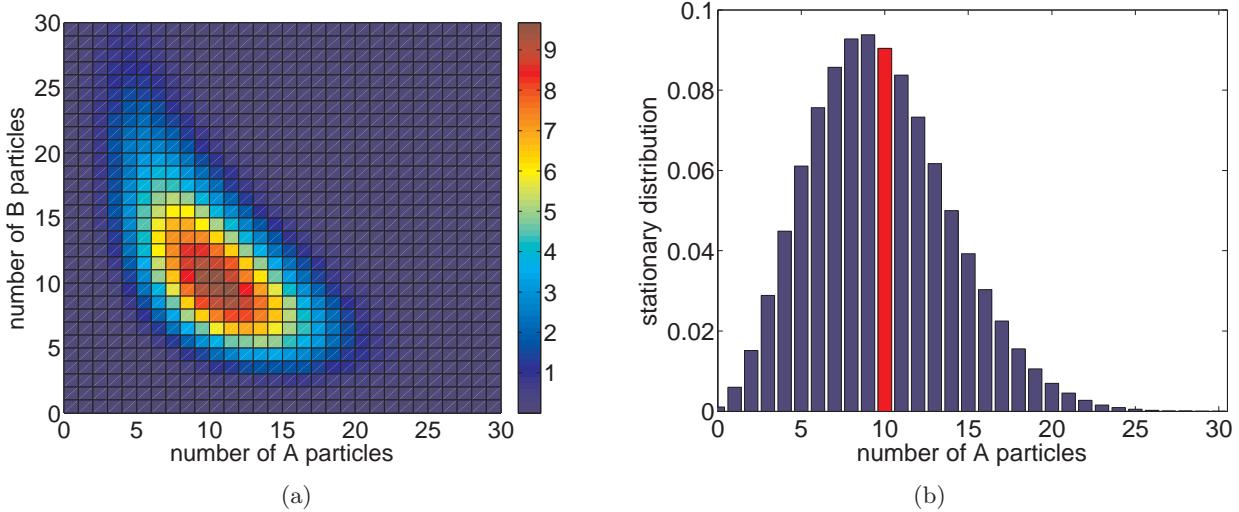


Figure 3.3: (a) Stationary distribution $\phi(n, m)$ obtained by long-time simulation of (4a)-(4d) for $k_1/\nu = 10^{-3} \text{ sec}^{-1}$, $k_2/\nu = 10^{-2} \text{ sec}^{-1}$, $k_3\nu = 1.2 \text{ sec}^{-1}$ and $k_4\nu = 1 \text{ sec}^{-1}$ (b) Stationary distribution of A obtained as in equation (3.22). The red bar highlights the steady state value $A_s = 10$ of the ODE system.

The stationary distribution is defined by

$$\phi(n, m) = \lim_{t \rightarrow \infty} p_{n, m}(t).$$

This can be computed by long-time simulations of SSA (4a)-(4d) and is plotted in Figure 3.3 (a). We see that there is a correlation between the values of A and B . This can also be observed in Figure 3.2. Looking at the blue realisations, we see that the values of $A(t)$ are below the average and the values of $B(t)$ are above the average, similarly for other realisations.

One can also define the marginal stationary distribution of A only by

$$\phi(n) = \sum_{m=0}^{\infty} \phi(n, m). \quad (3.22)$$

Summing the results of Figure 3.3 (a) over m , we obtain $\phi(n)$ which is plotted in Figure 3.3 (b) as a blue histogram. The red bar highlights the steady state value $\bar{A}_s = 10$ of the ODE system (3.20)-(3.21).

Lecture 4 - Approximate stochastic simulation algorithms

By simulating one reaction at a time, the Gillespie algorithm maps out the complete history of a system: for the many applications which do not require knowledge of such details, an alternative approach may be prudent. If we are interested in accelerating the time evolution of the Gillespie SSA, we can also use the so-called τ -leaping simulation method [8].

In Lecture 1 we introduced the naive SSA (1a)-(1b) which was based on taking small time-steps, Δt , and determining whether any reaction had taken place. We quickly improved the efficiency of the method by switching to an event-driven philosophy in which we jumped forward to the next reaction event using (1.5).

τ -leaping returns to the time-driven philosophy but considers a much larger time-step, τ . The idea is that many reactions take place in the interval t to $t+\tau$, but not so many that the propensity functions change “significantly”. The propensity functions are taken to be constant during the time interval, so that the number of reactions of each type is Poisson distributed. Thus, many reaction events are incorporated into every time-step [8]. τ -leaping is an approximate method which is reasonably accurate as long as no propensity function changes its value significantly during any time-step.

We will work with a general system of q chemical reactions and N chemical species. Let $\alpha_j(t)$ be the propensity function of the j -th reaction, $j = 1, \dots, q$, at time t . $\boldsymbol{\nu}_j = (\nu_{1j}, \dots, \nu_{Nj})$ for $j = 1, \dots, q$ is known as the stoichiometric vector corresponding to reaction j . ν_{ij} denotes the changes in species i due to reaction j . $\mathbf{X}(t) = (X_1(t), \dots, X_N(t))$ is the state vector which holds the values of each of the species at time t . A reaction of type j can be implemented by simply adding $\boldsymbol{\nu}_j$ to \mathbf{X} .

Here we ask if simulating a number of reactions per time step can improve performance whilst still generating reasonable stochastic trajectories. Gillespie [8] suggests the τ -leap approach as such an improvement. We first introduce the requisite theory.

4.1 Poisson processes

A Poisson process of rate λ is a time-indexed counting process, $Y = \{Y(t) : t \geq 0\}$, taking values in $S = \{0, 1, 2, 3, \dots\}$ such that [9]:

1. $Y(0) = 0$; if $s < t$ then $Y(s) \leq Y(t)$,

- 2.

$$P(Y(t + dt) = n + m | Y(t) = n) = \begin{cases} 1 - \lambda dt + O(dt^2), & \text{if } m = 0; \\ \lambda dt + O(dt^2), & \text{if } m = 1. \\ O(dt^2), & \text{if } m > 1; \end{cases} \quad (4.1)$$

3. if $s < t$, the number $Y(t) - Y(s)$ of firings in the interval $(s, t]$ is independent of the times of firings during $[0, s]$.

We speak of $Y(t)$ as the number of firings of a reaction channel. We are interested in the distribution of $Y(t)$.

Theorem 1 For a Poisson process Y , $Y(t)$ is Poisson distributed with mean λt i.e.

$$Y(t) \sim \text{Poisson}(\lambda \cdot t).$$

Specifically

$$P(Y(t) = n) = \frac{(\lambda t)^n}{n!} \exp(-\lambda t), \quad n = 0, 1, 2, \dots \quad (4.2)$$

Proof

Condition $Y(t + dt)$ on $Y(t)$ to obtain

$$\begin{aligned} P(Y(t + dt) = n) &= \sum_m P(Y(t) = m) P(Y(t + dt) = n | Y(t) = m) \\ &= \sum_m P(Y(t) = m) P(n - m \text{ arrivals in } (t, t + dt]) \\ &= P(Y(t) = n - 1) P(\text{one arrival}) + P(Y(t) = n) P(\text{no arrivals}) + O(dt^2). \end{aligned} \quad (4.3)$$

Thus $p_n(t) = P(Y(t) = n)$ satisfies the master equation

$$p_n(t + dt) = \lambda dt p_{n-1}(t) + (1 - \lambda dt)p_n(t) + O(dt^2), \quad (4.4)$$

where, as usual, by convention, $p_n = 0$ for $n < 0$. Subtracting $p_n(t)$ from both sides, dividing through by dt and passing to the limit $dt \rightarrow 0$ gives

$$\frac{dp_n}{dt} = \lambda p_{n-1} - \lambda p_n. \quad (4.5)$$

By point 1 of the definition of the Poisson process the initial condition is

$$p_n(0) = \delta_{n,0} = \begin{cases} 1, & \text{if } n = 0; \\ 0, & \text{if } n \neq 0. \end{cases} \quad (4.6)$$

Equation (4.5) is a family of differential equations for p_n , $n \geq 0$. They can be solved by at least two methods: induction (see Problem Sheet 2) or probability generating functions (see Lecture 5) to give the desired result (4.2).

In Lecture 1 we demonstrated that a process which satisfied property (4.1) had inter-arrival times (i.e. times between firings) which were exponentially distributed with mean $1/\lambda$. Theorem 1 demonstrates that the number of fixed-rate, exponentially distributed reaction firing events to occur in a fixed interval of time τ is Poisson distributed with mean $\lambda\tau$.

4.2 τ -leap algorithm

This idea provides a natural way for us to develop an approximate stochastic simulation algorithm using large fixed time increments. We now record the evolution, $\mathbf{X}(t)$ at successive times t_1, t_2, \dots until a terminal time, T , is reached. Starting from an initial time, $t = t_0$, we suppose that for a period of time $[t, t + \tau)$ the propensity functions of the underlying reactions can be assumed to be constant. Accurate τ -leaping requires the *leap condition* [8] to be met:

“ τ is required to be small enough that the change in the state during $[t, t + \tau)$ will be so slight that no propensity function will suffer an appreciable change in its value.”

If $K_j(\tau, \mathbf{x}, t)$ represents the number of times the reaction channel R_j fires in the interval $[t, t + \tau)$, generically we can write

$$\mathbf{X}(t + \tau) = \mathbf{X}(t) + \sum_{j=1}^q K_j(\tau, \mathbf{x}, t) \boldsymbol{\nu}_j. \quad (4.7)$$

Generating sample paths then reduces to generating values for the K_j in order to “leap” the system ahead from t to $t + \tau$. τ -leaping approximates the K_j terms as constant-rate Poisson processes; Theorem 1 then implies that we can approximate the number of firings of reaction j during the time interval $(t, t + \tau]$ as

$$K_j \sim \text{Poisson}(\alpha_j(\mathbf{X}(t)) \cdot \tau),$$

where α_j is the propensity function of the j^{th} reaction, fixed at time t .

The basic τ -leap algorithm is:

Algorithm 5

- (5a) Set $\mathbf{X}(t_0) = \mathbf{x}_0$, and $t = t_0$.
- (5b) Choose τ maximally so that the *leap condition* is met, but so that $t + \tau \leq T$.
- (5c) Generate Poisson random numbers p_j as sample values of $K_j(\tau, \mathbf{X}, t)$, for $j = 1, \dots, q$.
- (5d) Set $\mathbf{X} = \mathbf{X} + \sum_{j=1}^q p_j \boldsymbol{\nu}_j$, and $t = t + \tau$.
- (5e) Return to step (5b) if $t < T$.

The accuracy of this algorithm depends on the extent to which our choice of τ meets the leap condition [10].

4.3 A fixed choice of τ

A simple implementation of the τ -leap algorithm involves explicitly fixing τ throughout the simulation of each path. Whilst easy to implement and analyse, a cautious approach is necessary to ensure the leap condition is met throughout the simulation.

Using an approximate method like τ -leaping will inherently engender some “error” in the simulations. If we are trying to estimate a statistic such as the mean, then our estimator over a number of repeats of the stochastic simulation algorithm, $\hat{\mu}_\tau$, will have some discrepancy from the true mean, μ . Note that the estimator $\hat{\mu}_\tau$ is a random variable which will depend on the simulation method, whereas μ is a deterministic property of the reaction system.

We can characterise the error in our estimator using the mean-squared error (MSE):

$$\begin{aligned} \text{MSE}(\hat{\mu}_\tau) &= E[(\hat{\mu}_\tau - \mu)^2] \\ &= E[\hat{\mu}_\tau^2 - 2\mu\hat{\mu}_\tau + \mu^2] \\ &= E[\hat{\mu}_\tau^2] - E[\hat{\mu}_\tau]^2 + E[\hat{\mu}_\tau]^2 - 2\mu\hat{\mu}_\tau + \mu^2 \\ &= E[\hat{\mu}_\tau^2] - E[\hat{\mu}_\tau]^2 + E[\hat{\mu}_\tau]^2 - 2\mu E[\hat{\mu}_\tau] + \mu^2 \\ &= \text{var}(\hat{\mu}_\tau) + (E[\hat{\mu}_\tau] - \mu)^2. \end{aligned} \tag{4.8}$$

To get from the second line to the third we have simultaneously added and subtracted $E[\hat{\mu}_\tau]^2$. The MSE can thus be broken down into two parts:

- the *statistical error*, $\text{var}(\hat{\mu}_\tau)$;
- the square of the *bias*, $(E[\hat{\mu}_\tau] - \mu)^2$.

The statistical error will appear in estimates from any simulation algorithm, even Gillespie's exact SSA **(4a)-(4d)**. The statistical error will diminish when we average over more repeats of the simulation. The bias, however, will not. In exact stochastic simulation algorithms like Gillespie's algorithm the expected value of the estimated moments of the distribution (e.g. mean, variance etc) are given by the exact moments, hence we say the method is unbiased. In τ -leaping there will be a discrepancy between the expected value of the estimators and those of the true distribution. We say these methods are biased. The bias does not diminish as we average over more repeats. In τ -leaping, the larger the value of τ we take the more we invalidate the leap condition and simulate from the incorrect Poisson distributions and hence the larger the bias in the method.

Consider the simple first-order production reaction



We could consider this to be a very basic model of cell division (although see Lectures 11 and 12 for a more realistic model of cell division times).

The first-order nature of the system means that its mean behaviour obeys a linear ODE which we can solve exactly. First we write down the probability master equation. Let us denote by $p_n(t)$ the probability that there are n particles of A at time t in the system, i.e. $A(t) = n$. The corresponding chemical master equation is given by

$$\frac{dp_n}{dt} = k_1(n-1)p_{n-1} - k_1np_n, \quad (4.10)$$

with the usual convention that $p_n(t) = 0$ for $n < 0$. Our initial condition $A(0) = n_0$ implies that there are never fewer than n_0 particles in the system, so that $p_n \equiv 0$ for $n < n_0$. Equation (4.10) is an infinite system of ODEs, but may be solved using induction (see Problem Sheet 2) or generating functions to give the exact evolution of the probability distribution. Alternatively we may obtain equations directly for the moments of the system.

Recall the definition of the mean, $M(t)$,

$$M(t) = \sum_{n=0}^{\infty} np_n(t). \quad (4.11)$$

Multiplying (4.10) by n and summing over n , we obtain

$$\frac{d}{dt} \sum_{n=0}^{\infty} np_n = k_1 \sum_{n=0}^{\infty} n(n-1)p_{n-1} - k_1 \sum_{n=0}^{\infty} n^2 p_n$$

Using the definition of $M(t)$ this can be simplified (see Problem Sheet 2) to show that $M(t)$ evolves according to

$$\frac{dM}{dt} = k_1 M. \quad (4.12)$$

The solution of (4.12) with initial condition $M(0) = n_0$ is

$$M(t) = n_0 \exp(k_1 t), \quad (4.13)$$

and is plotted as a dashed black line in Figure 4.1 (b).

In Figure 4.1 we have simulated reaction (4.9) with a range of different values of τ in order to demonstrate the increasing bias in the algorithm as τ increases. The larger the value of τ the further the simulated mean number of particles deviates from the true expected value. For each value of τ we carry out sufficient repeats in order to ensure that the statistical error of each estimator is less than

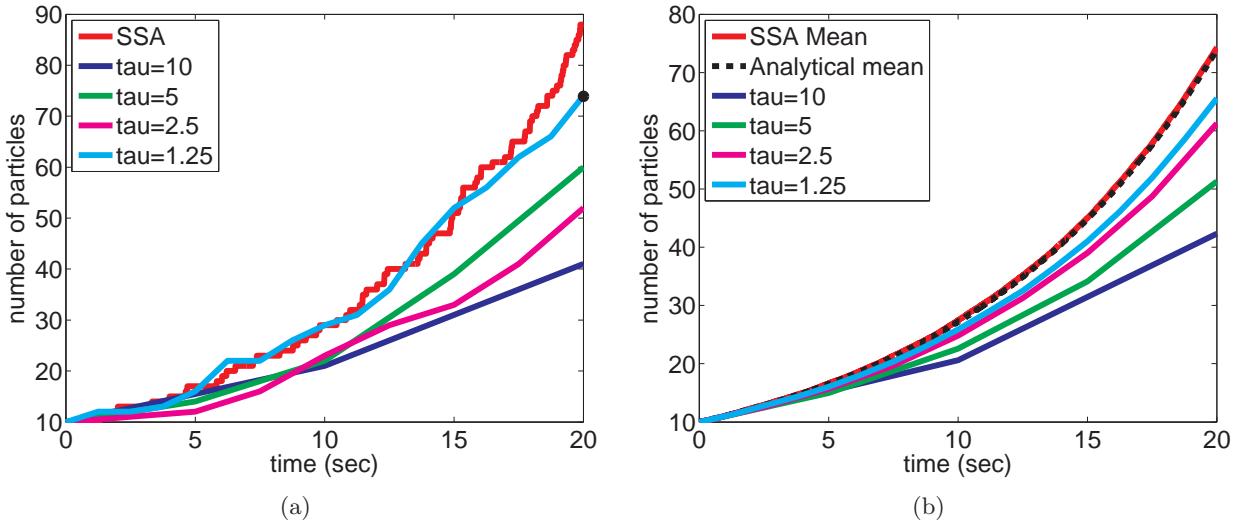


Figure 4.1: (a) One realisation of reaction (4.9) using the exact SSA (**4a**)-(**4d**) and using the τ -leaping algorithm (**5a**)-(**5e**) with a range of values of τ . (b) The average over several realisations of reaction (4.9) using the exact SSA and τ -leaping. The solution of (4.12) is given by the black dashed line. We use $A(0) = 10$, $k_1 = 10^{-1} \text{sec}^{-1}$ and simulate to $t = 20$ in both panels.

unity. That is to say

$$\text{var}(\hat{\mu}_\tau) = \frac{\sigma_\tau^2}{n} < 1, \quad (4.14)$$

where σ_τ^2 is the variance of the random variable simulated with time-step τ .

A comparison of timings, number of repeats and estimator values is given in Table 4.1. τ -leaping with $\tau = 10$ is more than an order of magnitude faster than using the exact SSA (4a)-(4d), but there is significant bias in the method. If we decrease the value of τ we can decrease the bias in the method at the expense of longer simulation times. With $\tau = 1.25$, τ -leaping takes time on the same order of magnitude as the exact SSA and still produces a significantly biased estimator.

τ	E[A(20)]	Paths	Time (sec)
SSA	75.14	450	2.94
1.25	66.17	343	1.24
2.5	60.21	231	0.52
5	50.44	156	0.24
10	39.54	37	0.11

Table 4.1: Simulation times for the first-order production reaction (4.9). Estimator values, number of paths required and time taken for the exact SSA (4a)-(4d) and the τ -leaping algorithm (5a)-(5e) for a range of values of τ . The true value of the expected number of particles is $\mu = 73.89$. Note that bounding $\text{var}(\hat{\mu}_\tau) < 1$ does not bound the estimator to be within one particles of the mean.

Consider a model of gene transcription, translation and protein-protein dimerisation [11]:



A single gene is transcribed to produce a particle of messenger RNA (mRNA). Molecules of mRNA can be translated into proteins and two protein molecules can form a dimer. Both mRNA and protein molecules degrade spontaneously. Write the numbers of mRNA, protein and dimer molecules at time t as $(M(t), P(t), D(t))$, respectively, with initial condition $(M(0), P(0), D(0)) = (0, 0, 0)$. We are interested in estimating the mean number of dimer molecules at time $t = 0.4$.

In Figure 4.2 the mean evolution of mRNA, proteins and dimers can be seen for a range of different values of τ . Although the τ -leaping methods appear to approximate the number of mRNA molecules well, there is a clear bias in the numbers of dimer molecules by $t = 0.4$. We simulate enough paths to ensure the statistical error in the number of dimers is less than unity in each case.

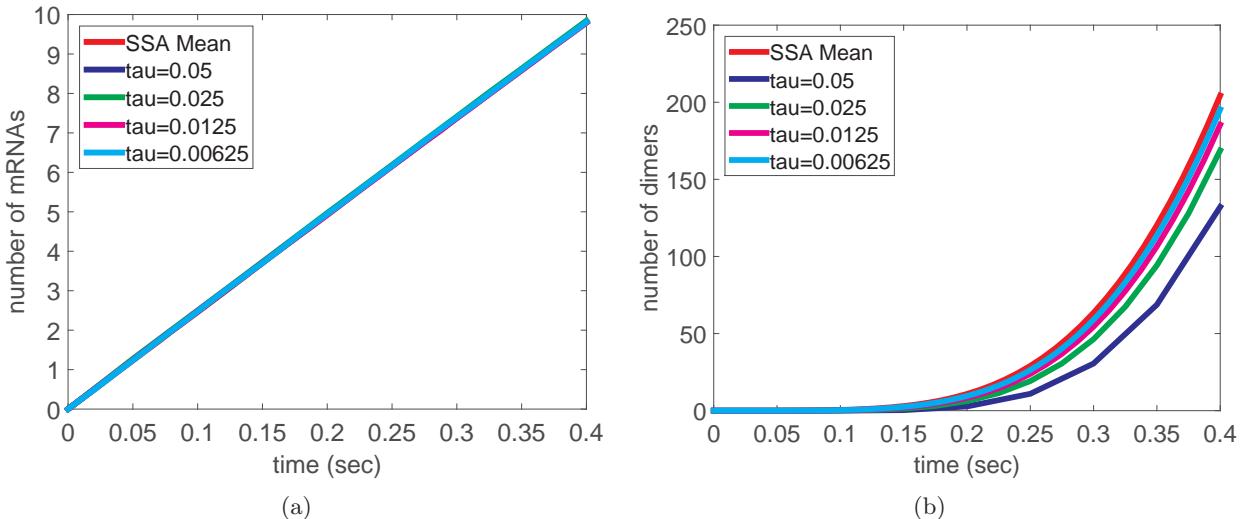


Figure 4.2: The average over several realisations of reaction system (4.15) using the exact SSA and τ -leaping with a range of different value of τ . (a) number of mRNA molecules. (b) Number of dimer molecules. We use $(M(0), P(0), D(0)) = (0, 0, 0)$, and simulate to $t = 0.4$ in both panels.

A comparison of timings, number of repeats and estimator values for the dimer population is given in Table 4.2. We see similar trends to those of the simple production reaction (4.9). Larger values of τ produce faster but less accurate estimators. Decreasing the value of τ we decreases the bias in the method at the expense of longer simulation times. With $\tau = 0.00625$, τ -leaping takes longer than the exact SSA and the estimator produced is still biased.

τ	$E[D(0.4)]$	Paths	Time (sec)
SSA	204.39	18303	335.59
0.00625	196.38	17213	453.41
0.0125	185.66	15812	230.18
0.025	169.44	13919	106.91
0.05	132.93	9980	35.58

Table 4.2: Simulation times for the gene expression system (4.15)-(4.19). Estimator values (mean number of dimers), number of paths required and time taken for the exact SSA (4a)-(4d) and the τ -leaping algorithm (5a)-(5e) for a range of values of τ .

4.4 Meeting the leap condition

Whilst we require τ to be sufficiently small to meet the leap condition, this method is of computational benefit, when compared with the SSA, if many reactions occur during the time interval [12]. At first sight, it may seem best to balance these competing demands by dynamically choosing a τ , performing a leap, and then redoing the leap with a smaller τ if necessary. However, this risks skewing the dynamics of the system as large, but correct, fluctuations are excluded from the dynamics [13]. A range of options for adaptively pre-selecting τ are given by Gillespie and Petzold [13], Cao et al. [14], and Gillespie [8].

4.5 Negative populations

Occasionally, the τ -leap approach generates negative population numbers. This is because the Poisson distribution is unbounded at the right hand side and multiple reactions can deplete a single species. Thus more reactants than are available may be consumed during a leap [15]. However, the usual cause of negative species populations is an inappropriately risky attitude to the choosing of τ [16]. The population numbers, initial conditions, reaction intensities and time frame should all affect the choice of τ .

We would like to avoid the inevitable dilemma we face when attempting to remedy a negative population in a sample path. If we exclude the offending path, or simply set the negative population to zero, we introduce a model bias, as disproportionately many paths with low populations are deleted, whilst paths reaching high populations are not.

Cao et al. [15] have suggested a method which attempts to avoid negative specie populations by bounding the time-step taken by the relative change in propensity functions and classifying reactions as critical and non-critical. Tian and Burrage [17] and Chatterjee et al. [16] have independently suggested that an appropriately parametrised binomial distribution should be used to avoid negative populations. An alternative approach is given in [18].

Lecture 5 - Deterministic versus stochastic modelling I

In the previous two lectures we found that for zero-order and first-order chemical reactions the evolution equations for the stochastic mean are exactly the ordinary differential equations (ODEs) describing the corresponding deterministic system. On the other hand, when we considered higher-order chemical reactions (3.12), for which the deterministic description is non-linear, we found that the deterministic ODEs do not provide an exact description of the stochastic mean. Nevertheless, the solution of these ODEs was still a reasonable approximation of the evolution of the stochastic mean: the results of the SSAs looked like “noisy solutions” of the corresponding deterministic ODE model. Thus for all the models presented so far one could use the deterministic description to obtain a reasonable description of the average behaviour of the system. In Section 5.1, we will analyse another chemical system of this type. We will use this example to introduce a different approach for the analysis of stochastic chemical systems. Then, in Section 6.1, we present an example where deterministic modelling fails to predict the average behaviour of the system. Other examples where SSAs give results which cannot be obtained from the corresponding deterministic models will be studied in Lectures 7 and 8.

5.1 Stochastic simulation of dimerisation

Systems with second-order (or higher-order) chemical reactions lead to chemical master equations which are more difficult to analyse (for example, the method presented in Section 2.1 does not work). In this section we use the example of dimerisation to introduce an alternative approach to solving the chemical master equation, the so-called *probability-generating function*.

We consider a chemical species A in a container of volume ν which is subject to the following following reversible chemical reaction:



The first reaction describes the dimerisation-degradation of the chemical A with the rate constant k_1 . We couple it with the second reaction which represents the first-order production of the chemical A with the rate constant k_2 . Using Table 3.1, we get the propensity function of the first and second reactions as $\alpha_1(t) = A(t)(A(t) - 1)k_1/\nu$ and $\alpha_2(t) = k_2A(t)$ respectively where, as usual, $A(t)$ is the number of particles of A at time t and ν is the system volume. To simulate the system of chemical reactions (5.1), we use the Gillespie SSA (4a)-(4d). That is, we perform the following four steps at time t (starting with $A(0) = n_0$ at time $t = 0$):

Algorithm 6

- (6a) Generate two random numbers r_1, r_2 uniformly distributed in $(0, 1)$.

(6b) Compute the propensity functions of both reactions:

$$\alpha_1 = A(t)(A(t) - 1)k_1/\nu \text{ and } \alpha_2 = k_2 A(t).$$

$$\text{Compute } \alpha_0 = \alpha_1 + \alpha_2.$$

(6c) Compute the time when the next chemical reaction takes place as $t + \tau$ where τ is given by equation (3.11).

(6d) Compute the number of particles at time $t + \tau$ by

$$A(t + \tau) = \begin{cases} A(t) - 1 & \text{if } r_2 < \alpha_1/\alpha_0; \\ A(t) + 1 & \text{if } r_2 \geq \alpha_1/\alpha_0. \end{cases}$$

Then continue with step (6a) for time $t + \tau$.

As before, the probability that any of the reactions in (5.1) takes place in the time interval $[t, t + dt]$ is equal to $\alpha_0 dt$ where α_0 is the sum of the propensity functions of both reactions. Formula (3.11) gives the time $t + \tau$ when the next reaction takes place. Once we know the time $t + \tau$, the first reaction in (5.1) takes place with probability α_1/α_0 . Otherwise, the second reaction in (5.1) occurs. The decision as to which reaction takes place is given in the step (6d) with the help of the second uniformly distributed random number r_2 . Ten realisations of the SSA (6a)-(6d) are presented in Figure 5.1 (a) as solid lines. We plot the number of particles of A as a function of time for $A(0) = 1$, $k_1/\nu = 0.1 \text{ sec}^{-1}$ and $k_2 = 0.4 \text{ sec}^{-1}$. We see that, after an initial transient, the number of particles $A(t)$ fluctuates around its mean value. To compute the mean and quantify the stochastic fluctuations, we will analyse the corresponding chemical master equation.

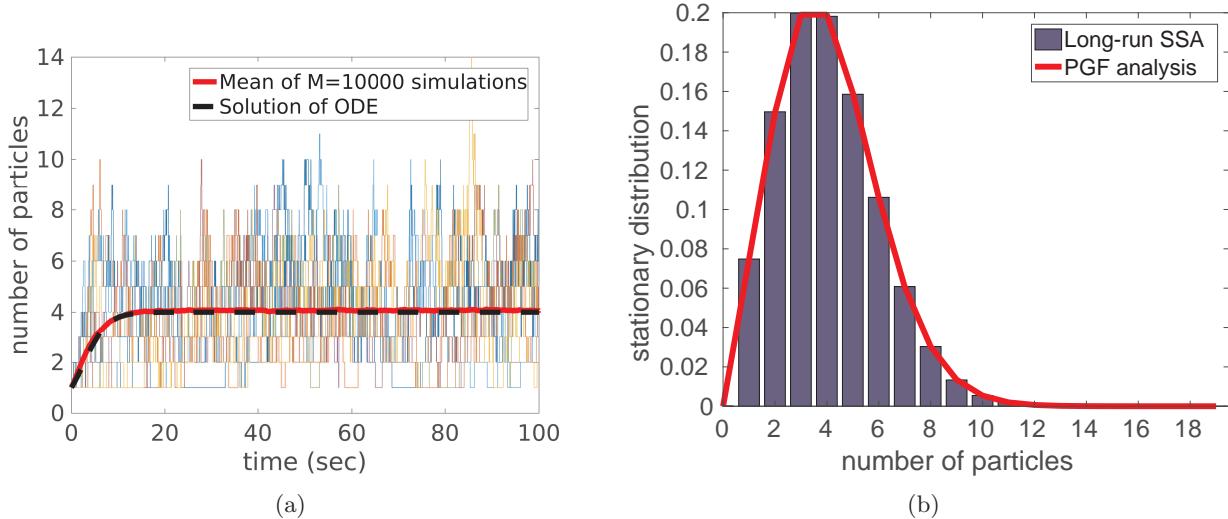


Figure 5.1: Stochastic simulation of the system of chemical reactions (5.1) for $A(0) = 1$, $k_1/\nu = 0.1 \text{ sec}^{-1}$ and $k_2 = 0.4 \text{ sec}^{-1}$. (a) $A(t)$ given by five realisations of the SSA (6a)-(6d) (thin solid coloured lines), the mean value of $A(t)$ averaged over 10,000 repeats (thick solid red line) and the solution of the ODE equation (5.26) (dashed black line). (b) Stationary distribution, $\phi(n)$, obtained by long-time simulation of the SSA (6a)-(6d) (blue histogram) and by formula (5.24) (red solid line).

Let us denote by $p_n(t)$ the probability that there are n particles of A at time t in the system, i.e. $A(t) = n$. Let us consider an (infinitesimally) small time-step dt chosen such that the probability that two reactions occur during $[t, t + dt]$ is negligible compared to the probability that only one reaction takes place during $[t, t + dt]$. Then there are three possible ways for $A(t + dt)$ to take the value n :

either $A(t) = n$ and no reaction occurred in $[t, t + dt]$, or $A(t) = n + 1$ and the first reaction in (5.1) occurred in $[t, t + dt]$, or $A(t) = n - 1$ and one particle was produced in $[t, t + dt]$ according to the second chemical reaction in (5.1). Hence

$$\begin{aligned} p_n(t + dt) &= p_n(t) \times \left(1 - \frac{k_1}{\nu}n(n-1)dt - k_2ndt\right) \\ &\quad + p_{n+1}(t) \times \frac{k_1}{\nu}(n+1)ndt + p_{n-1}(t) \times k_2(n-1)dt. \end{aligned}$$

A simple algebraic manipulation yields

$$\begin{aligned} \frac{p_n(t + dt) - p_n(t)}{dt} &= \frac{k_1}{\nu}(n+1)np_{n+1}(t) - \frac{k_1}{\nu}n(n-1)p_n(t) \\ &\quad + k_2(n-1)p_{n-1}(t) - k_2np_n(t). \end{aligned}$$

Passing to the limit $dt \rightarrow 0$, we obtain the chemical master equation for the chemical system (5.1) in the following form

$$\frac{dp_n}{dt} = \frac{k_1}{\nu}(n+1)np_{n+1}(t) - \frac{k_1}{\nu}n(n-1)p_n + k_2(n-1)p_{n-1} - k_2np_n, \quad (5.2)$$

where, for $n = 0$, the third term on the right hand side of equation (5.2) is missing; we again use the convention that $p_{-1} \equiv 0$.

The mean $M(t)$ and variance $V(t)$ are defined by equations (2.5).

$$M(t) = \sum_{n=0}^{\infty} np_n(t), \quad V(t) = \sum_{n=0}^{\infty} (n - M(t))^2 p_n(t). \quad (2.5)$$

In Section 2.1, we were able to derive the evolution equations for $M(t)$ and $V(t)$. Unfortunately, the approach of Section 2.1 is not applicable to chemical systems with the second-order (or higher-order) reactions. If we try to follow the method from Section 2.1 (i.e. multiply the chemical master equation (5.2) by n and sum over n), we do not obtain a closed evolution equation for $M(t)$ (although we can use moment closure to obtain an approximate set of equations as in Lectures 9 and 10). However, we may be interested in the stationary values M_s and V_s which are defined by (2.12) and the stationary distribution $\phi(n)$ which is defined by (2.13). Fortunately, it is possible to derive analytical formulae for M_s , V_s and $\phi(n)$. To do that, we define the *probability-generating function* $G : [-1, 1] \times [0, \infty) \rightarrow \mathbb{R}$ by

$$G(x, t) = E[x^n] = \sum_{n=0}^{\infty} x^n p_n(t). \quad (5.3)$$

Differentiating $G(x, t)$ with respect to x , we obtain

$$\frac{\partial G}{\partial x}(x, t) = \frac{\partial}{\partial x} \sum_{n=0}^{\infty} x^n p_n(t) = \sum_{n=1}^{\infty} nx^{n-1} p_n(t). \quad (5.4)$$

Differentiating with respect to x again, we deduce

$$\frac{\partial^2 G}{\partial x^2}(x, t) = \frac{\partial}{\partial x} \sum_{n=1}^{\infty} nx^{n-1} p_n(t) = \sum_{n=2}^{\infty} n(n-1)x^{n-2} p_n(t). \quad (5.5)$$

Substituting $x = 1$ into (5.4) and (5.5), we obtain the identities

$$\frac{\partial G}{\partial x}(1, t) = \sum_{n=0}^{\infty} np_n(t), \quad \frac{\partial^2 G}{\partial x^2}(1, t) = \sum_{n=0}^{\infty} n(n-1)p_n(t). \quad (5.6)$$

Using the definition of $M(t)$ given by (2.5),

$$M(t) = \sum_{n=0}^{\infty} np_n(t), \quad V(t) = \sum_{n=0}^{\infty} (n - M(t))^2 p_n(t).$$

we immediately conclude

$$M(t) = \frac{\partial G}{\partial x}(1, t). \quad (5.7)$$

Using the identity (2.10), we get

$$V(t) = \sum_{n=0}^{\infty} n^2 p_n(t) - M(t)^2 = \sum_{n=0}^{\infty} n(n-1)p_n(t) + M(t) - M(t)^2,$$

which together with (5.6) implies

$$\begin{aligned} V(t) &= \frac{\partial^2 G}{\partial x^2}(1, t) + M(t) - M(t)^2 \\ &= \frac{\partial^2 G}{\partial x^2}(1, t) + \frac{\partial G}{\partial x}(1, t) - \left(\frac{\partial G}{\partial x}(1, t) \right)^2. \end{aligned} \quad (5.8)$$

Substituting $x = 0$ into (5.3), (5.4) and (5.5), we obtain

$$p_0(t) = G(0, t), \quad p_1(t) = \frac{\partial G}{\partial x}(0, t), \quad p_2(t) = \frac{1}{2} \frac{\partial^2 G}{\partial x^2}(0, t).$$

Using mathematical induction (see Problem Sheet 3), we can prove the general formula

$$p_n(t) = \frac{1}{n!} \frac{\partial^n G}{\partial x^n}(0, t), \quad (5.9)$$

for any non-negative integer n . If we know the probability-generating function $G(x, t)$, we can use formulae (5.7), (5.8) and (5.9) to find $M(t)$, $V(t)$ and $p_n(t)$.

To find $G(x, t)$, for the dimerisation-degradation reaction system (5.1), we multiply the chemical master equation (5.2) by x^n and sum over n to deduce

$$\begin{aligned} \frac{\partial}{\partial t} \sum_{n=0}^{\infty} x^n p_n &= \frac{k_1}{\nu} \sum_{n=1}^{\infty} x^n (n+1) n p_{n+1} - \frac{k_1}{\nu} \sum_{n=2}^{\infty} x^n n (n-1) p_n \\ &\quad + k_2 \sum_{n=1}^{\infty} x^n (n-1) p_{n-1} - k_2 \sum_{n=0}^{\infty} x^n n p_n. \end{aligned}$$

Changing indices in the first and third sum on the right hand side, we obtain

$$\begin{aligned} \frac{\partial}{\partial t} \sum_{n=0}^{\infty} x^n p_n &= \frac{k_1}{\nu} x \sum_{n=2}^{\infty} x^{n-2} n (n-1) p_n - \frac{k_1}{\nu} x^2 \sum_{n=2}^{\infty} x^{n-2} n (n-1) p_n \\ &\quad + k_2 x^2 \sum_{n=0}^{\infty} n x^{n-1} p_n - k_2 x \sum_{n=0}^{\infty} n x^{n-1} p_n. \end{aligned}$$

Using (5.3) and (5.5), we obtain the partial differential equation for G in the following form

$$\frac{\partial G}{\partial t} = \frac{k_1}{\nu}(1-x)x\frac{\partial^2 G}{\partial x^2} + k_2(x-1)x\frac{\partial G}{\partial x}. \quad (5.10)$$

Solving this equation numerically, we could find $G(x, t)$ and, with the help of formulae (5.7), (5.8) and (5.9), we could also find $M(t)$, $V(t)$ and $p_n(t)$. Since we are mostly interested in the stationary values M_s , V_s and $\phi(n)$ which are defined by (2.12) and (2.13), we look for the stationary probability-generating function G_s : $[-1, 1] \rightarrow \mathbb{R}$ which is defined as

$$G_s(x) = \lim_{t \rightarrow \infty} G(x, t) = \sum_{n=0}^{\infty} x^n \phi(n). \quad (5.11)$$

Then formulae (5.7), (5.8) and (5.9) imply

$$M_s = \frac{dG_s}{dx}(1), \quad (5.12)$$

$$V_s = \frac{d^2G_s}{dx^2}(1) + M_s - M_s^2, \quad (5.13)$$

$$\phi(n) = \frac{1}{n!} \frac{d^n G_s}{dx^n}(0), \quad n = 0, 1, 2, 3, \dots, \quad (5.14)$$

i.e. M_s , V_s and $\phi(n)$ can be computed using the stationary probability-generating function, $G_s(x)$, which is also a solution of the stationary problem corresponding to the evolution equation (5.10), namely

$$0 = \frac{k_1}{\nu}(1-x)x\frac{d^2G_s}{dx^2} + k_2(x-1)x\frac{dG_s}{dx}.$$

For values of x in the interior region $x \in (0, 1)$ this simplifies to

$$0 = \frac{k_1}{\nu} \frac{d^2G_s}{dx^2} - k_2 \frac{dG_s}{dx}. \quad (5.15)$$

Integrating once with respect to x gives the first order differential equation

$$\frac{dG_s}{dx} - KG_s = C, \quad (5.16)$$

where $K = k_2\nu/k_1$ and C is an, as yet, undetermined constant. Using separation of variables the general solution to equation (5.16) is

$$G_s(x) = \frac{1}{K} \exp(D + Kx) - C, \quad (5.17)$$

where D is another, as yet, undetermined constant and C is not the same C as in equation (5.16). In order to determine the two undefined constants we need two boundary conditions for the ODE (5.15). The first comes from evaluating $G_s(x)$ at $x = 1$:

$$G_s(1) = \lim_{t \rightarrow \infty} \sum_{n=0}^{\infty} 1^n p_n(t) = \sum_{n=0}^{\infty} \phi_n = 1, \quad (5.18)$$

where the last equality is a consequence of equation (2.7). The second boundary condition comes from considering what happens with low molecule numbers. Providing $n(0) \neq 0$, $p_0 = 0$ since at least two molecules are required to reduce molecule numbers by one, so we can never go from one molecule to

zero. This implies:

$$G_s(0) = \lim_{t \rightarrow \infty} \sum_{n=0}^{\infty} 0^n p_n(t) = \phi_0 = 0. \quad (5.19)$$

Applying these two boundary conditions gives

$$G_s(x) = \frac{\exp(Kx) - 1}{\exp(K) - 1}. \quad (5.20)$$

Differentiating (5.20) with respect to x , we obtain

$$\frac{dG_s}{dx} = \frac{K \exp(Kx)}{\exp(K) - 1}.$$

Substituting $x = 1$ and using (5.12)

$$M_s = \frac{dG_s}{dx}(1) = \frac{K \exp(K)}{(\exp(K) - 1)}. \quad (5.21)$$

Substituting $x = 1$ into (5.15) and using (5.21), we obtain

$$\frac{d^2G_s}{dx^2}(1) = K \frac{dG_s}{dx}(1) = \frac{K^2 \exp(K)}{\exp(K) - 1}. \quad (5.22)$$

Consequently, formula (5.13) implies

$$\begin{aligned} V_s &= K \frac{dG_s}{dx}(1) + \frac{dG_s}{dx}(1) - \left(\frac{dG_s}{dx}(1) \right)^2 \\ &= \frac{K \exp(K)}{\exp(K) - 1} \left\{ K + 1 - \frac{K \exp(K)}{\exp(K) - 1} \right\}. \end{aligned} \quad (5.23)$$

Using (5.14), one can verify (see Problem Sheet 3) that the stationary distribution is given by

$$\phi(n) = \begin{cases} 0, & \text{for } n = 0; \\ \frac{1}{n!} \frac{K^n}{\exp(K) - 1}, & \text{for } n = 1, 2, 3, \dots \end{cases} \quad (5.24)$$

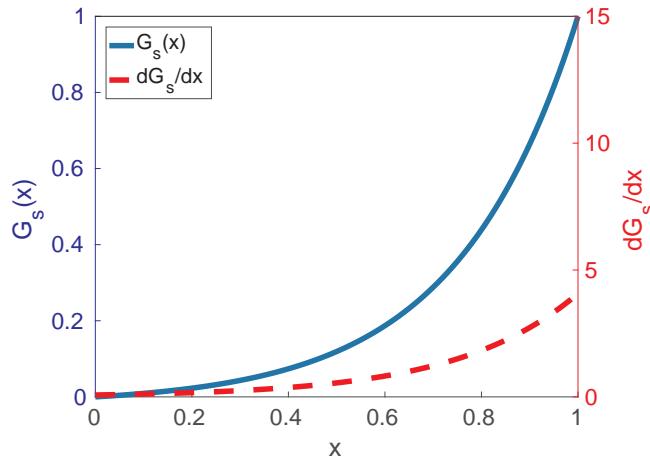


Figure 5.2: The solution $G_s(x)$ of (5.15) given by (5.20) for $k_1/\nu = 0.1 \text{ sec}^{-1}$, $k_2 = 0.4 \text{ sec}^{-1}$ (solid blue line). The derivative $G'_s(x)$ is also plotted (dashed red line).

Using our parameter values $k_1/\nu = 0.1 \text{ sec}^{-1}$ and $k_2 = 0.4 \text{ sec}^{-1}$, we plot $\phi(n)$ given by (5.24) in Figure 5.1 (b) as the red solid line. Comparison with the results of stochastic simulation is excellent.

The probability-generating function $G_s(x)$ together with its derivative $G'_s(x)$ is plotted in Figure 5.2. Formulae (5.21) and (5.23) give $M_s = 4.075$ and $V_s = 3.771$ which are also in excellent agreement with the results obtained by long stochastic simulation ($M_s^{sim} = 4.073$ and $V_s^{sim} = 3.769$).

Finally, let us consider a classical deterministic description of the chemical system (5.1). For the concentration, $a(t) = A(t)/\nu$, it is given by the following ODE

$$\frac{da}{dt} = -k_1 a^2 + k_2 a. \quad (5.25)$$

The solution of this ODE is presented in Figure 5.1 (a) for comparison. More precisely, we compare the results of the stochastic simulation with the function

$$\bar{A}(t) = a(t)\nu,$$

which gives the number of particles in the volume with concentration $a(t)$. Multiplying (5.25) by ν , we obtain that $\bar{A}(t)$ satisfies the ODE

$$\frac{d\bar{A}}{dt} = -k_1/\nu \bar{A}^2 + k_2 \bar{A}. \quad (5.26)$$

The solution of (5.26) is plotted for $k_1/\nu = 0.005 \text{ sec}^{-1}$, $k_2 = 0.4 \text{ sec}^{-1}$ and initial condition $\bar{A}(0) = 0$ in Figure 5.1 (a) as the black dashed line. We should note here that the equation (5.26) does not give us the time evolution of the stochastic mean $M(t)$, i.e. $\bar{A}(t) \neq M(t)$. To see that, let us consider the stationary value of $\bar{A}(t)$, which is given by

$$0 = -k_1/\nu \bar{A}_s^2 + k_2 \bar{A},$$

i.e.

$$\bar{A}_s = \frac{k_2 \nu}{k_1}.$$

Using our parameter values, we conclude that $\bar{A}_s = 4$. On the other hand, we have already found (using formula (5.21)) that $M_s = 4.075$. The difference between the exact value $M_s = 4.075$ and the deterministic approximation $\bar{A}_s = 4$ is not large. However, it can be still seen if we compute M_s as an average over many stochastic realisations. In the next Lectures 6-8, we will present examples of chemical systems where the difference between the results of stochastic simulation and the corresponding deterministic approximation is more significant. But it is worth noting that even in the case of a simple second-order chemical reaction presented here, the deterministic system of ODEs does not provide the exact description of the stochastic mean. Moreover, deterministic modelling does not give us a description of stochastic fluctuations. The size of fluctuations can be estimated as the mean standard deviation which is the square root of the variance V_s . Using (5.23), we find that $\sqrt{V_s} = 2.75$.

Lecture 6 - Deterministic versus stochastic modelling II

When we considered higher-order reactions in Lectures 3 and 5 (see equations (3.12), (3.13) and (5.1)), for which the deterministic description is non-linear, we found that the deterministic ODEs do not provide an exact description of the mean behaviour of the system. However, for the simple system we considered, the agreement was not terrible.

In this lecture we present an example of “stochastic focussing” in which deterministic modelling dramatically fails to predict the average behaviour of the system. Other examples where SSAs give results which cannot be obtained from the corresponding deterministic models will be studied in Lecture 7 and 8.

6.1 Stochastic focusing

We consider now three chemical species A, B and C in a container of volume ν . Following Paulsson et al. [19], we will refer to A as a signal and to B as a product. We assume that A, B and C are subject to the following six chemical reactions. The product B is produced from the intermediate C (which is itself produced from a source) and degraded:



The coupling between the production of the product B and the signal A is indirect: the signal A catalyses the degradation of the intermediate, so that



We will study how changes in the number of signal particles, A , influence the changes in the number of product particles, B . We will show that the mean value of B depends not only on the mean value of A but also on the fluctuations of the number of signal particles, A . To that end, we let the signal A evolve according to the production-degradation chemical system (2.1)-(2.2) which was studied in Section 2.1, i.e.



The approximate ODE description of the chemical system (6.1)-(6.3) can be written as follows:

$$\frac{da}{dt} = k_5 - k_6a, \quad (6.4)$$

$$\frac{db}{dt} = k_2c - k_3b, \quad (6.5)$$

$$\frac{dc}{dt} = k_1 - k_2c - k_4ac, \quad (6.6)$$

where $a(t)$, $b(t)$ and $c(t)$ are concentrations of A, B and C, respectively, in the reactor of volume ν . The system (6.4)-(6.6) can be rewritten as the ODE system for approximate numbers of particles in the reactor which are given by $\bar{A}(t) = a(t)\nu$, $\bar{B}(t) = b(t)\nu$ and $\bar{C}(t) = c(t)\nu$. We obtain

$$\frac{d\bar{A}}{dt} = k_5\nu - k_6\bar{A}, \quad (6.7)$$

$$\frac{d\bar{B}}{dt} = k_2\bar{C} - k_3\bar{B}, \quad (6.8)$$

$$\frac{d\bar{C}}{dt} = k_1\nu - k_2\bar{C} - \frac{k_4}{\nu}\bar{A}\bar{C}. \quad (6.9)$$

The rate constants of reactions (6.1)-(6.2) are chosen as

$$k_1\nu = 10^2 \text{ sec}^{-1}, k_2 = 10^3 \text{ sec}^{-1}, k_3 = 0.01 \text{ sec}^{-1}, \frac{k_4}{\nu} = 9900 \text{ sec}^{-1}. \quad (6.10)$$

We assume that the value of the rate constant k_5 is abruptly halved at time 10 minutes and choose the rate constants of the chemical reaction (6.3) by

$$k_5\nu = \begin{cases} 10^3 \text{ sec}^{-1}, & \text{for } t < 10 \text{ min;} \\ 5 \times 10^2 \text{ sec}^{-1}, & \text{for } t \geq 10 \text{ min;} \end{cases} \quad \text{and} \quad k_6 = 10^2 \text{ sec}^{-1}. \quad (6.11)$$

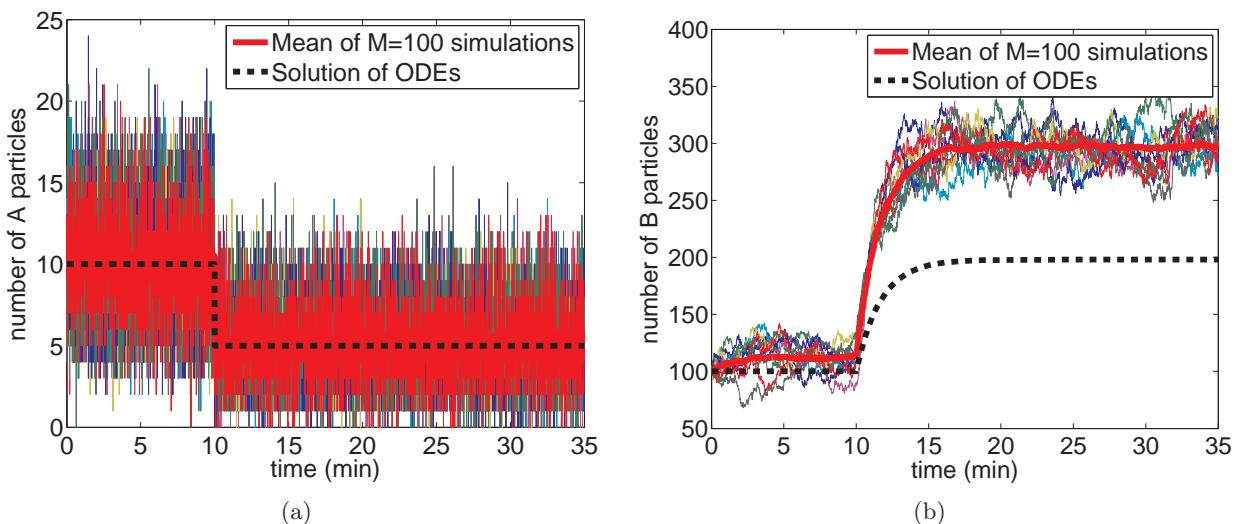


Figure 6.1: Ten realisations of the SSA (4a)-(4d) for the system of chemical reactions (6.1)-(6.3) (solid thin coloured lines), the mean value of $A(t)$ averaged over 100 repeats (thick solid red line) and the solution of the system of ODEs (6.7)-(6.9) (dashed black line). The parameters are given by (6.10) and (6.11). (a) Time evolution of the signal A . (b) Time evolution of the product B .

In Figure 6.1, we compare the time evolution of the (approximate) system of ODEs (6.7)-(6.9) with the results obtained by the Gillespie SSA (4a)-(4d). We use the same initial condition $[\bar{A}(0), \bar{B}(0), \bar{C}(0)] =$

$[10, 100, 0]$ for both models. We see that the average number of signal particles, A , halves at time 10 minutes because of the change of the production rate k_5 given by (6.11). However, the response of the number of product particles B to the change in the signal is not correctly predicted by the ODE model (6.7)-(6.9).

The steady state of the ODE system (6.7)-(6.9) can be found by solving the algebraic system

$$\begin{aligned} 0 &= k_5\nu - k_6\bar{A}_s, \\ 0 &= k_2\bar{C}_s - k_3\bar{B}_s, \\ 0 &= k_1\nu - k_2\bar{C}_s - \bar{A}_s\bar{C}_s k_4/\nu. \end{aligned}$$

Consequently,

$$\bar{A}_s = \frac{k_5\nu}{k_6}, \bar{B}_s = \frac{k_1\nu k_2 k_6}{k_3(k_2 k_6 + k_4 k_5)}, \bar{C}_s = \frac{k_1\nu k_6}{k_2 k_6 + k_4 k_5}. \quad (6.12)$$

Using the parameter values (6.10)-(6.11), we find that the deterministic steady state values \bar{A}_s and \bar{B}_s are

$$\bar{A}_s = \begin{cases} 10, & \text{for } t < 10 \text{ min;} \\ 5, & \text{for } t \geq 10 \text{ min;} \end{cases} \quad \bar{B}_s = \begin{cases} 100, & \text{for } t < 10 \text{ min;} \\ 198, & \text{for } t \geq 10 \text{ min.} \end{cases} \quad (6.13)$$

Thus the deterministic system predicts that the number of product particles at steady state, B_s , should, roughly speaking, double, while the number of signal particles, A_s should be halved. The numerical solution the approximate system of ODEs (6.7)-(6.9) (which is plotted in Figure 6.1 as the black line) confirms this observation. On the other hand, the stochastic model shows approximately three times increase in the number of the product particles at steady state, B_s , i.e. it is more sensitive to the change in the signal (see Figure 6.1 (b)).

To understand this behaviour, let $M_A(t)$ be the average number of A particles at time t . The number of A particles is only influenced by the chemical reactions (6.3). Such a process was studied in Section 2.1 as the chemical system (2.1)-(2.2). In particular, we were able to derive the evolution equation for the stochastic mean (see equation (2.8)). Using the notation of this section, the evolution equation for $M_A(t)$ can be written as

$$\frac{dM_A}{dt} = k_5\nu - k_6M_A. \quad (6.14)$$

The stationary value of the stochastic mean is $M_{A,s} = k_5\nu/k_6$. Using the parameter values (6.10)-(6.11), we have

$$M_{A,s} = \begin{cases} 10, & \text{for } t < 10 \text{ min;} \\ 5, & \text{for } t \geq 10 \text{ min.} \end{cases} \quad (6.15)$$

Thus the stationary values of the signal A are in agreement with the values (6.13) predicted by the ODE model (6.7)-(6.9).

We will show that it is the fluctuations in A which lead to the differences in B which we observe between the stochastic and deterministic models. To that end, we simulate the system of chemical reactions (6.1)-(6.2) with the rate constants given by (6.10). However, instead of simulating (6.3), we directly impose the time evolution of the signal $A(t)$ in the following form

$$A(t) = \begin{cases} 10, & \text{for } t < 10 \text{ min;} \\ 5, & \text{for } t \geq 10 \text{ min;} \end{cases} \quad (6.16)$$

i.e. we ignore the fluctuations in A and postulate that A is equal to the corresponding steady state

values (6.15). Then the deterministic ODE model (6.7)-(6.9) simplifies to

$$\frac{d\bar{B}}{dt} = k_2 \bar{C} - k_3 \bar{B}, \quad (6.17)$$

$$\frac{d\bar{C}}{dt} = k_1 \nu - k_2 \bar{C} - \frac{k_4}{\nu} A(t) \bar{C}. \quad (6.18)$$

where $A(t)$ is given by (6.16).

In Figure 6.2, we compare the time evolution of the system of ODEs (6.17)-(6.18) with the results obtained by the Gillespie SSA (4a)-(4d). We use the same initial condition $[\bar{B}(0), \bar{C}(0)] = [100, 0]$ for both models. We see that there are no significant differences between the evolution of the number of product particles B computed by the deterministic and stochastic models. Moreover, one can show (Problem Sheet 1) that the ODE system (6.17)-(6.18) is the exact description of the stochastic mean of B and C . In fact, the only non-linear reaction in the chemical system (6.1)-(6.3) is the second-order chemical reaction (6.2). It can be rewritten as



i.e. the degradation of C with the rate $k_4 A(t)$. If $A(t)$ is assumed to be constant (no fluctuations in A), then (6.19) is the first-order chemical reaction, degradation, which was studied in Section 1.1 as the chemical reaction (1.1). In particular, the linearity of the ODE model implies that it provides an exact description of the average behaviour of the chemical system. The same is true for the piecewise constant $A(t)$ given by (6.16).

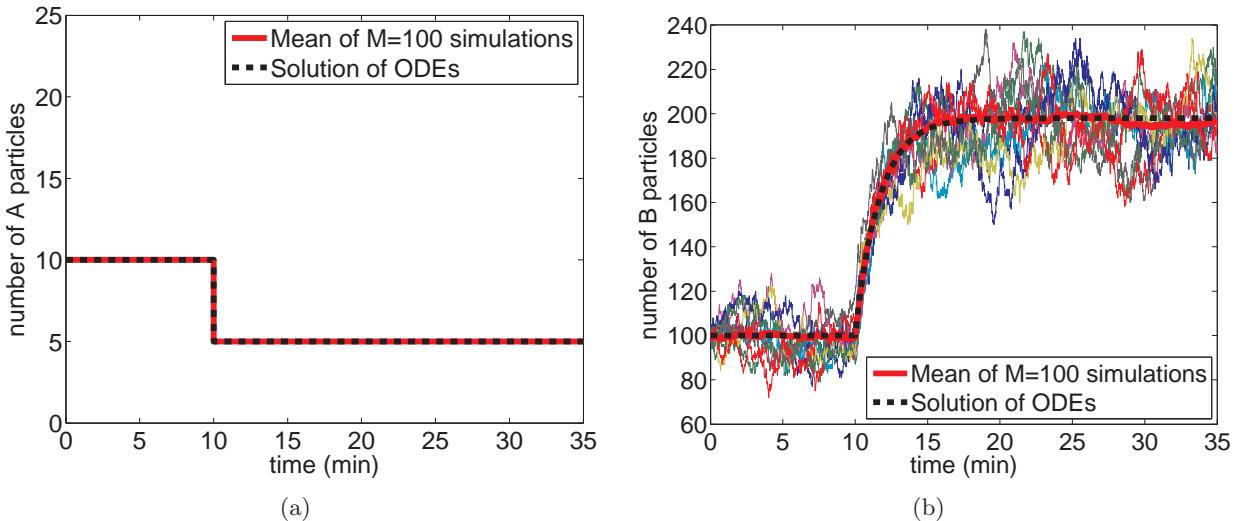


Figure 6.2: (a) Time evolution of the signal A given by (6.16) (black dashed line) and mean behaviour of A as an average over 100 repeats (solid red line). (b) The evolution of $B(t)$ for five realisations of the SSA (4a)-(4d) (thin solid coloured lines) for the system of chemical reactions (6.1)-(6.2) (6.16), the mean behaviour of B averaged over 100 repeats (solid thick red line) and the solution of the system of ODEs (6.17)-(6.18) (black dashed line).

The presence of fluctuations of A (or equivalently the second-order chemical reaction (6.2)) is not a sufficient condition to obtain such significant differences between the stochastic and deterministic descriptions as presented in Figure 6.1. Indeed, we have already presented non-linear chemical systems in Sections 5.1 and 3.3 for which the corresponding deterministic equations provide a reasonable (approximate) description of the average behaviour of the system. The extra feature of the model

(6.1)-(6.3) which makes the ODE model inapplicable is a very low number of the intermediate particles C in the model. Using steady state formulae (6.12) and the parameter values (6.10)-(6.11), we find that the steady state value \bar{C}_s is equal to 10^{-3} for $t < 10$ min, and 1.98×10^{-3} for $t \geq 10$ min. In particular, \bar{C}_s can be no longer successfully interpreted as the number of C particles in the reactor. The stochastic model predicts that either 0 or 1 particle of C is in the system (see Figure 6.4 (a) ; the probability that there is more than one particle of C in the system is very low). It is this low number of C particles, combined with its participation in a second-order reaction, which is the cause of the large differences in the number of B particles predicted by the stochastic and deterministic models. To illustrate this point, we consider the model (6.1)-(6.3) with the parameter values

$$k_1\nu = 10^2 \text{ sec}^{-1}, k_2 = 0.1 \text{ sec}^{-1}, k_3 = 0.01 \text{ sec}^{-1}, \frac{k_4}{\nu} = 0.99 \text{ sec}^{-1} \quad (6.20)$$

and the changing parameter values for k_5 given by (6.11).

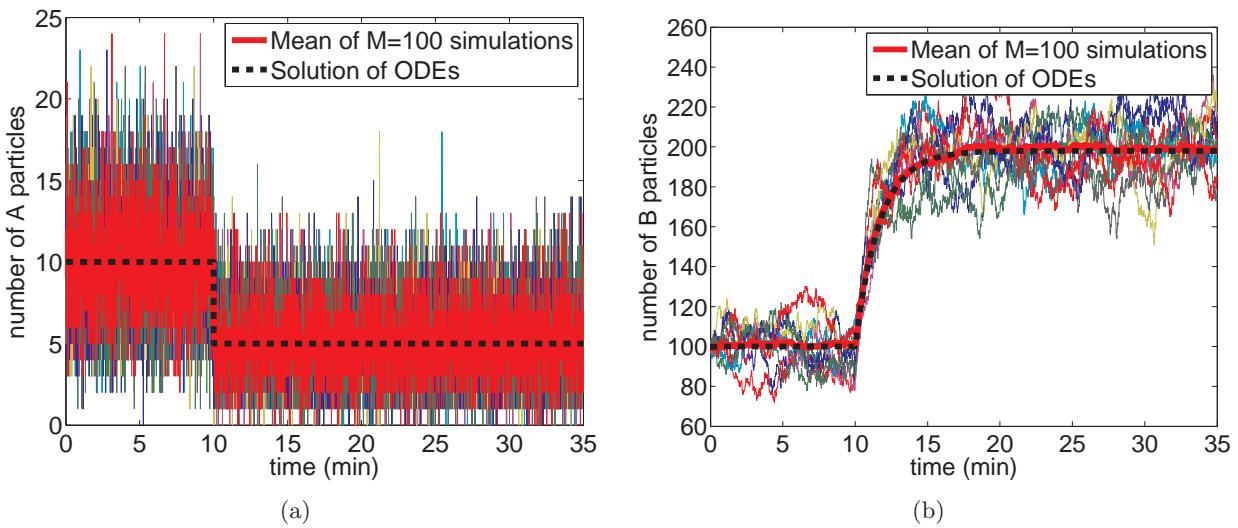


Figure 6.3: Ten realisations of the SSA (4a)-(4d) for the system of chemical reactions (6.1)-(6.3) (thin solid coloured lines), the mean behaviour averaged over 100 repeats (solid red line) and the solution of the system of ODEs (6.7)-(6.9) (black dashed line). The parameters are given by (6.20) and (6.11). (a) Time evolution of the signal A . (b) Time evolution of the product B .

In Figure 6.3, we compare the time evolution of the system given by the (approximate) system of ODEs (6.7)-(6.9) with the results obtained by the Gillespie SSA (4a)-(4d). We use the same initial condition $[\bar{A}(0), \bar{B}(0), \bar{C}(0)] = [10, 100, 10]$ for both models. Since we use the same values of the rate constants for the chemical reactions (6.3) as before, the time evolution of A is the same as in Figure 6.1 (a). However (unlike the case of Figure 6.1), the deterministic description of B matches the stochastic results very well.

The values of \bar{A}_s and \bar{B}_s , for the second set of parameter values (6.20), are still the same as for the rate constants (6.10), i.e. given by (6.13). The difference is in the values of \bar{C}_s . Using (6.12) together with the parameter values (6.20) and (6.11), we find that the steady state value \bar{C}_s is equal to 10 for $t < 10$ min, and 19.8 for $t \geq 10$ min. In particular, \bar{C}_s is already high enough that it can be approximately interpreted as the number of C particles in the reactor. In Figure 6.4 (b), we present the time evolution of C for the second set of parameter values (6.20) and (6.11). We observe that C fluctuates around the values predicted by the ODE model (6.7)-(6.9). This is in contrast with the results obtained by the original set of parameter values (6.10) and (6.11) (see Figure 6.4 (a)) where we plot the time evolution of C corresponding to the simulation presented in Figure 6.1. Where, the number of particles of C fluctuated between 0 and 1 and the ODE solution for \bar{C} had values of the

order 10^{-3} .

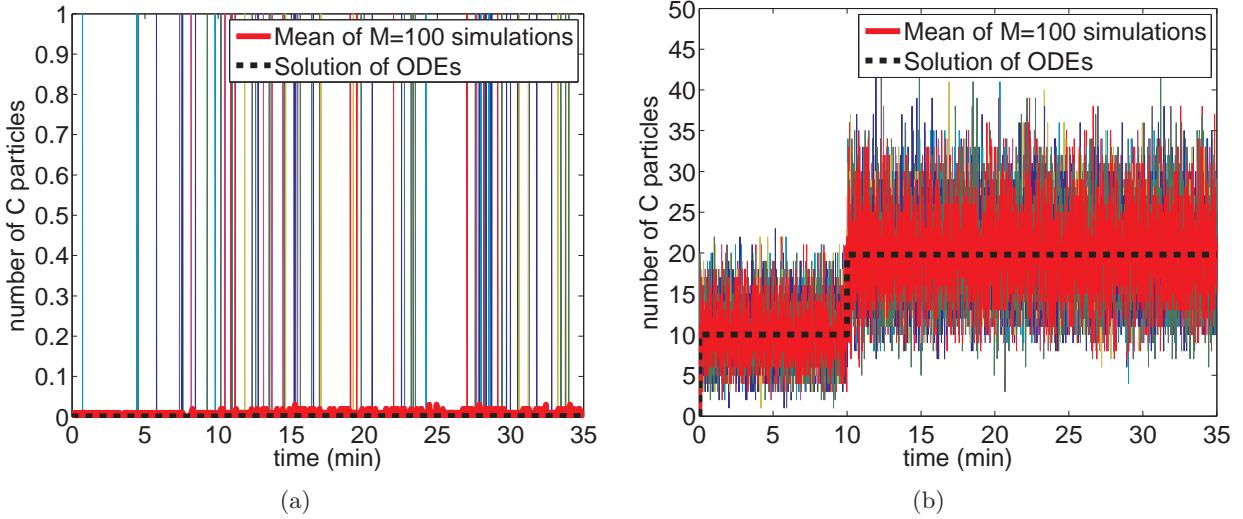


Figure 6.4: The time evolution of the number of particles of the intermediate C. Ten realisations of the SSA (4a)-(4d) for the system of chemical reactions (6.1)-(6.3) (solid thin coloured lines), the mean behaviour averaged over 100 repeats (thick red solid line) and the solution of the system of ODEs (6.7)-(6.9) (black dashed line). (a) The parameters are given by (6.10) and (6.11), *i.e.* we plot the time evolution of C for the simulation presented in Figure 6.1. (b) The parameters are given by (6.20) and (6.11), *i.e.* we plot the time evolution of C for the simulation presented in Figure 6.3.

Thus second-order (or higher-order) chemical reactions together with low copy numbers of some chemical species can lead to significant differences between the average behaviour given by the stochastic model and by the corresponding (approximate) deterministic ODEs. Note that in problem (6.1)-(6.3) the signal we observed was the product B, which had molecular abundances of the order of hundreds. However, we still reported significant differences between the stochastic and deterministic description of B because the intermediate chemical C was presented in low copy numbers.

If the rate constants k_5 and k_6 of the chemical reaction (6.3) are fixed, then the probability that there are n particles of A in the system can be estimated as

$$\phi(n) = \frac{1}{n!} (M_{A,s})^n \exp[-M_{A,s}], \quad (6.21)$$

which is the stationary distribution derived for the production-degradation system (2.1)-(2.2) as the formula (2.18). We know that for parameter values (6.10)-(6.11) that there are either 0 or 1 particle of C in the system during most of the simulation time (see Figure 6.4 (a)). If we consider that the number of particles of A is fixed and equal to n , then calculating the steady state value C_s , using equation (6.18), the probability that there is a C particle in the system can be approximated by

$$\frac{k_1\nu}{k_2 + nk_4/\nu}. \quad (6.22)$$

Combining with (6.21), we can estimate the average number of C particles in the system as

$$\sum_{n=0}^{\infty} \frac{k_1\nu}{k_2 + nk_4/\nu} \phi(n) = \sum_{n=0}^{\infty} \frac{k_1\nu}{k_2 + nk_4/\nu} \frac{1}{n!} (M_{A,s})^n \exp[-M_{A,s}]. \quad (6.23)$$

Using equation (6.17) we can relate the average number of particles of B at steady state to the average

number of particles of C at steady state by multiplying the RHS of equation (6.23) by k_2/k_3 :

$$M_{B,s} = \sum_{n=0}^{\infty} \frac{k_1 \nu k_2}{k_3(k_2 + nk_4/\nu)} \frac{1}{n!} (M_{A,s})^n \exp[-M_{A,s}].$$

Using equation (6.15) for the value of $M_{A,s}$, we obtain

$$M_{B,s} = \begin{cases} 113.0, & \text{for } t < 10 \text{ min;} \\ 316.7, & \text{for } t \geq 10 \text{ min;} \end{cases} \quad (6.24)$$

which is a good estimate of $M_{B,s}$ for the rate constants (6.10)-(6.11) (see Figure 6.1 (b)). The steady state values \bar{B}_s of the ODEs (6.7)-(6.9) are for the rate constants (6.10)-(6.11) given in (6.13). Comparing (6.13) and (6.24), we conclude that the estimate (6.24) provides much better approximation of the average behaviour of the system than the results obtained by the deterministic ODEs.

Lecture 7 - Deterministic versus stochastic modelling III

In the previous lecture, we presented the phenomenon of stochastic focusing by showing a simple system for which the stochastic model is more sensitive (in amplification of the signal) than its deterministic counterpart. In this lecture, we present another example where deterministic modelling fails and a stochastic approach is necessary; for this system, SSAs give results which cannot be obtained from the corresponding deterministic model. We illustrate a simple example of stochastic switching between favourable states of a system and then demonstrate how this model is used for modelling the collective behaviour of locusts.

7.1 Systems with multiple favourable states

The system (3.20)-(3.21) which we considered in Section 3.3 had only one non-negative (stable) steady state for our parameter values, namely $\bar{A}_s = \bar{B}_s = 10$. Solutions of (3.20)-(3.21) converge to \bar{A}_s and \bar{B}_s as $t \rightarrow \infty$ for all non-negative initial conditions (see Figure 3.2). Similarly the results of SSAs show fluctuation about the means, which are roughly equal to \bar{A}_s and \bar{B}_s (they are 9.6 particles for A and 12.2 particles for B).

The situation becomes more interesting when a chemical system has two or more favourable states, so that the corresponding ODEs have more than one non-negative stable steady state. The following system (first considered by Schlögl [20]) has two favourable states. Consider the chemical A in a container of volume ν which is subject to the four chemical reactions



The (approximate) deterministic ODE model can be written for the concentration $a(t)$ as

$$\frac{da}{dt} = -k_1 a^3 + k_2 a^2 - k_3 a + k_4. \quad (7.2)$$

We will compare results obtained by solving (7.2) with the results obtained by the Gillespie SSA (**4a**)-(**4d**). To do that, we rewrite equation (7.2) in terms of $\bar{A}(t) = a(t)\nu$ which is the deterministic approximation of the number of particles of A in the volume with concentration $a(t)$, to give

$$\frac{d\bar{A}}{dt} = -\frac{k_1}{\nu^2} \bar{A}^3 + \frac{k_2}{\nu} \bar{A}^2 - k_3 \bar{A} + k_4 \nu. \quad (7.3)$$

The steady states of (7.3) can be found by solving the cubic equation

$$-\frac{k_1}{\nu^2} \bar{A}^3 + \frac{k_2}{\nu} \bar{A}^2 - k_3 \bar{A} + k_4 \nu = 0. \quad (7.4)$$

We choose the rate constants to be $k_1/\nu^2 = 2.5 \times 10^{-4} \text{ min}^{-1}$, $k_2/\nu = 0.18 \text{ min}^{-1}$, $k_3 = 37.5 \text{ min}^{-1}$ and

$k_4\nu = 2200 \text{ min}^{-1}$. Then, solving (7.4), we find that the ODE (7.3) has the three steady states

$$\bar{A}_{s1} = 100, \quad \bar{A}_u = 220, \quad \text{and} \quad \bar{A}_{s2} = 400. \quad (7.5)$$

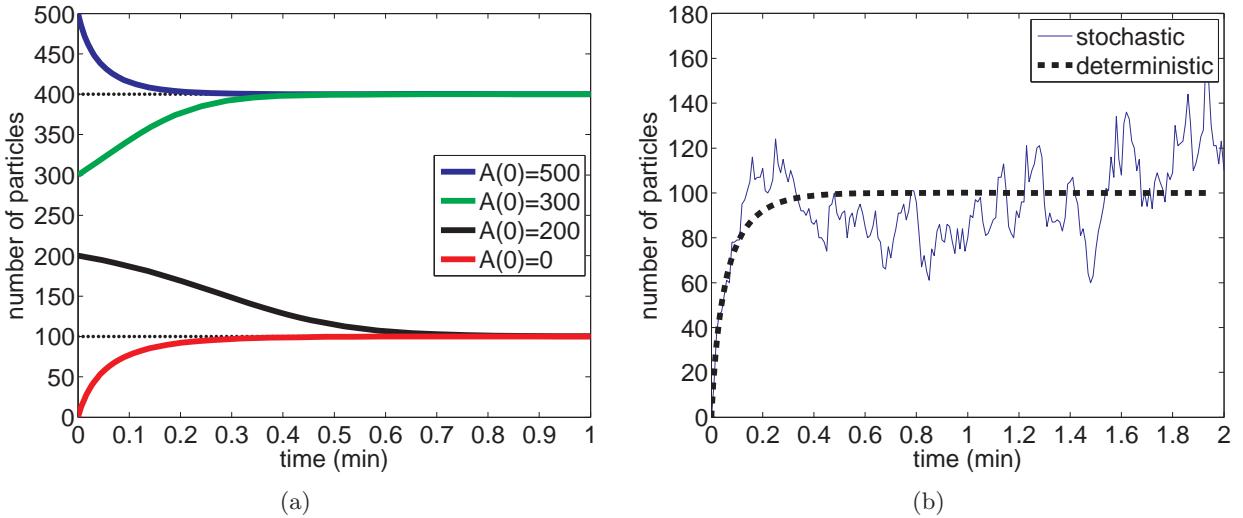


Figure 7.1: (a) Solution of the ODE (7.3) for four different initial conditions. The steady states (7.5) are denoted as dotted lines. (b) The number of particles of A as a function of time over the first two minutes of simulation of (7.1). One realisation of the SSA (4a)-(4d) for the system of chemical reactions (7.1) (blue line) and the solution of the deterministic ODE (7.3) (black dashed line).

We leave it to the reader (Problem Sheet 4) to verify that \bar{A}_{s1} and \bar{A}_{s2} are stable steady states and \bar{A}_u is an unstable steady state. The solution of (7.3) converges to one of the steady states, with the choice of the steady state dependent on the initial condition. If the initial condition $\bar{A}(0)$ satisfies $\bar{A}(0) \in [0, \bar{A}_u]$, then the solution of (7.3) converges to the stable steady state $\bar{A}_{s1} = 100$. If $\bar{A}(0) > \bar{A}_u$, then the solution of (7.3) converges to the second stable steady state $\bar{A}_{s2} = 400$ (see Figure 7.1 (a) where we plot the solution of (7.3) for four different initial conditions).

Now suppose that there are initially no particles of A in the system, i.e. $\bar{A}(0) = 0$. The solution of (7.3) is plotted in Figure 7.1 (a) as the red line. We see that the solution of (7.3) converges to the stable steady state $\bar{A}_{s1} = 100$. Next, we use the Gillespie SSA (4a)-(4d) to simulate the chemical system (7.1). Starting with no particles of A in the system, we plot one realisation of SSA (4a)-(4d) in Figure 7.1 (b) as a blue line. We see that the time evolution of A given by the SSA (4a)-(4d) initially (over the first 2 minutes) looks like the “noisy solution” of (7.3). However, we can find significant differences between the stochastic and deterministic model if we observe both models over sufficiently large times (see Figure 7.2 (a)) where we plot the time evolution of A over the first 100 minutes. As expected, the solution of the deterministic model (7.3) remains close to the stable steady state $\bar{A}_{s1} = 100$. The number of particles given by the stochastic model initially fluctuates around one of the favourable states of the system (which is close to $\bar{A}_{s1} = 100$). However, the fluctuations are sometimes so strong that the system spontaneously switches to another favourable state (which is close to $\bar{A}_{s2} = 400$). This random switching is missed by the deterministic description.

Let $p_n(t)$ be the probability that there are n particles of A at time t in the system, i.e. that

$A(t) = n$. The chemical master equation corresponding to (7.1) can be written as

$$\begin{aligned} \frac{dp_n}{dt} &= k_1/\nu^2(n+1)n(n-1)p_{n+1} - k_1/\nu^2n(n-1)(n-2)p_n \\ &\quad + k_2/\nu(n-1)(n-2)p_{n-1} - k_2/\nu n(n-1)p_n \\ &\quad + k_3(n+1)p_{n+1} - k_3np_n + k_4\nu p_{n-1} - k_4\nu p_n, \end{aligned} \quad (7.6)$$

for $n \geq 0$, where, as usual, we use the convention $p_{-1} = 0$. The stationary distribution $\phi(n)$ (which is defined by (2.13)) can be obtained as a solution of the equation $dp_n/dt = 0$, namely

$$\begin{aligned} 0 &= k_1/\nu^2(n+1)n(n-1)\phi(n+1) - k_1/\nu^2n(n-1)(n-2)\phi(n) \\ &\quad + k_2/\nu(n-1)(n-2)\phi(n-1) - k_2/\nu n(n-1)\phi(n) \\ &\quad + k_3(n+1)\phi(n+1) - k_3n\phi(n) + k_4\nu\phi(n-1) - k_4\nu\phi(n). \end{aligned} \quad (7.7)$$

Considering $n = 0$ in (7.7) many of the terms cancel and we can deduce

$$\phi(1) = \phi(0) \frac{k_4\nu}{k_3}.$$

Similarly, using $n = 1$ in (7.7), we obtain

$$\phi(2) = \phi(0) \frac{k_4\nu}{k_3} \frac{k_4\nu}{2k_3}.$$

Similarly, we can express all $\phi(n)$, $n \geq 1$ in terms of $\phi(0)$ and the rate constants. We leave it to the reader (Problem Sheet 4) to prove the following general formula

$$\phi(n) = \phi(0) \prod_{i=0}^{n-1} \frac{k_2/\nu i(i-1) + k_4\nu}{k_1/\nu^2(i+1)i(i-1) + k_3(i+1)}, \quad (7.8)$$

for $n \geq 1$. Using the normalisation condition (2.16), we can determine the value of $\phi(0)$. To compute $\phi(n)$ in practice we can set $\phi(0) = 1$ and calculate $\phi(n)$, for sufficiently many n , by (7.8). Then we divide $\phi(n)$, $n \geq 0$, by $\sum \phi(n)$ so that the normalisation condition (2.16) is satisfied. In Figure 7.2 (b), we plot the resulting stationary distribution as the red solid line. As expected, the results compare well with the results obtained by the long-time stochastic simulation - the blue histogram.

The most important characteristic of a system with two (or more) favourable states is the mean switching time between these states, that is, how long, on average, does the system spend on each visit to a favourable state. Such a quantity cannot be obtained from the deterministic ODE model because it depends crucially on the fluctuations.

By running an SSA for a long time, we can estimate how much time the system spends in each of its favourable states. However, it is also possible to analyse the model further analytically and estimate the mean switching time without doing long-time simulations. To do so, we will need some further theoretical concepts, which will be introduced in Lectures 13 and 15. In Section 17.1 we return to the model (7.1) and estimate the mean switching time analytically.

The model (7.1) is a pedagogical example, and is the simplest model (with one chemical species) which enables switching between multiple favourable states. However, we note that random switching between states has been found in more realistic models, including, for example, in gene regulatory networks [21–23] and models for the directional switching of locusts in a ring-shaped arena [24–28], as we will see in the next section.

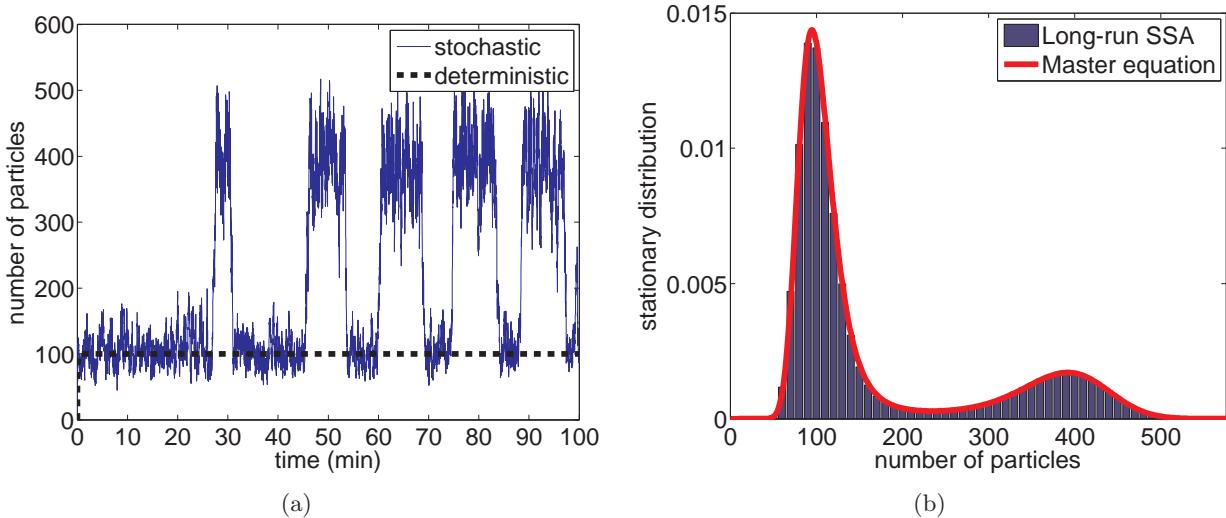


Figure 7.2: (a) Simulation of (7.1) over 100 minutes. One realisation of the SSA (4a)-(4d) for the system of chemical reactions (7.1) (blue line) and the solution of the deterministic ODE (7.3) (black dashed line). (b) Stationary distribution of (7.1) obtained by the long-time simulation of the SSA (4a)-(4d) (blue histogram). Red solid line is the solution of the stationary chemical master equation (see equation (7.8)).

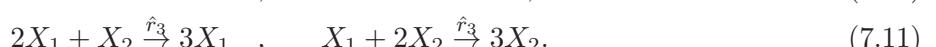
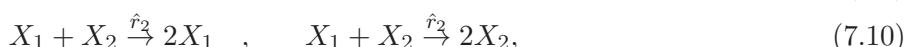
7.2 A research example - Modelling the collective behaviour of locusts

In the previous section we discussed a system for which the deterministic counterpart model has two stable steady states. Deterministic trajectories approach one or other of these stable states depending on the initial condition. However, stochastic trajectories switch between fluctuating about each of these steady states (see Figure 7.2 (a)).

Locusts and other migrating insects can form cohesive swarms at large population densities, which subsequently travel over huge distances and can have a devastating effect on agriculture. It is therefore important to understand the mechanisms governing how the population decides collectively on the direction of migration, and the population density at which this occurs. Trajectories similar to those generated by chemical reactions (7.1) are seen in the average direction of locusts marching together in a ring-shaped arena [24] (see Figure 7.3). Locusts travelling clockwise are assigned a direction value 1 whereas those travelling anticlockwise are assigned a direction value -1.

Consider a population of N individuals, split into clockwise-moving (X_1) and anticlockwise-moving (X_2) populations. The average direction, z , is found by subtracting the proportion of anticlockwise-moving individuals ($x_2 = X_2/N$) from the proportion of clockwise-moving individuals ($x_1 = X_1/N$).

A simple model allows individuals to change direction spontaneously, or to change direction as the result of interactions with one or two locusts travelling in the opposite direction. The model may be summarised in the following system of interactions:



Because the number of locusts, N , is conserved we can eliminate x_2 in favour of x_1 , (i.e. $x_1 = 1 - x_2$).

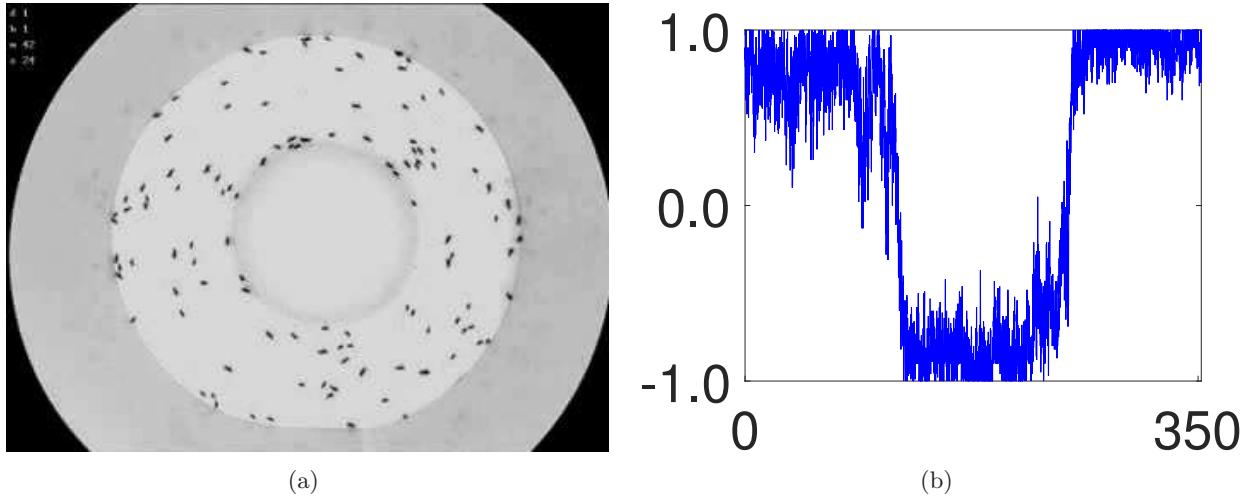


Figure 7.3: (a) An overhead view of the ring-shaped experimental arena around which locusts move. The positions of the locusts are recorded 5 times per second and individual locusts are tracked. Based on these tracks, directions (e.g. clockwise or anticlockwise) of locusts can be inferred. (b) The average direction of $N = 20$ locusts over time exhibits clear switching behaviour between largely clockwise (+1) and largely anticlockwise (-1) movement of the group of locusts.

Thus the rate of transitioning from state b to state a , $T(a|b)$ is given by

$$T^+(x_1) \equiv T\left(x_1 + \frac{1}{N} | x_1\right) = \sum_{i=1}^3 r_i x_1^{i-1} (1 - x_1), \quad (7.12)$$

$$T^-(x_1) \equiv T\left(x_1 - \frac{1}{N} | x_1\right) = \sum_{i=1}^3 r_i x_1 (1-x_1)^{i-1}, \quad (7.13)$$

where we have rescaled the rates $r_i = \hat{r}_i/N^i$ for $i = 1, 2, 3$ when converting between locust numbers, X_1 , and locust proportions, $x_1 = X_1/N$.

In Figure 7.4 (a) a typical trajectory of the above model (simulated using the Gillespie SSA) is presented. There are clear qualitative similarities between the model trajectory and that of the experimental data (see Figure 7.3 (b)).

We can calculate an approximation to the stationary probability density as

$$\phi(z) = C \left[4r_1 + (2r_2 + r_3)(1 - z^2) \right]^{\frac{4Nr_1(r_2+r_3)}{(2r_2+r_3)^2} - 1} e^{\frac{r_3z^2N}{2(2r_2+r_3)}}, \quad (7.14)$$

where C is a normalisation constant (see Lecture 16 for more details). This approximation is compared to the stationary distribution of (7.9)-(7.11) obtained by the long-time simulation of the SSA (4a)-(4d) in Figure 7.4 (b). The agreement is excellent.

The corresponding approximate ODEs for this model are as follows:

$$\frac{dx_1}{dt} = r_1(x_2 - x_1) + r_3 x_1 x_2 (x_1 - x_2), \quad (7.15)$$

$$\frac{dx_2}{dt} = r_1(x_1 - x_2) + r_3x_2x_1(x_2 - x_1). \quad (7.16)$$

Since the total number of locust does not change over time (i.e. $x_1 + x_2 = 1$) we can summarise the approximate system with a single ODE for an auxiliary variable $z = 2x_1 - 1$ (see Problem Sheet 4).

$$\frac{dz}{dt} = -2r_1 z + r_3 z(1 - z^2)/2. \quad (7.17)$$

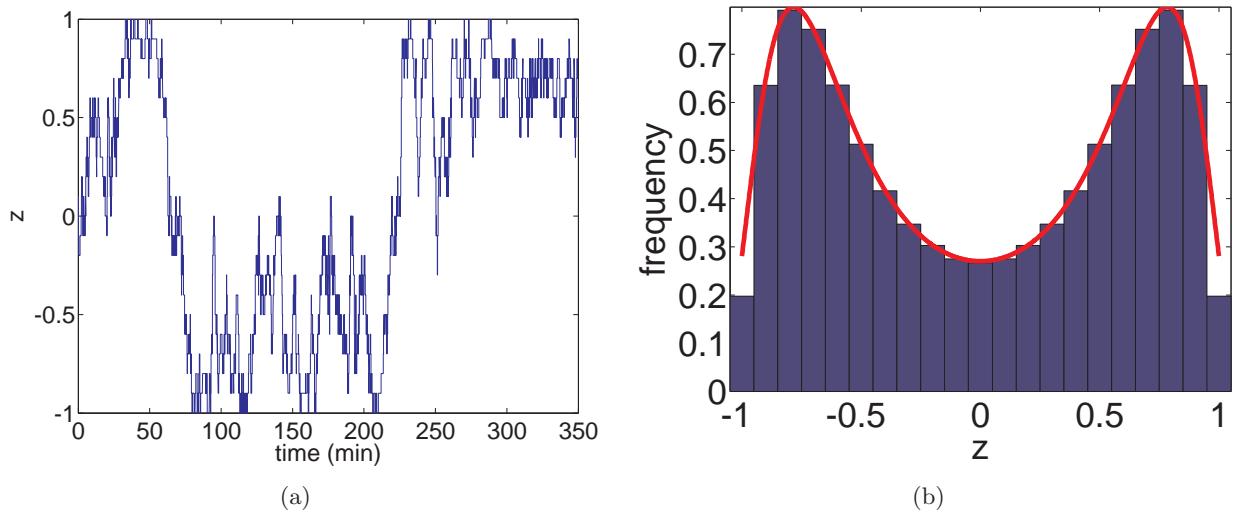


Figure 7.4: (a) The average direction of $N = 20$ model locusts exhibits similar switching behaviour to the experiments (see Fig 7.3 (b)) with average direction, z , between largely clockwise (+1) and largely anticlockwise (-1) movement of the group of locusts. (b) Stationary distribution of (7.9)-(7.11) obtained by the long-time simulation of the SSA (4a)-(4d) (blue histogram). Red solid line is the solution of the stationary chemical Fokker-Planck equation corresponding to this system (see Lecture 16). In both subfigures parameters (derived directly from the experimental data are $r_1 = 0.0225$, $r_2 = 0.0453$, $r_3 = 0.1664$.)

Undertaking basic stability analysis we see that there are three steady state values, $z_0 = 0$ and $z_{\pm} = \pm\sqrt{1 - 4r_1/r_3}$. The existence of the second two steady states relies on the condition $r_3 > 4r_1$ holding. Analysing the stability we find that the steady state z_0 is stable for $r_3 < 4r_1$. At $r_3 = 4r_1$ the system undergoes a supercritical pitchfork bifurcation and for $r_3 > 4r_1$, z_0 is unstable and the two steady states $z_{\pm} = \pm\sqrt{1 - 4r_1/r_3}$ are stable (see Fig 7.5 (a) for a bifurcation diagram).

Clearly, if we neglect third-order interactions between locusts (i.e. $r_3 = 0$), then we are in the situation where $r_3 < 4r_1$ and the system has a single stable steady state at $z_0 = 0$ in the deterministic system (and no unstable steady states). In the stochastic system, however, the value of z may continue to switch between two non-zero favourable states (see Fig 7.5 (b)), even when the deterministic system only exhibits a single stable steady state. We will explain this phenomenon in more detail in Lecture 16.

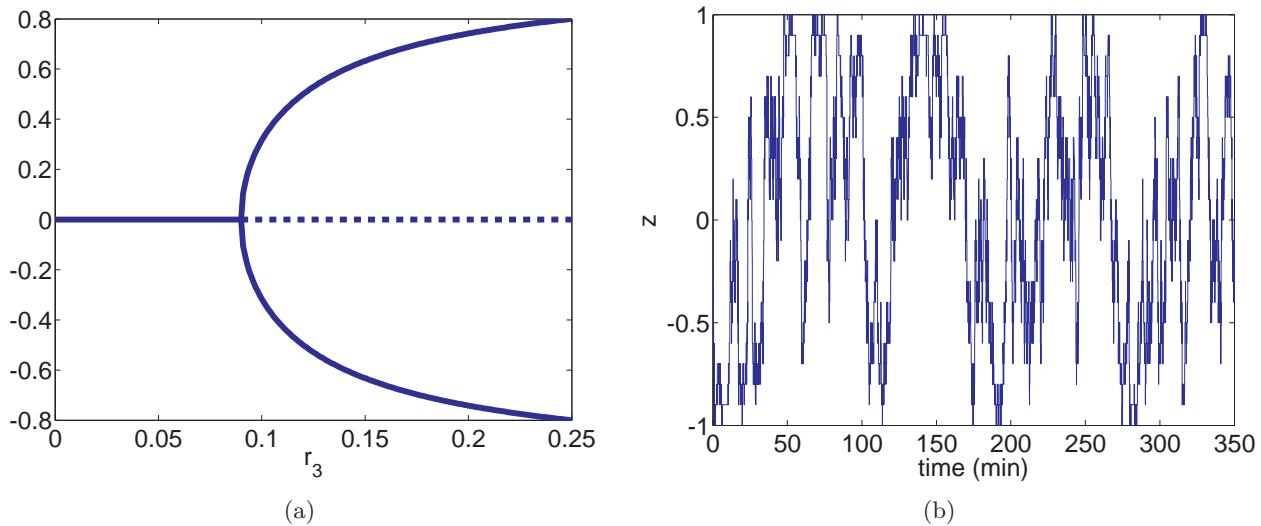


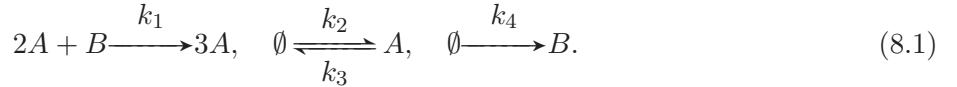
Figure 7.5: (a) The bifurcation diagram for the approximate deterministic system (7.15)-(7.16) exhibits a supercritical pitchfork bifurcation as parameter r_3 passes $4r_1$ for $r_1 = 0.0225$, $r_2 = 0.0453$. (b) Even when the deterministic system no longer exhibits bistability, the stochastic system may continue to switch between two favourable states. Parameter values are $r_1 = 0.0225$, $r_2 = 0.453$, $r_3 = 0$. Note r_2 is 10 times bigger than before.

Lecture 8 - Deterministic versus stochastic modelling IV

In the previous lecture we looked at systems with multiple favourable states in which the behaviour of the stochastic and deterministic systems differers dramatically due to the stochastic system exhibiting switching between the favourable states. Specifically we considered a simple system which can be used to model the collective behaviour of locusts. In that case the stochastic model had qualitatively different properties than its deterministic counterpart for some parameter regimes, that is, the stochastic model is not just “equal” to the “noisy solution” of the corresponding deterministic ODEs. Similarly in this lecture we present a simple system of chemical reactions for which the deterministic description converges to a steady state whereas the stochastic model of the same system of chemical reactions exhibits oscillatory behaviour.

8.1 Self-induced stochastic resonance

Our next example is a non-linear system of chemical equations for which the stochastic model has a qualitatively different behaviour than its deterministic counterpart in some parameter regimes. The phenomenon we present is sometimes called self-induced stochastic resonance [29, 30]. We consider a system of two chemical species A and B in a reactor of volume ν which are subject to the chemical reactions



Such a system was first studied by Schnakenberg [31]. The approximate deterministic description is given by the system of ODEs

$$\frac{da}{dt} = k_1 a^2 b + k_2 - k_3 a, \quad (8.2)$$

$$\frac{db}{dt} = -k_1 a^2 b + k_4, \quad (8.3)$$

where $a(t)$ and $b(t)$ are concentrations of A and B, respectively. Translating this into the (approximate) number of particles of A and B gives

$$\frac{d\bar{A}}{dt} = \frac{k_1}{\nu^2} \bar{A}^2 \bar{B} + k_2 \nu - k_3 \bar{A}, \quad (8.4)$$

$$\frac{d\bar{B}}{dt} = -\frac{k_1}{\nu^2} \bar{A}^2 \bar{B} + k_4 \nu, \quad (8.5)$$

where $\bar{A}(t) = a(t)\nu$ and $\bar{B}(t) = b(t)\nu$. We choose the rate constants

$$\frac{k_1}{\nu^2} = 4 \times 10^{-5} \text{ sec}^{-1}, k_2 \nu = 50 \text{ sec}^{-1}, k_3 = 10 \text{ sec}^{-1}, k_4 \nu = 25 \text{ sec}^{-1}. \quad (8.6)$$

We use the Gillespie SSA (4a)-(4d) to simulate the time evolution of this system. We also solve the deterministic system of ODEs (8.4)-(8.5). Using the same initial conditions $[A(0), B(0)] = [10, 10]$, we compare the results of the stochastic and deterministic models in Figure 8.1. We plot the time evolution of $A(t)$ in Figure 8.1 (a). We see that the solution of the deterministic equations converges to a steady state while the stochastic model has oscillatory solutions. Note that there is a *log scale* on the A -axis - numbers of A given by the (more precise) SSA vary between zero and ten thousand. In Figure 8.1 (b), we use a linear scale on the A -axis. On this scale the low molecular fluctuations are invisible and the solution of the SSA looks as if there were “almost deterministic oscillations” (which are not present in the deterministic model). We also plot in Figure 8.1 (b) the time evolution of the number of particles of B .

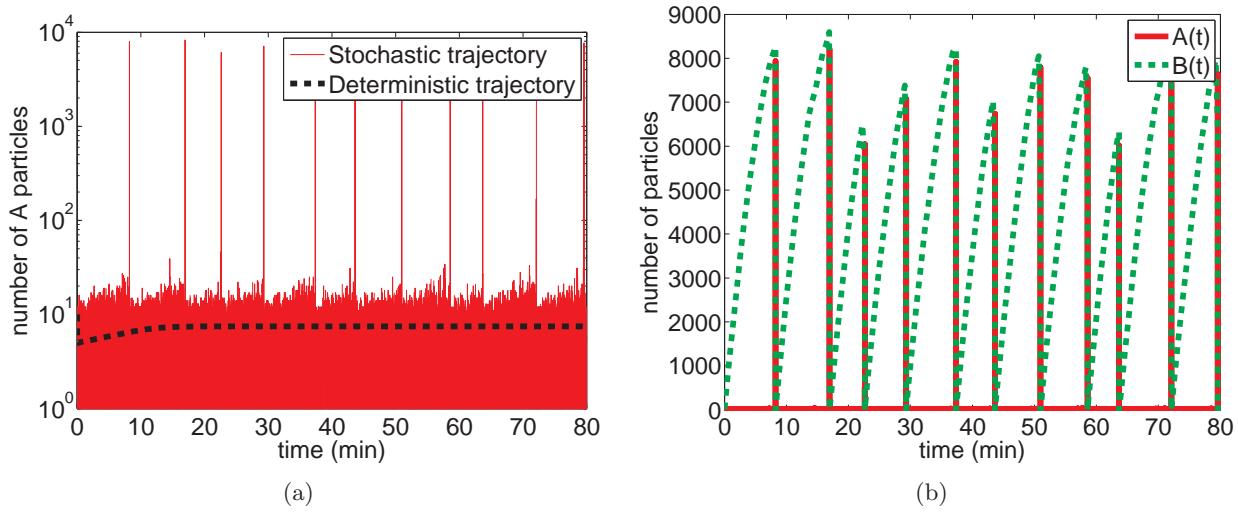


Figure 8.1: (a) Time evolution of $A(t)$ given by the Gillespie SSA (4a)-(4d) for the system of chemical reactions (8.1) with parameter values (8.6) (solid red line). Time evolution of $A(t)$ given by the system of ODEs (8.4)-(8.5) for parameter values (8.6) (black dashed line). Log scale on the A -axis. (b) Time evolution of $A(t)$ and $B(t)$ given by the Gillespie SSA (4a)-(4d) for the system of chemical reactions (8.1) for the values of rate constants (8.6).

The difference in behaviour between the stochastic and deterministic models of the chemical system (8.1) depends on the parameters. Changing the rate constants, we can find parameter regimes for which both systems have oscillatory solutions (although the periods of oscillations differ). To illustrate this point, we simulate the chemical system (8.1) with parameter values given by

$$\frac{k_1}{\nu^2} = 4 \times 10^{-5} \text{ sec}^{-1}, k_2\nu = 50 \text{ sec}^{-1}, k_3 = 10 \text{ sec}^{-1}, k_4\nu = 100 \text{ sec}^{-1}, \quad (8.7)$$

i.e. we change the value of the rate constant $k_4\nu$, keeping all other rate constants the same as in (8.6). In Figure 8.2, we compare the results obtained by the Gillespie SSA (4a)-(4d) with the solution of system of ODEs (8.4)-(8.5). We use the same initial conditions $[A(0), B(0)] = [10, 10]$ in both models. We see that both models have oscillatory solutions for the parameter values (8.7). However, the period of “stochastic oscillations” is shorter than the period of oscillations predicted by the ODEs.

To understand this behaviour better, we plot the stochastic and deterministic trajectories in the (A, B) -plane in Figure 8.3 (again we use a *log scale* on the A -axis). We include the nullclines of the deterministic system of ODEs (8.4)-(8.5) (green lines). In Figure 8.3 (a), we replot the trajectory from Figure 8.1 which was obtained for parameters (8.6). In Figure 8.3 (b), we replot the trajectory from Figure 8.2 which was obtained for parameters (8.7).

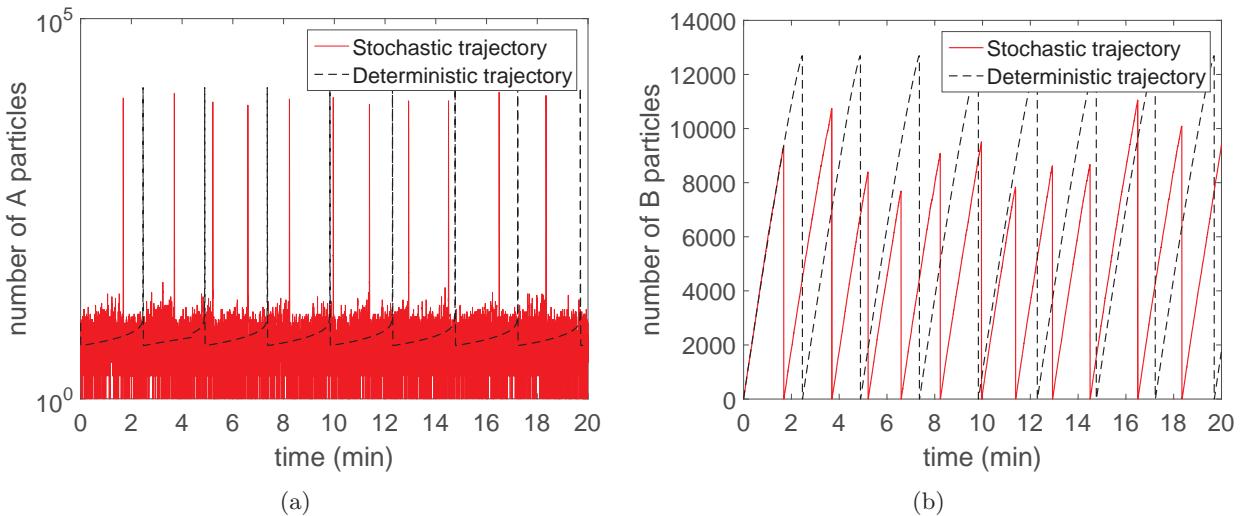


Figure 8.2: One realisation of SSA (4a)-(4d) for the system of chemical reactions (8.1) (solid red line) and the solution of the system of ODEs (8.4)-(8.5) (black dashed lines). Parameter values are given by (8.7). (a) The time evolution of $A(t)$. (b) The time evolution of $B(t)$.

The $[d\bar{A}/dt = 0]$ nullcline is given by

$$\overline{B} = \frac{k_3 \overline{A} - k_2 \nu}{\overline{A}^2 k_1 / \nu^2}. \quad (8.8)$$

We see that this is independent of $k_4\nu$, so that it appears as the same green curve in both panels of Figure 8.3. The maximum of this curve is attained at the point (see Problem Sheet 4)

$$\overline{A}_m = \frac{2k_2\nu}{k_3} = 10, \quad (8.9)$$

and it is equal to (see Problem Sheet 4)

$$\overline{B}_m = \frac{k_3^2}{4k_2\nu k_1/\nu^2} = 12500. \quad (8.10)$$

The $[d\bar{B}/dt = 0]$ nullcline is given by

$$\overline{B} = \frac{k_4 \nu}{\overline{A}^2 k_1 / \nu^2}. \quad (8.11)$$

This nullcline depends on $k_4\nu$. It intersects the nullcline (8.8) at the stable steady state (blue circle) for parameter values (8.6). Consequently, the deterministic system follows a stable nullcline into this steady state in Figure 8.3 (a). The stochastic model also initially “follows” the nullcline (8.8) (with some noise). However, occasionally the noise is enough to move the trajectory to the right of the right-hand branch of this nullcline. The deterministic trajectory starting from such a point exhibits a large excursion (sometimes called an action potential), before finally returning to the steady state. The stochastic solution now follows this new trajectory (again with some noise), before returning to the left-hand branch of the nullcline and repeating the procedure.

With parameter values (8.7) the nullcline (8.11) has shifted so that it intersects the other nullcline (8.8) on its unstable branch (as can be seen in Figure 8.3 (b)). The steady state (blue circle) is now unstable. The solution of ODEs (8.4)-(8.5) first follows the stable branch of the nullcline (8.8) to the top point $[\bar{A}_m, \bar{B}_m]$ (where the nullcline loses its stability) and then undergoes a large excursion before

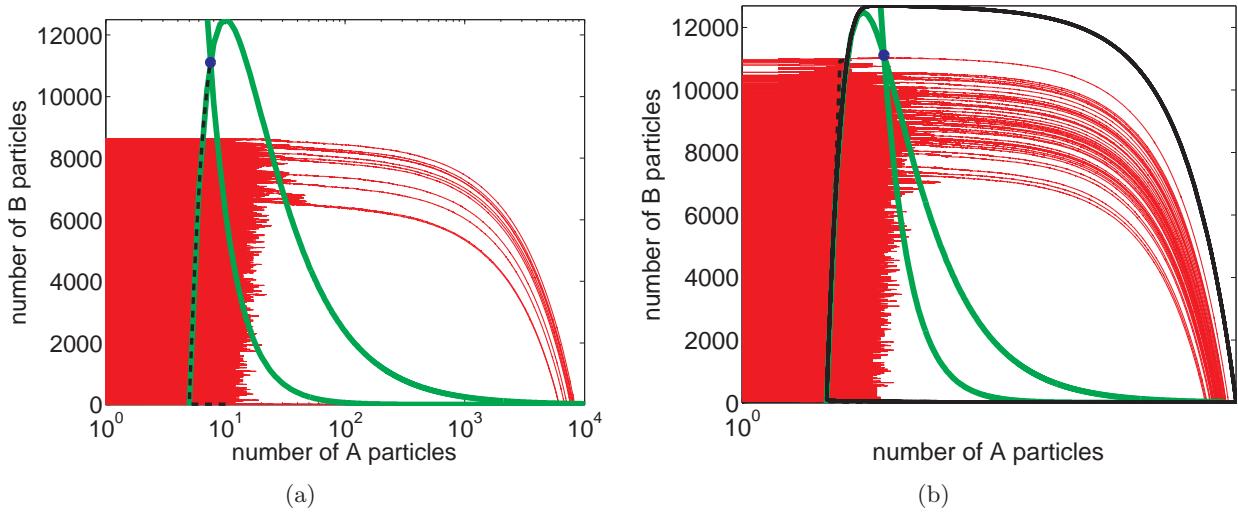


Figure 8.3: (a) One realisation of the SSA (4a)-(4d) for the system of chemical reactions (8.1) (solid red line) and the solution of the system of ODEs (8.4)-(8.5) (black dashed line). The parameters are given by (8.6). Nullclines of the deterministic ODEs are plotted as green lines. (b) The same plot for the parameters given by (8.7).

finally returning to this nullcline at a lower value of \bar{B} and repeating the procedure. The stochastic simulation follows a similar trajectory, but may start its excursion sooner if the noise takes it to the right of the unstable branch of the nullcline. Thus the (average) period of the stochastic oscillation is less than the period of the deterministic oscillation as seen anecdotally in Figure 8.2.

The system of ODEs (8.4)-(8.5) changes its qualitative behaviour dramatically as the parameter k_4 is varied from 25 sec^{-1} and 100 sec^{-1} . Such a change in qualitative behaviour is called a bifurcation. There exists a unique constant k_b (bifurcation value) in the interval $(25, 100) \text{ sec}^{-1}$ such that the solution of ODEs converges to the steady state for $k_4\nu < k_b$ and oscillates for $k_4\nu > k_b$. The value of the constant k_b can be identified by investigating the stability of the steady state of ODEs. We can also approximate k_b as the point for which the nullcline (8.11) intersects the other nullcline (8.8) at the top point $[\bar{A}_m, \bar{B}_m]$. Using (8.11), we obtain

$$\bar{B}_m = \frac{k_b}{\bar{A}_m^2 k_1 / \nu^2}.$$

Thus the bifurcation value k_b can be approximated as

$$k_b = \bar{A}_m^2 \bar{B}_m \frac{k_1}{\nu^2} = k_2 \nu = 50 \text{ sec}^{-1}.$$

The bifurcation of the deterministic system of ODEs (8.4)-(8.5) is a very sharp change in its qualitative behaviour. Choosing arbitrarily small ϵ , the solution of ODEs will oscillate for $k_4\nu = k_b + \epsilon$ and converge to the stable steady state for $k_4\nu = k_b - \epsilon$. The noise in the stochastic model makes it oscillate even for parameter values less than k_b (as we observed in Figures 8.1 and 8.3), smoothing the transition between two qualitatively different states of the system. This is generally true for any bifurcation [32]. Stochastic models will differ qualitatively from their ODE counterparts if the parameter values are close to the bifurcation values. Even for the case $k_4\nu > k_b$ where both models have oscillatory solutions, the periods of oscillations differ. It is interesting to investigate how the period of “stochastic oscillations” depends on the parameters. We will return to this problem in Section 17.2 after we develop a mathematical theory suitable for such an analysis.

Lecture 9 - Moment closure approximations I

In the previous lectures we have encountered systems for which the naive deterministic ODE model (which we might write down given the reaction system) is a poor description of the mean behaviour of the stochastic system. We have seen that the naive deterministic model system does a poor job when the system contains bimolecular (or higher order) reactions. In systems comprising only first or zeroth order reactions the naive deterministic ODE model is an exact description of the mean behaviour of the stochastic system (see Lectures 1 and 2).

In this lecture we will formalise the derivation of a naive deterministic model (usually known as the ‘mean-field’ model) for the mean behaviour of the system. Further, we will attempt to derive deterministic ODEs for higher moments. Typically, for systems which include higher (than first) order reactions, the ODEs for first moments will depend on the second (or higher order) moments. We may, therefore, attempt to write down an equation for the second moments. However, these in turn will depend on the third (or higher order) moments. Continuing this reasoning we will generate an infinite system of moment equations which we may not be able to solve analytically or numerically. In order to circumvent this problem, in this lecture and in Lecture 10 we will consider a number of different ways in which we might approximate the higher order moments in terms of the lower order moments in a process known as moment closure - effectively allowing us to derive a closed (but approximate) set of equations for some of the moments.

9.1 Cumulants

The cumulants provide an alternative to the moments. The cumulants determine the moments (and *vice versa*) in the sense that any two distributions with identical cumulants will also have identical moments. The cumulants of a random variable X are defined using the cumulant-generating function, $K(\theta)$, which is the natural log of the moment generating function (MGF):

$$K(\theta) = \log(M(\theta)), \quad (9.1)$$

where $M(\theta)$ is the MGF:

$$M(\theta) = \sum_{n=0}^{\infty} \exp(n\theta) p_n. \quad (9.2)$$

By repeated differentiation and evaluation at $\theta = 0$ the non-central moments can be evaluated from the MGF (hence the name). For example

$$\left. \frac{dM(\theta)}{d\theta} \right|_0 = \sum_{n=1}^{\infty} np_n = \langle n \rangle, \quad (9.3)$$

$$\left. \frac{d^2M(\theta)}{d\theta^2} \right|_0 = \sum_{n=1}^{\infty} n^2 p_n = \langle n^2 \rangle. \quad (9.4)$$

Similarly, by repeated differentiation and evaluation at $\theta = 0$, the first four cumulants, $\kappa_1, \kappa_2, \kappa_3$ and κ_4 , can be expressed in terms of the first four moments, $\mu_1 = \langle n \rangle, \mu_2 = \langle n^2 \rangle, \mu_3 = \langle n^3 \rangle$ and $\mu_4 = \langle n^4 \rangle$,

as follows:

$$\kappa_1 = \mu_1 = \langle n \rangle, \quad (9.5)$$

$$\kappa_2 = \mu_2 - \mu_1^2 = \langle n^2 \rangle - \langle n \rangle^2, \quad (9.6)$$

$$\kappa_3 = \mu_3 - 3\mu_2\mu_1 + 2\mu_1^3 = \langle n^3 \rangle - 3\langle n^2 \rangle \langle n \rangle + 2\langle n \rangle^3, \quad (9.7)$$

$$\begin{aligned} \kappa_4 &= \mu_4 - 4\mu_3\mu_1 - 3\mu_2^2 + 12\mu_2\mu_1^2 - 6\mu_1^4 \\ &= \langle n^4 \rangle - 4\langle n^3 \rangle \langle n \rangle - 3\langle n^2 \rangle^2 + 12\langle n^2 \rangle \langle n \rangle^2 - 6\langle n \rangle^4. \end{aligned} \quad (9.8)$$

Note here that the μ_i s denote the i th non-central moments not the central moments (i.e. mean, variance etc), although they can easily be calculated from each other.

9.2 Reversible Dimerisation Model

Consider chemical species A and B in a container of volume ν subject to the following two chemical reactions



The first reaction represents dimerisation of two molecules of A to the product B with rate constant k_1 . The second reaction represents the dissociation of the dimer B into two molecules of A with rate constant k_2 . This reaction system is similar to the dimerisation reaction system (5.1) described in Lecture 5, but distinct in that A and B must observe the following conservation relationship

$$A + 2B = A_0. \quad (9.10)$$

Here A_0 is the number of particles of A that would be present if they were fully dissociated.

Using this conservation relationship we can effectively eliminate B and write the system in terms of the monomer, A, alone. In particular, the propensity functions for the forwards and backwards reactions are

$$\alpha_1(t) = n(t)(n(t) - 1)k_1/\nu \text{ and } \alpha_2(t) = (A_0 - n(t))k_2, \quad (9.11)$$

respectively. Here $n(t)$ denotes the number of particles of A at time t . Note that in each propensity function we have absorbed the factor of 1/2 into the rate constant for notational convenience.

As standard, let us denote by $p_n(t)$ the probability that there are n particles of A at time t in the system. The chemical master equation for this system can be written in the following form

$$\frac{dp_n}{dt} = \frac{k_1}{\nu}(n+2)(n+1)p_{n+2}(t) - \frac{k_1}{\nu}n(n-1)p_n + k_2(n_0 - n+2)p_{n-2} - k_2(n_0 - n)p_n \quad (9.12)$$

where, we use the convention that $p_{-1} \equiv 0$, $p_{-2} \equiv 0$, $p_{n_0+1} \equiv 0$ and $p_{n_0+2} \equiv 0$ and $n_0 = A_0$.

The mean $M(t)$ and is defined by equation (2.5):

$$M(t) = \sum_{n=0}^{\infty} np_n(t). \quad (2.5)$$

Multiplying equation (9.12) by n and summing over n gives us the equation for the rate of change of the mean number of A particles (see Problem Sheet 5) as

$$\frac{d\langle n \rangle}{dt} = \frac{2k_1}{\nu} (\langle n \rangle - \langle n^2 \rangle) + 2k_2 (n_0 - \langle n \rangle). \quad (9.13)$$

Clearly we can see that the equation for the first moment, $\langle n \rangle$, depends upon the second moment, $\langle n^2 \rangle$, and consequently is not closed.

We can derive an equation for the second moment by multiplying equation (9.12) by n^2 and again summing over over the state space (see Problem sheet 5)

$$\frac{d\langle n^2 \rangle}{dt} = \frac{4k_1}{\nu} (-\langle n^3 \rangle + 2\langle n^2 \rangle - \langle n \rangle) + 4k_2 (n_0 (\langle n \rangle + 1) - \langle n^2 \rangle - \langle n \rangle). \quad (9.14)$$

This has not solved the problem, since the equation for the evolution of the second moment now depends on the third moment. We can go even further and derive an equation for the evolution of the third moment by multiplying equation (9.12) by n^3 and again summing over over the state space (not examinable)

$$\begin{aligned} \frac{d\langle n^3 \rangle}{dt} = & \frac{2k_1}{\nu} (4\langle n \rangle - 10\langle n^2 \rangle + 9\langle n^3 \rangle - 3\langle n^4 \rangle) \\ & + 2k_2 (4(n_0 - \langle n \rangle) + 6(n_0\langle n \rangle - \langle n^2 \rangle) + 3(n_0\langle n^2 \rangle - \langle n^3 \rangle)). \end{aligned} \quad (9.15)$$

Again, this equation is not closed, the right hand side depending on a yet a higher order moment ($\langle n^4 \rangle$). No matter how many equations we derive for higher and higher order moments, we will find that none of the equations are “closed” due to the non-linear propensity functions dictated by the non-linear nature of the reaction system. In order to “close” these equations we will make an approximation which allows us to express higher order moments in terms of lower order counterparts.

Note that the initial conditions for equations (9.13), (9.14) and (9.15) are $\langle n \rangle(0) = n_0$, $\langle n^2 \rangle(0) = n_0^2$ and $\langle n^3 \rangle(0) = n_0^3$, respectively.

9.3 Mean-field model

The mean-field model, is the resulting set of ODEs when we make the most basic possible closure for the second-order moments in terms of the first. Explicitly we write:

$$\langle n^2 \rangle = \langle n \rangle^2. \quad (9.16)$$

Recalling the definition of the variance from equation (2.5):

$$V(t) = \sum_{n=0}^{\infty} (n - M(t))^2 p_n(t) = \langle n^2 \rangle - \langle n \rangle^2, \quad (2.5)$$

this closure effectively assumes that the variance in the population is zero.

This is known as the *mean-field* moment closure and yields the *mean-field* model:

$$\frac{d\langle n \rangle}{dt} = -\frac{2k_1}{\nu} \langle n \rangle (\langle n \rangle - 1) + 2k_2 (n_0 - \langle n \rangle). \quad (9.17)$$

Figure 9.1 (a) compares the solution of the mean-field approximate ODE (9.17) to the mean of the stochastic process. The agreement is not terrible, but there is a noticeable discrepancy between the mean-field model and the mean of the stochastic simulations. Closing at a higher order will help to resolve these issues.

The mean-field approximation is the only sensible closure to make if we wish to close at first order. However, if we want to close at higher orders to obtain more accurate deterministic representations

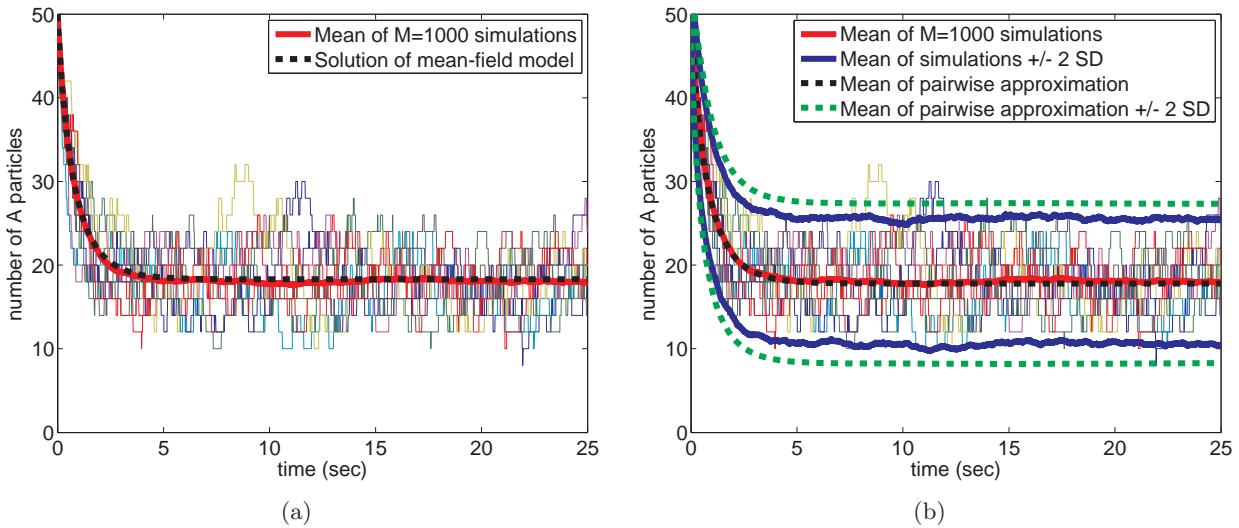


Figure 9.1: Simulations of reaction system (9.9) using the SSA (**4a**)-(**4d**). In both panels ten independent repeats for the number of particles of chemical species A are plotted as functions of time (thin solid coloured lines). The mean behaviour for A, averaged over 1000 repeats is given by the thick solid red line. In (a) the solutions of mean-field moment closure equation (9.17) is given by the black dashed lines. In (b) the first moment in the pairwise approximation is illustrated with the black-dashed line and the first moment ± 2 standard deviations is given by the solid thick blue lines for the stochastic simulation and by the dashed thick green lines for the pairwise moment closure approximations. In both panels we use $A(0) = 50$, $B(0) = 0$, $k_1/\nu = 0.01 \text{ sec}^{-1}$, $k_2 = 0.1 \text{ sec}^{-1}$.

then there are a number of different ways to achieve this.

9.4 Pairwise approximation

If we wish to close our deterministic equations at the next highest order then we need to find an expression for the third-order moments in terms of the second-order moments. One such expression is known as the pairwise approximation. For a single chemical species, the pairwise approximation takes the following form:

$$\langle n^3 \rangle = \frac{\langle n^2 \rangle \langle n^2 \rangle}{\langle n \rangle}. \quad (9.18)$$

Applying this approximation to equation (9.14) closes the system and leads to the following equation for the second moment

$$\frac{d\langle n^2 \rangle}{dt} = \frac{4k_1}{\nu} \left(-\frac{\langle n^2 \rangle^2}{\langle n \rangle} + 2\langle n^2 \rangle - \langle n \rangle \right) + 4k_2 \left(n_0 (\langle n \rangle + 1) - \langle n^2 \rangle - \langle n \rangle \right). \quad (9.19)$$

The pairwise approximation to the mean is plotted as the thick dashed black line in Figure 9.1 (b). This can be seen to be a slightly better approximation to the mean than for the basic mean-field closure (although if anything the mean is now slightly under-estimated rather than over-estimated). We also provide an approximate 95% confidence interval (± 2 standard deviations about the mean). It can be seen that the pairwise closure approximation to the second moment is still inaccurate. Closing at a higher order can correct this, but often the higher order moments we approximate will disagree with their true values.

9.5 Cumulant neglect

One popular and general method for moment closure involves cumulants. In the method of *cumulant neglect*, a level is chosen past which all cumulants are set to zero. This provides relationships between the moments which can be used to close the system of moment equations [33, 34]. Setting all third and higher order cumulants to zero means we have closed evolution equations for the first two moments. This is often known as the *normal* moment closure since the normal distribution is defined by its first and second-order cumulants (which correspond to the mean and the variance), but all higher order cumulants are zero (see Problem Sheet 5).

Setting the third cumulant to zero in equation (9.7) gives the normal moment closure for the third moment:

$$\langle n^3 \rangle = 3\langle n^2 \rangle \langle n \rangle - 2\langle n \rangle^3. \quad (9.20)$$

Substituting this expression into equation (9.14) allows us to close the system at second-order with the resultant equation for the second moment:

$$\frac{d\langle n^2 \rangle}{dt} = \frac{4k_1}{\nu} \left(2\langle n \rangle^3 - 3\langle n^2 \rangle \langle n \rangle + 2\langle n^2 \rangle - \langle n \rangle \right) + 4k_2 \left(n_0 (\langle n \rangle + 1) - \langle n^2 \rangle - \langle n \rangle \right). \quad (9.21)$$

The normal approximation to the mean, found by solving equations (9.13) and (9.21) is plotted as the thick dashed black line in Figure 9.2 (a). This is an even better approximation to the mean than either of the previous closures. The approximate 95% confidence intervals are also improved under the normal approximation. Figure 9.2 (b) demonstrates the good agreement between the stationary distribution, found by long term stochastic simulation, and the normal approximation to the stationary distribution, generated using moments from the solution of the stationary versions of equations (9.13) and (9.21). Specifically, the stationary mean, M_s , can be found as the solution of the cubic equation (see Problem sheet 5)

$$2M_s^3 + 3 \left(\frac{k_2\nu}{k_1} - 1 \right) M_s^2 + \left(\left(\frac{k_2\nu}{k_1} \right)^2 - 2\frac{k_2\nu}{k_1}(2+n_0) + 1 \right) M_s + n_0 \frac{k_2\nu}{k_1} \left(3 - \frac{k_2\nu}{k_1} \right) = 0, \quad (9.22)$$

which, for the parameter values given in Figure 9.1 can be evaluated numerically as $M_s \approx 17.99$. Consequently the stationary second moment, S_s , can be found from the relationship (see Problem sheet 5)

$$S_s = \frac{k_2\nu}{k_1} (n_0 - M_s) + M_s. \quad (9.23)$$

Similarly S_s can be evaluated numerically as $S_s = 338.05$.

The reason why this is such an effective closure in this case is that the distribution is well approximated by the normal distribution. Note, however, as we will see in Lecture 10 that this will not always be the case.

The cumulant neglect approach is more general than the other approaches mentioned thus far in that it is not specific to a particular order but is able to provide a closure at any order depending on the level above which cumulants are neglected. Setting all fourth and higher order cumulants to zero, for example, will provide closed evolution equations for the first three moments. In particular, setting the fourth cumulant equal to zero provides us with the following relationship between the fourth and lower order moments:

$$\langle n^4 \rangle = 4\langle n^3 \rangle \langle n \rangle + 3\langle n^2 \rangle^2 - 12\langle n^2 \rangle \langle n \rangle^2 + 6\langle n \rangle^4. \quad (9.24)$$

Substituting this expression into equation (9.15) allows us to close the system at third-order with the

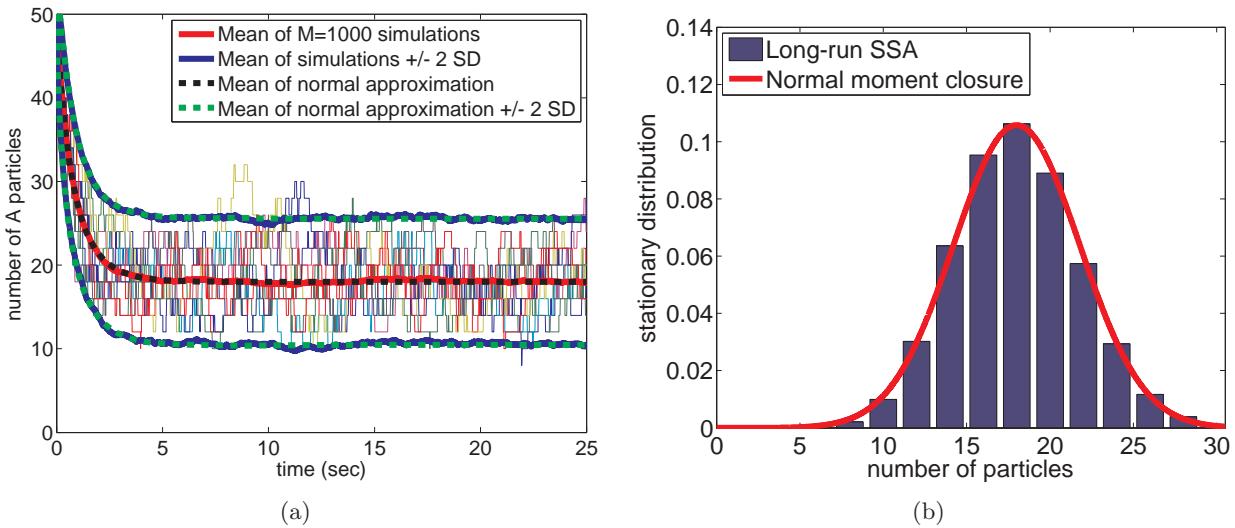


Figure 9.2: Simulations of reaction system (9.9) using the SSA (4a)-(4d). (a) Ten independent repeats for the number of particles of chemical species A are plotted as functions of time (thin solid coloured lines). The mean behaviour for A, averaged over 1000 repeats is given by the thick solid red line. The first moment in the normal approximation is illustrated with the black-dashed line and the first moment ± 2 standard deviations is given by the solid thick blue lines for the stochastic simulation and by the dashed thick green lines for the normal moment closure approximations. (b) Stationary distribution obtained by long-time simulation of the SSA (blue histogram) and by the normal approximation to the stationary distribution using moments from the solution of the stationary versions of equations (9.13) and (9.21) (thick red line). Note that the stochastic data is binned in boxes of width two since it is only possible for the value of A to take even values due to the nature of the reaction system. Initial conditions and parameters are as in Figure 9.1.

resultant equation for the third moment:

$$\begin{aligned} \frac{d\langle n^3 \rangle}{dt} = & \frac{2k_1}{\nu} \left(4\langle n \rangle - 10\langle n^2 \rangle + 9\langle n^3 \rangle - 12\langle n^3 \rangle \langle n \rangle - 9\langle n^2 \rangle^2 + 36\langle n^2 \rangle \langle n \rangle^2 - 18\langle n \rangle^4 \right) \\ & + 2k_2 \left(4(n_0 - \langle n \rangle) + 6(n_0 \langle n \rangle - \langle n^2 \rangle) + 3(n_0 \langle n^2 \rangle - \langle n^3 \rangle) \right). \end{aligned} \quad (9.25)$$

Figure 9.3 (a) demonstrates that the fourth-order cumulant neglect method provides an excellent approximation to the mean and the approximate 95% confidence intervals, although not considerably better than the simpler normal approximation. This is not particularly surprising since we have seen in Figure 9.2 (b) that the distribution is well approximated as normal. Incorporating an equation for a further moment will do little to improve the fit. Investigating higher-order closures in this way is a good way of determining the quality of a particular closure. Upon closing at successively higher moments, once the estimates stop changing sufficiently many moments have been taking into account.

9.6 Kirkwood superposition approximation

Another possibility in order to derive closed equations for the first and second moments is the Kirkwood superposition approximation. For a single variable this simplifies to the following relationship:

$$\langle n^3 \rangle = \frac{\langle n^2 \rangle^3}{\langle n \rangle^3}. \quad (9.26)$$

Combining this approximation with equation (9.14) gives the following equation for the second moment:

$$\frac{d\langle n^2 \rangle}{dt} = \frac{4k_1}{\nu} \left(-\frac{\langle n^2 \rangle^3}{\langle n \rangle^3} + 2\langle n^2 \rangle - \langle n \rangle \right) + 4k_2 \left(n_0 (\langle n \rangle + 1) - \langle n^2 \rangle - \langle n \rangle \right). \quad (9.27)$$

Figure 9.3 (b) demonstrates that the Kirkwood superposition approximation also provides a good approximation to the mean and the approximate 95% confidence intervals.

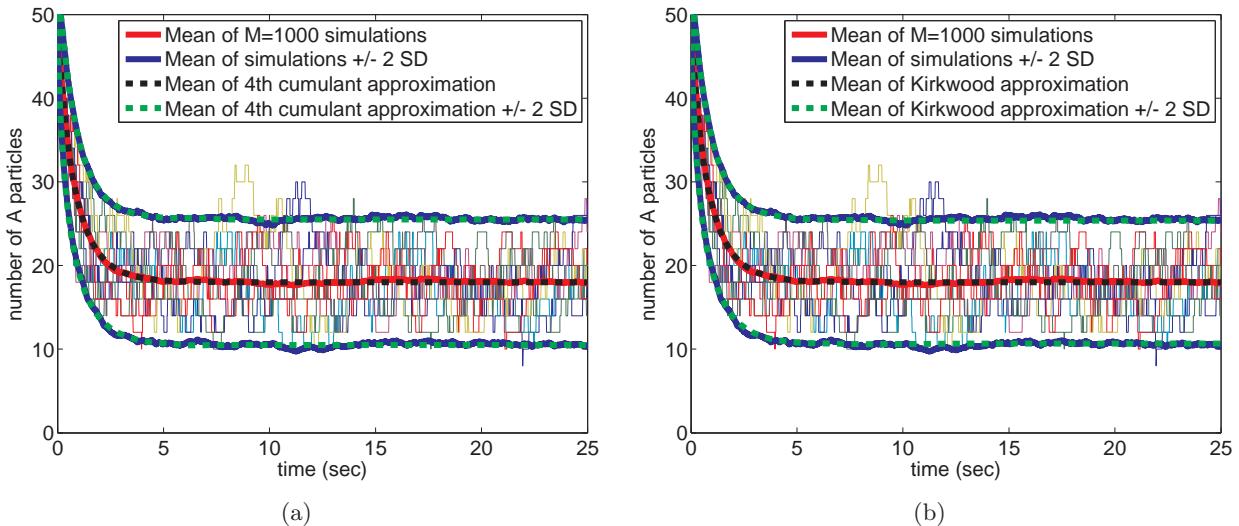


Figure 9.3: Simulations of reaction system (9.9) using the SSA (4a)-(4d). Ten independent repeats for the number of particles of chemical species A are plotted as functions of time (thin solid coloured lines). The mean behaviour for A, averaged over 1000 repeats is given by the thick solid red line. The first moment in (a) the fourth order cumulant neglect approximation and (b) the Kirkwood superposition approximation are illustrated with black-dashed lines. The first moment ± 2 standard deviations are given by the solid thick blue lines for the stochastic simulation and by the dashed thick green lines for (a) the fourth order cumulant neglect approximation and (b) the Kirkwood superposition approximation. Initial conditions and parameters are as in Figure 9.1

Lecture 10 - Moment closure approximations II

In the previous lecture we introduced several moment closure schemes for a model example with a single dependent variable. In this lecture we look at extending the definition of the various closure approximations to models with multiple dependent variables. In particular, we will consider a model of the spread of an epidemic in a fixed population.

10.1 SIR example

One of the most important mathematical models currently in use describes the spread of infectious diseases in a population. To do this, the population is compartmentalised according to the infection status. One simple version of this model considers a population of fixed size divided into subgroups: individuals that are susceptible to infection, individuals that are infected and individuals who have recovered from the disease at some time. Let

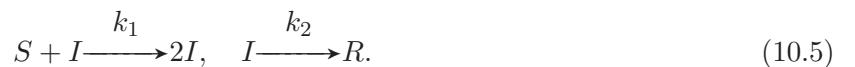
$$S(t) = \text{Numbers of susceptibles at time } t, \quad (10.1)$$

$$I(t) = \text{Numbers of infected individuals at time } t, \quad (10.2)$$

$$R(t) = \text{Numbers of removed individuals at time } t, \quad (10.3)$$

$$N = S(t) + I(t) + R(t), \quad (10.4)$$

so that $N(t)$ denotes the total population size at time t . The epidemic plays out in a region of size ν , through the following pair of interactions:



Since the total population is fixed we can express $R = N - I - S$ and need not account for R in the corresponding probability master equation. Let $p(n, m, t)$ be the probability that $S(t) = n$ and $I(t) = m$. The probability master equation can be written in the following form

$$\begin{aligned} \frac{dp(n, m)}{dt} &= \frac{k_1}{\nu}(n+1)(m-1)p(n+1, m-1) - \frac{k_1}{\nu}nmp(n, m) \\ &\quad + k_2(m+1)p(n, m+1) - k_2mp(n, m), \end{aligned} \quad (10.6)$$

for $n, m \geq 0$, with the convention that $p(n, m) \equiv 0$ if $n < 0$, $m < 0$, $n > N$ or $m > N$. Note that we have dropped the notation for the dependence of $p(n, m, t)$ on t for convenience. The mean number of susceptibles, $\langle S \rangle$, and the mean number of infectives, $\langle I \rangle$, are defined as

$$\langle S \rangle = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} np(n, m, t), \quad \langle I \rangle = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} mp(n, m, t). \quad (10.7)$$

Similarly, definitions of the second-order moments are as follows

$$\langle SI \rangle = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} nmp(n, m, t), \quad (10.8)$$

$$\langle S^2 \rangle = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} n^2 p(n, m, t), \quad (10.9)$$

$$\langle I^2 \rangle = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} m^2 p(n, m, t), \quad (10.10)$$

with higher order moments defined similarly. Multiplying master equation (10.6) by n and summing over n and m gives us the equation for the rate of change of the mean number of susceptibles (see Problem Sheet 5) as

$$\frac{d\langle S \rangle}{dt} = -\frac{k_1}{\nu} \langle SI \rangle. \quad (10.11)$$

Multiplying master equation (10.6) by m and summing over n and m gives us the equation for the rate of change of the mean number of infectives (see Problem Sheet 5) as

$$\frac{d\langle I \rangle}{dt} = \frac{k_1}{\nu} \langle SI \rangle - k_2 \langle I \rangle. \quad (10.12)$$

Clearly we can see that the equation for the first moments, $\langle S \rangle$ and $\langle I \rangle$, depend upon one of the second moments, $\langle SI \rangle$, and consequently the equations are not closed. We must derive an equation for the evolution of the second-order moment $\langle SI \rangle$. Multiplying master equation (10.6) by nm and summing over n and m gives us the relevant equation (see Problem Sheet 5):

$$\frac{d\langle SI \rangle}{dt} = \frac{k_1}{\nu} \left(\langle S^2 I \rangle - \langle SI^2 \rangle - \langle SI \rangle \right) - k_2 \langle SI \rangle. \quad (10.13)$$

As expected this equation is not closed either, containing expressions for higher moments.

If we wish to close this equation at second-order, most closure schemes will require equations for the evolution of $\langle S^2 \rangle$ and $\langle I^2 \rangle$ in addition to $\langle SI \rangle$. These can be derived in a similar manner to equation (10.13) (see Problem Sheet 5) as

$$\frac{d\langle S^2 \rangle}{dt} = \frac{k_1}{\nu} \left(\langle SI \rangle - 2\langle S^2 I \rangle \right), \quad (10.14)$$

$$\frac{d\langle I^2 \rangle}{dt} = \frac{k_1}{\nu} \left(2\langle SI^2 \rangle + \langle SI \rangle \right) + k_2 \left(\langle I \rangle - 2\langle I^2 \rangle \right). \quad (10.15)$$

10.2 Mean-field model

In the general case, the mean-field model results from the assumption that the mean of a product can be expressed as the product of means. Specifically, in the case of the SIR model, these assumptions are manifested in the following equations:

$$\langle SI \rangle = \langle S \rangle \langle I \rangle, \quad (10.16)$$

$$\langle S^2 \rangle = \langle S \rangle^2, \quad (10.17)$$

$$\langle I^2 \rangle = \langle I \rangle^2. \quad (10.18)$$

Note that these assumptions imply that the variance of S and I , as well as the covariance between S and I are all zero. Substituting expressions (10.16)-(10.18) into equations (10.11) and (10.12) gives

the mean-field model:

$$\frac{d\langle S \rangle}{dt} = -\frac{k_1}{\nu} \langle S \rangle \langle I \rangle, \quad (10.19)$$

$$\frac{d\langle I \rangle}{dt} = \frac{k_1}{\nu} \langle S \rangle \langle I \rangle - k_2 \langle I \rangle. \quad (10.20)$$

Figure 10.1 compares the mean-field approximate ODEs (10.19)-(10.20) to the means, for S and I respectively, of the stochastic process. Although the qualitative behaviour of the mean-field model mimics the mean of the stochastic model, the quantitative agreement is not perfect. Closure at a higher order may help to improve the agreement.

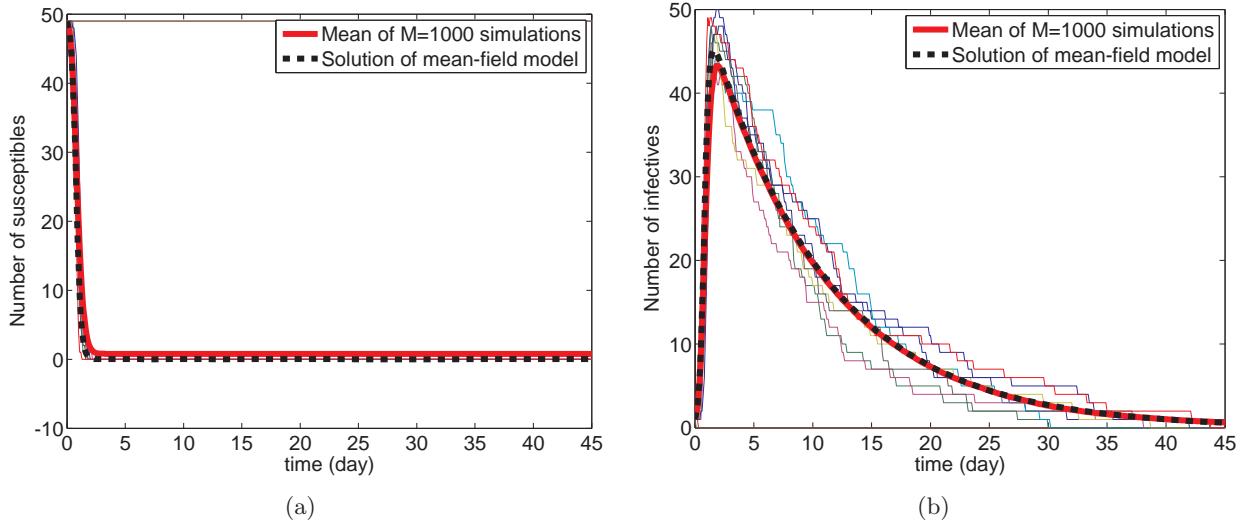


Figure 10.1: Mean-field moment closure. Simulations of the SIR model of disease spread (10.5) using the SSA (4a)-(4d). In panel (a) ten independent repeats for the number of susceptibles, S , are plotted as functions of time (thin solid coloured lines). The mean behaviour for S , averaged over 1000 repeats is given by the thick solid red line. The solution of mean-field moment closure equation (10.19) is given by the black dashed line. In (b) we plot the same quantities for the number of infectives, I , using mean-field moment closure equation (10.20). Initial conditions are $S(0) = 49$, $I(0) = 1$ with parameter values $k_1/\nu = 0.1 \text{ day}^{-1}$, $k_2 = 0.1 \text{ day}^{-1}$.

10.3 Pairwise approximation

For multiple dependent variables A , B and C , the pairwise approximation takes the form

$$\langle ABC \rangle = \frac{\langle AB \rangle \langle BC \rangle}{\langle B \rangle}. \quad (10.21)$$

This definition is not unique since species A or C could readily be swapped with B in the above definition. For $A, B, C \in \{S, I\}$ we use the closures

$$\langle S^2 I \rangle = \frac{\langle S^2 \rangle \langle SI \rangle}{\langle S \rangle} \quad \text{and} \quad (10.22)$$

$$\langle SI^2 \rangle = \frac{\langle SI \rangle \langle I^2 \rangle}{\langle I \rangle}. \quad (10.23)$$

Applying these approximations to equations (10.13)-(10.15) leads to the following closed equations for the second moments

$$\frac{d\langle SI \rangle}{dt} = \frac{k_1}{\nu} \left(\frac{\langle SI \rangle \langle S^2 \rangle}{\langle S \rangle} - \frac{\langle SI \rangle \langle I^2 \rangle}{\langle I \rangle} - \langle SI \rangle \right) - k_2 \langle SI \rangle, \quad (10.24)$$

$$\frac{d\langle S^2 \rangle}{dt} = \frac{k_1}{\nu} \left(\langle SI \rangle - 2 \frac{\langle SI \rangle \langle S^2 \rangle}{\langle I \rangle} \right), \quad (10.25)$$

$$\frac{d\langle I^2 \rangle}{dt} = \frac{k_1}{\nu} \left(2 \frac{\langle SI \rangle \langle I^2 \rangle}{\langle I \rangle} + \langle SI \rangle \right) + k_2 \left(\langle I \rangle - 2 \langle I^2 \rangle \right). \quad (10.26)$$

The pairwise approximations to the mean of S and I are plotted as the thick dashed black line in Figure 10.2 (a) and (b), respectively. The agreement between the mean of the stochastic model and the pairwise approximation is poor (much poorer than the mean-field model). We also plot ± 1 standard deviations about the mean in order to give a sense of the variation. It can be seen that the pairwise closure approximation to the second moment is also inaccurate.

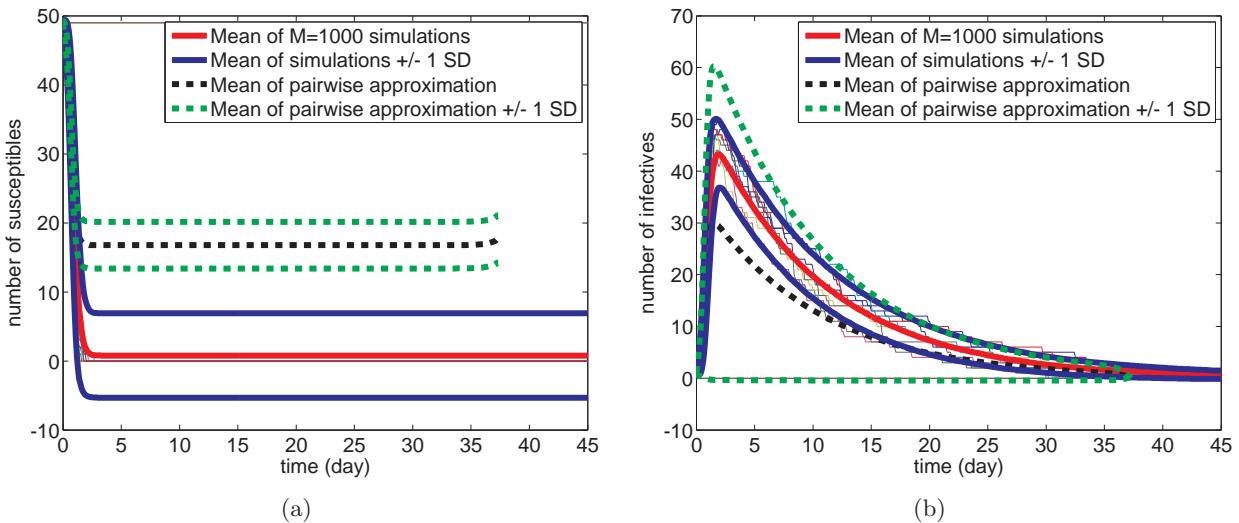


Figure 10.2: Pairwise moment closure. (a) Susceptibles and (b) infectives. In both panels, the first moment ± 1 standard deviation is given by the solid thick blue lines for the stochastic simulation and by the dashed thick green lines for the pairwise moment closure approximations. Otherwise, figure descriptions, parameter values and initial conditions are as in Figure 10.1.

10.4 Cumulant neglect

In order to use the method of cumulant neglect, we need to find expressions for the cumulants we wish to neglect in terms of the moments. One way to find such relationships is to use the bivariate cumulant generating function:

$$K(\theta_S, \theta_I) = \log(M(\theta_S, \theta_I)), \quad (10.27)$$

where $M(\theta_S, \theta_I)$ is the bivariate moment generating function:

$$M(\theta_S, \theta_I) = \mathbb{E}(\exp(\theta_S S) \exp(\theta_I I)) = \sum_{S=0}^{\infty} \sum_{I=0}^{\infty} p(S, I) \exp(\theta_S S) \exp(\theta_I I), \quad (10.28)$$

where for simplicity we have dropped the dependence on time. Recall that in order to find the general moment, μ_{n_S, n_I} , from the moment generating function, we must differentiate n_S times with respect to

θ_S, n_I times with respect to θ_I and evaluate the resulting derivative at $(\theta_S, \theta_I) = (0, 0)$:

$$\mu_{n_S, n_I} = \left. \frac{\partial^{n_S + n_I} M(\theta_S, \theta_I)}{\partial \theta_S^{n_S} \partial \theta_I^{n_I}} \right|_{0,0}. \quad (10.29)$$

A similar definition holds true for evaluating the cumulants from the cumulant generating function:

$$\kappa_{n_S, n_I} = \left. \frac{\partial^{n_S + n_I} K(\theta_S, \theta_I)}{\partial \theta_S^{n_S} \partial \theta_I^{n_I}} \right|_{0,0}. \quad (10.30)$$

To find the cumulants in terms of the moments we simply differentiate equation (10.27) with respect to the relevant variables using the chain rule and evaluate the resulting expression at $\theta_S = \theta_I = 0$. For example, we find an expression for the first cumulants of S in terms of the first moments as follows:

$$\kappa_{1,0} = \left. \frac{\partial K(\theta_S, \theta_I)}{\partial \theta_S} \right|_{0,0} = \frac{1}{M} \left. \frac{\partial M(\theta_S, \theta_I)}{\partial \theta_S} \right|_{0,0} = \mu_{1,0} = \langle S \rangle. \quad (10.31)$$

Similarly, we find expression for the cumulant, $\kappa_{1,1}$ as follows:

$$\kappa_{1,1} = \left. \frac{\partial^2 K(\theta_S, \theta_I)}{\partial \theta_S \partial \theta_I} \right|_{0,0} = \frac{1}{M} \left. \frac{\partial^2 M}{\partial \theta_S \partial \theta_I} \right|_{0,0} - \frac{1}{M^2} \left. \frac{\partial M}{\partial \theta_S} \frac{\partial M}{\partial \theta_I} \right|_{0,0} = \mu_{1,1} - \mu_{1,0} \mu_{0,1} = \langle SI \rangle - \langle S \rangle \langle I \rangle. \quad (10.32)$$

In order to achieve the normal closure we need to neglect third-order cumulants and use the resulting relationships to substitute for the third-order moments in equations (10.13)-(10.15). The third-order cumulants are given by (see Problem Sheet 5)

$$\kappa_{3,0} = \mu_{3,0} + 2\mu_{1,0}^3 - 3\mu_{1,0}\mu_{2,0} = \langle S^3 \rangle + 2\langle S \rangle^3 - 3\langle S \rangle \langle S^2 \rangle, \quad (10.33)$$

$$\kappa_{2,1} = \mu_{2,1} + 2\mu_{0,1}\mu_{1,0}^2 - \mu_{2,0}\mu_{0,1} - 2\mu_{1,0}\mu_{1,1} = \langle S^2 I \rangle + 2\langle I \rangle \langle S \rangle^2 - \langle S^2 \rangle \langle I \rangle - 2\langle S \rangle \langle SI \rangle, \quad (10.34)$$

$$\kappa_{1,2} = \mu_{1,2} + 2\mu_{1,0}\mu_{0,1}^2 - \mu_{0,2}\mu_{1,0} - 2\mu_{0,1}\mu_{1,1} = \langle I^2 S \rangle + 2\langle S \rangle \langle I \rangle^2 - \langle I^2 \rangle \langle S \rangle - 2\langle I \rangle \langle SI \rangle, \quad (10.35)$$

$$\kappa_{0,3} = \mu_{0,3} + 2\mu_{0,1}^3 - 3\mu_{0,1}\mu_{0,2} = \langle I^3 \rangle + 2\langle I \rangle^3 - 3\langle I \rangle \langle I^2 \rangle. \quad (10.36)$$

Upon setting the third-order cumulants equal to zero we find expressions for the third-order moments in terms of the lower order moments. This is the bi-variate normal moment closure scheme. The resulting closed moment equations are

$$\begin{aligned} \frac{d\langle SI \rangle}{dt} &= \frac{k_1}{\nu} \left(\langle S^2 \rangle \langle I \rangle + 2\langle S \rangle \langle SI \rangle - 2\langle I \rangle \langle S \rangle^2 - \langle S \rangle \langle I^2 \rangle - 2\langle I \rangle \langle SI \rangle + 2\langle S \rangle \langle I \rangle^2 - \langle SI \rangle \right) \\ &\quad - k_2 \langle SI \rangle, \end{aligned} \quad (10.37)$$

$$\frac{d\langle S^2 \rangle}{dt} = \frac{k_1}{\nu} \left(\langle SI \rangle - 2\langle S^2 \rangle \langle I \rangle - 4\langle S \rangle \langle SI \rangle + 4\langle I \rangle \langle S \rangle^2 \right), \quad (10.38)$$

$$\frac{d\langle I^2 \rangle}{dt} = \frac{k_1}{\nu} \left(2\langle S \rangle \langle I^2 \rangle + 4\langle I \rangle \langle SI \rangle - 4\langle S \rangle \langle I \rangle^2 + \langle SI \rangle \right) + k_2 \left(\langle I \rangle - 2\langle I^2 \rangle \right). \quad (10.39)$$

The normal approximation to the means for S and I are plotted as the thick dashed black lines in Figure 10.3 (a) and (b), respectively. Although the normal approximation is clearly better than the pairwise moment closure, it is arguable as to whether it is superior to the mean-field moment closure. In part this is because the joint distribution of the number of susceptibles and infectives is poorly approximated by a bivariate normal distribution. In Figure 10.4 (a) and (b) (respectively) the marginal distributions of the number of susceptibles and infectives (respectively) at time $t = 1$ days are compared with the corresponding normal distributions. The agreement is poor since the underlying marginal distributions of susceptibles and infectives is clearly not normal. Despite choosing parameter

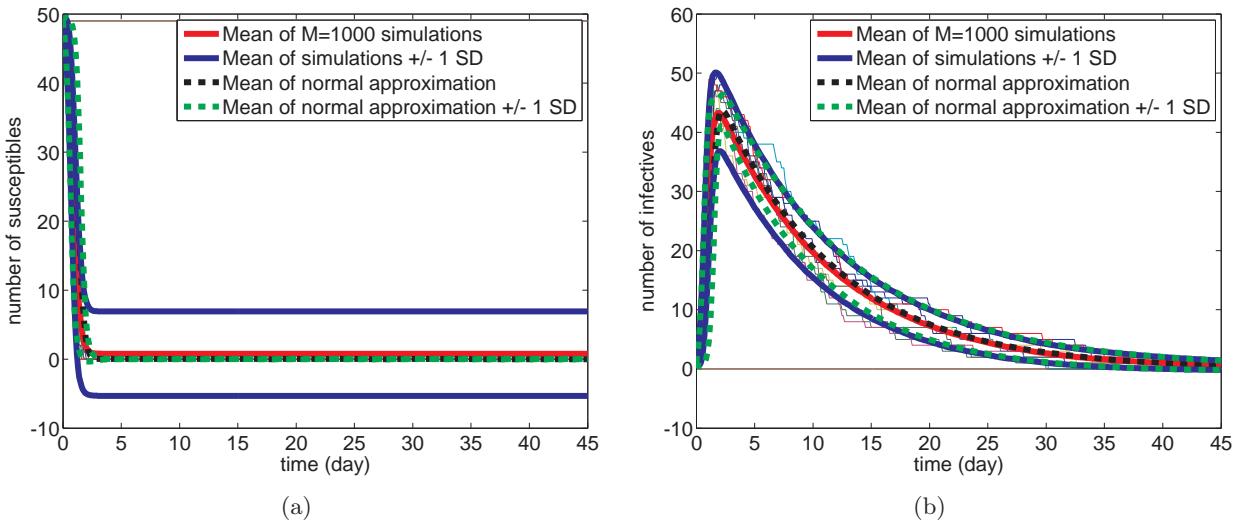


Figure 10.3: Normal moment closure. (a) Susceptibles and (b) infectives. Figure descriptions, parameter values and initial conditions are as in Figure 10.2.

values which give $R_0 (= N_0 k_1 / \nu k_2) > 1$, in the stochastic simulations the infection does not always break out where it would in the deterministic mean-field model. There are a significant number of simulations for which susceptible numbers remain at 49 and infective numbers at 1. This bimodality in the distributions further hampers the normal approximation.

Using cumulant neglect at higher orders may allow the better approximation of these distributions, although deriving the closed equations rapidly becomes complicated, especially for these multivariate problems. Closing at fourth order by neglecting fourth order cumulants, for example, does little to improve the agreement and is omitted here.

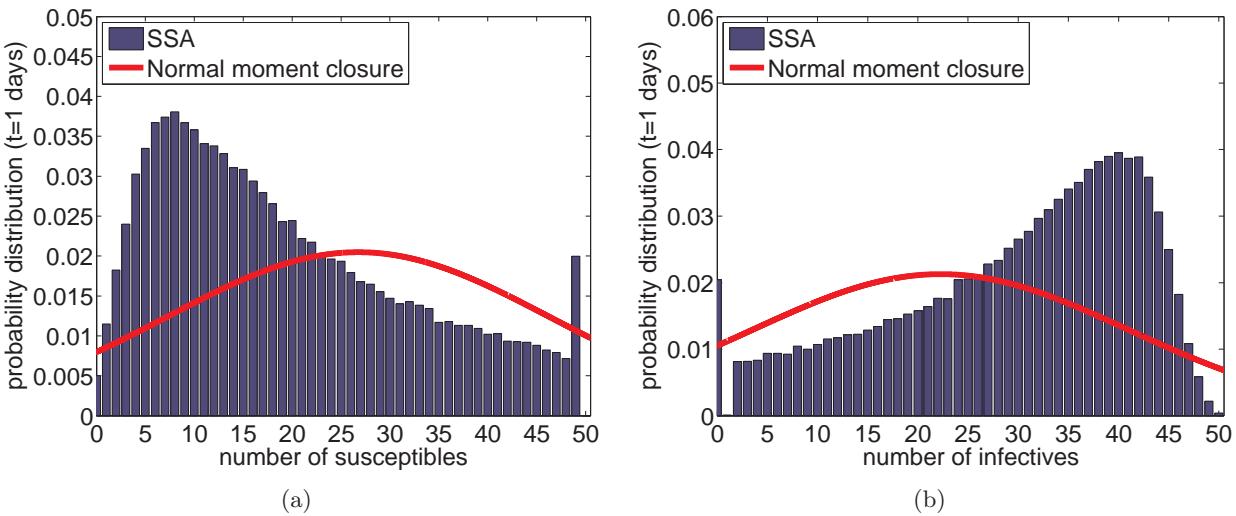


Figure 10.4: Probability distribution for (a) $S(1)$ and (b) $I(1)$ at $t = 1$ obtained by multiple repeats (100,000) of the SSA (blue histogram) and by the normal approximation using moments from the solution of the stationary versions of equations (10.37)-(10.39) (thick red line). Parameter values and initial conditions are as in Figure 10.1.

10.5 Kirkwood superposition approximation

As before, the final moment closure approximation we consider is the Kirkwood superposition approximation. For multiple dependent variables A, B and C , the Kirkwood superposition approximation takes the form

$$\langle ABC \rangle = \frac{\langle AB \rangle \langle BC \rangle \langle AC \rangle}{\langle A \rangle \langle B \rangle \langle C \rangle}. \quad (10.40)$$

In contrast to the pairwise approximation this definition is unique. In the bivariate case, for $A, B, C \in \{S, I\}$, the closure takes the form

$$\langle S^3 \rangle = \frac{\langle S^2 \rangle^3}{\langle S \rangle^3}, \quad (10.41)$$

$$\langle S^2 I \rangle = \frac{\langle SI \rangle^2 \langle S^2 \rangle}{\langle S \rangle^2 \langle I \rangle}, \quad (10.42)$$

$$\langle SI^2 \rangle = \frac{\langle SI \rangle^2 \langle I^2 \rangle}{\langle I \rangle^2 \langle S \rangle}, \quad (10.43)$$

$$\langle I^3 \rangle = \frac{\langle I^2 \rangle^3}{\langle I \rangle^3}, \quad (10.44)$$

leading to the following closed equations for the second moments:

$$\frac{d \langle SI \rangle}{dt} = \frac{k_1}{\nu} \left(\frac{\langle SI \rangle^2 \langle S^2 \rangle}{\langle S \rangle^2 \langle I \rangle} - \frac{\langle SI \rangle^2 \langle I^2 \rangle}{\langle S \rangle \langle I \rangle^2} - \langle SI \rangle \right) - k_2 \langle SI \rangle, \quad (10.45)$$

$$\frac{d \langle S^2 \rangle}{dt} = \frac{k_1}{\nu} \left(\langle SI \rangle - 2 \frac{\langle SI \rangle^2 \langle S^2 \rangle}{\langle S \rangle^2 \langle I \rangle} \right), \quad (10.46)$$

$$\frac{d \langle I^2 \rangle}{dt} = \frac{k_1}{\nu} \left(2 \frac{\langle SI \rangle^2 \langle I^2 \rangle}{\langle S \rangle \langle I \rangle^2} + \langle SI \rangle \right) + k_2 \left(\langle I \rangle - 2 \langle I^2 \rangle \right). \quad (10.47)$$

Figure 10.5 demonstrates that the Kirkwood superposition approximation provides estimates for the means and the standard deviations which are comparable to the normal approximation whilst being considerably less complicated to obtain.

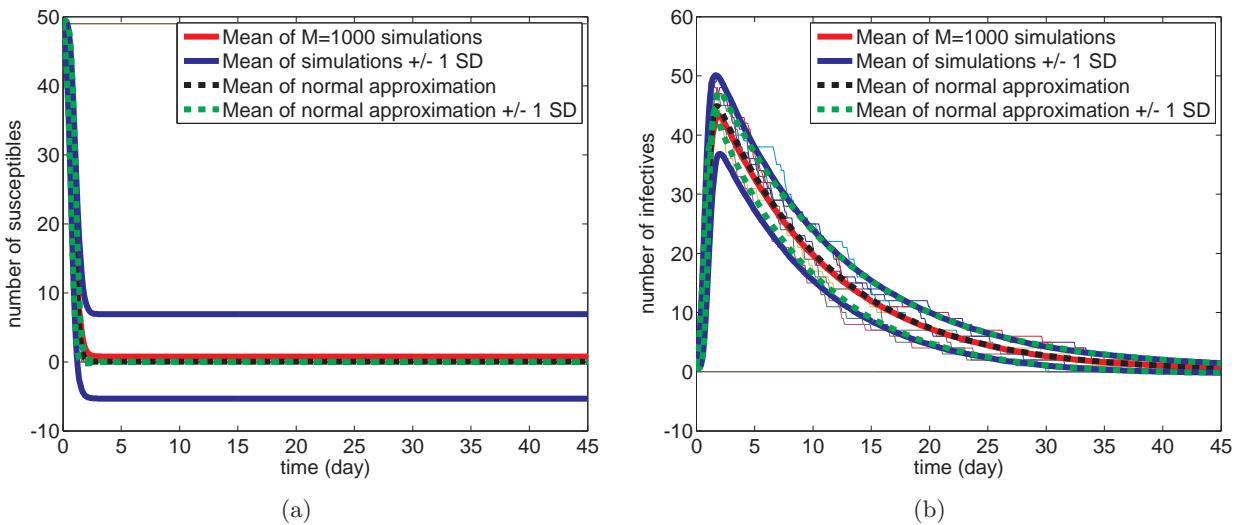


Figure 10.5: Kirkwood moment closure. (a) Susceptibles and (b) infectives. Figure descriptions, parameter values and initial conditions are as in Figure 10.2.

Lecture 11 - A realistic model of cell proliferation I

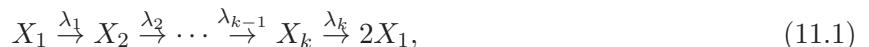
In Lectures 1-10 we have considered models of well-mixed chemical systems. The inter-arrival times of the corresponding “chemical reaction” events can be modelled as being exponentially distributed. However, there are non-Markovian stochastic events for which the exponential distribution does not accurately characterise inter-arrival times. One such example is cellular proliferation [35–41]. In Lecture 4 we considered a simple first order production reaction (4.9). This reaction is often used to represent process of cellular proliferation and assumes that the cell cycle is exponentially distributed. In reality this is not the case.

The cell cycle comprises a series of events which take place inside a cell leading ultimately to its division and consequently the production of two daughter cells (see Figure 11.1 (a)). In Eukaryotes (organisms whose cells contain a nucleus and other organelles enclosed within membranes - one of the three taxa (along with prokaryotes and archaea) in the tree of life) the cell cycle is divided into three phases: interphase, mitotic and cytokinesis. During interphase the cell grows and accumulates nutrients needed for mitosis and duplicates its DNA. During the mitotic phase, the chromosomes separate. During cytokinesis the material of the cell separates into two new cells. Interphase is itself usually further broken down into three other phases: G_1 , S and G_2 . In S -phase the DNA replicates. The two phases G_1 and G_2 are checkpoints which ensure everything is ready for DNA synthesis and mitosis, respectively.

11.1 Multi-stage model of the cell cycle

The distribution of cell cycle times (CCTs) is invariably peaked with positive skewness (see Fig. 11.1 (b)). Hypoexponential distributions have been suggested to accurately represent phases of the cell cycle Weber et al. [43], Smith and Martin [44]. These distributions are made up of a series of k independent exponential distributions, each with its own rate, λ_i , in series. However, if k is large, then we have a large number of parameters and face issues of parameter identifiability when fitting these distributions to data. Two special cases of the general hypoexponential distribution, the Erlang and exponentially modified Erlang distribution, have been suggested to be good models of the cell cycle time distribution (CCTDs) whilst being characterised by a small numbers of parameters Golubev [45].

The transition through the cell cycle can be thought of as being separated into k distinct stages. We divide the cell cycle (with mean length C) into k stages¹ The time to progress through each of these stages is exponentially distributed with mean μ_i . We can represent the progression through these stages of the cell cycle as the following chain of ‘reactions’



¹Note, these stages do not necessarily correspond to the traditional (G_1 S G_2 M) phases of the cell cycle. Rather they are arbitrary division of the cell cycle which will allow us to recreate the correct CCTD.

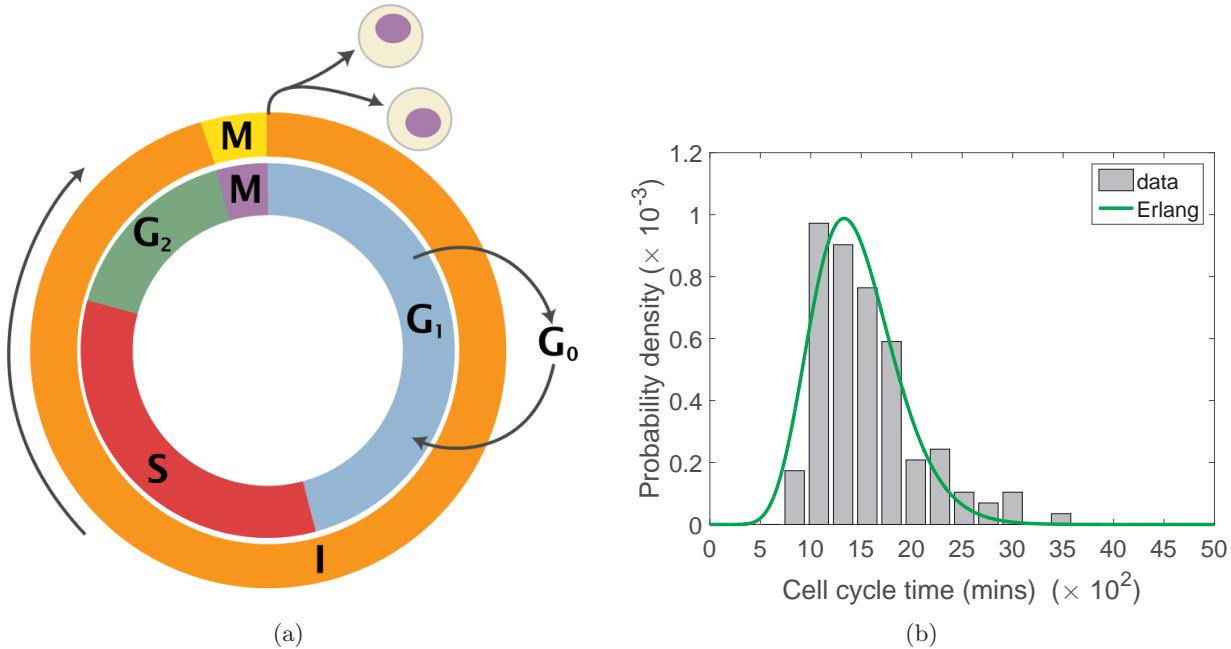


Figure 11.1: (a) Schematic of the cell cycle. The outer ring depicts interphase (I) in orange, mitosis (M) in yellow and cytokinesis occurring at the end of mitosis. The inner ring depicts: the first gap phase (G_1), the synthesis phase (S) in red, the second gap phase (G_2) in green and mitosis (M) in purple. It also denotes an optional phase G_0 , the resting phase which is entered into if conditions are not ideal for the progression cell cycle. (b) A typical example of the peaked distribution of experimentally determined inter-division times in NIH 3T3 mouse embryonic fibroblasts grown *in vitro* (grey histograms). The green curve shows a fit to the experimental data using an Erlang distribution with $k = 12$ stages and $\lambda = 0.0083$. Taken from Yates et al. [42].

where $\lambda_i = 1/\mu_i$.

If we assume that all transition rates, λ_i , are identically equal to λ in our general hypoexponential model then the time to progress through the whole cell cycle is distributed according to the sum of k identically exponentially distributed random variables. It is straightforward to show (using moment generating functions or convolutions - see Problem Sheet 6) that the cell cycle times are Gamma distributed with scale parameter $\mu = C/k$ (equivalently rate parameter λ) and shape parameter² k . For integer k (as we have in this model) this is commonly known as the Erlang distribution:

$$f(x) = \frac{\lambda^k x^{k-1} e^{-\lambda x}}{(k-1)!}. \quad (11.2)$$

If we decrease μ and simultaneously increase k so that $\mu k = C$ remains constant, the Erlang distribution approaches the Dirac delta distribution centred on C . This demonstrates that we can arbitrarily reduce the variance to match the distribution we are trying to model.

We can analyse the cell cycle reaction chain (11.1) further by considering the associated chemical master equation. Let $P(n_1, n_2, \dots, n_k, t)$ be shorthand for the probability that there are n_1 cells in

²Note that for $k = 1$ the Gamma distribution simplifies to the exponential distribution and we return to our simple model (4.9) of cellular proliferation from Lecture 4

stage one, n_2 in stage two and so on. The resulting chemical master equation is

$$\begin{aligned} \frac{dP(n_1, n_2, \dots, n_k, t)}{dt} = & \sum_{i=1}^{k-1} P(n_1, \dots, n_i + 1, n_{i+1} - 1, \dots, n_k, t)(n_i + 1)\lambda \\ & + P(n_1 - 2, n_2, \dots, n_k + 1, t)(n_k + 1)\lambda \\ & - \sum_{i=1}^k P(n_1, \dots, n_i, n_{i+1}, \dots, n_k, t)n_i\lambda. \end{aligned} \quad (11.3)$$

The first thing we might think to do, since we have a purely first order reaction system, is to attempt to solve the chemical master equation in generality using moment generating functions. With multiple independent variables the definition of the probability generating function changes from equation (5.3) in Lecture 5 to

$$G(x_1, x_2, \dots, x_k, t) = \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \cdots \sum_{n_k=0}^{\infty} x_1^{n_1} x_2^{n_2} \cdots x_k^{n_k} p(n_1, \dots, n_i, n_{i+1}, \dots, n_k, t).$$

Multiply the chemical master equation (11.3) by $\prod_{j=1}^k x_j^{n_j}$ and sum over all possible values of $\mathbf{n} = (n_1, n_2, \dots, n_k)$ to deduce

$$\begin{aligned} \frac{\partial G}{\partial t} = & \sum_{i=1}^{k-1} \sum_{\mathbf{n}} \prod_{j=1}^k x_j^{n_j} P(n_1, \dots, n_i + 1, n_{i+1} - 1, \dots, n_k, t)(n_i + 1)\lambda \\ & + \sum_{\mathbf{n}} \prod_{j=1}^k x_j^{n_j} P(n_1 - 2, n_2, \dots, n_k + 1, t)(n_k + 1)\lambda \\ & - \sum_{i=1}^k \sum_{\mathbf{n}} \prod_{j=1}^k x_j^{n_j} P(n_1, \dots, n_i, n_{i+1}, \dots, n_k, t)n_i\lambda \end{aligned} \quad (11.4)$$

where $\sum_{\mathbf{n}}$ is shorthand for $\sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \cdots \sum_{n_k=0}^{\infty}$. Changing the indices in the first and second terms and using the properties of the derivative of the generating function derived in Lecture 5, we can show (see Problem Sheet 6) that the generating function satisfies the following first order partial differential equation:

$$\frac{1}{\lambda} \frac{\partial G}{\partial t} = \sum_{i=1}^{k-1} x_{i+1} \frac{\partial G}{\partial x_i} + x_1^2 \frac{\partial G}{\partial x_k} - \sum_{i=1}^k x_i \frac{\partial G}{\partial x_i}. \quad (11.5)$$

It is convenient to introduce a rescaled time $\theta = \lambda t$ at this point. Unfortunately this system is intractable for values of $k > 1$. However, with $k = 1$ we can use the usual method of characteristics to solve the reduced version of (11.5):

$$\frac{\partial G}{\partial \theta} = x^2 \frac{\partial G}{\partial x} - x \frac{\partial G}{\partial x}, \quad (11.6)$$

for the generating function and hence the evolution of the probability distribution of the number of cells. Recall that the solution, $G(x, \theta)$, can be thought of as a surface over the (x, θ) -plane with initial data given on the x axis. The surface is made up of a number of characteristic curves parametrised by a dummy variable s , which emanate from the initial data.

In order to find the characteristic equations for this problem, apply the chain rule to calculate the derivative

$$\frac{dG}{ds} = \frac{d\theta}{ds} \frac{\partial G}{\partial \theta} + \frac{dx}{ds} \frac{\partial G}{\partial x}. \quad (11.7)$$

Comparing coefficients with the simplified (for $k = 1$) PDE (11.6) gives the following characteristic

equations

$$\frac{dG}{ds} = 0, \quad (11.8)$$

$$\frac{d\theta}{ds} = 1, \quad (11.9)$$

$$\frac{dx}{ds} = x - x^2. \quad (11.10)$$

If we assume there is initially one cell in the first phase, i.e. $p(n, 0) = \delta_{1,n}$ this translates to initial condition

$$G(x, 0) = G_0(x) = \sum_{n=0}^{\infty} x^n \delta_{1,n} = x. \quad (11.11)$$

for the generating function.

Solving equation (11.8) for G in terms of s and applying the initial condition gives

$$G(s) = G_0(x(0)) = x(0), \quad (11.12)$$

where the argument of x now refers to the dummy variable s . Solving equation (11.9) gives

$$\theta(s) = s + \theta(0) = s, \quad (11.13)$$

since $s = 0$ characterises the initial condition (i.e. when $s = 0$, $\theta = 0$). Finally we solve equation (11.10) for x in terms of $x(0)$:

$$\frac{x}{1-x} = \frac{x(0)}{1-x(0)} \exp(s). \quad (11.14)$$

Solving for $x(0)$ and substituting into equation (11.12) gives

$$G(x, t) = \frac{x \exp(-\lambda t)}{1 - x(1 - \exp(-\lambda t))}, \quad (11.15)$$

where we have substituted $s = \theta = \lambda t$.

Consequently it can be shown (see Problem Sheet 6) that the probability distribution is given by the well known [46] geometric distribution

$$P(n, t) = \exp(-\lambda t)(1 - \exp(-\lambda t))^{n-1} \text{ for } n \geq 1. \quad (11.16)$$

From this we can demonstrate (using the definitions of $M(t)$ and $V(t)$) from Lecture 2 (see Problem Sheet 6) that

$$M(t) = \exp(\lambda t) \quad \text{and} \quad V(t) = \exp(\lambda t)(\exp(\lambda t) - 1), \quad (11.17)$$

in the case of a single stage model of the cell cycle (i.e. $k = 1$).

In Figure 11.2 (a) we can see that the mean of the stochastic simulations in the case $k = 1$ matches well with the analytically predicted mean given in equation (11.17) for a simulation starting with just a single cell. However, as expected, in Figure 11.2 (b) we see that in the case $k = 1$, the CCTs are exponentially distributed, rather than peaked.

Although it is not possible to solve equation (11.3) explicitly for $p(\mathbf{n}, t)$, in the next lecture, we will derive evolution equations for the mean behaviour of the number of cells in each phase of the general multi-stage model which we are able to solve explicitly.

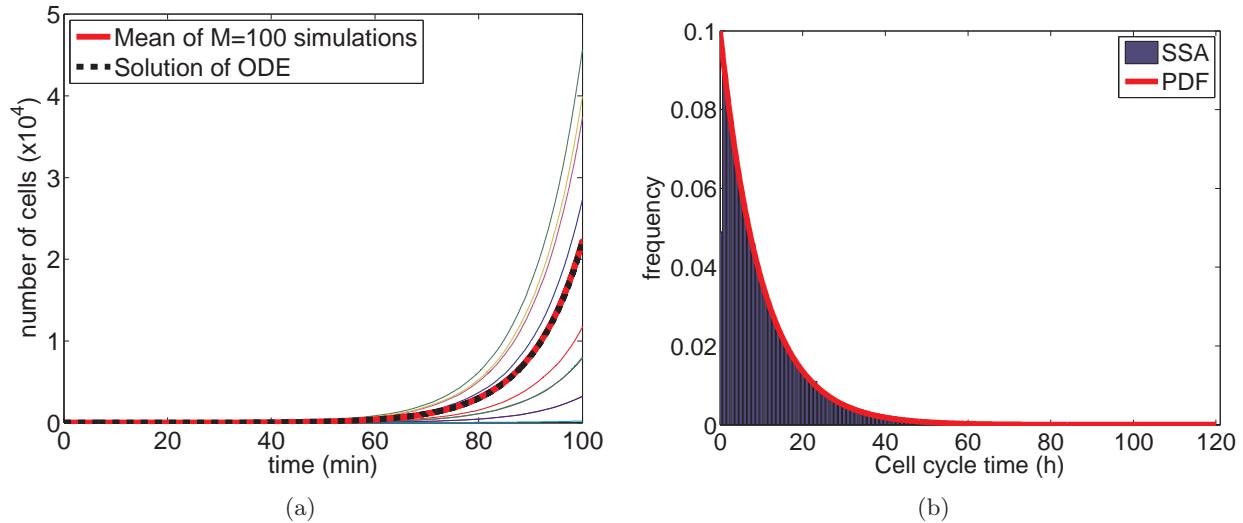


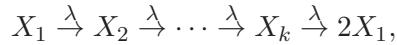
Figure 11.2: (a) A stochastic simulation of the system of reactions (11.1) for $k = 1$ representing a single-stage model of the cell cycle. Ten realisations of the SSA (thin solid coloured lines), mean behaviour averaged over 100 repeats (solid thick red line) and mean behaviour given by equation (11.17) (dashed black line). (b) The distribution of cell cycle times given by stochastic simulation of 100,000 cell division events (blue histogram) and by the PDF of the exponential distribution (11.2) (solid red line).

Lecture 12 - A realistic model of cell proliferation II

In the previous lecture we introduced a multi-stage model of cellular proliferation which we argued would make the cell cycle time distribution (CCTD) more realistic. We analyse a simple version of this model in which the rates of transitioning between the different phases were all the same. However, even this simplified model was too complicated to solve for the evolution of the full probability distribution. In this lecture we will derive and solve equations for the evolution of the mean number of cells in each stage.

12.1 Mean behaviour

Recall the chain of reactions (11.1) which characterised our multi-stage model of cellular proliferation:



where, as in the previous chapter, we have set all transition rates to be equal. By multiplying the chemical master equation (11.3) by n_j and summing over the state space we can find the evolution of the mean number of cells, $M_j(t) = \sum_n n_j P$, in each stage. Upon simplification we find the following evolution equations for the mean number of cells in each stage

$$\frac{dM_j}{d\theta} = \begin{cases} 2M_k - M_1, & \text{for } j = 1, \\ M_{j-1} - M_j, & \text{for } j \neq 1, \end{cases} \quad (12.1)$$

where we recall that time has been rescaled as $\theta = \lambda t$. In order to solve these equations it is convenient to define $\rho_j = M_j e^\theta$. This transformation can easily be seen to leave us with the following equations for ρ_j , $j = 1, \dots, k$:

$$\frac{d\rho_j}{d\theta} = \begin{cases} 2\rho_k, & \text{for } j = 1, \\ \rho_{j-1}, & \text{for } j \neq 1. \end{cases} \quad (12.2)$$

By repetitive substitution we can show that ρ_k satisfies the following closed equation:

$$\frac{d^k \rho_k}{d\theta^k} = 2\rho_k, \quad (12.3)$$

and that the other ρ_j s can be found from ρ_k with the appropriate amount of differentiation:

$$\rho_j = \left(\frac{d}{d\theta} \right)^{k-j} \rho_k. \quad (12.4)$$

By considering the auxiliary equation ($\sigma^k = 2$) we can show that the general solution of equation (12.3) can be expressed as the following linear combination

$$\rho_k = \sum_{r=0}^{k-1} A_r \exp(2^{1/k} \omega^r \theta), \quad (12.5)$$

where $\omega = \exp(2\pi i/k)$ is the primitive k^{th} root of unity and the A_r are arbitrary constants to be determined by the initial condition. Consequently, using equation (12.4), we can determine each of the ρ_j , for $j \neq k$ by differentiation (see Problem Sheet 6):

$$\rho_j = 2^{1-j/k} \sum_{r=0}^{k-1} A_r \omega^{-jr} \exp(2^{1/k} \omega^r \theta). \quad (12.6)$$

Assuming the IC of a single cell in the first phase at $t = 0$ ($P(\mathbf{n}, 0) = \delta_{\mathbf{n},(1,0,0,\dots,0)}$) implies $M_j(0) = \delta_{1,j}$. We can show (by substitution, see Problem Sheet 6) that

$$A_r = \frac{1}{2k} 2^{1/k} \omega^r. \quad (12.7)$$

Therefore, the number of cells in the j^{th} stage of our cell cycle model (in unscaled time) is given by

$$M_j = \rho_j \exp(-\lambda t) = \frac{1}{k} 2^{(1-j)/k} \sum_{r=0}^{k-1} \omega^{(1-j)r} \exp((2^{1/k} \omega^r - 1)\lambda t), \text{ for } j = 1, \dots, k. \quad (12.8)$$

Although, generally this expression is quite complicated, for small values of k it is possible to express the formulae for the expected population sizes in closed form. For example, when $k = 4$ (for $k = 2$ see Problem Sheet 6)

$$M_1 = \frac{\exp(-4t/C)}{2} \left\{ \cosh\left(\frac{2^{9/4}t}{C}\right) + \cos\left(\frac{2^{9/4}t}{C}\right) \right\}, \quad (12.9)$$

$$M_2 = \frac{\exp(-4t/C)}{2^{5/4}} \left\{ \sinh\left(\frac{2^{9/4}t}{C}\right) + \sin\left(\frac{2^{9/4}t}{C}\right) \right\}, \quad (12.10)$$

$$M_3 = \frac{\exp(-4t/C)}{2^{3/2}} \left\{ \cosh\left(\frac{2^{9/4}t}{C}\right) - \cos\left(\frac{2^{9/4}t}{C}\right) \right\}, \quad (12.11)$$

$$M_4 = \frac{\exp(-4t/C)}{2^{7/4}} \left\{ \sinh\left(\frac{2^{9/4}t}{C}\right) - \sin\left(\frac{2^{9/4}t}{C}\right) \right\}. \quad (12.12)$$

A comparison between these analytical solutions and the average behaviour of the stochastic simulations, for the $k = 4$ stage model is verified in Figure 12.1 (a). In Figure 12.1 (b) we also see that the inter-division time distribution for the 4 stage model is now peaked and better matches the real inter-division time of NIH 3T3 mouse embryonic fibroblast cells (see Figure 11.1 (b) in Lecture 11).

12.2 Limiting behaviour

Since, in a biological experiment, there is no way to determine which stage a particular cell is in (since our stages do not correspond to biologically meaningful phases of the cell cycle) it seems sensible to consider the behaviour of total number of cells, which is easily determined from an experiment.

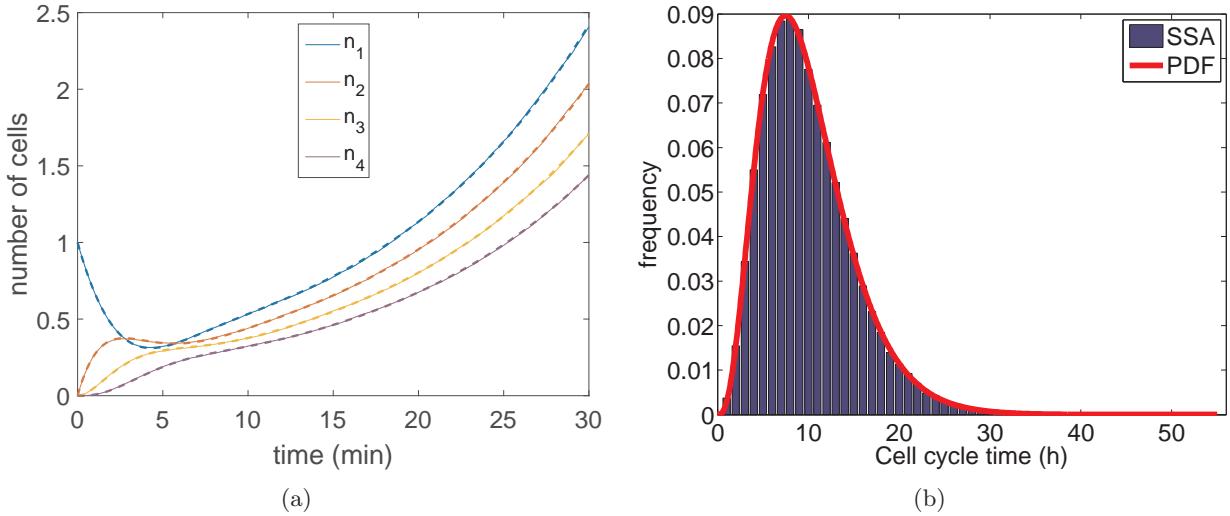


Figure 12.1: (a) Mean behaviour of the number of cells in each of the $k = 4$ different phases of the multi-stage model. The average calculated over 100,000 repeat simulations are plotted as thin solid lines for each of the $k = 4$ stages and the analytical solution given by equations (12.9)-(12.12) are plotted as thick dashed lines. (b) The distribution of cell cycle times given by stochastic simulation of 100,000 cell division events (blue histogram) and by the PDF of the Erlang distribution (11.2) (solid red line).

Summing equation (12.8) over all stages (see Problem Sheet 6) gives the total number of cells

$$M = \frac{2^{1/k}}{2k} \sum_{r=0}^{k-1} \frac{\omega^r}{2^{1/k}\omega^r - 1} \exp((2^{1/k}\omega^r - 1)\lambda t). \quad (12.13)$$

By expanding the exponential as an infinite sum and further simplifying (see Problem Sheet 6) this can be expressed alternately as

$$M = e^{-\lambda t} \sum_{m=0}^{\infty} 2^m \left\{ \frac{(\lambda t)^{mk}}{(mk)!} + \frac{(\lambda t)^{mk+1}}{(mk+1)!} + \cdots + \frac{(\lambda t)^{mk+k-1}}{(mk+k-1)!} \right\}. \quad (12.14)$$

In particular, one way to attempt to establish the parameters of our model from experimental data might be to consider the long time behaviour of the model and compare it to the long time behaviour of an experiment. For large t , the dominant term in equations (12.8) is always the first term, corresponding to $r = 0$. Indeed for $2 \leq k \leq 28$ the real part of the exponent, $\text{Re}(2^{1/k}\omega^r - 1) = 2^{1/k} \cos(2\pi/k) - 1 < 0$, meaning that the remaining terms correspond to damped oscillations. For larger values of k , not all terms tend to zero, but they are still of lower order than the first term. This means we can express the limiting behaviour of the number of cells in each stage as

$$\lim_{t \rightarrow \infty} M_j = \frac{2^{(1-j)/k}}{k} \exp(\lambda t \alpha_k) \text{ for } j = 1, \dots, k, \quad (12.15)$$

where $\alpha_k = (2^{1/k} - 1)$. Summing these limits over all k stages (See Problem Sheet 6) gives the limiting behaviour of the total cell population:

$$\lim_{t \rightarrow \infty} M = \frac{2^{1/k}}{2k\alpha_k} \exp(\lambda t \alpha_k). \quad (12.16)$$

A comparison of equation (12.16) and the mean stochastic behaviour for a $k = 4$ stage model is displayed in Figure 12.2 (a). Even for very small values of t the limiting behaviour is a good fit to the

mean stochastic behaviour.

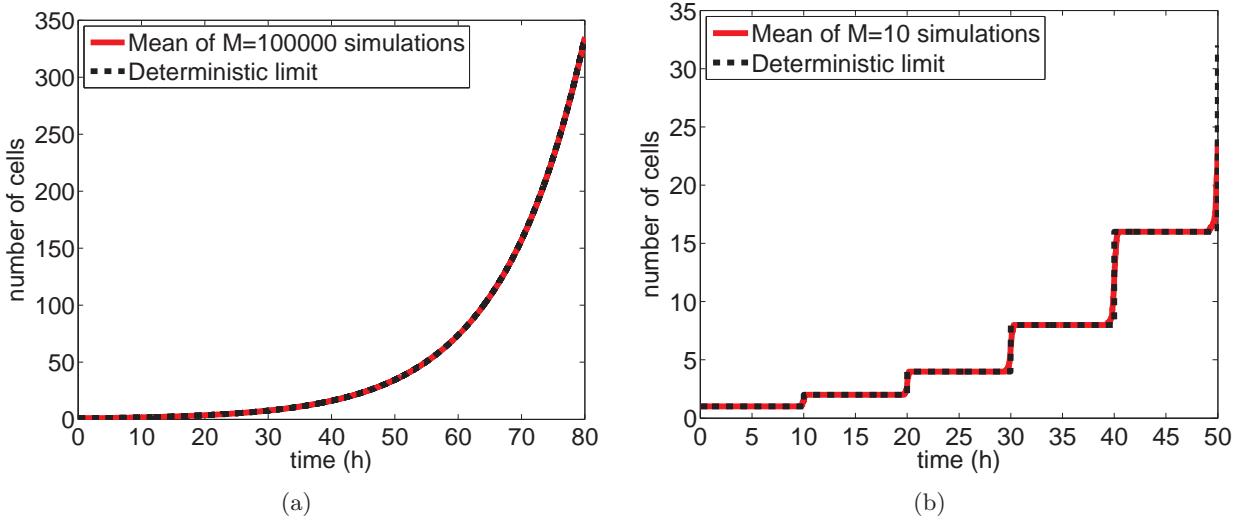


Figure 12.2: (a) The evolution of the mean total number of cells from 100,000 repeats of the stochastic simulation (thick red line) compared to the limiting behaviour of the analytically derived behaviour for the mean number of cells, in the limit $t \rightarrow \infty$ (black dashed line) (see equation (12.16)). (b) The evolution of the mean total number of cells from 10 repeats of the stochastic simulation with $k = 10,000$ stages (thick red line) compared to the limiting behaviour of the analytically derived expression for the mean number of cells, in the limit $k \rightarrow \infty$ (black dashed line) (see equation (12.17)).

Alternatively we might be interested in the “deterministic liming behaviour” of the mean number of cells. That is, the behaviour as $k \rightarrow \infty$ and the CCTD approaches a δ -function. From equation (12.14) it can be shown that M is equal to the expected value of $2^{\lfloor N/k \rfloor}$, where N is a Poisson random variable of mean value λt and $\lfloor x \rfloor$ represents the integer part of x . From this it can be shown that

$$\lim_{k \rightarrow \infty} M = \begin{cases} 2^{\lfloor \lambda t / k \rfloor}, & \text{for non-integer arguments,} \\ \frac{3}{4} 2^{\lambda t / k}, & \text{for integer arguments} \end{cases} \quad (12.17)$$

The proof of this is complicated and will not be given in these lectures. This formula captures the deterministic nature of the cell cycle process in the large k limit. All cells divide synchronously. The mean of several example simulation paths with large k is compared to this deterministic formula in Figure 12.2 (b).

In this lecture we have analysed the mean-field behaviour of the stochastic model of cellular proliferation introduced in Lecture 11. Since the model is linear we would expect the mean behaviour to correspond exactly to the behaviour of the evolution of the deterministic system of equations (12.1). That we can not solve for the probability master equation in generality is disappointing, but we have still gained insight into the behaviour of the system by analysing the mean behaviour. In particular by looking at limiting behaviour of the model it may be possible to fit model parameters to data.

Lecture 13 - Stochastic differential equations I

The classical (approximate) deterministic description of chemical reactions is based on ordinary differential equations (ODEs). Such equations provide a reasonable description of the system in some cases (as we saw in Lectures 1 and 3), while they are inapplicable in some other cases (as shown in Lectures 5, 6, 7 and 8).

In this lecture, we introduce stochastic differential equations (SDEs) which are, roughly speaking, ODEs with some additional noise terms. Such equations can be used as approximate models for some of the stochastic chemical systems we saw in Lectures 7 and 8. They are easier to solve and analyse (from the mathematical point view) than the (exact) chemical master equation, and they provide a better approximation of a stochastic chemical system than ODEs.

Although we will introduce SDEs with reference to the stochastic approximation of chemical systems, they have many other applications. In particular they can also be used for the stochastic modelling of molecular diffusion, which we will discuss in more detail in Lecture 18.

We begin by introducing SDEs from the computational point of view, presenting several examples to illustrate the computational definition of the SDE which is used throughout the course. We then introduce the Fokker-Planck and backward Kolmogorov equations, which correspond in some sense to the chemical master equation. We will use these equations to compute the mean transition time between favourable states of SDEs. We then apply the SDE formalism to a chemical system by deriving the chemical Fokker-Planck equation and the corresponding chemical Langevin equation. We use these to further analyse the chemical system from Section 7.1, and then to analyse the self-induced stochastic resonance which was introduced in Section 8.1.

13.1 A computational definition of an SDE

Consider first a variable $x \equiv x(t) \in \mathbb{R}$ which evolves according to the ODE

$$\frac{dx}{dt} = f(x, t) \tag{13.1}$$

where $f : \mathbb{R} \times [0, \infty) \rightarrow \mathbb{R}$ is a given function. Given the initial condition $x(0) = x_0$, one can use the ODE (13.1) to find the values of $x(t)$ at times $t > 0$, provided that the function f is “sufficiently nice”. Although, it is possible to find conditions on f which guarantees existence and uniqueness of the solution of (13.1), we will not pursue such a rigorous mathematical treatment of ODEs in this course. Instead, we rewrite the ODE (13.1) formally as

$$dx = f(x, t)dt. \tag{13.2}$$

Intuitively, the equation (13.2) specifies the infinitesimal change

$$dx \equiv dx(t) = x(t + dt) - x(t)$$

of the x variable in the infinitesimally small time interval $[t, t + dt]$. Thus, the equation (13.2) can also be written as

$$x(t + dt) = x(t) + f(x(t), t)dt. \quad (13.3)$$

This equation can be used as a simple way to compute the solution of the original ODE (13.1). Namely, we choose a small time-step Δt and we compute the value of $x(t + \Delta t)$ from the value of $x(t)$ by

$$x(t + \Delta t) = x(t) + f(x(t), t)\Delta t. \quad (13.4)$$

Given the initial condition $x(0) = x_0$, we can use formula (13.4) to find the values of $x(t)$ at times $t > 0$, iteratively. This iterative approach is called the Forward Euler Method for solving the ODE (13.1). If f is a “sufficiently nice” function, the Forward Euler Method will yield an approximate solution of the ODE (13.1). The error of the approximation can be made smaller by choosing the time-step, Δt , smaller. Indeed, the formula (13.4) is exactly equal to (13.3) if we replace the (small) time-step Δt by the infinitesimally small time-step dt .

Roughly speaking, an SDE is an ODE such as (13.1) with an additional noise term describing stochastic fluctuations. If we consider (13.4) to be a “computational definition” of the ODE (13.1), we may write the computational definition of the corresponding SDE with Gaussian noise as

$$X(t + \Delta t) = X(t) + f(X(t), t)\Delta t + g(X(t), t)\sqrt{\Delta t}\xi, \quad (13.5)$$

where $g : \mathbb{R} \times [0, \infty) \rightarrow \mathbb{R}$ is a given strength of the noise and ξ is a random number which is sampled from the normal distribution with zero mean and unit variance. Although other noise terms are possible, we will only be concerned with this so-called “Gaussian” noise.

Normally distributed random numbers with zero mean and unit variance can be generated by many standard computer languages (e.g. the function `randn` in Matlab). If a computer language does not include a routine for generating normally distributed random numbers, one can use a generator of uniformly distributed random numbers together with a suitable transformation (see Problem Sheet 7) to generate normally distributed random numbers.

Given the initial condition $X(0) = x_0$, we can use formula (13.5) to find the values of $X(t)$ at times $t > 0$ iteratively, by performing two steps at time t :

Algorithm 7

(7a) Generate a normally distributed (with zero mean and unit variance) random number ξ .

(7b) Compute $X(t + \Delta t)$ from $X(t)$ by (13.5).

Then continue with step (7a) for time $t + \Delta t$.

The application of the SSA (7a)-(7b) is shown in Section 13.2 where several examples of SDEs are presented. The SSA (7a)-(7b) gives the approximate solution of the SDE which can be formally written in the following form

$$X(t + dt) = X(t) + f(X(t), t)dt + g(X(t), t)dW, \quad (13.6)$$

where dW is the so-called white noise (or differential of the Wiener process). If we wanted to provide a rigorous mathematical treatment of SDEs, we would have to rigorously define dW which would require a formal definition of stochastic integration [47, 48]. However, for the purposes of this course, it is

sufficient to assume that the meaning of the SDE (13.6) is given by the corresponding computational definition (13.5). In fact, this is all we need to know to simulate SDEs numerically and to use them for the analysis of reaction-diffusion processes. Consequently, whenever we write SDEs in the form (13.6), we understand them in terms of the computational definition (13.5), i.e. we replace dW by the product of $\sqrt{\Delta t}$ and the random number ξ which is sampled from normal distribution with zero mean and unit variance. The SSA (7a)-(7b) is often called the Euler-Maruyama method for solving the SDE (13.6).

If $g(x, t) \equiv 0$, then the SDE (13.6) reduces to the ODE (13.3) which can be equivalently written as (13.1). If $g(x, t) \not\equiv 0$, then the SDE (13.6) and the ODE (13.3) differ by the extra noise term $g(X(t), t)dW$. There are many other ways one could add noise to ODEs. However, the noise term $g(X(t), t)dW$ is in some sense the most natural choice which appears in applications. We will make this point clear in the following section where we present several examples.

13.2 Examples of SDEs

We consider that both time, t , and variable, $X(t)$, are dimensionless, i.e. they have no physical units. Let us start by choosing $f(x, t) \equiv 0$ and $g(x, t) \equiv 1$. Then (13.6) reads as follows

$$X(t + dt) = X(t) + dW, \quad (13.7)$$

which, using the computational definition (13.5), can be interpreted as

$$X(t + \Delta t) = X(t) + \sqrt{\Delta t}\xi, \quad (13.8)$$

where Δt is the (small) time-step. We choose $\Delta t = 10^{-3}$. Given the initial condition $X(0) = 0$, we compute the time evolution of $X(t)$ by the SSA (7a)-(7b). Six illustrative realisations of the SSA (7a)-(7b) are presented in Figure 13.1 (a).

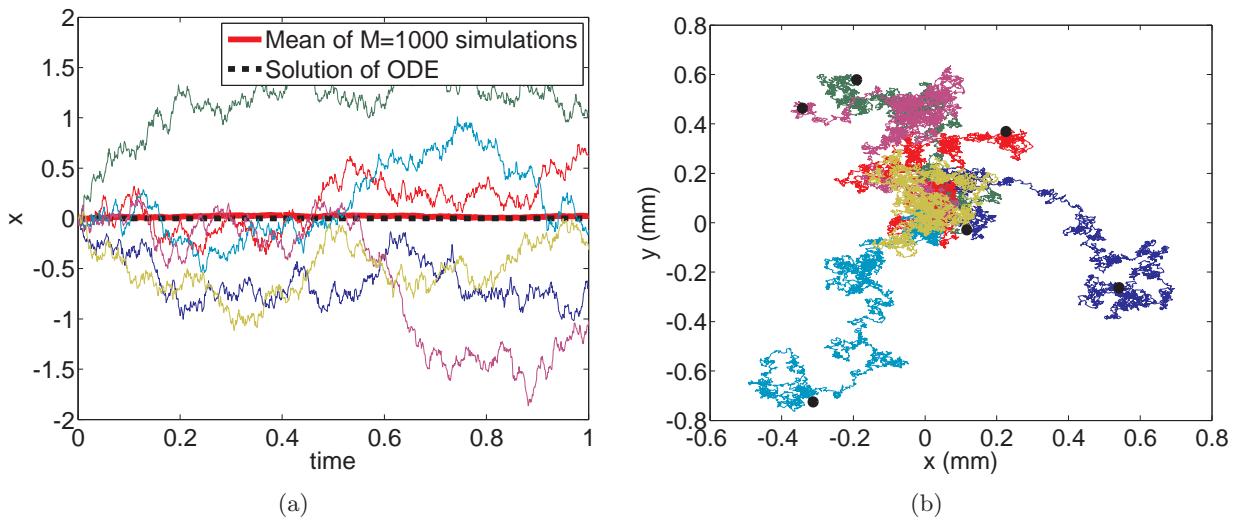


Figure 13.1: (a) Stochastic simulation of the SDE (13.7). Six realisations computed by the SSA (7a)-(7b). The red curve is the stochastic mean, $M(t)$, averaged over 1000 realisations. The black dashed curve is the solution of the corresponding ODE (13.1). We use $\Delta t = 10^{-3}$ and the initial condition $X(0) = 0$. (b) Solution of (13.18)-(13.19). Six trajectories obtained by the SSA (7a)-(7b) for $D = 10^{-4} \text{ mm}^2 \text{ sec}^{-1}$ and $\Delta t = 0.1 \text{ sec}$. Trajectories were started at the origin and followed for 10 minutes. Black dots mark the end-points of the trajectories.

Let $E[\cdot]$ denote the average over (infinitely) many realisations of the SSA (7a)-(7b). Let $M(t)$ be

the mean value of $X(t)$ and let $V(t)$ be the variance of $X(t)$, defined by

$$M(t) = E[X(t)], \quad (13.9)$$

$$V(t) = E[(X(t) - M(t))^2] = E[X(t)^2] - M(t)^2. \quad (13.10)$$

Using the formula (13.8), we can compute the average value of $M(t + \Delta t)$ by

$$\begin{aligned} M(t + \Delta t) &= E[X(t + \Delta t)] = E[X(t) + \sqrt{\Delta t}\xi] \\ &= E[X(t)] + \sqrt{\Delta t}E[\xi] = E[X(t)] = M(t). \end{aligned} \quad (13.11)$$

where we used the fact that ξ is sampled from the normal distribution with zero mean, i.e. $E[\xi] = 0$. The initial condition $X(0) = 0$ implies $M(0) = 0$. Thus, (13.11) implies

$$M(t) = 0. \quad (13.12)$$

Using (13.7), (13.10) and (13.12), we obtain

$$\begin{aligned} V(t + \Delta t) &= E[X(t + \Delta t)^2] - M(t + \Delta t)^2 = E[X(t + \Delta t)^2] \\ &= E[(X(t) + \sqrt{\Delta t}\xi)^2] = E[X(t)^2 + 2X(t)\sqrt{\Delta t}\xi + \Delta t\xi^2] \\ &= E[X(t)^2] + 2E[X(t)]\sqrt{\Delta t}E[\xi] + \Delta tE[\xi^2] \\ &= E[X(t)^2] + \Delta t \\ &= E[X(t)^2] - M(t)^2 + \Delta t \\ &= V(t) + \Delta t. \end{aligned} \quad (13.13)$$

where we have used the fact that $E[\xi] = 0$ and $E[\xi^2] = 1$. Since all realisations start at $X(0) = 0$, we have $V(0) = 0$. Thus, (13.13) implies

$$V(t) = t. \quad (13.14)$$

Consequently, we see that $M(t)$ and $V(t)$ are independent of the time-step Δt . Using the fact that ξ is sampled from the normal distribution, one can actually show that any moment $E[X(t)^k]$, $k = 1, 2, 3, \dots$, is independent of the time-step Δt (see Problem Sheet 7). This is one of the reasons we chose the noise term in the form $\sqrt{\Delta t}\xi$ in the computational definitions (13.5) and (13.7). If the noise term did not scale with the time-step as $\sqrt{\Delta t}$, then $V(t)$ would depend on the time-step Δt . If ξ was not sampled from the normal distribution, then the third moment $E[X(t)^3]$, would be dependent on the time-step Δt .

The main goal of example (13.7) was to show that solving an SDE of the form (13.6) is relatively straightforward through its computational definition (13.5) and the corresponding SSA (7a)-(7b). In Figure 13.1 (a), we had no real application in mind. However, it is also straightforward to add a physical meaning to our computation. All we need to do is to give appropriate physical units to the variable $X(t)$ and time t . For example, assuming that $X(t)$ has units of length we can interpret the SDE (13.7) as a simple evolution equation of the x -coordinate of a diffusing particle. To be more specific, let us consider a protein particle diffusing in aqueous environment. Its position can be described by the three dimensional vector $[X(t), Y(t), Z(t)]$. In Lecture 18, we show that the position of the protein particle at time $t + \Delta t$ can be computed from its position at time t by

$$X(t + \Delta t) = X(t) + \sqrt{2D\Delta t}\xi_x, \quad (13.15)$$

$$Y(t + \Delta t) = Y(t) + \sqrt{2D\Delta t}\xi_y, \quad (13.16)$$

$$Z(t + \Delta t) = Z(t) + \sqrt{2D\Delta t}\xi_z, \quad (13.17)$$

where D is the diffusion constant and ξ_x, ξ_y, ξ_z are random numbers which are sampled from the normal distribution with zero mean and unit variance. Thus, following our discussion from Section 13.1, we can say that the position of the diffusing protein evolves according to the system of SDEs

$$X(t + dt) = X(t) + \sqrt{2D}dW_x, \quad (13.18)$$

$$Y(t + dt) = Y(t) + \sqrt{2D}dW_y, \quad (13.19)$$

$$Z(t + dt) = Z(t) + \sqrt{2D}dW_z, \quad (13.20)$$

where the subscripts in dW_x, dW_y, dW_z emphasise the fact that the white noises are not correlated, i.e. we need three different random numbers ξ_x, ξ_y, ξ_z to simulate the system of SDEs (13.18)-(13.20) at each time-step by (13.15)-(13.17). In particular, the equations (13.15)-(13.17) are not coupled and we can use the SSA (7a)-(7b) to compute example solution trajectories. Choosing $D = 10^{-4}\text{mm}^2\text{sec}^{-1}$ (diffusion constant of a typical protein particle), $\Delta t = 0.1\text{ sec}$ and $[X(0), Y(0), Z(0)] = [0, 0, 0]$, we plot six realisations of the SSA (7a)-(7b) in Figure 13.1 (b). We plot only the x and y coordinates. We follow the diffusing particle for 10 minutes. The position of the particle at time $t = 10\text{ min}$ is denoted as a black circle for each trajectory. We clearly see the typical picture of the trajectory of Brownian motion as shown in school physics textbooks.

The function $f(x, t)$ in (13.6) gives the deterministic part of the dynamics. It is often called the *drift coefficient*. Let us now return to dimensionless $X(t)$ and t and add a non-zero drift to our SDE. We consider the case $f(x, t) \equiv 1$ and $g(x, t) \equiv 1$. Then (13.6) reads as follows

$$X(t + dt) = X(t) + dt + dW, \quad (13.21)$$

which, using the computational definition (13.5), can be interpreted as

$$X(t + \Delta t) = X(t) + \Delta t + \sqrt{\Delta t}\xi, \quad (13.22)$$

where Δt is the (small) time-step. We again use $\Delta t = 10^{-3}$ and the initial condition $X(0) = 0$, and compute the time evolution of $X(t)$ by the SSA (7a)-(7b). Six illustrative realisations of the SSA (7a)-(7b) are presented in Figure 13.2 (a).

Let $M(t)$ be the mean value of $X(t)$ and let $V(t)$ be the variance of $X(t)$, again defined by (13.9) and (13.10). Using the formula (13.22), we find

$$\begin{aligned} M(t + \Delta t) &= E[X(t + \Delta t)] = E[X(t) + \Delta t + \sqrt{\Delta t}\xi] \\ &= E[X(t)] + \Delta t + \sqrt{\Delta t}E[\xi] = E[X(t)] + \Delta t = M(t) + \Delta t, \end{aligned} \quad (13.23)$$

where we again used $E[\xi] = 0$. Using the initial condition $X(0) = 0$, we obtain $M(t) = t$. This line is plotted in Figure 13.2 (a) as the black dashed line. We see that the solution of (13.21) fluctuates around the mean value $M(t) = t$ which depends only on the drift coefficient. One can also show that $V(t) = t$ (see Problem Sheet 7).

Our final SDE example has two favourable states, and is motivated by the ODE (7.2). We choose

$$f(x, t) \equiv f(x) = -k_1x^3 + k_2x^2 - k_3x + k_4, \quad (13.24)$$

$$g(x, t) \equiv g(x) = k_5, \quad (13.25)$$

where k_1, k_2, k_3, k_4 and k_5 are constants. Then (13.6) reads as follows

$$X(t + dt) = X(t) + (-k_1X(t)^3 + k_2X(t)^2 - k_3X(t) + k_4)dt + k_5dW. \quad (13.26)$$

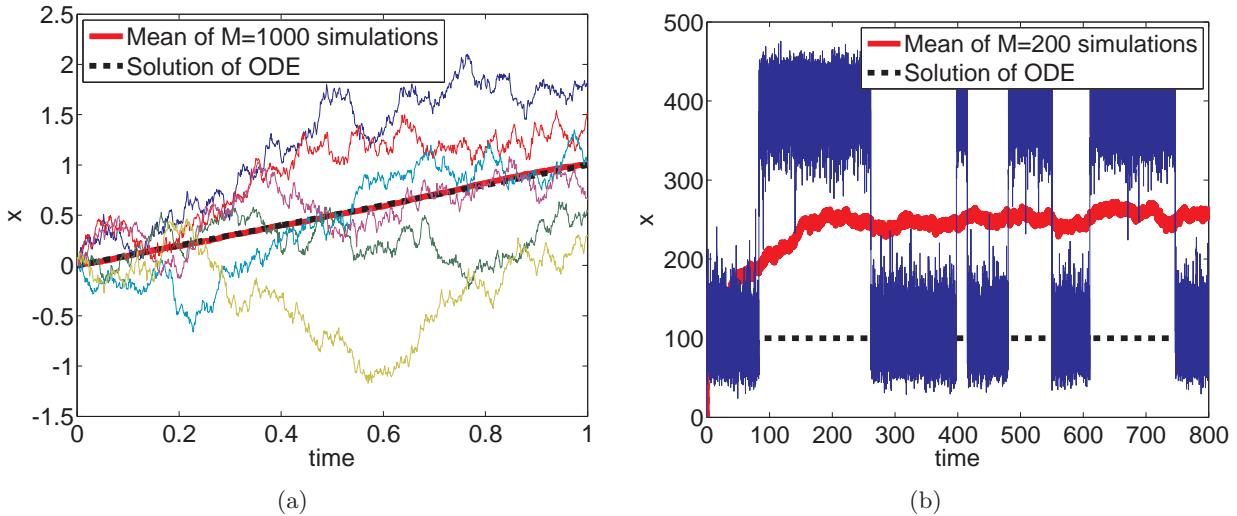


Figure 13.2: (a) Stochastic simulation of the SDE (13.21). Six realisations computed by the SSA (7a)-(7b) (coloured lines). We use $\Delta t = 10^{-3}$ and $X(0) = 0$. The time evolution of the stochastic mean, $M(t)$, is plotted as red line and the solution of the corresponding ODE (13.1) is a black dashed line. (b) Solution of the SDE (13.26) obtained by the SSA (7a)-(7b) for $k_1 = 10^{-3}, k_2 = 0.75, k_3 = 165, k_4 = 10^4, k_5 = 200$ and the initial condition $X(0) = 0$ (blue line). The stochastic mean, $M(t)$, over 200 repeats (red line) and the solution of the ODE (13.28) for the same parameter values (black dashed line).

If $k_5 = 0$ then (13.26) is equal to

$$X(t + dt) = X(t) + (-k_1 X(t)^3 + k_2 X(t)^2 - k_3 X(t) + k_4)dt. \quad (13.27)$$

There is no noise in this equation, and $X(t)$ is no longer random variable. It is a deterministic variable which we will denote as $x(t)$. The equation (13.27) is an ODE which can be equivalently written as

$$\frac{dx}{dt} = -k_1x^3 + k_2x^2 - k_3x + k_4. \quad (13.28)$$

It has the same form as the ODE (7.2) which was studied in Section 7.1. We choose the rate constants as $k_1 = 10^{-3}$, $k_2 = 0.75$, $k_3 = 165$ and $k_4 = 10^4$. Then the ODE (13.28) has three steady states

$$x_{s1} = 100, \quad x_u = 250, \quad x_{s2} = 400. \quad (13.29)$$

The steady states x_{s1} and x_{s2} are stable and the steady state x_u is unstable. Thus we are in similar situation as in the case of the ODE (7.2) from Section 7.1. Starting with $x(0) = 0$, the solution of the ODE (13.28) is plotted in Figure 13.2 (b) as the red line. We see that it converges to the stable steady state x_{s1} . Next, we consider the original SDE (13.26). We use the same values of the parameters k_1, k_2, k_3 and k_4 as before, but now we set $k_5 = 200$. Starting with $x(0) = 0$, we compute the solution of the SDE (13.26) by the SSA (7a)-(7b). The result is plotted in Figure 13.2 (b) as the blue line.

We observe that the solution of the SDE (13.26) fluctuates around the stable steady states x_{s1} and x_{s2} and occasionally switches between them. The situation is qualitatively similar to the behaviour of the chemical system (7.1) which was described in Section 7.1 (see Figure 7.2 (a)). In Section 15.2, we present a calculation of the average switching time between the favourable states of the SDE (13.26). Such a theory will be applicable to chemical systems too because the chemical systems can be approximately described by SDEs under some conditions (as we will see in Lecture 16). Before we proceed we need to introduce the Fokker-Planck equation and its consequences.

Lecture 14 - Stochastic differential equations II

- Fokker-Planck equation

In the previous lecture, we introduced stochastic differential equations (SDEs) both formally and from a computational point of view. We used the computational definition to study some basic examples of SDEs and to deduce the evolution of their mean and variance.

In this Lecture we derive the “so-called” Fokker-Planck equation which is the continuous-state space analogue of the probability master equation for discrete chemical systems). The Fokker-Planck equation describes the evolution of the probability of being in a particular state.

Let us suppose that $X(t)$ evolves according to the SDE (13.6):

$$X(t + dt) = X(t) + f(X(t), t)dt + g(X(t), t)dW. \quad (13.6)$$

We define its probability distribution function, $p(x, t)$, so that $p(x, t)dx$ is the probability that $X(t) \in [x, x + dx]$. Since $X(t)$ must be somewhere on the real line, the function $p(x, t)$ satisfies the normalisation condition

$$\int_{\mathbb{R}} p(x, t)dx = 1, \quad (14.1)$$

for any time t . Roughly speaking, $p(x, t)$ quantifies the chance that the trajectory of the SDE is around the point x at time t . Such information can be computed by many repetitions of the SSA (7a)-(7b). For example, let us consider the SDE example (13.7):

$$X(t + dt) = X(t) + dW, \quad (13.7)$$

and let us compute $p(x, 1)$. To do that we choose the space step Δx and we partition (a sufficiently large subset of) the real line into the bins $[i\Delta x, (i + 1)\Delta x]$ where $i \in \mathbb{Z}$. We run many realisations of the SSA (7a)-(7b), starting at the initial condition $X(0) = 0$ until the time $t = 1$. For each bin, we compute the number of realisations which arrived at this bin at time $t = 1$. We divide the resulting spatial histogram by the number of realisations and the bin size Δx to ensure that the normalisation condition (14.1) is satisfied. Using $\Delta x = 0.2$, we plot the result in Figure 14.1 (a) as the blue histogram. To compute the histogram, we used 10^5 realisations of the SSA (7a)-(7b) with the time-step $\Delta t = 10^{-3}$. If we used fewer realisations, then our results would be more noisy.

Let us now see how the probability distribution function $p(x, t)$ can be computed more efficiently, without performing any stochastic simulation. We will show below that $p(x, t)$ evolves according to the following partial differential equation (PDE)

$$\frac{\partial p}{\partial t}(x, t) = \frac{\partial^2}{\partial x^2} \left(\frac{g^2(x, t)}{2} p(x, t) \right) - \frac{\partial}{\partial x} (f(x, t)p(x, t)). \quad (14.2)$$

This PDE is often called the Fokker-Planck equation in the literature [49] and we will use this name in this course; occasionally it is called the Smoluchowski equation or the forward Kolmogorov equation. We will derive the Fokker-Planck equation later in this section. First, we show its applications.

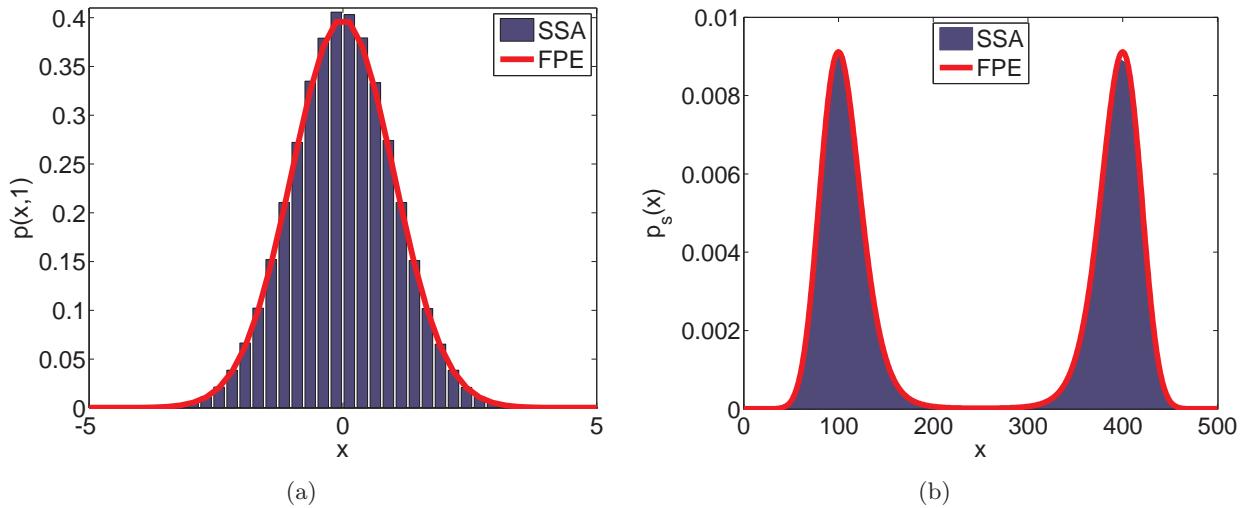


Figure 14.1: (a) The probability distribution function $p(x, 1)$ for the SDE (13.7) computed by the SSA (7a)-(7b) (grey histogram) and by solving the Fokker-Planck equation (red line) given by (14.4). (b) The stationary distribution $p_s(x)$ of the SDE (13.26) computed by the SSA (7a)-(7b) for $k_1 = 10^{-3}$, $k_2 = 0.75$, $k_3 = 165$, $k_4 = 10^4$, $k_5 = 200$ (grey histogram). The solution (14.10) of the stationary Fokker-Planck equation (red line).

For the SDE (13.7) we have $f(x, t) \equiv 0$ and $g(x, t) \equiv 1$. Thus the corresponding Fokker-Planck equation (14.2) is

$$\frac{\partial p}{\partial t}(x, t) = \frac{1}{2} \frac{\partial^2 p}{\partial x^2}(x, t). \quad (14.3)$$

We imposed previously the initial condition $X(0) = 0$ for the SDE (13.7). This implies that the initial condition of (14.3) is in the form of the Dirac delta function $p(x, 0) = \delta(x)$. Solving (14.3) together with the initial condition $p(x, 0) = \delta(x)$, we obtain (see Problem Sheet 7)

$$p(x, t) = \frac{1}{\sqrt{2\pi t}} \exp\left[-\frac{x^2}{2t}\right]. \quad (14.4)$$

The function $p(x, 1)$ is plotted in Figure 14.1 (a) for comparison. As expected, it agrees well with the results obtained by many realisations of the SSA (7a)-(7b).

Suppose now that the coefficients f and g of the SDE (13.6) are arbitrary functions of x , but do not depend on time, i.e. $f(x, t) \equiv f(x)$ and $g(x, t) \equiv g(x)$. The long-time behaviour of such SDEs is reasonably characterised by the stationary distribution

$$p_s(x) = \lim_{t \rightarrow \infty} p(x, t). \quad (14.5)$$

This function can be obtained as a solution of the stationary problem corresponding to (14.2), which is the ODE

$$\frac{d^2}{dx^2} \left(\frac{g^2(x)}{2} p_s(x) \right) - \frac{d}{dx} (f(x) p_s(x)) = 0. \quad (14.6)$$

Solving this equation (see Problem Sheet 7), we obtain

$$p_s(x) = \frac{C}{g^2(x)} \exp \left[\int_{-\infty}^x \frac{2f(y)}{g^2(y)} dy \right], \quad (14.7)$$

where C is a real constant. To determine the value of C , let us note that (14.1) together with (14.5)

imply the normalisation condition

$$\int_{\mathbb{R}} p_s(x) dx = 1. \quad (14.8)$$

Thus

$$C = \left(\int_{\mathbb{R}} \frac{1}{g^2(x)} \exp \left[\int_{-\infty}^x \frac{2f(y)}{g^2(y)} dy \right] dx \right)^{-1}. \quad (14.9)$$

The stationary distribution $p_s(x)$ is a very useful description for SDEs with multiple favourable states, such as SDE (13.26). Substituting (13.24)-(13.25) into (14.7) and integrating over y , we find

$$p_s(x) = \bar{C} \exp \left[\frac{-3k_1 x^4 + 4k_2 x^3 - 6k_3 x^2 + 12k_4 x}{6k_5^2} \right], \quad (14.10)$$

where

$$\bar{C} = \left(\int_{\mathbb{R}} \exp \left[\frac{-3k_1 x^4 + 4k_2 x^3 - 6k_3 x^2 + 12k_4 x}{6k_5^2} \right] dx \right)^{-1}.$$

The function (14.10) is plotted in Figure 14.1 (b) as the red line. We compare it with the grey histogram which is obtained by the SSA (7a)-(7b). To compute this histogram, we chose the space step $\Delta x = 1$ and we partitioned the real line into the bins $[i\Delta x, (i+1)\Delta x]$ where $i \in \mathbb{N}$. Starting at $X(0) = 100$, we ran the SSA (7a)-(7b) for (long) time 10^7 . Recording the values of $X(t)$ at equal time intervals (which are time 10^{-1} apart), we computed how many times the simulation visited each bin. To satisfy the normalisation condition (14.8), we divided the histogram by Δx and by the number of recordings (which is 10^8 in our case).

Having seen, in Figure 14.1, the numerical evidence that the Fokker-Planck equation (14.2) provides a good description of the behaviour of our model SDEs, let us now return to its derivation. We let $p(x, t|y, s)dx$ be the probability that $X(t) \in [x, x+dx]$ under the condition that $X(s) = y$, where $s < t$ are arbitrary time points.

Now consider the value of X at time $t + \Delta t$. The probability that the variable X is in the interval $[z, z + dz]$ at time $t + \Delta t$ (i.e. $X(t + \Delta t) \in [z, z + dz]$) is equal to $p(z, t + \Delta t|y, s)dz$. We now divide this time interval from s to $t + \Delta t$ into a time interval from s to t , and an interval from t to $t + \Delta t$. At time t the variable X takes on a value in the interval $[x, x+dx]$ with probability $p(x, t|y, s)dx$. It then moves from $[x, x+dx]$ to the interval $[z, z + dz]$ over the time interval t to $t + \Delta t$ with probability $p(z, t + \Delta t|x, t)dz$. To calculate the probability that X moves from the value y at time s to $[z, z + dz]$ at time $t + \Delta t$ we need to integrate over all possible intermediate points x . In doing so we obtain the so-called Chapman-Kolmogorov equation

$$p(z, t + \Delta t|y, s) = \int_{\mathbb{R}} p(z, t + \Delta t|x, t)p(x, t|y, s)dx, \quad (14.11)$$

where $s < t$. This equation is valid for all Δt , not just small Δt . To derive the Fokker-Planck equation we now take the limit as $\Delta t \rightarrow 0$. We first multiply both sides by a smooth test function $\varphi(z)$ (with finite support) and integrate over z , to give

$$\int_{\mathbb{R}} p(z, t + \Delta t|y, s)\varphi(z)dz = \int_{\mathbb{R}} \left[\int_{\mathbb{R}} p(z, t + \Delta t|x, t)\varphi(z)dz \right] p(x, t|y, s)dx.$$

We rename the integrating variable z to x on the left hand side to get

$$\int_{\mathbb{R}} p(x, t + \Delta t|y, s)\varphi(x)dx = \int_{\mathbb{R}} \left[\int_{\mathbb{R}} p(z, t + \Delta t|x, t)\varphi(z)dz \right] p(x, t|y, s)dx.$$

Now using the Taylor expansion of $\varphi(z)$ around the point x on the right hand side, we get

$$\begin{aligned} \int_{\mathbb{R}} p(x, t + \Delta t | y, s) \varphi(x) dx &= \int_{\mathbb{R}} \left[\int_{\mathbb{R}} p(z, t + \Delta t | x, t) \right. \\ &\quad \times \left(\varphi(x) + \varphi'(x)(z - x) + \varphi''(x) \frac{(z - x)^2}{2} + o((z - x)^2) \right) dz \Big] p(x, t | y, s) dx \\ &= \int_{\mathbb{R}} \left[\varphi(x) \int_{\mathbb{R}} p(z, t + \Delta t | x, t) dz + \varphi'(x) \int_{\mathbb{R}} (z - x) p(z, t + \Delta t | x, t) dz \right. \\ &\quad \left. + \frac{1}{2} \varphi''(x) \int_{\mathbb{R}} (z - x)^2 p(z, t + \Delta t | x, t) dz + \int_{\mathbb{R}} o((z - x)^2) p(z, t + \Delta t | x, t) dz \right] \\ &\quad \times p(x, t | y, s) dx. \end{aligned} \tag{14.12}$$

Next, we simplify the integrals on the right hand side. Starting at $X(t) = x$, the probability that the trajectory arrives somewhere at time $t + \Delta t$ is equal to 1 (we do not lose the trajectory). Thus

$$\int_{\mathbb{R}} p(z, t + \Delta t | x, t) dz = 1. \tag{14.13}$$

The average value of $X(t + \Delta t) - x$ provided that $X(t) = x$ is given by the integral

$$E[X(t + \Delta t) - x | X(t) = x] = \int_{\mathbb{R}} (z - x) p(z, t + \Delta t | x, t) dz. \tag{14.14}$$

Using the computational definition (13.5), we obtain

$$\begin{aligned} E[X(t + \Delta t) - x | X(t) = x] &= E[f(x, t)\Delta t + g(x, t)\sqrt{\Delta t}\xi] \\ &= f(x, t)\Delta t + g(x, t)\sqrt{\Delta t}E[\xi] \\ &= f(x, t)\Delta t, \end{aligned}$$

where we have used the fact that $E[\xi] = 0$. Thus (14.14) implies

$$\int_{\mathbb{R}} (z - x) p(z, t + \Delta t | x, t) dz = f(x, t)\Delta t. \tag{14.15}$$

The average value of $(X(t + \Delta t) - x)^2$ provided that $X(t) = x$ is given by the integral

$$E[(X(t + \Delta t) - x)^2 | X(t) = x] = \int_{\mathbb{R}} (z - x)^2 p(z, t + \Delta t | x, t) dz. \tag{14.16}$$

Using the computational definition (13.5), we obtain

$$\begin{aligned} E[(X(t + \Delta t) - x)^2 | X(t) = x] &= E[(f(x, t)\Delta t + g(x, t)\sqrt{\Delta t}\xi)^2] \\ &= f^2(x, t)\Delta t^2 + 2f(x, t)g(x, t)\Delta t^{3/2}E[\xi] + g^2(x, t)\Delta t E[\xi^2] \\ &= g^2(x, t)\Delta t + o(\Delta t). \end{aligned}$$

where we have used $E[\xi] = 0$ and $E[\xi^2] = 1$. Consequently, (14.16) yields

$$\int_{\mathbb{R}} (z - x)^2 p(z, t + \Delta t | x, t) dz = g^2(x, t)\Delta t + o(\Delta t). \tag{14.17}$$

Substituting (14.13), (14.15) and (14.17) into (14.12) we obtain

$$\begin{aligned} \int_{\mathbb{R}} p(x, t + \Delta t | y, s) \varphi(x) dx = \\ \int_{\mathbb{R}} \left[\varphi(x) + \varphi'(x) f(x, t) \Delta t + \varphi''(x) \frac{g^2(x, t)}{2} \Delta t \right] p(x, t | y, s) dx + o(\Delta t). \end{aligned}$$

A simple algebraic manipulation yields

$$\begin{aligned} \int_{\mathbb{R}} \frac{p(x, t + \Delta t | y, s) - p(x, t | y, s)}{\Delta t} \varphi(x) dx = \\ \int_{\mathbb{R}} \varphi'(x) f(x, t) p(x, t | y, s) dx + \int_{\mathbb{R}} \varphi''(x) \frac{g^2(x, t)}{2} p(x, t | y, s) dx + o(1). \end{aligned}$$

Using integration by parts on the right hand side, we get

$$\begin{aligned} \int_{\mathbb{R}} \frac{p(x, t + \Delta t | y, s) - p(x, t | y, s)}{\Delta t} \varphi(x) dx = \\ - \int_{\mathbb{R}} \varphi(x) \frac{\partial}{\partial x} (f(x, t) p(x, t | y, s)) dx \\ + \int_{\mathbb{R}} \varphi(x) \frac{\partial^2}{\partial x^2} \left(\frac{g^2(x, t)}{2} p(x, t | y, s) \right) dx + o(1). \end{aligned}$$

This can be rewritten as one integral

$$\begin{aligned} 0 = \int_{\mathbb{R}} \varphi(x) \times \left\{ - \frac{p(x, t + \Delta t | y, s) - p(x, t | y, s)}{\Delta t} \right. \\ \left. - \frac{\partial}{\partial x} (f(x, t) p(x, t | y, s)) + \frac{\partial^2}{\partial x^2} \left(\frac{g^2(x, t)}{2} p(x, t | y, s) \right) \right\} dx + o(1). \end{aligned}$$

Since the test function $\varphi(x)$ is arbitrary, the term inside curly brackets must be zero. Thus

$$\begin{aligned} \frac{p(x, t + \Delta t | y, s) - p(x, t | y, s)}{\Delta t} = \\ \frac{\partial^2}{\partial x^2} \left(\frac{g^2(x, t)}{2} p(x, t | y, s) \right) - \frac{\partial}{\partial x} (f(x, t) p(x, t | y, s)) + o(1). \end{aligned}$$

Passing to the limit $\Delta t \rightarrow 0$, we obtain the Fokker-Planck equation in the form

$$\frac{\partial}{\partial t} p(x, t | y, s) = \frac{\partial^2}{\partial x^2} \left(\frac{g^2(x, t)}{2} p(x, t | y, s) \right) - \frac{\partial}{\partial x} (f(x, t) p(x, t | y, s)). \quad (14.18)$$

Lecture 15 - Stochastic differential equations III

In the previous lecture, we introduced stochastic differential equations (SDEs) both formally and from a computational point of view. We used the computational definition to study some basic examples of SDEs and to deduce the evolution of their mean and variance. In Lecture 14 we derived the Fokker-Planck equation (in analogy to the probability master equation for discrete chemical systems) which describes the evolution of the probability of being in a particular state. We used the FPE to derive an equation for the stationary probability distribution (which describes the long-time behaviour of the system). We saw that this agreed with the stationary probability distribution generated by long-time simulation in the case of systems with multiple favourable steady states.

In this lecture we again consider systems with multiple favourable states and develop tools to answer questions relating to the mean dwell-time in each of the steady states and the time to switch between these states.

15.1 The backward Kolmogorov equation

Sometimes we want to know how the likelihood of ending up in a given state depends on the starting state. In such cases the end position is known, and the starting position is undetermined: in a sense we are thinking about the problem backwards. It is then very useful to be able to see how the conditional probability distribution function $p(x, t|y, s)$ depends on y and s . Using a slightly modified approach to that which led to the Fokker-Planck equation, we can derive an evolution equation for the distribution function $p(x, t|y, s)$ in terms of the initial time s and initial value y . To do that, we rename variables in the Chapman-Kolmogorov equation (14.11) to obtain

$$p(x, t|y, s - \Delta s) = \int_{\mathbb{R}} p(x, t|z, s)p(z, s|y, s - \Delta s)dz. \quad (15.1)$$

As before, this equation is valid for any Δs , but we will take the limit in which $\Delta s \rightarrow 0$. Using the Taylor expansion about the point $z = y$ we rewrite the term $p(x, t|z, s)$ as

$$\begin{aligned} p(x, t|z, s) &= p(x, t|y, s) + (z - y) \frac{\partial p}{\partial y}(x, t|y, s) \\ &\quad + \frac{(z - y)^2}{2} \frac{\partial^2 p}{\partial y^2}(x, t|y, s) + o((z - y)^2). \end{aligned}$$

Substituting into the right hand side of (15.1), we get

$$\begin{aligned} p(x, t|y, s - \Delta s) &= p(x, t|y, s) \times \int_{\mathbb{R}} p(z, s|y, s - \Delta s)dz \\ &\quad + \frac{\partial p}{\partial y}(x, t|y, s) \times \int_{\mathbb{R}} (z - y)p(z, s|y, s - \Delta s)dz \\ &\quad + \frac{\partial^2 p}{\partial y^2}(x, t|y, s) \times \int_{\mathbb{R}} \frac{(z - y)^2}{2} p(z, s|y, s - \Delta s)dz + O(\Delta s^2). \end{aligned}$$

Using (14.13), (14.14) and (14.17), we obtain

$$\begin{aligned} \frac{p(x, t|y, s - \Delta s) - p(x, t|y, s)}{\Delta s} &= \\ f(y, s) \frac{\partial p}{\partial y}(x, t|y, s) + \frac{g^2(y, s)}{2} \frac{\partial^2 p}{\partial y^2}(x, t|y, s) + O(\Delta s). \end{aligned}$$

Passing to the limit $\Delta s \rightarrow 0$, we derive the so-called backward Kolmogorov equation

$$-\frac{\partial p}{\partial s}(x, t|y, s) = f(y, s) \frac{\partial p}{\partial y}(x, t|y, s) + \frac{g^2(y, s)}{2} \frac{\partial^2 p}{\partial y^2}(x, t|y, s). \quad (15.2)$$

Both the Fokker-Planck equation (14.18) and the backward Kolmogorov equation (15.2) provide an exact description of $p(x, t|y, s)$ corresponding to the SDE (13.6). To simplify the notation, we define the *diffusion coefficient* by

$$d(y, s) = \frac{g^2(y, s)}{2}. \quad (15.3)$$

Then the Fokker-Planck equation (14.2) reads as follows

$$\frac{\partial p}{\partial t}(x, t) = \frac{\partial^2}{\partial x^2}(d(x, t)p(x, t)) - \frac{\partial}{\partial x}(f(x, t)p(x, t)), \quad (15.4)$$

the backward Kolmogorov equation (15.2) can be rewritten as

$$-\frac{\partial p}{\partial s}(x, t|y, s) = f(y, s) \frac{\partial p}{\partial y}(x, t|y, s) + d(y, s) \frac{\partial^2 p}{\partial y^2}(x, t|y, s) \quad (15.5)$$

and the stationary distribution (14.7) is given by

$$p_s(x) = \frac{C}{d(x)} \exp \left[\int_0^x \frac{f(y)}{d(y)} dy \right] \quad (15.6)$$

where the normalisation constant is

$$C = \left(\int_{\mathbb{R}} \frac{1}{d(x)} \exp \left[\int_0^x \frac{f(y)}{d(y)} dy \right] dx \right)^{-1}.$$

15.2 SDEs with multiple favourable states

The SDE (13.26) is a simple example of an SDE with two favourable states. As seen in Figure 13.2 (b), the value of $X(t)$ stays either close to x_{s1} or close to x_{s2} where x_{s1} and x_{s2} (together with x_u) are given by (13.29). One of the most important characteristics of a system with multiple favourable states is the average time it takes to switch between two of those states. Having the SDE example (13.26) in mind, we focus on the switch from x_{s1} to x_{s2} . While it is clear to the eye from looking at Figure 13.2 which state the system is in at any given time, it is not immediately clear how to define *mathematically* when the system is in a given state. In the deterministic version of the equation the system will sit exactly at $x = x_{s1}$ or $x = x_{s2}$, but in the stochastic model the system will jump around near one state or the other, and will rarely be *exactly* at x_{s1} or x_{s2} . We could define a neighbourhood of each state, and define the system to be in state 1 when it lies in the neighbourhood of x_{s1} . We would hope that the transition time was not too sensitive to how large we chose the neighbourhood, which would be the case if the noise is not too large. Here we adopt a slightly different view. We denote by $\bar{\tau}$ the average time for the simulation to reach the point x_u provided that it started at $X(0) = x_{s1}$. We have seen that this is the point which separates the deterministic trajectories from those which will settle in x_{s1} and those which will settle in x_{s2} . If the simulation reaches x_u , there is

a 50% chance that it will return back to x_{s1} and a 50% chance to continue on to x_{s2} . Thus the mean switching time from x_{s1} to x_{s2} is simply $2\bar{\tau}$.

Using many realisations of the SSA (7a)-(7b), we can estimate $\bar{\tau}$ as follows. We start each realisation of the SSA (7a)-(7b) at $X(0) = x_{s1}$ and we wait until the first time the simulation leaves the interval $(-\infty, x_u)$. We compute $\bar{\tau}$ as an average of the simulation time over all realisations of the SSA (7a)-(7b). Using the parameter values from Figure 13.2 (b), $\Delta t = 10^{-3}$ and averaging over 10^5 realisations, we obtain that $\bar{\tau}$ is approximately equal to 64.7. We denote this estimate of $\bar{\tau}$ as $\bar{\tau}_{sim} = 64.7$.

Next, we derive an analytical formula for $\bar{\tau}$. This formula will give us the value of $\bar{\tau}$ without the necessity of performing (computationally intensive) stochastic simulation. We will derive a general formula for any SDE of the form (13.6) where $f(x, t) \equiv f(x)$ and $g(x, t) \equiv g(x)$. Let $h(y, t)$ be the probability that $X(t') \in (-\infty, x_u)$ for all times $0 < t' < t$ given that it started at $X(0) = y \in (-\infty, x_u)$. $h(y, t)$ is known as the *survival probability*. Then

$$h(y, t) = \int_{-\infty}^{x_u} p(x, t|y, 0)dx, \quad (15.7)$$

where $p(x, t|y, s)dx$ represents the probability that the trajectory remains in $(-\infty, x_u)$ and lies in the interval $[x, x+dx]$ at time t given that it started at position y at time $s < t$. For this reason p is known as the *occupancy probability*; p satisfies the Fokker-Planck and backward Kolmogorov equations with the boundary conditions $p(x_u, t|y, s) = p(x, t|x_u, s) = 0$, so that $p(x, t|y, s) = 0$ if $y \geq x_u$ or $x \geq x_u$. Since we assume that the coefficients $f(x)$ and $g(x)$ of the SDE (13.6) do not depend explicitly on time, we can shift time in p in (15.7) to give

$$h(y, t) = \int_{-\infty}^{x_u} p(x, 0|y, -t)dx. \quad (15.8)$$

Evaluate backward Kolmogorov equation (15.5) at $t = 0$. Using the transformation $s = -t$ in the resulting equation, we obtain

$$\frac{\partial p}{\partial t}(x, 0|y, -t) = f(y)\frac{\partial p}{\partial y}(x, 0|y, -t) + d(y)\frac{\partial^2 p}{\partial y^2}(x, 0|y, -t),$$

where the diffusion coefficient $d(y)$ is defined by (15.3). Integrating over x and using (15.8), we obtain

$$\frac{\partial h}{\partial t}(y, t) = f(y)\frac{\partial h}{\partial y}(y, t) + d(y)\frac{\partial^2 h}{\partial y^2}(y, t). \quad (15.9)$$

Let $\tau(y)$ be the average time to first leave the interval $(-\infty, x_u)$ given that initially $X(0) = y$. The probability that X first leaves the interval $(-\infty, x_u)$ during the time interval $[t, t + dt]$ (under the condition that $X(0) = y$) is given by

$$h(y, t) - h(y, t + dt) \approx -\frac{\partial h}{\partial t}(y, t)dt.$$

Thus $\tau(y)$ can be computed as follows

$$\tau(y) = - \int_0^\infty t \frac{\partial h}{\partial t}(y, t)dt = -[th(y, t)]_0^\infty + \int_0^\infty h(y, t)dt = \int_0^\infty h(y, t)dt,$$

where we have used integration by parts on the right-hand side; the boundary term at $t = \infty$ disappears under the assumption that trajectories must eventually cross x_u for any given initial position (i.e. $h(y, \infty) = 0$). Integrating equation (15.9) over t and using the fact that $h(y, 0) = 1$ since $y \in (-\infty, x_u)$) and $h(y, \infty) = 0$ (since the trajectory must cross x_u eventually), we obtain the following ODE for the

average time to first leave the interval $(-\infty, x_u)$:

$$-1 = f(y) \frac{d\tau}{dy}(y) + d(y) \frac{d^2\tau}{dy^2}(y) \text{ for } y \in (-\infty, x_u). \quad (15.10)$$

To solve this equation we need to impose some boundary conditions. If we start the trajectory further and further to the left we expect that the exit time does not depend on the starting position: the trajectories in this region are far from equilibrium and are dominated by the deterministic potential. Thus we impose

$$\frac{d\tau}{dy}(-\infty) = 0. \quad (15.11)$$

The second boundary condition comes from the fact that $p(x, t|x_u, s) = 0$ which implies $h(x_u, t) = 0$ which gives

$$\tau(x_u) = 0. \quad (15.12)$$

This is to be expected: if the trajectory starts at $x = x_u$ it leaves the region $(-\infty, x_u)$ in zero time. Using an integrating factor to integrate (15.10) subject to (15.11) we obtain

$$\begin{aligned} \frac{d\tau}{dy}(y) &= -\exp\left[-\int_0^y \frac{f(z)}{d(z)} dz\right] \int_{-\infty}^y \frac{1}{d(x)} \exp\left[\int_0^x \frac{f(z)}{d(z)} dz\right] dx \\ &= -\frac{1}{d(y)p_s(y)} \int_{-\infty}^y p_s(x) dx, \end{aligned}$$

where the stationary distribution p_s is given by (15.6). Now integrating over y and using (15.12) we obtain

$$\tau(y) = \int_y^{x_u} \frac{1}{d(z)p_s(z)} \int_{-\infty}^z p_s(x) dx dz. \quad (15.13)$$

Thus $\bar{\tau}$ can be computed as

$$\bar{\tau} = \tau(x_{s1}) = \int_{x_{s1}}^{x_u} \frac{1}{d(z)p_s(z)} \int_{-\infty}^z p_s(x) dx dz. \quad (15.14)$$

The stationary distribution p_s for the SDE example (13.26) was computed as (14.10). Substituting p_s into (15.14) and evaluating the resulting integrals numerically, we obtain $\bar{\tau} = 59.45$. We see that there is an error (approximately 9%) between the theoretical value of $\bar{\tau}$ and the value $\bar{\tau}_{sim} = 64.7$ estimated previously using the SSA (7a)-(7b). The main contribution to the error is due to the fact that we simulated the SDE (13.26) using the SSA (7a)-(7b) with the finite time-step $\Delta t = 10^{-3}$. Decreasing the time-step Δt improves the accuracy of the SSA (7a)-(7b).

In Figure 15.1 (a), we plot an example of the stochastic trajectory with $\Delta t = 10^{-5}$ (blue line). This trajectory crosses the boundary $x_u = 250$ several times. However, if we consider the same trajectory using the coarser time-step $\Delta t = 10^{-3}$ (red line), then the coarser trajectory leaves the domain $(-\infty, x_u)$ only once. This explains why the SSA (7a)-(7b) with $\Delta t = 10^{-3}$ overestimates the value of $\bar{\tau}$. The stochastic trajectories (which should be terminated) stay in the domain $(-\infty, x_u)$ for a longer time if the time-step is larger.

Decreasing the time-step Δt is one possible way to improve the accuracy of the estimation of $\bar{\tau}$ by the SSA (7a)-(7b). However, it makes the simulation computationally intensive. Another option is to consider that there is a non-zero probability that the simulation left the domain $(-\infty, x_u)$ during the time interval $(t, t+\Delta t)$ even if $X(t+\Delta t) < x_u$ and $X(t) < x_u$. To estimate this probability we suppose that during this time interval the particle is diffusing only, with a diffusion constant $D = k_5^2/2$. This is a good approximation close to the point x_u , where the drift coefficient is zero. We can then estimate

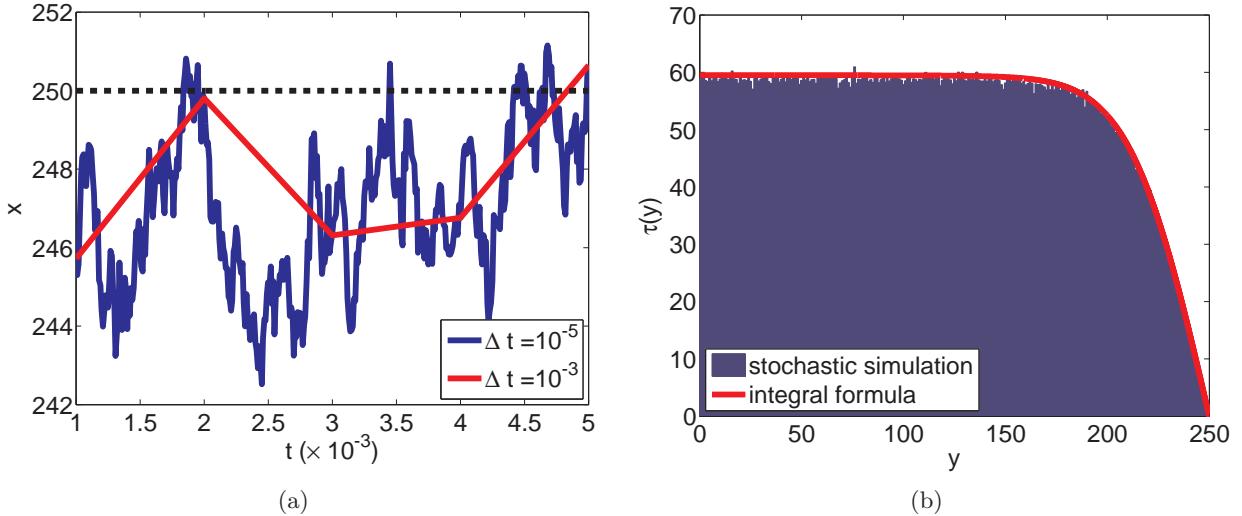


Figure 15.1: (a) The trajectory of (13.26) computed by the SSA (7a)-(7b) for $\Delta t = 10^{-5}$ (blue line). The coarser trajectory with $\Delta t = 10^{-3}$ (red line). (b) The exit time $\tau(y)$ obtained by the integral formula (15.13) (red line) and by the SSA (7a)-(7b) (blue histogram). The parameter values are $k_1 = 10^{-3}, k_2 = 0.75, k_3 = 165, k_4 = 10^4$ and $k_5 = 200$.

the desired probability as (see Problem Sheet 8)

$$\exp \left[-\frac{(X(t) - x_u)(X(t + \Delta t) - x_u)}{D\Delta t} \right]. \quad (15.15)$$

Thus we can estimate $\bar{\tau}$ by executing many realisations of the SSA (7a)-(7b) as follows. We start each realisation of the SSA (7a)-(7b) at $X(0) = x_{s1}$. If $X(t + \Delta t) \geq x_u$, we terminate the trajectory. If $X(t + \Delta t) < x_u$, then we generate a random number uniformly distributed in $(0, 1)$. If this number is less than (15.15), we terminate the trajectory. We compute $\bar{\tau}$ as an average of the simulation time over all realisations. Using the parameter values from Figure 13.2 (b), $\Delta t = 10^{-3}$ and averaging over 10^5 realisations, we obtain that $\bar{\tau}$ is approximately equal to 59.2 which compares well with the theoretical value $\bar{\tau}$ obtained by formula (15.14).

In Figure 15.1 (b), we plot the function $\tau(y)$ given by the formula (15.13). It compares well with the results obtained by stochastic simulation (blue histogram). To get the blue histogram, we estimated $\tau(y)$ for each integer value y as an average of 1000 realisations of the SSA (7a)-(7b). We used $\Delta t = 10^{-3}$ and we computed the simulation results with the help of the correction probability (15.15) just described.

Lecture 16 - Chemical Fokker-Planck equation I

In this lecture, we show how SDEs such as (13.6) can be used to give us an approximate description of discrete-state systems. The methods developed in this lecture can then be used to determine, for example, the mean switching time between favourable states.

16.1 Derivation of the Chemical Fokker-Planck equation

Let us reconsider example (2.1)-(2.2) from Section 2.1, that is, we consider a chemical species A in a container of volume ν , which is subject to the following two chemical reactions



The chemical master equation for this problem is given as (2.4), and is an infinite system of ODEs for $p_n(t)$, $n = 0, 1, 2, 3, \dots$, where $p_n(t)$ is the probability that $A(t) = n$. We can rewrite (2.4) as

$$\frac{dp_n}{dt}(t) = h_1(n+1, t) - h_1(n, t) + h_2(n-1, t) - h_2(n, t) \quad (16.2)$$

where

$$h_1(n, t) = k_1 np_n(t), \quad h_2(n, t) = k_2 \nu p_n(t). \quad (16.3)$$

Now let us suppose that we are interested in what happens for large values of n . We quantify this by introducing a fixed large number ω , so that we are interested in the values of p_n for n not too far from ω . We now write $n = \eta\omega$ and $p_n(t) = p(\eta, t)$, treating η as a continuous variable. The same change of variables in (16.3) gives the real valued functions

$$h_1(\eta, t) = k_1 \eta \omega p(\eta, t), \quad h_2(\eta, t) = k_2 \nu p(\eta, t), \quad (16.4)$$

and (16.2) reads as follows

$$\frac{\partial p}{\partial t}(\eta, t) = h_1\left(\eta + \frac{1}{\omega}, t\right) - h_1(\eta, t) + h_2\left(\eta - \frac{1}{\omega}, t\right) - h_2(\eta, t). \quad (16.5)$$

Since ω is large we can approximate the right-hand side by a Taylor series about the point (η, t) ,

$$h\left(\eta \pm \frac{1}{\omega}, t\right) = h(\eta, t) \pm \frac{1}{\omega} \frac{\partial h}{\partial \eta}(\eta, t) + \frac{1}{2\omega^2} \frac{\partial^2 h}{\partial \eta^2}(\eta, t) + O\left(\frac{h}{\omega^3}\right)$$

Substituting into (16.5), and truncating at $O(1/\omega^2)$, we get

$$\frac{\partial p}{\partial t}(\eta, t) = \frac{1}{\omega} \frac{\partial h_1}{\partial \eta}(\eta, t) + \frac{1}{2\omega^2} \frac{\partial^2 h_1}{\partial \eta^2}(\eta, t) - \frac{1}{\omega} \frac{\partial h_2}{\partial \eta}(\eta, t) + \frac{1}{2\omega^2} \frac{\partial^2 h_2}{\partial \eta^2}(\eta, t).$$

The rescaling of n with ω when defining η allowed us to asymptotically expand in inverse powers of the large number ω . However, now that we have truncated the expansion it is often more convenient

to write the equations in terms of $x = \eta\omega$, which can be thought of as an extension of n to non-integer values. This gives

$$\frac{\partial p}{\partial t}(x, t) = \frac{\partial h_1}{\partial x}(x, t) + \frac{1}{2} \frac{\partial^2 h_1}{\partial x^2}(x, t) - \frac{\partial h_2}{\partial x}(x, t) + \frac{1}{2} \frac{\partial^2 h_2}{\partial x^2}(x, t). \quad (16.6)$$

We need to be careful to remember that this equation is only valid for large values of x . Using (16.4) this can be written as the Fokker-Planck equation

$$\frac{\partial p}{\partial t}(x, t) = \frac{\partial^2}{\partial x^2}(d(x)p(x, t)) - \frac{\partial}{\partial x}(f(x)p(x, t)), \quad (16.7)$$

where the drift and diffusion coefficients are given by

$$f(x) = -k_1x + k_2\nu, \quad (16.8)$$

$$d(x) = (k_1x + k_2\nu)/2. \quad (16.9)$$

Now that we have obtained the Fokker-Planck equation (16.7) for the discrete-state system (16.1), we can apply the theory we developed earlier in Lectures 13 and 15 to it. We found previously that the stationary distribution is given by (15.6). Substituting (16.8)-(16.9) into (15.6), we obtain

$$\begin{aligned} p_s(x) &= \frac{2C}{k_1x + k_2\nu} \exp \left[2 \int_0^x \frac{-k_1y + k_2\nu}{k_1y + k_2\nu} dy \right] \\ &= \frac{2C}{k_1x + k_2\nu} \exp \left[-2x + 4k_2\nu \int_0^x \frac{1}{k_1y + k_2\nu} dy \right] \\ &= \frac{2C}{k_1x + k_2\nu} \exp \left[-2x + \frac{4k_2\nu}{k_1} \log(k_1x + k_2\nu) \right] \\ &= 2C \exp \left[-2x + \left(\frac{4k_2\nu}{k_1} - 1 \right) \log(k_1x + k_2\nu) \right], \end{aligned} \quad (16.10)$$

where

$$2C = \left(\int_{\mathbb{R}} \exp \left[-2x + \left(\frac{4k_2\nu}{k_1} - 1 \right) \log(k_1x + k_2\nu) \right] dx \right)^{-1}.$$

The function $p_s(x)$ given by (16.10) is plotted as the red line in Figure 16.1 (a). It compares well with the results obtained by the Gillespie SSA (3a)-(3d) (the blue histogram), which we computed earlier in Section 2.1 (see Figure 2.1 (b)).

Let us see how well the chemical Fokker-Planck equation does at estimating mean transition times. Define $\tau_{sim}(n)$, $n = 1, 2, \dots, 18$, to be the average time the Gillespie SSA (3a)-(3d) requires to leave the interval $(-\infty, 18]$ when the initial condition is $A(0) = n$. The values of $\tau_{sim}(n)$ estimated as the averages of 10^3 realisations (which were started as $A(0) = n$ and continued until the first time, t , at which $A(t) = 19$ is satisfied) are plotted in Figure 16.1 (b) as the blue histogram. The number $\tau(n)$ can also be approximated by the formula (15.13) which was derived from the Fokker-Planck equation (16.7). Using $x_u = 19$ in (15.13) (i.e. we are looking for the average time to leave the interval $(-\infty, 19)$), we get

$$\tau(y) = \int_y^{19} \frac{1}{d(z)p_s(z)} \int_{-\infty}^z p_s(x) dx dz, \quad (16.11)$$

where $p_s(z)$ is given by (16.10) and $d(z)$ is given by (16.9). Evaluating the integrals in (16.11) numerically, we present the function $\tau(y)$ in Figure 16.1 (b), as the red solid line, for comparison. We see that the formula (16.11) is a good approximation of $\tau(n)$.

So far we have seen the chemical Fokker-Planck equation only for the specific chemical system (16.1). Let us now consider a general well-stirred mixture of $N \geq 1$ molecular species that chemically interact through $q \geq 1$ chemical reactions. Recall that the state of this system is described by the

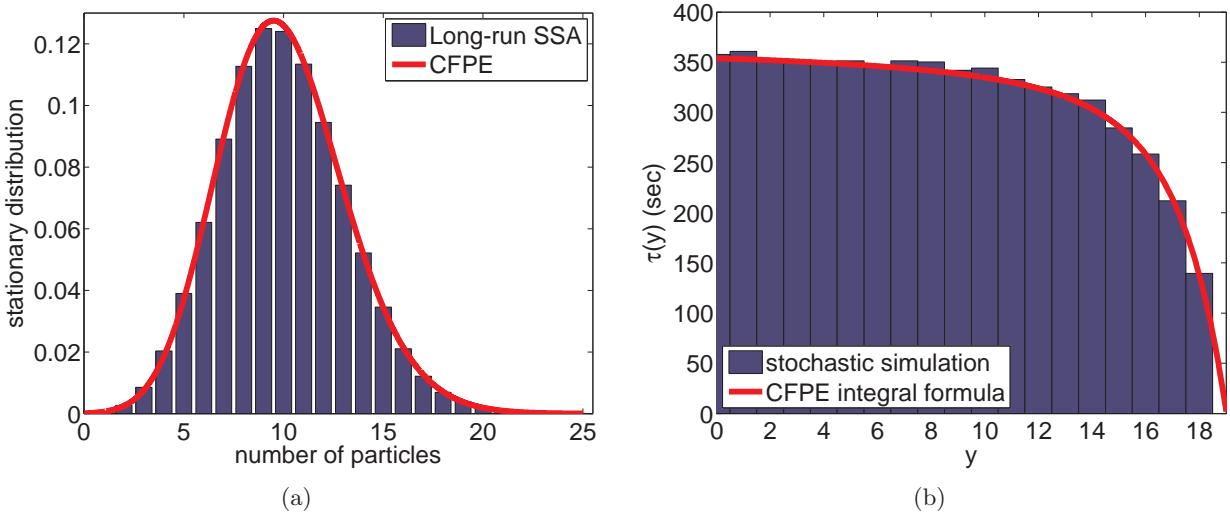


Figure 16.1: System of chemical reactions (16.1) for $k_1 = 0.1 \text{ sec}^{-1}$ and $k_2\nu = 1 \text{ sec}^{-1}$ (a) Stationary distribution obtained by the formula (16.10) (red solid line) and by long-time simulation of the SSA (3a)-(3d) (blue histogram). (b) The mean exit time, $\tau(y)$, obtained by the integral formula (16.11) (red solid line) and by the Gillespie SSA (3a)-(3d) (blue histogram).

state vector, $\mathbf{X} = [X_1, X_2, \dots, X_N]$ where $X_i \equiv X_i(t)$ is the number of particles of the i -th chemical species, $i = 1, 2, \dots, N$. Let $\alpha_j(\mathbf{x})$, $j = 1, 2, \dots, q$, be the propensity function of the j -th chemical reaction. Recall the stoichiometric vectors ν_j for $j = 1, 2, \dots, q$, which give the change in the state vector caused by one occurrence of the j -th reaction. Then the chemical system can be described by the general chemical master equation

$$\frac{\partial}{\partial t} p(\mathbf{x}, t) = \sum_{j=1}^q [\alpha_j (\mathbf{x} - \boldsymbol{\nu}_j) p(\mathbf{x} - \boldsymbol{\nu}_j, t) - \alpha_j (\mathbf{x}) p(\mathbf{x}, t)]. \quad (16.12)$$

The corresponding chemical Fokker-Planck equation is

$$\begin{aligned} \frac{\partial}{\partial t} p(\mathbf{x}, t) = & \sum_{i=1}^N \frac{\partial}{\partial x_i} \left[- \left(\sum_{j=1}^q \nu_{ji} \alpha_j(\mathbf{x}) \right) p(\mathbf{x}, t) \right] \\ & + \frac{1}{2} \sum_{i=1}^N \frac{\partial^2}{\partial x_i^2} \left[\left(\sum_{j=1}^q \nu_{ji}^2 \alpha_j(\mathbf{x}) \right) p(\mathbf{x}, t) \right] \\ & + \sum_{i=1}^N \sum_{k=1}^{i-1} \frac{\partial^2}{\partial x_i \partial x_k} \left[\left(\sum_{j=1}^q \nu_{ji} \nu_{jk} \alpha_j(\mathbf{x}) \right) p(\mathbf{x}, t) \right]. \end{aligned} \quad (16.13)$$

We derive the chemical Fokker-Planck equation (16.13) by asymptotically expanding the chemical master equation when the number of particles is large.

An alternative approach is to approximate the stochastic process which leads to the chemical master equation directly. Such an approach was carried out by Gillespie [50], and leads to the so-called chemical Langevin equation:

$$\boldsymbol{X}(t + \Delta t) = \boldsymbol{X}(t) + \left(\sum_{j=1}^q \boldsymbol{\nu}_j \alpha_j(\boldsymbol{X}(t)) \right) dt + \sum_{j=1}^q \boldsymbol{\nu}_j \sqrt{\alpha_j(\boldsymbol{X}(t))} dW_j. \quad (16.14)$$

The idea behind this approximation is similar to that of τ -leaping [8] (see Lecture 4). We suppose that

there are enough particles of each species that during the small but finite time-step Δt the propensity functions do not change appreciably, even though there are many reactions which take place. Since the reaction rates are then constant for this time interval, the number of times that each reaction occurs is drawn from a Poisson distribution. The state vector can be updated as in equation (4.7):

$$\mathbf{X}(t + \Delta t) = \mathbf{X}(t) + \sum_{j=1}^q K_j(\Delta t, \mathbf{x}, t) \boldsymbol{\nu}_j, \quad (4.7)$$

where $K_j(\Delta t, \mathbf{x}, t)$ represents the Poisson distributed number of times the reaction channel R_j fires in the interval $[t, t + \Delta t]$:

$$K_j \sim \text{Poisson}(\alpha_j(\mathbf{X}(t)) \cdot \Delta t).$$

If we now assume that each reaction occurs many times then this Poisson distribution is well-approximated by a normal distribution and equation (4.7) can be approximated by

$$\mathbf{X}(t + \Delta t) = \mathbf{X}(t) + \left(\sum_{j=1}^q \boldsymbol{\nu}_j \alpha_j(\mathbf{X}(t)) \right) \Delta t + \sum_{j=1}^q \boldsymbol{\nu}_j \sqrt{\alpha_j(\mathbf{X}(t))} \xi_j, \quad (16.15)$$

where ξ_j are random numbers which are sampled from the normal distribution with zero mean and unit variance. Equation (16.14) is recovered in the limit $\Delta t \rightarrow 0$. In fact, we can consider (16.15) as the computational definition of (16.14). The chemical Langevin equation for the model problem (16.1), can be written as the SDE

$$X(t + dt) = X(t) + (-k_1 X(t) + k_2 \nu) dt - \sqrt{k_1 X(t)} dW_1 + \sqrt{k_2 \nu} dW_2. \quad (16.16)$$

This equation is slightly different from the SDE (13.6) studied previously, in that it has two independent white noises, dW_1 and dW_2 . Fortunately, having two independent noise terms in the SDE does not complicate the derivation of the corresponding Fokker-Planck equation. In fact, one can prove (see Problem Sheet 8) that the Fokker-Planck equation corresponding to the SDE (16.16) is given by (16.7).

Bibliography

- [1] D.T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.*, 81(25):2340–2361, 1977.
- [2] M.A. Gibson and J. Bruck. Efficient exact stochastic simulation of chemical systems with many species and many channels. *J. Phys. Chem. A.*, 104(9):1876–1889, 2000.
- [3] Y. Cao, H. Li, and L. Petzold. Efficient formulation of the stochastic simulation algorithm for chemically reacting systems. *J. Chem. Phys.*, 121(9):4059–4067, 2004.
- [4] C.A. Yates and G. Klingsbeil. Recycling random numbers in the stochastic simulation algorithm. *J. Chem. Phys.*, 138(9):094103, 2013.
- [5] J.M. McCollum, G.D. Peterson, C.D. Cox, M.L. Simpson, and N.F. Samatova. The sorting direct method for stochastic simulation of biochemical systems with varying reaction execution behavior. *Comput. Biol. Chem.*, 30(1):39–49, 2006.
- [6] H. Li and L. Petzold. Logarithmic direct method for discrete stochastic simulation of chemically reacting systems. Technical report, University of California Santa Barbara, 2006.
- [7] C. Gadgil, Chang H. Lee, and H.G. Othmer. A stochastic analysis of first-order reaction networks. *Bull. Math. Biol.*, 67(5):901–946, 2005.
- [8] D.T. Gillespie. Approximate accelerated stochastic simulation of chemically reacting systems. *J. Chem. Phys.*, 115(4):1716–1733, 2001.
- [9] G. Grimmett and D. Stirzaker. *Probability and random processes*. Oxford university press, 2001.
- [10] H. Li, Y. Cao, L.R. Petzold, and D.T. Gillespie. Algorithms and software for stochastic simulation of biochemical reacting systems. *Biotechnol. Progr.*, 24(1):56–61, 2008.
- [11] D.F. Anderson and D.J. Higham. Multi-level monte carlo for continuous time markov chains, with applications in biochemical kinetics. *Multiscale. Model. Sim.*, 10(1):146–179, 2012.
- [12] D.T. Gillespie and L. Petzold. *Numerical simulation for biochemical kinetics*, pages 331–354. The MIT press, 2006.
- [13] D.T. Gillespie and L.R. Petzold. Improved leap-size selection for accelerated stochastic simulation. *J. Chem. Phys.*, 119(16):8229–8234, 2003.
- [14] Y. Cao, D.T. Gillespie, and L.R. Petzold. Efficient step size selection for the tau-leaping simulation method. *J. Chem. Phys.*, 124(4):044109, 2006.
- [15] Y. Cao, D.T. Gillespie, and L.R. Petzold. Avoiding negative populations in explicit Poisson tau-leaping. *J. Chem. Phys.*, 123(5):054104, 2005.

- [16] A. Chatterjee, D.G. Vlachos, and M.A. Katsoulakis. Binomial distribution based τ -leap accelerated stochastic simulation. *J. Chem. Phys.*, 122(2):024112, 2004.
- [17] T. Tian and K. Burrage. Binomial leap methods for simulating stochastic chemical kinetics. *J. Chem. Phys.*, 121(21):10356–10364, 2004.
- [18] C.A. Yates and K. Burrage. Look before you leap: A confidence-based method for selecting species criticality while avoiding negative populations in τ -leaping. *J. Chem. Phys.*, 134(8):084109, 2011.
- [19] J. Paulsson, O.G. Berg, and M. Ehrenberg. Stochastic focusing: fluctuation-enhanced sensitivity of intracellular regulation. *Proc. Natl. Acad. Sci. USA*, 97(13):7148–7153, 2000.
- [20] F. Schlögl. Chemical reaction models for non-equilibrium phase transitions. *Zeitschrift für Physik*, 253(2):147–161, 1972.
- [21] T.S. Gardner, C.R. Cantor, and J.J. Collins. Construction of a genetic toggle switch in *Escherichia coli*. *Nature*, 403(6767):339–342, 2000.
- [22] J. Hasty, D. McMillen, F. Isaacs, and J.J. Collins. Computational studies of gene regulatory networks: in numero molecular biology. *Nat. Rev. Genet.*, 2(4):268–279, 2001.
- [23] T.B. Kepler and T.C. Elston. Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations. *Biophys. J.*, 81(6):3116–3136, 2001.
- [24] J. Buhl, D.J.T. Sumpter, I.D. Couzin, J. Hale, E. Despland, E. Miller, and S.J. Simpson. From disorder to order in marching locusts. *Science*, 312(5778):1402–1406, 2006.
- [25] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet. Novel type of phase transition in a system of self-driven particles. *Phys. Rev. Lett.*, 75(6):1226–1229, 1995.
- [26] A. Czirók, A.L. Barabási, and T. Vicsek. Collective motion of self-propelled particles: kinetic phase transition in one dimension. *Phys. Rev. Lett.*, 82(1):209–212, 1999.
- [27] C.A. Yates, R. Erban, C. Escudero, I.D. Couzin, J. Buhl, I.G. Kevrekidis, P.K. Maini, and D.J.T. Sumpter. Inherent noise can facilitate coherence in collective swarm motion. *Proc. Natl. Acad. Sci. USA*, 106(14):5464–5469, 2009.
- [28] L. Dyson, C.A. Yates, J. Buhl, and A. McKane. A minimal model captures the onset of collective motion in locusts. *Phys. Rev. E*, 2015.
- [29] R.E.L. DeVille, C.B. Muratov, and E. Vanden-Eijnden. Non-meanfield deterministic limits in chemical reaction kinetics. *The Journal of chemical physics*, 124(23):231102, 2006.
- [30] C.B. Muratov, E. Vanden-Eijnden, and W. E. Self-induced stochastic resonance in excitable systems. *Physica D*, 210(3):227–240, 2005.
- [31] J. Schnakenberg. Simple chemical reaction systems with limit cycle behaviour. *J. Theor. Biol.*, 81(3):389–400, 1979.
- [32] R. Erban, S.J. Chapman, I.G. Kevrekidis, and T. Vejchodský. Analysis of a stochastic chemical system close to a sniper bifurcation of its mean-field model. *SIAM J. Appl. Math.*, 70(3):984–1016, 2009.
- [33] P. Whittle. On the use of the normal approximation in the treatment of stochastic processes. *J. Roy. Stat. Soc. B Met.*, pages 268–281, 1957.

- [34] T.I. Matis and I.G. Guardiola. Achieving moment closure through cumulant neglect. *The Mathematica Journal*, 12, 2010.
- [35] M. Baar, L. Coquille, H. Mayer, M. Hözel, M. Rogava, T. Tüting, and A. Bovier. A stochastic model for immunotherapy of cancer. *Sci. Rep.*, 6:24169, 2016.
- [36] A. Castellanos-Moreno, A. Castellanos-Jaramillo, A. Corella-Madueño, S. Gutiérrez-López, and R. Rosas-Burgos. Stochastic model for computer simulation of the number of cancer cells and lymphocytes in homogeneous sections of cancer tumors. *arXiv preprint arXiv:1410.3768*, 2014.
- [37] M. Ryser, W. Lee, N. Ready, K. Leder, and J. Foo. Quantifying the dynamics of field cancerization in hpv-negative head and neck cancer: A multiscale modeling approach. *Cancer research*, 76(24):7078–7088, 2016.
- [38] G.P. Figueredo, P.-O. Siebers, M.R. Owen, J. Reps, and U. Aickelin. Comparing stochastic differential equations and agent-based modelling and simulation for early-stage cancer. *PLoS one*, 9(4):e95150, 2014.
- [39] C. Turner, A.R. Stinchcombe, M. Kohandel, S. Singh, and S. Sivaloganathan. Characterization of brain cancer stem cells: a mathematical approach. *Cell proliferation*, 42(4):529–540, 2009.
- [40] R.L. Mort, R.J.H. Ross, K.J. Hainey, O.J. Harrison, M.A. Keighren, G. Landini, R.E. Baker, K.J. Painter, I.J. Jackson, and C.A. Yates. Reconciling diverse mammalian pigmentation patterns with a fundamental mathematical model. *Nat. Commun.*, 7:10288, 2016.
- [41] M. Zaider and G.N. Minerbo. Tumour control probability: a formulation applicable to any temporal protocol of dose delivery. *Phys. Med. Biol.*, 45(2):279, 2000.
- [42] C.A. Yates, M.J. Ford, and R.L. Mort. A multi-stage representation of cell proliferation as a markov process. *Bull. Math. Biol.*, 2017.
- [43] T.S. Weber, I. Jaehnert, C. Schichor, M. Or-Guil, and J. Carneiro. Quantifying the length and variance of the eukaryotic cell cycle phases by a stochastic model and dual nucleoside pulse labelling. *PLoS. Comput. Biol.*, 10(7):e1003616, 2014.
- [44] J.A. Smith and L. Martin. Do cells cycle? *Proc. Natl. Acad. Sci. USA*, 70(4):1263–1267, 1973.
- [45] A. Golubev. Applications and implications of the exponentially modified gamma distribution as a model for time variabilities related to cell proliferation and gene expression. *J. Theor. Biol.*, 393:203–217, 2016.
- [46] W. Feller. On the integro-differential equations of purely discontinuous Markoff processes. *T. Am. Math. Soc.*, 48(3):488–515, 1940.
- [47] L. Arnold. Stochastic differential equations: theory and applications. *New York*, 1974.
- [48] C.W. Gardiner. *Handbook of stochastic methods for physics, chemistry and the natural sciences*. Springer Series in Synergetics, fourth edition, 2009.
- [49] H. Risken. *The Fokker-Planck Equation. Methods of Solution and Applications*. 1989.
- [50] D.T. Gillespie. The chemical Langevin equation. *J. Chem. Phys.*, 113(1):297–306, 2000.
- [51] P. Hänggi, P. Talkner, and M. Borkovec. Reaction-rate theory: fifty years after kramers. *Rev. Mod. Phys.*, 62(2):251, 1990.

- [52] Albert Einstein. Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen. *Ann. Phys.-Leipzig.*, 322(8):549–560, 1905.
- [53] H.C. Berg. *Random Walks in Biology*. Princeton University Press, 1993.
- [54] H.G. Othmer, S.R. Dunbar, and W. Alt. Models of dispersal in biological systems. *J. Math. Biol.*, 26(3):263–298, 1988.
- [55] M. Kac. A stochastic model related to the telegrapher’s equation. *Rocky Mountain Journal of Mathematics*, 4:000, 1974.
- [56] E. Zauderer. *Partial differential equations of applied mathematics*, volume 71. John Wiley & Sons, New York, 2011.
- [57] G. Karch. Selfsimilar profiles in large time asymptotics of solutions to damped wave equations. *Stud. Math.*, 143(2):175–197, 2000.
- [58] S. Chandrasekhar. Stochastic problems in physics and astronomy. *Rev. Mod. Phys.*, 15(1):1–89, 1943.
- [59] H.C. Berg and E.M. Purcell. Physics of chemoreception. *Biophys. J.*, 20(2):193–219, 1977.
- [60] D. Shoup and A. Szabo. Role of diffusion in ligand binding to macromolecules and cell-bound receptors. *Biophys. J.*, 40(1):33–39, 1982.
- [61] R. Erban and S.J. Chapman. Time scale of random sequential adsorption. *Phys. Rev. E*, 75(4):041116, 2007.
- [62] R. Erban and S.J. Chapman. Reactive boundary conditions for stochastic simulations of reaction–diffusion processes. *Phys. Biol.*, 4(1):16–28, 2007.
- [63] P.I. Zhuravlev and G.A. Papoian. Molecular noise of capping protein binding induces macroscopic instability in filopodial dynamics. *Proc. Natl. Acad. Sci. USA*, 106(28):11570–11575, 2009.
- [64] M. Howard. How to build a robust intracellular concentration gradient. *Trends Cell Biol.*, 22(6):311–317, 2012.
- [65] K. Lipkow, S.S. Andrews, and D. Bray. Simulated diffusion of phosphorylated chey through the cytoplasm of *Escherichia coli*. *J. Bacteriol.*, 187(1):45–53, 2005.
- [66] D. Fange and J. Elf. Noise-induced min phenotypes in e. coli. *PLoS. Comput. Biol.*, 2(6):e80, 2006.
- [67] K. Takahashi, S. Tănase-Nicola, and P. R. Ten Wolde. Spatio-temporal correlations can drastically change the response of a mapk pathway. *Proc. Natl. Acad. Sci. USA.*, 107(6):2473–2478, 2010.
- [68] J. Hattne, D. Fange, and J. Elf. Stochastic reaction-diffusion simulation with MesoRD. *Method. Biochem. Anal.*, 21(12):2923–2924, 2005.
- [69] S. Engblom, L. Ferm, A. Hellander, and P. Lötstedt. Simulation of stochastic reaction-diffusion processes on unstructured meshes. *SIAM J. Sci. Comput.*, 31(3):1774–1797, 2009.
- [70] S. Wils and E. De Schutter. Steps: modeling and simulating complex reaction-diffusion systems with python. *Front. Neuroinfo.*, 3, 2009.

- [71] M. Ander, P. Beltrao, B. Di Ventura, J. Ferkinghoff-Borg, M. Foglierini, A. Kaplan, C. Lemerle, I. Tomas-Oliveira, and Serrano. SmartCell, a framework to simulate cellular processes that combines stochastic approximation with diffusion and localisation: analysis of simple networks. *Syst. Biol.*, 1(1):129–138, 2004.
- [72] R. Erban and S.J. Chapman. Stochastic modelling of reaction–diffusion processes: algorithms for bimolecular reactions. *Phys. Biol.*, 6(4):1–18, 2009.
- [73] S.A. Isaacson. The reaction-diffusion master equation as an asymptotic approximation of diffusion to a small target. *SIAM J. Appl. Math.*, 70(1):77–111, 2009.
- [74] S. Isaacson and C. Peskin. Incorporating diffusion in complex geometries into stochastic chemical kinetics simulations. *SIAM J. Sci. Comput.*, 28(1):47–74, 2006.
- [75] M.B. Elowitz, A.J. Levine, E.D. Siggia, and P.S. Swain. Stochastic gene expression in a single cell. *Sci. Signal.*, 297(5584):1183, 2002.
- [76] O. Shimmi, D. Umulis, H. Othmer, and M.B. O’Connor. Facilitated transport of a dpp/scw heterodimer by sog/tsg leads to robust patterning of the drosophila blastoderm embryo. *Cell*, 120(6):873–886, 2005.
- [77] G.T. Reeves, R. Kalifa, D.E. Klein, M.A. Lemmon, and S.Y. Shvartsman. Computational analysis of egfr inhibition by argos. *Dev. Biol.*, 284(2):523–535, 2005.
- [78] L. Wolpert, R. Beddington, T. Jessell, P. Lawrence, E. Meyerowitz, and J. Smith. *Principles of Development*. Oxford University Press, New York, 3rd edition, 2006.
- [79] F. Tostevin, P. R. Ten Wolde, and M. Howard. Fundamental limits to position determination by concentration gradients. *PLoS. Comput. Biol.*, 3(4):e78, 2007.
- [80] A.M. Turing. The chemical basis of morphogenesis. *Phil. Trans. R. Soc. B.*, 237(641):37–72, 1952.
- [81] A. Gierer and H. Meinhardt. A theory of biological pattern formation. *Kybernetik*, 12(1):30–39, 1972.
- [82] J.D. Murray. *Mathematical Biology I: An Introduction*, volume 1 of *Interdisciplinary Mathematics*. Springer, New York, 2002.
- [83] S. Sick, S. Reinker, J. Timmer, and T. Schlake. Wnt and dkk determine hair follicle spacing through a reaction-diffusion mechanism. *Science*, 314(5804):1447–1450, 2006.
- [84] L. Qiao, R. Erban, C.T. Kelley, and I.G. Kevrekidis. Spatially distributed stochastic systems: Equation-free and equation-assisted preconditioned computations. *The Journal of chemical physics*, 125(20):204108, 2006.
- [85] T. Biancalani, D. Fanelli, and F. Di Patti. Stochastic Turing patterns in the Brusselator model. *Phys. Rev. E*, 81(4):046215, 2010.
- [86] T.E. Woolley, R.E. Baker, E.A. Gaffney, and P.K. Maini. Stochastic reaction and diffusion on growing domains: Understanding the breakdown of robust pattern formation. *Phys. Rev. E*, 84(4):046216, 2011.
- [87] S.A. Andrews and D. Bray. Stochastic simulation of chemical reactions with spatial resolution and single molecule detail. *Phys. Biol.*, 1(3-4):137–151, 2004.

- [88] J. Stiles and T. Bartol. *Monte Carlo Methods for Simulating Realistic Synaptic Microphysiology Using MCell*, in, “*Computational Neuroscience, Realistic Modelling for Experimentalists*”, chapter 4, pages 87–127. CRC Press, 2001.
- [89] J.S. van Zon and P.R. ten Wolde. Green’s-function reaction dynamics: a particle-based approach for simulating biochemical networks in time and space. *J. Chem. Phys.*, 123(23):234910, 2005.
- [90] M. Smoluchowski. Versuch einer mathematischen theorie der koagulationskinetik kolloider lösungen. *Z. Phys. Chem.*, 92(129-168):9, 1917.
- [91] J. Lipková, K.C. Zygalakis, S.J. Chapman, and R. Erban. Analysis of brownian dynamics simulations of reversible bimolecular reactions. *SIAM J. Appl. Math.*, 71(3):714–730, 2011.
- [92] Radek Erban and Hans G Othmer. From signal transduction to spatial pattern formation in *e. coli*: A paradigm for multiscale modeling in biology. *Multiscale Modeling & Simulation*, 3(2):362–394, 2005.
- [93] N. Barkai and S. Leibler. Robustness in simple biochemical networks. *Nature*, 387(6636):913–917, 1997.
- [94] P.A. Spiro, J.S. Parkinson, and H.G. Othmer. A model of excitation and adaptation in bacterial chemotaxis. *Proceedings of the National Academy of Sciences*, 94(14):7263–7268, 1997.
- [95] U. Alon, M.G. Surette, N. Barkai, and S. Leibler. Robustness in bacterial chemotaxis. *Nature*, 397(6715):168–171, 1999.
- [96] R. Erban and H.G. Othmer. From individual to collective behavior in bacterial chemotaxis. *SIAM J. Appl. Math.*, 65(2):361–391, 2004.
- [97] H.G. Othmer and P. Schaap. Oscillatory cAMP signaling in the development of *Dictyostelium discoideum*. *Comments on Theoretical Biology*, 5:175–282, 1998.
- [98] T. Hillen. Hyperbolic models for chemosensitive movement. *Math. Models Methods Appl. Sci.*, 12(07):1007–1034, 2002.
- [99] T. Hillen and H.G. Othmer. The diffusion limit of transport equations derived from velocity-jump processes. *SIAM J. Appl. Math.*, 61(3):751–775, 2000.
- [100] H.G. Othmer and T. Hillen. The diffusion limit of transport equations II: Chemotaxis equations. *SIAM J. Appl. Math.*, 62(4):1222–1250, 2002.
- [101] G. Gerisch. Chemotaxis in *dictyostelium*. *Annu. Rev. Physiol.*, 44(1):535–552, 1982.
- [102] R. Erban and H.G. Othmer. Taxis equations for amoeboid cells. *J. Math. Biol.*, 54(6):847–885, 2007.
- [103] E.F. Keller and L.A. Segel. Model for chemotaxis. *J. Theor. Biol.*, 30(2):225–234, 1971.
- [104] E.F. Keller and L.A. Segel. Traveling bands of chemotactic bacteria: a theoretical analysis. *Journal of Theoretical Biology*, 30(2):235–248, 1971.
- [105] M. Wooldridge. An introduction to multiagent systems. 2009.
- [106] B. Franz and R. Erban. *Hybrid modelling of individual movement and collective behaviour*. Springer, 2011.
- [107] D.J.T. Sumpter. Collective animal behavior. 2010.

- [108] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *J. Chem. Phys.*, 21(6):1087–1092, 1953.
- [109] S. Chib and E. Greenberg. Understanding the metropolis-hastings algorithm. *Am. Stat.*, 49(4):327–335, 1995.
- [110] Y. Cao, D.T. Gillespie, and L.R. Petzold. The slow-scale stochastic simulation algorithm. *J. Chem. Phys.*, 122:014116, 2005.
- [111] Y. Cao and L. Petzold. Accuracy limitations and the measurement of errors in the stochastic simulation of chemically reacting systems. *J. Comput. Phys.*, 212(1):6–24, 2005.
- [112] W. E, D. Liu, and E. Vanden-Eijnden. Nested stochastic simulation algorithm for chemical kinetic systems with disparate rates. *J. Chem. Phys.*, 123(19):194107:1:8, 2005.
- [113] R. Erban, I.G. Kevrekidis, D. Adalsteinsson, and T.C. Elston. Gene regulatory networks: a coarse-grained, equation-free approach to multiscale computation. *J. Chem. Phys.*, 124(8):084106:1–17, 2006.
- [114] S.L. Cotter, K.C. Zygalakis, I.G. Kevrekidis, and R. Erban. A constrained approach to multiscale stochastic simulation of chemically reacting systems. *J. Chem. Phys.*, 135(9):094102, 2011.
- [115] A. Arkin, J. Ross, and H.H. McAdams. Stochastic kinetic analysis of developmental pathway bifurcation in phage λ -infected *Escherichia coli* cells. *Genetics*, 149(4):1633–1648, 1998.
- [116] S. Kar, W.T. Baumann, M.R. Paul, and J.J. Tyson. Exploring the roles of noise in the eukaryotic cell cycle. *Proc. Natl. Acad. Sci. USA*, 106(16):6471–6476, 2009.
- [117] M.B. Flegg, S.J. Chapman, and R. Erban. The two-regime method for optimizing stochastic reaction-diffusion simulations. *J. Roy. Soc. Interface*, 9(70):859–868, 2012.
- [118] G. Moy, B. Corry, S. Kuyucak, and S.H. Chung. Tests of continuum theories as models of ion channels. I. Poisson- Boltzmann theory versus Brownian dynamics. *Biophys. J.*, 78(5):2349–2363, 2000.
- [119] B. Corry, S. Kuyucak, and S.-H. Chung. Tests of continuum theories as models of ion channels. ii. poisson–nernst–planck theory versus brownian dynamics. *Biophys. J.*, 78(5):2364–2381, 2000.
- [120] E. Moro. Hybrid method for simulating front propagation in reaction-diffusion systems. *Phys. Rev. E*, 69(6):060101, 2004.
- [121] B. Franz, M.B. Flegg, S.J. Chapman, and R. Erban. Multiscale reaction-diffusion algorithms: PDE-assisted Brownian dynamics. *SIAM J. Appl. Math.*, 73(3):1224–1247, 2013.
- [122] L. Ferm, A. Hellander, and P. Lötstedt. An adaptive algorithm for simulation of stochastic reaction-diffusion processes. *J. Comput. Phys.*, 229(2):343–360, 2010.
- [123] F.J. Alexander, A.L. Garcia, and D.M. Tartakovsky. Algorithm refinement for stochastic partial differential equations: I. Linear diffusion. *J. Comput. Phys.*, 182(1):47–66, 2002.
- [124] E.G. Flekkøy, J. Feder, and G. Wagner. Coupling particles and fields in a diffusive hybrid model. *Phys. Rev. E*, 64(6):066302, 2001.
- [125] G. Wagner and E.G. Flekkoy. Hybrid computations with flux exchange. *Phil. Trans. R. Soc. A*, 362:1655–1666, 2004.

- [126] J.C. Dallon and H.G. Othmer. A discrete cell model with adaptive signalling for aggregation of *Dictyostelium discoideum*. *Phil. Trans. R. Soc. B*, 352(1351):391, 1997.
- [127] A.A. Patel, E.T. Gawlinski, S.K. Lemieux, and R.A. Gatenby. A cellular automaton model of early tumor growth and invasion: the effects of native tissue vascularity and increased anaerobic tumor metabolism. *J. Theor. Biol.*, 213(3):315–331, 2001.
- [128] K. Smallbone, R.A. Gatenby, R.J. Gillies, P.K. Maini, and D.J. Gavaghan. Metabolic changes during carcinogenesis: potential impact on invasiveness. *J. Theor. Biol.*, 244(4):703–713, 2007.
- [129] A.R.A. Anderson. A hybrid mathematical model of solid tumour invasion: the importance of cell adhesion. *Math. Med. Biol.*, 22(2):163–186, 2005.
- [130] P. Gerlee and A.R.A. Anderson. An evolutionary hybrid cellular automaton model of solid tumour growth. *J. Theor. Biol.*, 246(4):583–603, 2007.