Stochastic Differential Equation Models in Population Dynamics

Marina Filipa Amado Ferreira

Stochastic Differential Equation Models in Population Dynamics

Marina Filipa Amado Ferreira

Dissertação para a obtenção do Grau de **Mestre em Matemática**

Área de Especialização em **Estatística, Otimização e Matemática Financeira**

Júri

Presidente:	Prof. Dr. Paulo Eduardo Aragão Aleixo e Neves de Oliveira
Orientador:	Prof. Dra. Ana Cristina Martins Rosa
Vogais:	Prof. Dra. Ercília Cristina da Costa e Sousa

Data: Setembro de 2014

Resumo

O cálculo estocástico tem adquirido uma grande dimensão durante as últimas décadas, tanto do ponto de vista teórico como prático. Em particular, no ramo da biologia teórica tem contribuído para a dedução de modelos mais realistas, mas ao mesmo tempo mais complexos. O estudo destes modelos reduz-se, por isso, muitas vezes a uma abordagem numérica. Os principais objetivos desta tese passam por motivar, estabelecer e analisar modelos estocásticos de dinâmica de populações que resultam de dois tipos diferentes de aleatoriedade: ambiental e demográfica. De modo a analisarmos a primeira classe de modelos, assim como a justificar o procedimento usado para deduzir a segunda classe, é previamente brevemente apresentada a teoria dos processos de difusão e dos sistemas de equações diferenciais estocásticas. Adicionalmente, apresentamos também um breve estudo numérico dos modelos que é feito recorrendo ao método de Euler-Maruyama.

Palavras Chave: processos de difusão, sistemas de equações diferenciais estocásticas de Itô, modelos de crescimento populacional, modelo de epidemias

Abstract

Stochastic calculus has gained a great expansion in the last decades both from theoretical and practical perspectives. In particular, in the field of theoretical biology, it has contributing for the deduction of more realistic but also less tractable models. Due to the complexity of the mathematical theory behind stochastic calculus, the approaches followed are often limited to numerical technics. The principal aims of this thesis are the motivation, establishment and discussion of classic stochastic differential equation models in population dynamics, resulting from two different types of randomness: environmental and demographic. In order to analyse the first class of models, as well as rigorously justify the procedure used in the second class, we deal with the theory of diffusion processes and Itô stochastic differential systems. Additionally, a numerical study by implementing the Euler-Maruyama method is made.

Keywords: diffusion processes, system of Itô stochastic differential equations, growth population models, epidemic model

Agradecimentos

Contents

Chapter 1

Introduction

The study of Markov processes, and particularly of diffusion processes, marked undoubtedly the origin of the research concerning continuous-time stochastic processes. Besides the contributions of N. Wiener (1923) and Lévy-Khinchin (1937), we may say that the modern theory of Markov processes was initiated by A. Kolmogorov (1931), who characterized the transition probability distribution of a diffusion process as a solution of a parabolic partial differential equation involving its drift and diffusion coefficients, extended afterwards by Feller (1936) to the characterization of a Markov process admitting jumps.

K. Itô (1940) entered in the scene right after this:

"I saw a powerful analytic method to study the transition probabilities of the process, namely Kolmogorov's parabolic equation and its extension by Feller. But I wanted to study the paths of Markov processes (...)".

Itô's original goal was thus to complement the study of diffusion processes (see [1] [2]), in order to permit the explicit construction of their trajectories. This goal led him to conceive a new concept of integral, named *stochastic integral*, giving rise to the famous stochastic calculus and stochastic differential equations (SDEs), which revealed to be a powerful tool for the construction and analysis of mathematical models. In fact, during the last decades, SDEs gained a remarkable popularity by finding multiple applications in different areas of science such as finance, chemistry, physics, engineering, population dynamics and, more generally, in systems biology (see, for instance, [3], [4], [5], [6], [7], [8] and [9]). They are particularly important in modelling complicated systems where randomness can not be ignored. On the other hand, these applications have also motivated theoretical developments conducting to new kinds of stochastic integrals (see, for example, [1], [5]).

This thesis aims to motivate, establish, compare and analyse, essentially from a

numerical point of view, some classic models of stochastic differential equations in the field of life sciences, supported by a rigorous mathematical justification, as far as possible. We begin, in Chapter 2, by briefly describing the probabilistic background needed to understand the models that will be considered and the techniques supporting their manipulation. We then proceed by presenting the class of diffusion processes and the corresponding Kolmogorov's forward differential equation, in order to establish the link between SDE solutions and diffusion processes. Later on, in Chapter 3, two different classes of stochastic models are analised and several examples are given. In the first class, models are obtained directly from a deterministic ordinary differential equation (ODE) model, by adding a random perturbation, which pretends to represent environmental stochasticity. The stochastic Malthusian growth model (Black-Scholes model), the stochastic logistic growth model as well as the stochastic Lotka-Volterra model, are then settled and slightly explored. In the second class of models, the stochasticity arises intrinsically from the underlying phenomena. A general time-continuous Markov chain model is first constructed, leading to the establishment of a SDE by diffusion approximation procedures (see [10] [11] [12] [13]). A few examples are subsequently given, including a second version of the stochastic logistic growth model, a two species stochastic Lotka-Volterra model and, finally, a stochastic epidemic model.

In order to gain some insight about the models' behavior and, consequently, about the underlying biological phenomenon, the Euler-Maruyama method was implemented for approximating trajectories of SDE solutions. Although there exist higher order convergence methods (see [3]), this method is computationally simple and often proposed in the literature. In this way, it was possible to obtain some sample paths for the solutions of nonlinear models, for which an analytical resolution procedure is in general not available. As in the scope of PDEs, some analytical approaches exist and others continue being developed, in order to complement the numerical study. However, because of the complex mathematical tools they encompass, we opted by not bringing very deeply this analysis to this thesis, instead, we suggest appropriate references. Since this is yet a quite recent and complex field of research, when compared to deterministic partial differential equations (PDEs), still many questions are far from being answered and new techniques must be developed. Finally, together with the models' descriptions, we present a short discussion on the comparison between deterministic and stochastic models.

Chapter 2

Diffusion processes and solutions of Stochastic Differential Equations

2.1. Preliminaries

In this thesis, we consider the complete[1] probability measure space $(\Omega, \mathcal{A}, P)$ where all real random vectors and real vector-valued stochastic processes are defined. The set of d-dimensional real vectors, as well as any subset of it, is equipped with the Borel σ-algebra, $\mathcal{B}_C$, $C \subseteq \mathbb{R}^d$, $d \in \mathbb{N}$. The time interval on which all stochastic processes are defined is $[0, T]$, for some $T > 0$, which may be allowed to be infinite.

Vectors will be written in bold to distinguish from scalars. The Euclidean vector-norm will be denoted by $|.|$ and the *Hilbert-Schmidt* norm, $\|.\|$, for an $m \times n$ real matrix $A = [a_{ij}]_{i=1,\dots,m,j=1,\dots,n}$, $m, n \in \mathbb{N}$, is defined by

$$\|A\| = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2 \right)^{1/2}.$$

The Wiener process is a basic tool in stochastic calculus. Among several definitions that can be found in the literature, we consider the following one.(see [14])

Definition 1. **(i)** *Given a complete[2] filtration $(G_t, t \in [0, T])$ of $(\Omega, \mathcal{A}, P)$, the stochastic process $W = (W_t, t \in [0, T])$ is called a* **scalar Wiener process** *with respect to $(G_t, t \in [0, T])$ if it satisfies the following properties:*

- $X_0 = 0$ *a.s.,*

- W_t *is adapted to G_t, $\forall t \in [0, T]$*

- $\forall\ t \in [0, T],\ \forall\ h > 0,\ with\ 0 \leq t < t + h \leq T,\ W_{t+h} - W_t \sim \mathcal{N}(0, \sigma^2 h),$ *where σ^2 is a real positive constant,*

---

[1]A probability measure space is complete if its σ-algebra is complete. In turn, a σ-algebra, $\mathcal{F}$, is called complete if for each $A \in \mathcal{F}$ such that $P(A) = 0$, one has $B \in \mathcal{F}$ for any $B \subset A$. The complete framework is used in this thesis for technical convenience.

[2]Notice that a filtration is complete if all σ-algebras which constitute it are complete.

- $\forall t \in [0, T]$, $\forall h > 0$, with $0 \le t < t + h \le T$, $W_{t+h} - W_t$ is independent of G_t,

- almost all sample paths of W are continuous functions.

(ii) A d-**dimensional Wiener process** $\mathbf{W}$ is defined by $\mathbf{W} = (\mathbf{W}_t, t \in [0, T])$, where $\mathbf{W}_t = (W_t^{(1)}, ..., W_t^{(d)})$ and $(W_t^{(i)}, t \in [0, T])$, $i = 1, ..., d$, $d \in \mathbb{N}$, are d independent scalar Wiener processes.

If $\sigma^2 = 1$, W is called the *standard Wiener process.*

We denote by $(\mathcal{F}_t, \ t \in [0, T])$ the complete filtration of $(\Omega, \mathcal{A}, P)$ generated by the Wiener process $\mathbf{W}$.

For $T \in (0, T]$, we denote by $L^i([0, T] \times \Omega)$, $i = 1, 2$, the spaces of functions $f : [0, T] \times \Omega \to \mathbb{R}^d$ such that

- f is measurable regarding the tensor-product σ-algebra $\mathcal{B}_{[0,T]} \otimes \mathcal{A}$,

- f is adapted to the filtration $(\mathcal{F}_t, \ t \in [0, T])$,

- $\int_0^T |f|^i dt < \infty$ a.s., $i = 1, 2$.

We next present a central tool in stochastic calculus, namely, the famous Itô's lemma, which is formulated for a specific class of processes, called Itô processes.

Definition 2. *Let $\mathbf{W} = (\mathbf{W}_t, t \in [0, T])$ be an m-dimensional Wiener process.*

(i) *A scalar stochastic process $X = (X_t, t \in [0, T])$ is called a* **scalar Itô process** *driven by an m-dimensional Wiener process $\mathbf{W}$ if*

  - X_0 *is $\mathcal{F}_0$-measurable,*

  - X *can be expressed in the form*

$$X_t = X_0 + \int_0^t B_s ds + \sum_{j=1}^m \int_0^t C_s^{(j)} dW_s^{(j)} \ \ a.s., \ t \in [0, T], \qquad (2.1)$$

  where $(B_t, t \in [0, T]) \in L^1(\Omega \times [0, T])$ and $(C_t^{(j)}, t \in [0, T]) \in L^2(\Omega \times [0, T])$, $j = 1, ..., m$.

(ii) *An n-dimensional stochastic process $\mathbf{X} = (\mathbf{X}_t, t \in [0, T])$, with $\mathbf{X}_t = (X_t^{(1)}, ..., X_t^{(n)})$ is called an* **Itô process** *driven by $\mathbf{W}$ if, for each $i = 1, ..., n$, $X_t^{(i)}$ is a scalar Itô process, defined by*

$$X_t^{(i)} = X_0^{(i)} + \int_0^t B_s^{(i)} ds + \sum_{j=1}^m \int_0^t C_s^{(i,j)} dW_s^{(j)} \ \ a.s., \ t \in [0, T] \qquad (2.2)$$

where $(B_t^{(i)}, t \in [0,T]) \in L^1(\Omega \times [0,T])$ and $(C_t^{(i,j)}, t \in [0,T]) \in L^2(\Omega \times [0,T])$, $i = 1,...,n$ $j = 1,...,m$.

Moreover, rewriting (2.2) *in the differential form, we obtain the* **stochastic differential** *of* **X**, *namely,*

$$d\mathbf{X}_t = \mathbf{B}_t dt + \mathbf{C}_t d\mathbf{W}_t, \ t \in [0,T],$$

with

$$\mathbf{C}_t = \begin{bmatrix} C_t^{(1,1)} & \cdots & C_t^{(1,m)} \\ \vdots & \ddots & \vdots \\ C_t^{(n,1)} & \cdots & C_t^{(n,m)} \end{bmatrix}, \ \mathbf{B}_t = \begin{bmatrix} B_t^{(1)} \\ \vdots \\ B_t^{(n)} \end{bmatrix}.$$

An Itô process thus results from a linear combination of a Lesbesgue and a stochastic integral.

The Itô's lemma for n-dimensional Itô processes is next presented. A proof can be found in [3].

Lemma 1. *Let $(\mathbf{X}_t, t \in [0,T])$ be an n-dimensional Itô process with stochastic differential $d\mathbf{X}_t = \mathbf{B}_t dt + \mathbf{C}_t d\mathbf{W}_t, \ t \in [0,T]$. Let also $f(t,\mathbf{x})$, $(t,\mathbf{x}) \in [0,T] \times \mathbb{R}^n$, be a continuous real function with continuous partial derivatives $\frac{\partial f}{\partial t}$, $\frac{\partial f}{\partial x_i}$ and $\frac{\partial^2 f}{\partial x_i x_j}$, for $i,j = 1,...,n$. Then $f(t,\mathbf{X}_t)$ is also an Itô process satisfying the* **Itô's formula***:*

$$f(t,\mathbf{X}_t) = f(0,\mathbf{X}_0) +$$
$$\int_0^t \left[\frac{\partial f}{\partial t}(s,\mathbf{X}_s) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(s,\mathbf{X}_s) B_s^{(i)} + \sum_{k=1}^m \sum_{i,j=1}^n \frac{1}{2} \frac{\partial^2 f}{\partial x_i x_j}(s,\mathbf{X}_s) C_s^{(i,k)} C_s^{(j,k)} \right] ds +$$
$$\sum_{k=1}^m \sum_{i=1}^n \int_0^t \frac{\partial f}{\partial x_i}(s,\mathbf{X}_s) C_s^{(i,k)} dW_s^{(k)} \ a.s.. \ (2.3)$$

In the scalar case, the Itô's formula for $f(t,X_t)$, where $X = (X_t, t \in [0,T])$ is a scalar Itô process with stochastic differential $dX_t = B_t dt + C_t dW_t$, takes the form

$$f(t,X_t) = f(0,X_0) + \int_0^t \left[\frac{\partial f}{\partial t}(s,X_s) + \frac{\partial f}{\partial x}(s,X_s) B_s + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(s,X_s) C_s^2 \right] ds +$$
$$\int_0^t \frac{\partial f}{\partial x}(s,X_s) C_s dW_s \ a.s., \ t \in [0,T]. \ (2.4)$$

Notice that (2.4) corresponds to a stochastic chain rule. If f is linear in x, then $\frac{\partial^2 f}{\partial x^2} = 0$ and the Itô formula reduces to the usual chain rule. Itô's lemma is particularly useful in the analytical calculation of the solution of some specific stochastic differential equations, as we shall see later.

From (2.3) we can deduce the Itô product formula. Consider $f(x, y) = xy$. Let $X = (X_t, t \in [0, T])$ and $Y = (Y_t, t \in [0, T])$ be two scalar Itô processes with stochastic differentials $dX_t = A_t dt + B_t dW_t$ and $dY_t = C_t dt + D_t dW_t$. Therefore, $\frac{\partial f}{\partial x}(x, y) = y$, $\frac{\partial f}{\partial y}(x, y) = x$, $\frac{\partial^2 f}{\partial xy}(x, y) = 1$ and $\frac{\partial^2 f}{\partial x^2} = \frac{\partial^2 f}{\partial y^2} = 0$. Applying (2.3) to $f(X_t, Y_t)$, we obtain the *Itô product formula*,

$$X_t Y_t = X_0 Y_0 + \int_0^t \left(Y_s A_s + X_s C_s + B_s D_s \right) ds + \int_0^t \left(Y_s B_s + X_s D_s \right) dW_s \ a.s, \ t \in [0, T]$$

$$(2.5)$$

or, in the differential form,

$$dX_t Y_t = \left(Y_t A_t + X_t C_t + B_t D_t \right) dt + \left(Y_t B_t + X_t D_t \right) dW_t, \quad t \in [0, T].$$

2.2. Diffusion processes

In this section we begin by considering real-valued processes. At the end, appropriate generalizations shall then be made.

In order to develop stochastic models in biology, we first present a brief exposition on the general concept of Markov processes.

Markov processes constitute a widely studied class of stochastic processes with the property that future probabilities only depend on the most recent values of the process, i.e., the process "forgets" the past. The formal definition is presented below.

Definition 3. *Let $X = (X_t, \ t \in [0, T])$ be a real-valued stochastic process. X is called a **Markov process** if the following so-called Markov property is satisfied: for any $0 \leq t_1 < ... < t_n \leq s \leq T$, the equality*

$$P(X_s \leq y / X_{t_1}, ..., X_{t_n}) = P(X_s \leq y / X_{t_n}) \ a.s.$$

holds for all $y \in \mathbb{R}$.

We find in the literature (see e.g. [2] and [15]) various equivalent formulations of the Markov property, from which we highlight the following one: for any $0 \leq t \leq s \leq T$ and $B \in \mathcal{B}_\mathbb{R}$,

$$P(X_s \in B / G_t) = P(X_s \in B / X_t) \ a.s.,$$

where $G_t := \sigma(X_s, s \in [0, t])$, i.e., G_t is the σ-algebra generated by the variables X_s.

These processes are used to model random systems that change according to a transition rule that only depends on the current state. Their finite distributions

are, in fact, characterized by the distribution of X_0 and the family of transition probability functions,

$$\mathcal{P} = \{P(t, x, s, B),\ 0 \le t \le s \le T,\ x \in \mathbb{R}\ B \in \mathcal{B}_{\mathbb{R}}\},$$

where the function P has the following properties:

(i) for fixed $t \le s$ and $B \in \mathcal{B}_{\mathbb{R}}$, we have

$$P(t, X, s, B) = P(X_s \in / X_t)\ a.s.,$$

(ii) for fixed $t \le s$ and $x \in \mathbb{R}$, $P(t, x, s, \cdot)$ is a probability on $\mathcal{B}_{\mathbb{R}}$,

(iii) for fixed $t \le s$ and $B \in \mathcal{B}_{\mathbb{R}}$ $P(t, x, s, \cdot)$ is adapted to $\mathcal{B}_{\mathbb{R}}$,

(iv) for $t \le u \le s$ and for all $B \in \mathcal{B}_{\mathbb{R}}$, (with a possible exception of a set $A \in \mathcal{B}_{\mathbb{R}}$ such that $P(X_t \in A) = 0$), the so-called **Chapman-Kolmogorov equation**

$$P(t, x, s, B) = \int_{\mathbb{R}} P(u, y, s, B) P(t, x, u, dy), \tag{2.6}$$

(v) for all $t \in [0, T]$ and $B \in \mathcal{B}_{\mathbb{R}}$ $P(t, x, t, B) = \mathbb{1}_B(x)$.

We shall use the notation $P(t, x, s, B) = P(X_s \in B / X_t \in x) = P_{t,x}(s, B)$.[3] For instance, if X_s represents the position of an object at time s, then $P_{t,x}(s, B)$ represents the probability that this object, being at x at time t, will be in B at time s.

If, for all t and s, $0 \le t < s \le T$, and $x \in \mathbb{R}$, the transition probability function $P_{t,x}(s, \cdot)$ is absolutely continuous, then it admits a density $f_{t,x}(s, \cdot)$, which is called a *transition density function* of X. In this case, the Chapman-Kolmogorov equation takes the form,

$$f_{t,x}(s, y) = \int_{\mathbb{R}} f_{u,z}(s, y) f_{t,x}(u, z) dz,$$

for $0 \le t < u < s \le T$.

One classical example of a Markov process is the Wiener process, which describes the behavior (with friction neglected) of the well-known *Brownian motion*, the chaotic motion of a grain of pollen on a water surface, first noticed by the British botanist Brown in 1826. For further details on the modellation of the Brownian motion using the Wiener process (see [16]). A first rigorous proof of its (mathematical) existence was given by Norbert Wiener in 1923 (see [1]).

---

[3]Note that the number $P(X_s \in B / X_t = x)$ is well defined, even though $P(X_t = x) = 0$.

We introduce now a particular subclass of Markov processes that represents a form of conditional probabilistic equilibrium, in the sense that the transition probabilities do not depend explicitly on time, they are stationary. Formally speaking, we have the following definition.

Definition 4. *Let X be a Markov process with transition probabilities $\mathcal{P} = \{P_{t,x}, \ x \in \mathbb{R}, \ t \in [0, T]\}$. X is called* **time-homogeneous** *if*

$$P_{t+u,x}(s + u, B) = P_{t,x}(s, B),$$

for all $0 \leq t \leq s \leq T$ and $0 \leq t + u \leq s + u \leq T$.

In this case, P is thus a function of x, $s - t$ and B. Hence, we can write, $P_{t,x}(s, B) = P_{0,x}(s - t, B)$, $0 \leq s - t \leq T$.

In natural systems, homogeneous Markov processes are quite common (see [17]). Some illustrations will be given in Chapter 3. The Wiener process is an example of a time-homogeneous Markov process.

In 1931 Kolmogorov found out that there are two types of time-continuous Markov processes: diffusion and jump processes, which differ on the behavior over small time intervals. Roughly speaking, suppose that in a small time interval there is an overwhelming probability that the state will remain unchanged, however if it changes, the change can be either radical or small, leading to a jump or a diffusion process, respectively. Kolmogorov then derived the so-called forward and backward equations for each kind of process referred before.

Our interest will remain on the Kolmogorov forward equation for diffusion processes, also known as Fokker-Plank equation, since this equation will be crucial for the deduction of mathematical models, as we shall see.

We present in the following the definition of diffusion process.

Definition 5. *Let $X = (X_t, t \in [0, T])$ be a real-valued Markov process with almost certainly continuous paths. X is said to be a **diffusion process** if there exist functions $m : [0, T] \times \mathbb{R} \to \mathbb{R}$ and $Q : [0, T] \times \mathbb{R} \to \mathbb{R}^+$, such that[4],*

   (i) $P(|X_{t+h} - X_t| \geq c/X_t = x) = o(h),$

   (ii) $\int_{|y-x|<c}(y - x)P_{t,x}(t + h, dy) = m(t, x)h + o(h),$

   (iii) $\int_{|y-x|<c}(y - x)^2 P_{t,x}(t + h, dy) = Q(t, x)h + o(h),$

---

[4]As usually, o denotes a function of h such that $\lim_{h \to 0} \frac{o(h)}{h} = 0$.

for all $c \in \mathbb{R}^+$ and $(t, x) \in [0, T] \times \mathbb{R}$.

*The functions m and Q are called the **drift** and **diffusion** coefficients of the diffusion process X, respectively.*

Remark 1. *In the literature, one can find different a more intuitive definition of diffusion process (see [18]), where the condition (i) is omitted, the following condition is added:*

(i)' *for some $\delta > 0$, $E\left(|X_{t+h} - X_t|^{2+\delta} / X_t = x\right) = o(h)$,*

and the truncated moments in conditions (ii) and (iii) are replaced by the non-truncated moments, namely

(ii)' $E(X_{t+h} - X_t / X_t = x) = m(t, x)h + o(h)$,

(iii)' $E\left((X_{t+h} - X_t)^2 / X_t = x\right) = Q(t, x)h + o(h)$.

These conditions are stronger, once they require the process to have finite condicional moments.

To prove that $(i)'$, $(ii)'$ and $(iii)'$ imply (i), (ii) and (iii) it is enough to show that

$$\int_{|y-x| \geq c} |y - x|^k P_{t,x}(t + h, dy) = o(h),$$

for $k = 0, 1, 2$, which is a consequence of $(i)'$ and the inequality

$$\int_{|y-x| \geq c} |y - x|^k P_{t,x}(t + h, dy) \leq \frac{1}{c^{2+\delta-k}} \int_{|y-x| \geq c} |y - x|^{2+\delta} P_{t,x}(t + h, dy).$$

Notice that condition (i) says that large displacements are very unlikely in small time intervals. This fact can be viewed as a formalization of the process sample paths' continuity. A diffusion process is thus a Markov process with no instantaneous jumps. Moreover, the functions m and Q correspond to the infinitesimal mean and variance of X, respectively, since m represents the instantaneous rate of change in the conditional mean of the process and Q measures the average size of instantaneous conditional fluctuations of the diffusion process.

The standard scalar Wiener process and the Ornstein-Uhlenbeck process, defined by $X_t = g(t)W_{f(t)}$, $t \in \mathbb{R}_0^+$, with $g(t) = e^{-\alpha t}$ and $f(t) = \frac{\sigma^2 e^{2\alpha t}}{2\alpha}$, α, σ^2 positive constants, are basic examples of diffusion processes. To prove this, note that W is a Markov process, as it was referred above, and, consequently, X is a Markov process too. Furthermore, for all $t, t + h \in [0, T]$, we know that $W_{t+h} - W_t \sim \mathcal{N}(0, h)$, by definition, which together with the independence of increments implies that

$$E(W_{t+h} - W_t) = 0, \quad E((W_{t+h} - W_t)^2) = h,$$

showing that W is a diffusion process with drift $m(t,x) = 0$ and diffusion coefficient $Q(t,x) = 1$.

It follows that for the Ornstein-Uhlenbeck process, the conditional distribution is also gaussian with mean $E(X_{t+h}/X_t = x) = xe^{-\alpha h}$ and variance $V(X_{t+h}/X_t = x) = \left(\frac{1-e^{-2\alpha h}}{2\alpha}\right)\sigma^2$. Therefore, X is a diffusion process with coefficients $m(t,x) = -\alpha x$ and $Q(t,x) = \sigma^2$.

This process arose as an improvement of the Wiener process model for the diffusion phenomenon, which has into account the friction effect inherent to the movement of the particles of pollen on a water surface.

We are now ready to deduce the Kolmogorov forward equation for diffusion processes. As we will see in the next theorem, the term forward is motivated by the fact that the time variable s moves forward from t, contrarily to the backward equation.

Theorem 1. *Let X be a real-valued diffusion process with drift m and diffusion coefficient Q, such that the limits (in definition 5) are uniform with respect to x. Moreover, assume that the transition density functions of X, $f_{t,x}$, $(t,x) \in [0,T] \times \mathbb{R}$, exist and that the derivatives*

$$\frac{\partial f_{t,x}}{\partial s}, \quad \frac{\partial (mf)_{t,x}}{\partial y}, \quad \frac{\partial (Qf_{t,x})}{\partial y} \quad and \quad \frac{\partial^2 (Qf_{t,x})}{\partial y^2}$$

exist and are continuous functions on $[0,T] \times \mathbb{R}$. Then, for each $(t,x) \in [0,T] \times \mathbb{R}$, $f_{t,x}$ satisfies the **Kolmogorov forward equation:**

$$\frac{\partial}{\partial s} f_{t,x} = -\frac{\partial}{\partial y}(mf_{t,x}) + \frac{1}{2}\frac{\partial^2}{\partial y^2}(Qf_{t,x}), \tag{2.7}$$

evaluated at $(s,y) \in (t,T] \times \mathbb{R}$, with the initial condition[5]

$$lim_{s \to t^+} f_{t,x}(s,y) = \delta_x(y). \tag{2.8}$$

Proof. Let $t \in [0,T)$ and $x \in \mathbb{R}$ be fixed and consider $f_{t,x}$ as a function of $s \in (t,T]$ and $y \in \mathbb{R}$. For any smooth real-valued function[6], ϕ, such that ϕ has compact support C, we define the function θ as

$$\theta(s) := \int_{\mathbb{R}} \phi(z) f_{t,x}(s,z)dz = \int_C \phi(z) f_{t,x}(s,z)dz, \ s \in (t,T]. \tag{2.9}$$

Notice that ϕ', ϕ'' and ϕ''' vanish outside C. For $h \in \mathbb{R}^+$, we may write, by the Chapman-Kolmogorov equation,

$$f_{t,x}(s+h,z) = \int_{\mathbb{R}} f_{s,y}(s+h,z) f_{t,x}(s,y)dy,$$

---

[5]δ_x denotes the Dirac delta function at x.

[6]For example, for any $a,b,c \in \mathbb{R}$, with $a < b$ and $c > 0$, define ϕ as

$\phi(z) = (z-a)^3 (b-z)^3 e^{-cz^2} \mathbb{1}_{[a,b]}(z)$.

which implies, from (2.9), that

$$\frac{\theta(s+h) - \theta(s)}{h} = \frac{1}{h}\int_{\mathbb{R}}\phi(z)\left[\int_{\mathbb{R}}f_{s,y}(s+h,z)f_{t,x}(s,y)dy\right]dz - \frac{1}{h}\int_{\mathbb{R}}\phi(y)f_{t,x}(s,y)dy.$$

Using Fubini's theorem and rearranging terms, we obtain

$$\frac{\theta(s+h) - \theta(s)}{h} = \frac{1}{h}\int_{\mathbb{R}}\left(\int_{\mathbb{R}}(\phi(z) - \phi(y))f_{s,y}(s+h,z)dz\right)f_{t,x}(s,y)dy. \qquad (2.10)$$

For any positive constant c, the integral with respect to the variable z can be written as a sum of two integrals as follows

$$\int_{|y-z|<c}(\phi(z) - \phi(y))f_{s,y}(s+h,z)dz + \int_{|y-z|\geq c}(\phi(z) - \phi(y))f_{s,y}(s+h,z)dz.$$

Once ϕ is bounded, from definition 5 the second term of the previous expression is an $o(h)$ uniformly with respect to y, by hypothesis. Moreover, considering $c > 0$ sufficiently small, we may use the Taylor expansion of ϕ at y with Lagrange reminder to get for the first term

$$\int_{|z-y|<c}\left(\phi'(y)(z-y) + \frac{1}{2}\phi''(y)(z-y)^2 + \frac{1}{6}\phi'''(\xi)(z-y)^3\right)f_{s,y}(s+h,z)dz, \quad (2.11)$$

for some real number ξ between y and z. Notice that if we divide (2.11) by h, and take the limit as h approaches zero, we obtain (see definition 5)

$$\phi'(y)m(s,y) + \frac{1}{2}\phi''(y)Q(s,y) + \lim_{h\to 0^+}\frac{1}{h}\frac{1}{6}\phi'''(\xi)\int_{|z-y|<c}(z-y)^3 f_{s,y}(s+h,z)dz.$$

By the Lebesgue's dominated convergence theorem, the limit of (2.10) as h approaches zero conduces to the expression

$$
\begin{aligned}
\theta'(s) &= \int_{\mathbb{R}}\left(\phi'(y)m(s,y) + \frac{1}{2}\phi''(y)Q(s,y)\right)f_{t,x}(s,y)dy \\
&+ \lim_{h\to 0^+}\frac{1}{h}\frac{1}{6}\int_{\mathbb{R}}\phi'''(\xi)\int_{|z-y|<c}(z-y)^3 f_{s,y}(s+h,z)dz f_{t,x}(s,y)dy.
\end{aligned}
$$

The second term of the previous expression turns out to be an $O(c)$ (recall that ϕ''' is null outside C), therefore it must be null. Once ϕ' and ϕ'' have support C, it follows that

$$\theta'(s) = \int_C\left(\phi'(y)m(s,y) + \frac{1}{2}\phi''(y)Q(s,y)\right)f_{t,x}(s,y)dy. \qquad (2.12)$$

The integration by parts leads to

$$\theta'(s) = \int_C\left[-\phi(y)\frac{\partial}{\partial y}(m(s,y)f_{t,x}(s,y)) + \frac{1}{2}\phi(y)\frac{\partial^2}{\partial y^2}(Q(s,y)f_{t,x}(s,y))\right]dy.$$

On the other hand, from (2.9), $\theta'(s)$ is obtained by bringing the derivative inside the integral

$$\theta'(s) = \int_C \phi(y) \left(\frac{\partial}{\partial s} f_{t,x}(s,y) \right) dy.$$

The equality of the last two integrals for all functions ϕ specified above implies that the integrand functions are equal too, which ends the proof. $\qquad \square$

Notice that, whenever the solution is unique, the Kolmogorov forward equation represents a useful tool to obtain the probabilistic characterization of a diffusion process X satisfying the conditions of the previous theorem, because it reduces the probabilistic problem of finding the transition density functions of the diffusion process to a deterministic problem of solving a PDE.

The simplest example is again given by the standard scalar Wiener process. Recall that the drift and diffusion coefficients of this process are 0 and 1, respectively. The Kolmogorov forward equation for the transition probability from x at time t to y at time s of this process is then

$$\frac{\partial}{\partial t} f_{t,x}(s,y) = \frac{1}{2} \frac{\partial^2}{\partial y^2} f_{t,x}(s,y), \;\; 0 \le t < s \le T, \; y \in \mathbb{R} \qquad (2.13)$$

$$\lim_{s \to t^+} f_{t,x}(s,y) = \delta_x(y).$$

This equation belongs to the class of diffusion equations (also called heat equations) which consist of well-known deterministic models for describing diffusion phenomena. In fact, their solutions represent the density of Brownian particles at point y and at time s, such that all particles were concentrated at a point x at time t. On the other hand, in the context of SDEs, the solution of this equation represents the probability of a particle being in a position y at time s, if it was in a position x at time t. In fact, the solution to (2.13) is given by

$$f_{t,x}(s,y) = \frac{1}{\sqrt{2\pi(s-t)}} exp \left(-\frac{(y-x)^2}{2(s-t)} \right), \; 0 \le t < s \le T, \; y \in \mathbb{R},$$

which corresponds to the probability density function of the Gaussian distribution with mean value x and variance $s-t$. We now understand why both models are used to model Brownian motion, following different interpretations, and that the Wiener process represents a basic example of a diffusion process.

The existence and uniqueness of the solution of (2.7)-(2.8), which is a partial differential equation of parabolic type, is ensured under certain conditions on the coefficients m and Q, described in the next theorem.

Theorem 2. *Let m, Q and the partial derivatives $\frac{\partial m}{\partial x}$, $\frac{\partial Q}{\partial x}$ and $\frac{\partial^2 Q}{\partial x^2}$ be continuous functions defined on $[0,T] \times \mathbb{R}$ and satisfying the Lipschitz and the linear growth conditions in x, namely, $\exists k_1, k_2 > 0 : \forall t \in [0,T], \ x, \ y \in \mathbb{R}$,*

Lipschitz condition $|m(t,x) - m(t,y)| + |Q(t,x) - Q(t,y)| \leq k_1 |x - y|$,

Growth condition $|m(t,x)|^2 + |Q(t,x)|^2 \leq k_2(1 + x^2)$.

Additionally, assume that there exists a positive constant c such that $Q \geq c$. Then, for each $x \in \mathbb{R}$ and $t \in [0,T)$, (2.7)- (2.8)) has exactly one solution.

A proof of this theorem can be found in [19].

Before proceeding to the introduction of systems of stochastic differential equations, we need to see how to generalize the previous concepts and results for vector-valued stochastic processes. Let $\mathbf{X} = (\mathbf{X}_t, t \in [0,T])$ be a $\mathbb{R}^n$-valued stochastic process. The definitions of Markov process, transition probability function and time-homogeneous Markov process are identical. The definition of diffusion process is similar too, but $\mathbf{m}$ is now an n-dimensional vector, $\mathbf{Q}$ a symmetric and positive definite $n \times m$-dimensional matrix, in (iii) and $(iii)'$ the square is replaced by the product $(\mathbf{X}_{t+h} - \mathbf{X}_t)(\mathbf{X}_{t+h} - \mathbf{X}_t)^T$. The Kolmogorov forward equation takes the form

$$\frac{\partial}{\partial s} f_{t,\mathbf{x}} = -\sum_{i=1}^{n} \frac{\partial}{\partial y_i}(m_i f_{t,\mathbf{x}}) + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2}{\partial y_i \partial y_j}(Q_{ij} f_{t,\mathbf{x}}) \tag{2.14}$$

evaluated at $(s, \mathbf{y}) \in (t, T] \times \mathbb{R}^n$, with the initial condition

$$\lim_{s \to t^+} f_{t,\mathbf{x}}(s, \mathbf{y}) = \delta_{\mathbf{x}}(\mathbf{y}). \tag{2.15}$$

In theorem 2 the inequality for Q takes the form

$$\forall \mathbf{v} \in \mathbb{R}^n, \ \mathbf{v}^T \mathbf{Q} \mathbf{v} \geq c|\mathbf{v}|^2,$$

for some positive constant c and the Lipschitz and growth conditions are now given by

$$|\mathbf{m}(t,\mathbf{x}) - \mathbf{m}(t,\mathbf{y})| + \|\mathbf{Q}(t,\mathbf{x}) - \mathbf{Q}(t,\mathbf{y})\| \leq k_1 |\mathbf{x} - \mathbf{y}|,$$

$$|\mathbf{m}(t,\mathbf{x})|^2 + \|\mathbf{Q}(t,\mathbf{x})\|^2 \leq k_2(1 + |\mathbf{x}|^2).$$

respectively.

Finally, we stress that if we consider an unbounded time domain, i.e., if $t \in [0,\infty)$, all results remain valid (see [20]) under the following modification in the definition of the space $L^i([0,\infty))$, $i = 1, 2$: the third condition must be replaced by

$$\forall t \in [0,\infty) \int_0^t |f|^i dt < \infty \ a.s., \ i = 1, 2.$$

2.3. Systems of stochastic differential equations

Stochastic differential equations are becoming increasingly important in modelling biological phenomena, by their capacity of including random aspects, which are intrinsic to real situations. Population dynamics, oncogenesis, epidemic and genetics are some examples where SDE models can constitute popular generalizations of ordinary differential equation models (see [17], [5], [4] and [21]).

We shall now briefly introduce stochastic differential equations and, subsequently, systems of stochastic differential equations, as they will be our main tool for establishing the models in the next chapter.

Let $X = (X_t, t \in [0, T])$ be a real-valued stochastic process and ρ and σ be real functions defined on $[0, T] \times \mathbb{R}$ and measurable with respect to $\mathcal{B}_{[0,T]} \otimes \mathcal{B}_{\mathbb{R}}$. A stochastic differential equation, with initial condition $X_0 = \xi$ a.s., $\mathcal{F}_0$-measurable, is in fact an integral equation involving a Lebesgue and a stochastic integral. Symbolically, it is of the form

$$dX_t = \rho(t, X_t)dt + \sigma(t, X_t)dW_t, \ t \in [0, T],$$

which should be interpreted, mathematically, as

$$X_t = \xi + \int_0^t \rho(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s \ a.s., \ t \in [0, T]$$

where ρ and σ are the coefficients of the equation and must satisfy appropriate conditions in order to ensure the existence of these integrals. The notation used to represent SDEs pretends to make the analogy with classic differential calculus.

Likewise, let $(W_t^{(j)}, \ t \in [0, T])$, $j = 1, ..., m$, $m \in \mathbb{N}$, be m independent Wiener processes and, for $i = 1, ..., n$ and $j = 1, ..., m$, let $\sigma_{i,j}$ and ρ_i be real functions on $[0, T] \times \mathbb{R}^n$, which we assume to be measurable regarding $\mathcal{B}_{[0,T]} \otimes \mathcal{B}_{\mathbb{R}^n}$. A system of n stochastic differential equations and n variables with the initial condition $X_0^{(i)} = \xi^{(i)}$ a.s., $i = 1, ..., n$, $\mathcal{F}_0$-measurable, is naturally given by

$$dX_t^{(i)} = \rho_i\left(t, X_t^{(1)}, \ldots, X_t^{(n)}\right)dt + \sum_{j=1}^m \sigma_{ij}\left(t, X_t^{(1)}, \ldots, X_t^{(n)}\right)dW_t^{(j)}, \ 1 \le i \le n,$$

which should be interpreted as

$$X_t^{(i)} = \xi_i + \int_0^t \rho_i\left(s, X_s^{(1)}, \ldots, X_s^{(n)}\right)ds + \sum_{j=1}^m \int_0^t \sigma_{ij}\left(s, X_s^{(1)}, \ldots, X_s^{(n)}\right)dW_s^{(j)} \ a.s., \ i = 1, ..., n.$$

Using matrix notation, this system can be rewritten as the matrix SDE

$$d\mathbf{X}_t = \boldsymbol{\rho}(t, \mathbf{X}_t)dt + \boldsymbol{\sigma}(t, \mathbf{X}_t)d\mathbf{W}_t, \ t \in [0, T], \quad \mathbf{X}_0 = \boldsymbol{\xi} \ a.s. \tag{2.16}$$

with

$$\mathbf{X}_t = \begin{bmatrix} X_t^{(1)} \\ \vdots \\ X_t^{(n)} \end{bmatrix}, \ \boldsymbol{\sigma} = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1m} \\ \vdots & \ddots & \vdots \\ \sigma_{n1} & \cdots & \sigma_{nm} \end{bmatrix}, \ \mathbf{W}_t = \begin{bmatrix} W_t^{(1)} \\ \vdots \\ W_t^{(m)} \end{bmatrix}, \ \boldsymbol{\rho} = \begin{bmatrix} \rho_1 \\ \vdots \\ \rho_n \end{bmatrix}, \ \boldsymbol{\xi} = \begin{bmatrix} \xi^{(1)} \\ \vdots \\ \xi^{(n)} \end{bmatrix},$$

where $\boldsymbol{\rho}$ and $\boldsymbol{\sigma}$ are the coefficients of the equation and, again, must satisfy appropriate conditions in order to ensure the existence of these integrals.

If $\boldsymbol{\rho}$ and $\boldsymbol{\sigma}$ are linear with respect to the second variable, we say (2.16) is a *system of linear stochastic differential equations.* It is important to notice that in the linear regime it is possible to find the exact solution of the system using analytical procedures (see section 2.3.1). However, in the general setting, there is no way of solving such equations, except in some particular cases, so that we must resort to numerical approximations, using the Euler-Maruyama or Mildstein methods, for instance.(see [22] and [3])

Since we are interested in formulating mathematical models in biology, we want to ensure certain properties on the system so that our model is consistent with the reality we want to model, namely, the existence and uniqueness of the solution and the continuous dependency on initial conditions. A problem satisfying these three properties is said to be *well-posed* in the sense of Hadamard. Notice that in biological models the initial conditions must frequently be estimated or measured with a certain degree of precision. For this reason, it is important to be sure that small perturbations on the initial conditions will not lead to very different solutions.

The well-posedness is also important for numerical simulations, as it ensures that an accurate solution can be found when using a stable algorithm.

Before checking these properties, we shall present the concept of strong solution of a system of SDEs.

Definition 6. *A $\mathbb{R}^n$-valued stochastic process $\boldsymbol{X} = (\boldsymbol{X}_t, t \in [0, T])$, with $\boldsymbol{X}_t = (X_t^{(1)}, ..., X_t^{(n)})$, is called a **strong solution** to the matrix SDE (2.16) if*

- *almost all sample paths of $\mathbf{X}$ are continuous functions,*

- *$\mathbf{X}$ is adapted to the filtration $(\mathcal{F}_t)_{t \in [0, T]}$,*

- *$(\boldsymbol{\rho}(t, \mathbf{X}_t), t \in [0, T]) \in L^1(\Omega \times [0, T])$,*

- $(\boldsymbol{\sigma}(t, \mathbf{X}_t), t \in [0, T]) \in L^2(\Omega \times [0, T])$,

- $\mathbf{X}$ *satisfies* $\mathbf{X}_0 = \boldsymbol{\xi}$ *a.s. and*

$$\mathbf{X}_t = \mathbf{X}_0 + \int_0^t \boldsymbol{\rho}(s, \mathbf{X}_s)ds + \int_0^t \boldsymbol{\sigma}(s, \mathbf{X}_s)d\mathbf{W}_s \ a.s., \ t \in [0, T].$$

The literature refers two types of solutions: strong and weak solutions. In the later case, the integrator Wiener process is not previously specified, making part of the solution. A weak solution consists thus in a pair $(\mathbf{X}_t, \ t \in [0, T])$ and $(\bar{\mathbf{W}}_t, \ t \in [0, T])$ such that

$$\mathbf{X}_t = \mathbf{X}_0 + \int_0^t \boldsymbol{\rho}(s, \mathbf{X}_s)ds + \int_0^t \boldsymbol{\sigma}(s, \mathbf{X}_s)d\bar{\mathbf{W}}_s \ a.s., \ t \in [0, T].$$

The next result assures the existence and uniqueness of a strong solution to (2.16) under appropriate conditions. As in the case of deterministic PDE's (see theorem 2), we shall require the Lipschitz continuity of the coefficients in $\mathbf{x}$, so that the solution exists and is unique in a finite time interval, as well as a growing condition, in order to safeguard the boundedness of the solution for all time $t > 0$, when we are working in $[0, +\infty)$. A detailed proof can be found in [14], for example.

Theorem 3. *Let $\boldsymbol{\xi}$ be a $\mathcal{F}_0$-measurable random vector such that $E(|\boldsymbol{\xi}|^2) < \infty$. Let $\boldsymbol{\rho}$ and $\boldsymbol{\sigma}$ be functions satisfying the Lipschitz and growth conditions in $\mathbf{x}$. Then, the stochastic differential equation (2.16) has exactly one strong solution, $\mathbf{X}$.*

The solution is unique in the sense of pathwise uniqueness, i.e, if two processes $\mathbf{X}$ and $\mathbf{Y}$ solve (2.16) in the conditions of theorem 3, then almost all their sample paths coincide. In this case, we say that the processes $\mathbf{X}$ and $\mathbf{Y}$ are indistinguishable.

It is possible to guarantee the existence of a unique solution under weaker hypothesis on the coefficients, such as local Lipschitz continuity, instead of global Lipschitz continuity (see [23]). By the mean-value theorem, any continuously differentiable function satisfies this condition. The growing condition is also not essential, but its absence may allow the unbounded growth of the sample paths within a finite time interval. Hence, we encounter in this context a similar behaviour as in ODE's context.

The strong solution to an SDE, $\mathbf{X}$, satisfies three properties that we shall address next. The most important property of the strong solution of an SDE states that, under certain conditions, $(\mathbf{X}_t, \ t \in [0, T])$ is a Markov process and, moreover, it is a diffusion process. As a consequence of this result, we can apply the powerful

analytical tools that have been developed for diffusion processes to the solutions of stochastic differential equations, which will assume a crucial role in the next chapter. In order to establish this property, we need the following auxiliary properties.

Lemma 2. *Let $\boldsymbol{\rho}$ and $\boldsymbol{\sigma}$ be functions as described in theorem 3, with Lipschitz constant k_1 and growth constant k_2. Let $\boldsymbol{\xi}$ be a $\mathcal{F}_0$-measurable random vector, such that $E(|\boldsymbol{\xi}|^{2n}) < \infty$, for some positive integer n. Then the solution $(\mathbf{X}_t, t \in [0, T])$ to the system (2.16) satisfies the inequalities:*

$$
\begin{aligned}
E(|\mathbf{X}_t|^{2n}) &\leq \left(1 + E(|\boldsymbol{\xi}|^{2n})\right) e^{at} \\
E(|\mathbf{X}_t - \boldsymbol{\xi}|^{2n}) &\leq b\left(1 + E(|\boldsymbol{\xi}|^{2n})\right) t^n e^{at}, \ t \in (0, T]
\end{aligned}
\tag{2.17}
$$

with $a = 2n(n+1)max(k_1, k_2)$ and b is a positive constant depending only on k_1, k_2 and T.

A proof of this lemma can be found in [6].

Lemma 3. *Let $\mathcal{H}$ be a sub-σ-algebra of $\mathcal{A}$ and $\mathbf{X}$ and $\mathbf{Y}$ $\mathbb{R}^n-$valued random vectors defined on $(\Omega, \mathcal{A}, P)$ such that $\mathbf{X}$ is $\mathcal{H}$-measurable and $\mathbf{Y}$ is independent of $\mathcal{G}$. Then, for all measurable and limited functions $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, we have*

$$
E(g(\mathbf{X}, \mathbf{Y}/\mathcal{G})) = \phi(\mathbf{X}) \ a.s.,
$$

with $\phi(\mathbf{x}) = E[g(\mathbf{x}, \mathbf{y})], \ \mathbf{x} \in \mathbb{R}^n$.

Proof. Due to $\phi(\mathbf{x}) = \int_{\mathbb{R}} g(\mathbf{x}, \mathbf{y})dP_Y(\mathbf{y}), \ \mathbf{x} \in \mathbb{R}^n$, this function is limited and measurable. Hence, for all random variable Z, $\mathcal{H}$-measurable and limited, we have, by Fubini's theorem,

$$
\begin{aligned}
E\left[g(\mathbf{X}, \mathbf{Y})Z\right] &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^{n+1}} g(\mathbf{x}, \mathbf{y})z dP_{(\mathbf{X}, \mathbf{Z})}(\mathbf{x}, \mathbf{z}) dP_{\mathbf{Y}}(\mathbf{y}) \\
&= \int_{\mathbb{R}^{n+1}} \left(\int_{\mathbb{R}^n} g(\mathbf{x}, \mathbf{y})dP_{\mathbf{Y}}(\mathbf{y}) \right) z dP_{(\mathbf{X}, Z)}(\mathbf{x}, z) \\
&= \int_{\mathbb{R}^{n+1}} \phi(\mathbf{x})z dP_{(\mathbf{X}, Z)}(s, z) \\
&= E(\phi(\mathbf{X})Z)
\end{aligned}
\tag{2.18}
$$

which, by definition of conditional probability, ends the proof. $\square$

Theorem 4. *Let $\boldsymbol{\rho}$ and $\boldsymbol{\sigma}$ be functions as described in theorem 3, and assume that they are continuous on $[0, T] \times \mathbb{R}^n$. Let $\boldsymbol{\xi}$ be a $\mathcal{F}_0$-measurable random vector such that $E(|\boldsymbol{\xi}|^4) < \infty$. Then the unique solution $\mathbf{X}$ to the system*

$$
d\boldsymbol{X}_t = \boldsymbol{\rho}(t, \boldsymbol{X}_t)dt + \boldsymbol{\sigma}(t, \boldsymbol{X}_t)d\boldsymbol{W}_t, \ t \in [0, T], \ \mathbf{X}_0 = \boldsymbol{\xi} \ a.s.
\tag{2.19}
$$

is a diffusion process with diffusion coefficient $\mathbf{Q} = \boldsymbol{\sigma}\boldsymbol{\sigma}^T$ *and drift* $\mathbf{m} = \boldsymbol{\rho}$.

Proof. We begin by noticing that under the conditions of this theorem, the processes $\boldsymbol{\rho}(t, \mathbf{X}_t)$ and $\boldsymbol{\sigma}(t, \mathbf{X}_t)$ satisfy

$$\int_0^t E\left(|\boldsymbol{\rho}(u, \mathbf{X}_u)|^2\right) du < \infty \tag{2.20}$$

and

$$\int_0^t E\left(|\boldsymbol{\sigma}(u, \mathbf{X}_u)|^2\right) du < \infty, \tag{2.21}$$

because the linear growth condition implies

$$\int_0^t E\left(|\boldsymbol{\rho}(u, \mathbf{X}_u)|^2\right) du \le \int_0^t E\left(|k_1(1 + \mathbf{X}_u^2)|\right) du$$

and

$$\int_0^t E\left(|\boldsymbol{\sigma}(u, \mathbf{X}_u)|^2\right) du \le \int_0^t E\left(|k_1(1 + \mathbf{X}_u^2)|\right) du$$

which is finite because

$$\int_0^t E(|\mathbf{X}_u|^2) du < \infty,$$

due to the upper bound of $E(|\mathbf{X}_t|^2)$ presented in lemma 2.

We proceed to prove that the only solution $\mathbf{X}$ to (2.16) is a Markov process with initial distribution $P_{\boldsymbol{\xi}}$ and transition probabilities given by

$$\forall B \in \mathcal{B}_{\mathbb{R}^n}, \ \forall \mathbf{x} \in \mathbb{R}^n, \forall s, t \in [0, T], s \le t, \ P(\mathbf{X}_t \in B / \mathbf{X}_s = \mathbf{x}) = P(\mathbf{X}_t^{s\mathbf{x}} \in B),$$

where $(\mathbf{X}_t^{s\mathbf{x}}, \ t \in [s, T])$ is the solution to the equation

$$\mathbf{X}_t^{s\mathbf{x}} = \mathbf{x} + \int_s^t \boldsymbol{\rho}(u, \mathbf{X}_u^{s\mathbf{x}}) du + \int_s^t \boldsymbol{\sigma}(u, \mathbf{X}_u^{s\mathbf{x}}) d\mathbf{W}_u \ a.s., \ t \in [s, T], \tag{2.22}$$

with the initial condition $\mathbf{X}_s = \mathbf{x} \in \mathbb{R}^n$ a.s..

Let $0 \le t_1 < ... < t_n < t \le T$ be arbitrarily fixed. Since $\mathbf{X}_{t_i}, \ i = 1, ..., n$ are measurable regarding $\mathcal{F}_{t_n}$, we have $\boldsymbol{\sigma}(\mathbf{X}_{t_1}, ..., \mathbf{X}_{t_n}) \subseteq \mathcal{F}_{t_n}$, therefore,

$$\forall B \in \mathcal{B}_{\mathbb{R}^n}, \ P(\mathbf{X}_t \in B / \mathbf{X}_{t_1}, ..., \mathbf{X}_{t_n}) = E(P(\mathbf{X}_t \in B / \mathcal{F}_{t_n}) / \mathbf{X}_{t_1}, ..., \mathbf{X}_{t_n}) \ a.s..$$

Consequently, the Markov property is valid if

$$P(\mathbf{X}_t \in B / \mathcal{F}_{t_n}) = P(\mathbf{X}_t \in B / \mathbf{X}_{t_n})$$

holds almost surely. To prove this, note that for all $t \in [t_n, T]$, $\mathbf{X}_t^{t_n \mathbf{x}}$ is independent os $\mathcal{F}_{t_n}$ since, by virtue of the initial value at time s being constant, this variable

is measurable with respect to the σ-algebra generated by $\mathbf{W}_t - \mathbf{W}_{t_n}$, for $t \geq t_n$. Therefore, $\mathbf{X}_t^{t_n \mathbf{x}}$, for $t \in [t_n, T]$, is independent of $\mathcal{F}_u$, $0 \leq u \leq t_n$, by the properties of the Wiener process. On the other hand, $(\mathbf{X}_t, t \in [0, T])$ also satisfies the equation

$$\mathbf{X}_t = \mathbf{X}_{t_n} + \int_{t_n}^t \boldsymbol{\rho}(u, \mathbf{X}_u) du + \int_{t_n}^t \boldsymbol{\sigma}(u, \mathbf{X}_u) d\mathbf{W}_u \ a.s., \ t \in [t_n, T].$$

Then, by the uniqueness of the solution, it follows that $\mathbf{X}_t = \mathbf{X}_t^{t_n \mathbf{X}_{t_n}}$, $t \in [t_n, T]$ a.s., and therefore, $\mathbb{1}_B \left(\mathbf{X}_t^{t_n \mathbf{X}_{t_n}} \right) = \mathbb{1}_B(\mathbf{X}_t)$, $t \in [t_n, T]$ a.s.. As $\mathbf{X}_{t_n}$ is $\mathcal{F}_{t_n}$-measurable, lemma 3 allows us to write

$$E \left(\mathbb{1}_B \left(\mathbf{X}_t^{t_n \mathbf{X}_{t_n}} \right) / \mathcal{F}_{t_n} \right) = \phi(\mathbf{X}_{t_n}) \ a.s., \ t \in [t_n, T]$$

with $\phi(x) = E \left(\mathbb{1}_B \left(\mathbf{X}_t^{t_n x} \right) \right)$, $x \in \mathbb{R}^n$, by a similar reasoning. Thus, we conclude that

$$P \left(\mathbf{X}_t \in B / \mathcal{F}_{t_n} \right) = P \left(\mathbb{1}_B \left(\mathbf{X}_t^{t_n \mathbf{X}_{t_n}} \in B / \mathcal{F}_{t_n} \right) \right) =$$

$$= P \left(\mathbb{1}_B \left(\mathbf{X}_t^{t \mathbf{X}_{t_n}} \in B / \mathbf{X}_{t_n} \right) \right) = P(\mathbf{X}_t \in B / \mathbf{X}_{t_n}) \ a.s..$$

Finally, we stress that

$$P(\mathbf{X}_t \in B / \mathbf{X}_{t_n} = \mathbf{x}) = \phi(\mathbf{x}) = P \left(\mathbf{X}_t^{t_n \mathbf{x}} \right),$$

which shows the the transition probabilities of the process are of the form referred above. We have thus proved that $\mathbf{X}$ is a Markov process.

We shall prove now that $\mathbf{X}$ satisfies the conditions (i), (ii) and (iii) in definition 5. Let $(\mathbf{X}_t^{s\mathbf{x}}, \ t \in [s, t])$ be the solution to (2.22) with the initial condition $\mathbf{X}_s = \mathbf{x}$, a.s.. Then, for any $0 \leq s < t \leq T$ and for $\mathbf{x} \in \mathbb{R}^n$, we have, by the inequality (2.17)

$$\begin{aligned} E(|\mathbf{X}_t - \mathbf{X}_s|^4 / \mathbf{X}_s = \mathbf{x}) &= E(|\mathbf{X}_t^{sx} - \mathbf{x}|^4) \\ &\leq b(1 + |\mathbf{x}|^4)(t - s)^2 e^{a(t-s)} = o(t - s), \end{aligned}$$

where a and b are the constants referred in theorem 2. The condition (iii) in definition 5 is then verified, with $\delta = 2$.

In order to prove the existence of the drift function, we notice that

$$\mathbf{X}_t^{sx} = \mathbf{x} + \int_s^t \boldsymbol{\rho}(u, \mathbf{X}_u^{sx}) du + \int_s^t \boldsymbol{\sigma}(u, \mathbf{X}_u^{sx}) d\mathbf{W}_u, \ a.s., \ t \in (s, T]. \tag{2.23}$$

Using the relation (2.20) and the fact that the mean value of a stochastic integral is 0, one gets, by applying Fubbini's theorem,

$$\begin{aligned} E(\mathbf{X}_t - \mathbf{x} / \mathbf{X}_s = \mathbf{x}) &= E(\mathbf{X}_t^{sx} - \mathbf{x}) \\ &= E \left(\int_s^t \boldsymbol{\rho}(u, \mathbf{X}_u^{sx}) du \right) \\ &= \int_s^t E(\boldsymbol{\rho}(u, \mathbf{X}_u^{sx})) du. \end{aligned}$$

Dividing by $t - s$ and taking the limit when $t \to s^+$ we obtain the drift coefficient

$$
\begin{aligned}
\lim_{t \to s^+} \frac{1}{t-s} E(\mathbf{X}_t - \mathbf{x} / \mathbf{X}_s = \mathbf{x}) &= \lim_{t \to s^+} \frac{1}{t-s} \int_s^t E(\boldsymbol{\rho}(u, \mathbf{X}_u^{sx})) du \\
&= E(\boldsymbol{\rho}(s, \mathbf{X}_s^{sx})) \hspace{3cm} (2.24) \\
&= \boldsymbol{\rho}(s, \mathbf{x})
\end{aligned}
$$

since $\boldsymbol{\rho}$ is continuous, by assumption, $E(\boldsymbol{\rho}(s, \mathbf{X}_s^{tx}))$ is continuous too and we can use the fundamental theorem of calculus to justify the step (2.24).

In what concerns the existence of the diffusion coefficient, we just need to prove that

$$
\lim_{t \to s^+} \frac{1}{t-s} E\left[\left((X_t^{sx})^{(i)} - x_i \right) \left((X_t^{sx})^{(j)} - x_j \right) \right] = \sigma_{ij}(s, x),
$$

for $i = 1, ..., n$ and $j = 1, ..., m$. We begin by applying the Itô product formula (2.5),

$$
\begin{aligned}
&\lim_{t \to s^+} \frac{1}{t-s} E\left((X_t^{sx})^{(i)} (X_t^{sx})^{(j)} - x_i x_j \right) \\
=\;& \lim_{t \to s^+} \frac{1}{t-s} \int_s^t E\left[(X_u^{sx})^{(i)} \rho_j(u, X_u^{sx}) + (X_u^{sx})^{(j)} \rho_i(u, X_u^{sx}) + \sigma_{ij}(u, X_u^{sx}) \right] du \\
=\;& x_i \rho_j(u, x) + x_j \rho_i(u, x) + \sigma_{ij}(u, x).
\end{aligned}
$$

Then,

$$
\begin{aligned}
&\lim_{t \to s^+} \frac{1}{t-s} E\left[\left((X_t^{sx})^{(i)} - x \right) \left((X_t^{sx})^{(j)} - x \right) \right] \\
=\;& x_i \rho_j(s, x) + x_j \rho_i(s, x) + \sigma_{ij}(s, x) - x_i \lim_{t \to s^+} \frac{1}{t-s} E\left((X_t^{sx})^{(j)} - x_j \right) - x_j \lim_{t \to s^+} \frac{1}{t-s} E\left((X_t^{sx})^{(i)} - x_i \right) \\
=\;& \sigma_{ij}(s, x).
\end{aligned}
$$

Hence, $(\mathbf{X}_t, t \in [0, T])$ is a diffusion process with diffusion coefficient $Q = \boldsymbol{\sigma}\boldsymbol{\sigma}^T$ and drift $m = \boldsymbol{\rho}$.

$\square$

Notice that if the coefficients of the system of SDEs do not depend on time, then the strong solution is an homogeneous Markov process. In this case, we say that the system of SDEs is autonomous.

It can be proved (see [20]) that if the drift and diffusion coefficients of a diffusion process $\mathbf{X}$ satisfy the conditions of theorem 2, then $\mathbf{X}$ has transition densities $f_{t,\mathbf{x}}$, $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^n$, which, by theorem 1, satisfy the Kolmogorov forward equations (2.14)-(2.15) and, by theorem 2, constitute their only solutions.

We can now think on the reverse of this theorem, i.e., if any diffusion process can constitute a solution to some stochastic differential equation. Let $\mathbf{X}$ be a diffusion

process with coefficients $\boldsymbol{\rho}$ and $\mathbf{Q}$ and $\mathbf{Y}$ the solution of the SDE with the corresponding coefficients $\boldsymbol{\rho}$ and $\boldsymbol{\sigma}$, such that $\mathbf{Q} = \boldsymbol{\sigma}^{\mathbf{T}}\boldsymbol{\sigma}$. If the initial distribution of both processes is the same and they both satisfy the conditions of theorem 2 then their transition densities will be the only solutions of the the same Kolmogorov forward equations, therefore, $\mathbf{X}$ and $\mathbf{Y}$ are identical in distribution.

Finally, we stress that all results in this section can be generalized for $t \in [0, \infty)$ with the additional requirement that in definition 6, for all $t > 0$, $(\boldsymbol{\rho}(t, \mathbf{X}_s), s \in [0, t]) \in L^1(\Omega \times [0, t])$, and $(\boldsymbol{\sigma}(t, \mathbf{X}_s), s \in [0, t]) \in L^2(\Omega \times [0, t])$.

2.3.1. Systems of linear stochastic differential equations

As mentioned in section 2, in the particular case of linear systems of SDEs, it is possible to find an analytical solution. That solution is unique, because the coefficients satisfy the Lipschitz and growth conditions in the second variable. Let $(W_t^{(j)}, \ t \in [0, T])$, $j = 1, ..., m$, $m \in \mathbb{N}$, be m independent Wiener processes. The general form of a system of n linear SDEs and n variables, with the initial condition $\mathbf{X}_0 = (\xi^{(1)}, ..., \xi^{(n)})$ a.s., $\mathcal{F}_0$-measurable, is given by

$$d\mathbf{X}_t = (\mathbf{A}(t)\mathbf{X}_t + \mathbf{a}(t))dt + \sum_{j=1}^{m}(\mathbf{B_j}(t)\mathbf{X}_t + \mathbf{b_j}(t))d\mathbf{W}_t^{(j)}, \ t \in [0, T], \ \mathbf{X}_0 = \boldsymbol{\xi} \ a.s., \quad (2.25)$$

where $\mathbf{A}(t)$, $\mathbf{B_1}(t), ..., \mathbf{B_m}(t)$ are $n \times n$-dimensional matrix functions and $\mathbf{a}(t)$, $\mathbf{b_1}(t)$, ..., $\mathbf{b_m}(t)$ are n-dimensional vector functions. When $\mathbf{a}(t)$, $\mathbf{b_1}(t), ..., \mathbf{b_m}(t)$ are identically zero we call it a *homogeneous system of SDEs*. As in the case of ODEs, the solution to the general system (2.25) is given in function of the fundamental matrix satisfying the homogeneous system, which is obtained by making all $\mathbf{a}(t)$, $\mathbf{b_1}(t), ..., \mathbf{b_m}(t)$ identically zero, as follows

$$d\boldsymbol{\Phi}_t = \mathbf{A}(t)\boldsymbol{\Phi}_t dt + \sum_{j=1}^{m}\mathbf{B}_j(t)\boldsymbol{\Phi}_t d\mathbf{W}_t^{(j)}, \ t \in [0, T], \quad \boldsymbol{\Phi}_0 = \mathbf{I} \ a.s., \quad (2.26)$$

where $\mathbf{I}$ is the identity matrix. Unfortunately, except for the scalar case, an explicit expression cannot be given in general, even if all the matrices are constant. However, if they are constant and pairwise commute, then it is possible to obtain the explicit solution by Itô's formula (see [24]), which is given by

$$\boldsymbol{\Phi}(t) = exp\left[\left(\mathbf{A}(t) - \frac{1}{2}\sum_{j=1}^{m}\mathbf{B}_j^2(t)\right)t + \sum_{j=1}^{m}\mathbf{B}_j(t)W^{(j)}(t)\right], \ t \in [0, T]. \quad (2.27)$$

The general solution to (2.25) can likewise be deduced as in the scalar case (see [22] [25]). It is given by

$$
\mathbf{X}_t = \boldsymbol{\Phi}(t) \left[\boldsymbol{\xi} + \int_0^t \boldsymbol{\Phi}^{-1}(s) \left(\mathbf{a}(s) - \sum_{j=1}^m \mathbf{B}_j(s)\mathbf{b_j}(\mathbf{s}) \right) ds \right.
$$

$$
\left. + \int_0^t \boldsymbol{\Phi}^{-1}(s) \sum_{j=1}^m \mathbf{b_j}(\mathbf{s}) dW^{(j)}(s) \right] .
\tag{2.28}
$$

An example of a linear SDE is given in the next Chapter. However, in general, it is not possible to construct the explicit solution to an SDE nor even to the Kolmogorov forward equation. It may still be possible to obtain invariant distributions, which do not depend on time, and therefore they must satisfy the Kolmogorov forward equation after taking out the derivative with respect to time, which is null, i.e.,

$$
0 = -\sum_{i=1}^n \frac{\partial}{\partial y_i} \left(m_i f_{t,\mathbf{x}} \right) + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2}{\partial y_i \partial y_j} \left(Q_{ij} f_{t,\mathbf{x}} \right)
\tag{2.29}
$$

evaluated at $\mathbf{x} \in \mathbb{R}^n$. Despite this equation being slightly easier to solve, it does not tell us whether a given invariant distribution is a stationary distribution or not, in the sense that its distribution will eventually converge to that distribution, as $t \to +\infty$. For more details, please see [24] and [4].

Chapter 3

Modelling the evolution of populations

Biological phenomena may involve frequently one or more different species that interact with each other making all the system to evolve across time; we can think, for example, in a growing population of cells or animals, the spread of diseases or the migration of species. The state of the systems may evolve depending just on the current state and not on the past, thus reflecting the Markovian property of the system. This assumption is frequently used in this context, either in deterministic and stochastic models.

In this Chapter we will present and motivate two classes of SDE models in biology. The first one arises as a generalization of a deterministic ordinary differential equation model, considering environmental randomness, which may affect the parameters of the model. The second class of models results from the natural demographic randomness, existing even when the growing rate is constant and the environmental disturbances are disregarded. As we will see, the models belonging to this latter class are intrinsically non-linear, for we shall resort to numerical technics, in order to gain some insight about their behavior.

We stress that the systems arising in the models are autonomous and satisfy the local Lipschitz condition, therefore, the solution exists and is unique until an eventual moment of explosion. In case the model being linear, we know already that the solution exists for all $t > 0$. In some cases considered in the next section, it is possible to obtain the explicit expression for the solution.

3.1. Environmental randomness

3.1.1. Stochastic Malthusian linear model

Consider a population of animals or cells growing at a constant rate per capita, r, in a bounded habitat with no food shortage. If no other aspects are relevant for the evolution of the population and if the habitat area is big enough, we may

approximately describe the population density, $x(t)$, on each time $t \in [0, T]$, using the continuous deterministic Malthusian growth model (also called exponential growth model),

$$\frac{dx}{dt}(t) = rx(t), \quad t \in [0, T] \tag{3.1}$$

with the initial condition $x(0) = x_0$, $x_0 \in \mathbb{R}^+$. However, if the growing rate per capita, r, is affected by small non-negligible random environmental pressures, such as the temperature, humidity, light or diseases, we should replace the constant r by a random variable R. As usually, we will consider that r is affected by a time dependent gaussian noise, i.e., $R := r + \sigma W_t$ a.s., $\sigma \in \mathbb{R}^+$. This gives rise to the following linear autonomous stochastic differential equation

$$dX_t = rX_t dt + \sigma X_t dW_t, \quad t \in [0, T], \quad X_0 = x_0. \tag{3.2}$$

This model constitutes a stochastic version of the deterministic Malthusian model and is called the *Black-Scholes model*[1]

The Black-Scholes equation has exactly one strong solution $X = (X_t, t \in [0, T])$, which is an homogeneous diffusion process with drift $m(t, x) = rx$ and diffusion coefficient $Q(t, x) = \sigma^2 x^2$.

From (2.28), the solution $X = (X_t, t \in [0, T])$ is thus given by

$$X_t = x_0 exp\left(\left(r - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right) \quad a.s. \; t \in [0, T] \tag{3.3}$$

which is known as the geometric Brownian motion, while the process $Z = (Z_t, \; t \in [0, T])$, defined by $X_t = x_0 exp(Z_t)$, is a Wiener process with drift such that $Z_t \sim \mathcal{N}\left((r - \frac{1}{2}\sigma^2)t, \; \sigma\sqrt{t}\right)$. Obviously, X is adapted to $(\mathcal{F}_t)_{t \in [0,T]}$ and has continuous sample paths almost surely, consequently, both integrals $\int_0^t r|X_s|ds$ and $\int_0^t \sigma^2|X_s|^2 ds$, $t > 0$, are almost surely finite. In view of definition 6, we conclude that X given in (3.3) is in fact the only strong solution of (3.2).

The parameter r represents the mean growing rate per capita and σ the intensity of the random effects on the growing rate. In fact, the mean function $E(X_t) = u(t)$, $t \in [0, T]$ verifies the ODE

$$\frac{du}{dt}(t) = ru(t), \; u(0) = x_0 \tag{3.4}$$

---

[1]The Black-Scholes model was first used in the context of finances by Black and Scholes in 1973 (see [26] [1]).

We conclude that the solution of the deterministic model corresponds to the mean function of the solution of the stochastic model. Naturally, the stochastic model gives more insight about the underlying biological phenomenon, but it needs more information to be settled because one more parameter must be estimated.
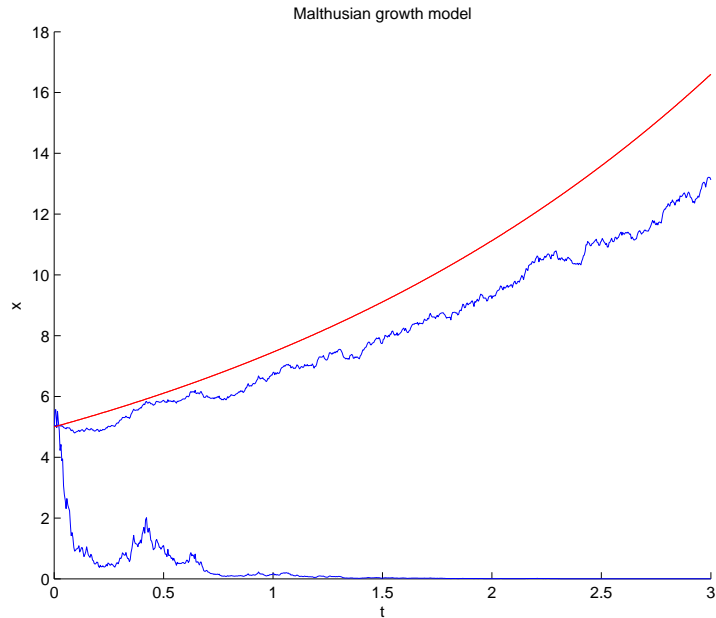

Figure 3.1: Two sample paths of the stochastic Malthusian growth model (blue lines) with the same underlying sample path of the Wiener process, illustrating the two different types of behavior of the solutions on the time interval $[0, 3]$. The time-step is 2^{-8}. The initial condition is $X_0 = 5$ and the parameter values are $r = 0.4$, $\sigma = 0.1$ (upper curve) and $\sigma = 2$ (lower curve). The red line represents the solution of the corresponding deterministic model.

Two sample paths of X as well as the solution of the deterministic Malthusian model are presented in figure 3.1. As expected, the sample paths are positive functions. The upper curve seems to have an approximately positive exponential growth, which is in accordance with the deterministic model. However, the lower curve apparently tends to zero, which means that the extinction of the population may occur. This behavior is different from the one presented by the deterministic model, on which the extinction never occurs for r positive. In fact, once $\frac{W_t}{t}$ converges to zero a.s. (see [20]) as t goes to infinity, then

$$\lim_{t \to +\infty} t(r - \frac{1}{2}\sigma^2 + \sigma \frac{W_t}{t}) = -\infty \ a.s.$$

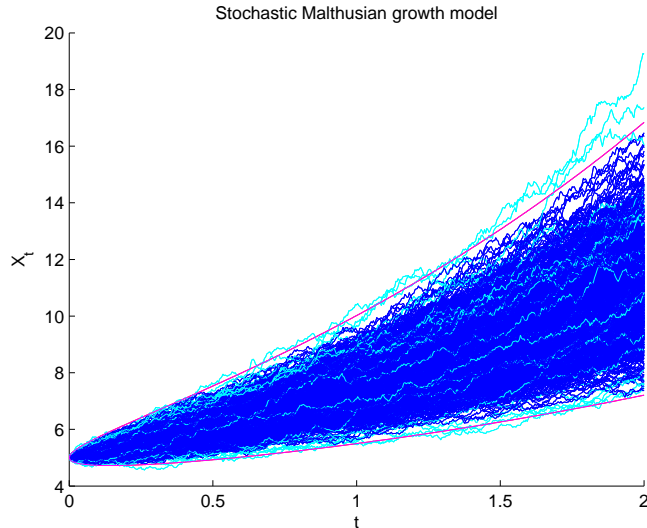whenever $r - \frac{1}{2}\sigma^2$ is negative, implying that X_t converges to zero a.s.. This feature

Figure 3.2: One thousand of sample paths of the stochastic Malthusian growth model on the time interval $[0, 2]$ are presented. Magenta lines represent the two pointwise bounds of the confidence interval. Cyan lines correspond to the paths that went out of the confident region at some time. The time-step is 2^{-8}. The initial condition is $X_0 = 5$ and the parameter values are $r = 0.4$, $\sigma = 0.1$.

of the stochastic model is realist, because even in the case that the growth rate per capita is positive, if the random effects of the medium have a relatively strong impact on the population growth, then it may get extinct.

To see how the dispersion evolves with time, we calculate the standard deviation of X_t, $t \in [0, T]$, either by using (3.3) or by applying the Itô's formula to $f(x) = x^2$ in order to obtain a SDE for the process $X^2 = (X_t^2, t \in [0, T])$. A straightforward computation leads to $\sqrt{V(X_t^2)} = x_0 e^{rt} \sqrt{e^{\sigma^2 t} - 1}$. This function is strictly increasing, which means that the model loses accuracy over time.

In order to gain a better understanding about how the model evolves with time we will obtain a interval where the X_t is likely to occur, for each $t > 0$. Recalling that $Z_t \sim \mathcal{N}\left((r - \frac{1}{2}\sigma^2)t, \ \sigma\sqrt{t}\right)$, we may write

$$
\begin{aligned}
0.9973 \ &= \ P\left(\left(r - \frac{\sigma^2}{2}\right)t - 3\sigma\sqrt{t} \leq Z_t \leq \left(r - \frac{\sigma^2}{2}\right)t + 3\sigma\sqrt{t}\right) \\
&= \ P\left(x_0 exp\left(\left(r - \frac{\sigma^2}{2}\right)t - 3\sigma\sqrt{t}\right) \leq X_t \leq x_0 exp\left(\left(r - \frac{\sigma^2}{2}\right)t + 3\sigma\sqrt{t}\right)\right)
\end{aligned}
$$

One thousand sample paths of $(X_t, t \in [0, 2])$ were generated in order to illustrate their position regarding the confidence band. For each t, an average of 99.84% of the generated values, $X_t(\omega)$, lie within the region referred above, which is near the

26

theoretical value, as expected. In figure 3.2, we can see the generated sample paths, as well as the confidence region. Despite being quite simple to calculate the pointwise confidence band, it would be more worthwhile to obtain a confidence band uniformly in t. However there is not a closed formula for obtaining that band (see [27]). In this simulation, only 95.5% of the sample paths presented in figure 3.2 lie entirely within the considered region. Those are presented in blue, while the remaining 4.5% are presented in cyan. If we consider a bigger time interval, these percentages get worse.

3.1.2. Stochastic logistic growth model I

One assumption underlying the previous model is the lack of any limitation of resources, which allows the unbounded growth of the population whenever the growing rate satisfies a certain condition. This situation is realistic when the size of the population is far bellow the carrying capacity of the habitat. However, as the population grows, a different dynamic may occur: the system will eventually enter a region in which negative feedback factors become important. The lack of resources makes the growing rate to decrease and the size of the population will eventually reach an equilibrium. A basic model that incorporates these two phenomena is the well-known deterministic logistic model,

$$\frac{dx}{dt}(t) = rx(t)\left(1 - \frac{x(t)}{K}\right), \ x(0) = x_0, \ t \in [0, T],$$

where $x(t)$ represents the density of the population at time t, $r \in \mathbb{R}^+$ the growth rate and $K \in \mathbb{R}^+$ the carrying capacity. The variation per capita is now given by the quotient $\frac{dx}{dt}/x = r_1(x) := r\left(1 - \frac{x}{K}\right)$. Notice that r is now the maximum rate. For $x_0 < K$, $\frac{dx}{dt}$ is positive and as the population increases, r_1 decreases until vanishing. At this moment, the derivative $\frac{dx}{dt}$ is null, which means that the population density reaches an equilibrium with the habitat: $lim_{t\to\infty}x(t) = K$. On the other hand, if $x_0 > K$, then $\frac{dx}{dt}$ is negative and the population decreases until it reaches again K. For $r > 0$, K is then a globally asymptotic stable equilibrium point, while 0 is an unstable equilibrium point.

As in the study of the Malthusian model, we will introduce a gaussian perturbation on the growth rate per capita, $R := r_1 + \sigma W_t \ a.s.$, $\sigma > 0$, leading to the following logistic stochastic differential equation,

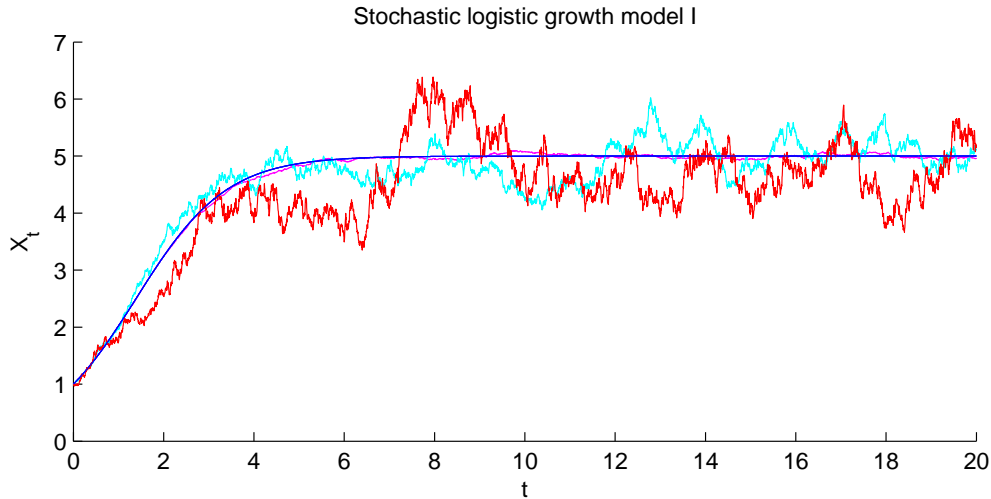$$dX_t = rX_t\left(1 - \frac{X_t}{K}\right)dt + \sigma X_t dW_t, \ X_0 = x_0, \ t \in [0, T]. \tag{3.5}$$

Figure 3.3: Three sample paths of the stochastic logistic growth model on the time interval $[0, 20]$, corresponding to three different values for σ: 0.01 (magenta), 0.1 (cyan) and 0.2 (red). The time-step is 2^{-8}, the initial condition is $X_0 = 1$ and the values of the remaining parameters are $r = 1$, $k = 5$. The blue line represents the solution of the corresponding deterministic model with the same time-step and the same parameters' values.

Since the coefficients are not linear, a unique strong solution, $X = (X_t, t \in [0, T])$, of (3.5) exists until a time of explosion, which turns out to be infinite (see [7]). The solution X is a homogeneous diffusion process with drift coefficient $m(t, x) = rx \left(1 - \frac{x}{K}\right)$ and diffusion coefficient $Q(t, x) = \sigma^2 x^2$.

The Euler-Maruyama method was implemented (see [6] [22]).[2] Some sample paths are plotted in figures 3.3 and 3.4.

We see that the sample paths are positive. In figure 3.3, the sample paths behave accordingly to the deterministic solution: they stay in the vicinity of the carrying capacity, $K = 5$, which suggests that the solution of the SDE has a stationary distribution with mean value K. This is in fact true and can be formally proved (see [7]). We also observe that as σ increases, the variance of X_t increases too, as was to be expected from (3.5). In figure 3.4, we see a different behavior: the sample paths differ from the deterministic model, they apparently do not follow any specific rule, except the sample path (c), which tends to zero (extinction). This behavior is justified by the big oscillations due to the value of σ being bigger then one, which causes the random term to be bigger then X_t. As X_t increases governed by the deterministic

---

[2]This model can also be solved analytically by performing an appropriate change of variable (see [3]).

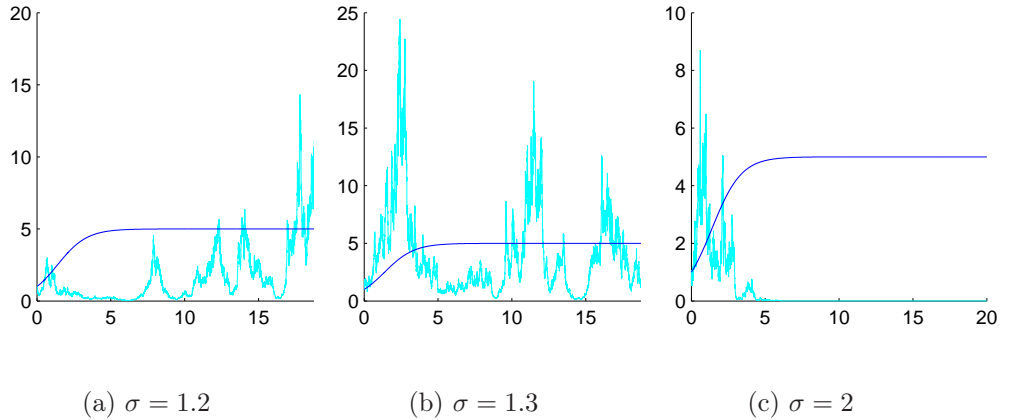(a) $\sigma = 1.2$ (b) $\sigma = 1.3$ (c) $\sigma = 2$

Figure 3.4: Three sample paths (cyan) of the stochastic logistic growth model on the time interval $[0, 20]$. The time-step is 2^{-8}, the initial condition is $X_0 = 1$ and the values of the remaining parameters are $r = 1$, $k = 5$. The blue line represents the solution of the corresponding deterministic model with the same time-step and the same parameters' values.

part of the equation, the random term increases too, making the sample path to vary with such a big amplitude that it nearly reaches zero. At this moment, the random term is small again and X_t remains near zero making small "jumps" only, until it eventually goes up and we return to the initial situation that may be repeated over time. This feature of the model is in accordance with population growth mechanisms observed in real life: if the randomness of the environment is big, then the changes on the population size are more unpredictable and may result in dramatic changes or even extinction in a short period of time. We can think, for example, in a population which is infected by a contagious disease: when the population is small, individuals do not have so much contact with each other, therefore, the disease does not propagate easily and the growth of the population is approximately logistic. When the density of the population gets bigger, then two situations may occur: the contacts do not increase and the size of the population may remain following a logistic growth or the contacts may increase so that many individuals get infected and may dye. In this example, the randomness of the environment can be associated to the contacts between individuals.

We shall next find an approximation for the mean function of the solution by applying the Monte Carlo method. The mean function was obtained from 10, 100 and 1000 generated sample paths. The resulting estimations were then compared to the solution, $\hat{u}$, of the deterministic logistic growth equation obtained numerically.

Number of generated sample paths	Error
10	9.6613
100	2.5444
1000	1.6085

Table 3.1: Euclidian norm of the difference between the vectors obtained in *matlab*, corresponding to the approximation of the mean function and the solution of the deterministic logistic model within the time interval $[0, 20]$ for $r = 1$, $k = 5$, $\sigma = 0.1$.

The outcome is outlined in table 3.1, where the Error corresponds to the euclidian norm of the difference between the two vectors referred before. We see that the mean function gets closer to $\hat{u}$ as the number of sample paths increase.

3.1.3. Stochastic predator-prey model I

We shall consider now two populations acting as predator and prey. The model for describing the interaction between them is then given by a system of two equations. The well-known deterministic predator-prey model, also called Lotka-Volterra model, is given by

$$
\begin{cases}
\frac{dx}{dt} = f(x, y) = ax - bxy \\
\frac{dy}{dt} = g(x, y) = -cy + dxy
\end{cases}, \quad x(0) = x_0, \ y(0) = y_0, \quad (3.6)
$$

where $x_0, y_0, a, b, c, d > 0$, x represents the density of the prey (for example, rabbits) and y represents the density of the predator (for example, foxes). The dynamics underlying this model is the following:

1. in the absence of predator, $\frac{dx}{dt} = ax$, the population of preys grows at a constant rate per capita a (exponential growth),

2. in the absence of preys, $\frac{dy}{dt} = -cy$, the population of foxes decreases until reaching extinction,

3. the number of meetings between the two species is proportional to the product of the population densities, contributing to the decrease of the prey population and to the increase the predator population.

Before we introduce a random perturbation in the parameters, we shall simplify the deterministic model through linearization about a steady state, (x^*, y^*), which yields a simple and reasonably good approximation, whenever the solution stays in

the vicinity of (x^*, y^*). The steady states of (3.6) are $(0,0)$ and $(c/d, a/b)$, which were easily obtained by making f and g equal to zero. On the one hand, a linearization around $(0,0)$ allows us to draw conclusions about the predation phenomenon when both prey and predator populations have a small number of individuals. On the other hand, a linearization around $(c/d, a/b)$ gives us insight about the phenomenon when the prey and predator populations have a density about a/b and c/d, respectively. In the first case, it turns out that the population growth is exponential for both populations, which conduces us to a Malthusian model already discussed in the beginning of this section. We will then see what happens in the second case. By applying the first order Taylor polynomial at $(c/d, a/b)$ we obtain a linear approximation for f and g, which leads to the following linear system,

$$\begin{cases} \frac{du}{dt} & = & -\frac{bc}{d}v \\ \frac{dv}{dt} & = & \frac{ad}{b}u \end{cases}, \quad u(0) = u_0, \ v(0) = v_0 \quad (3.7)$$

with $u = x - \frac{c}{d}$, $v = y - \frac{a}{b}$ and u_0 and v_0 in the vicinity of the steady state $(u^*, v^*) = (0, 0)$.

Introducing now a gaussian perturbation on the parameters c and a of the form $c + \sigma_1 W_t$ and $a + \sigma_2 W_t$, with $\sigma_i \in \mathbb{R}^+$, $i = 1, 2$, we obtain the following stochastic linear system,

$$\begin{cases} dU_t & = & -\frac{bc}{d}V_t dt - \sigma_1 \frac{b}{d}V_t dW_t \\ dV_t & = & \frac{ad}{b}U_t dt + \sigma_2 \frac{d}{b}U_t dW_t, \end{cases}, \quad U_0 = u_0, \ V_0 = v_0 \ a.s.$$

or, in the matrix form,

$$d\mathbf{Z}_t = \mathbf{A}\mathbf{Z}_t dt + \mathbf{B}\mathbf{Z}_t dW_t,$$

with $\mathbf{Z}_t = (U_t, V_t)$,

$$\mathbf{A} = \begin{bmatrix} 0 & -\frac{bc}{d} \\ \frac{ad}{b} & 0 \end{bmatrix}, \ \mathbf{B} = \begin{bmatrix} 0 & -\sigma_1 \frac{b}{d} \\ \sigma_2 \frac{d}{b} & 0 \end{bmatrix}.$$

Suppose that the parameters of the model satisfy $\frac{c}{a} = \frac{\sigma_1}{\sigma_2}$, which may be plausible, once the parameters responsible for the randomness, σ_1 and σ_2, do not need to be very specific, only the order of magnitude of the noise is important. In this case, we have that $\mathbf{A}$ and $\mathbf{B}$ are constants and verify $\mathbf{AB} = \mathbf{BA}$. We may then obtain an explicit expression for the solution to this system given in (2.28), namely,

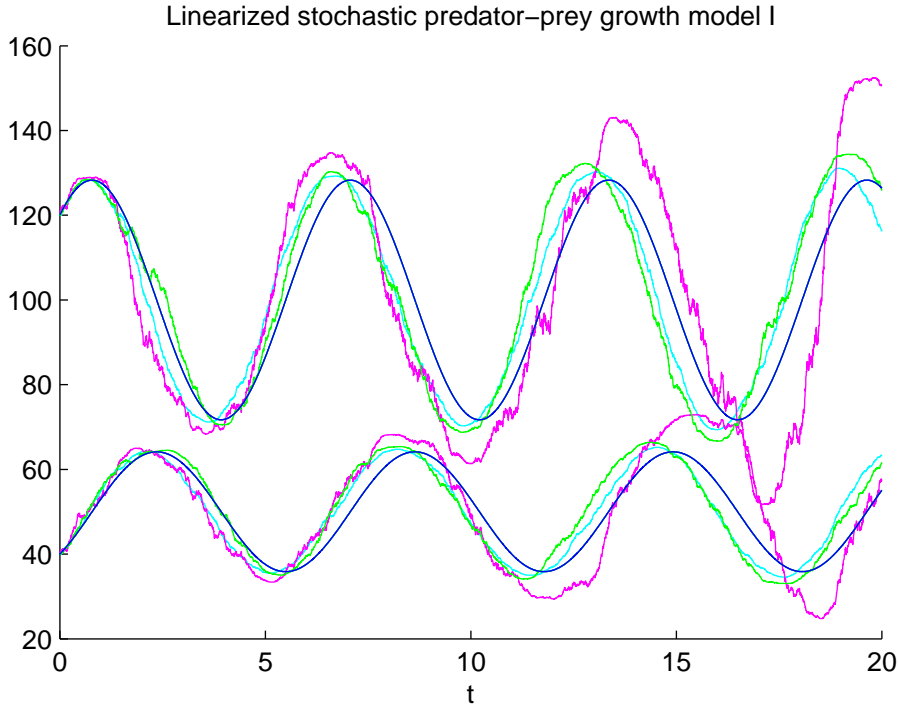$$\mathbf{Z}_t = \mathbf{\Phi}(t)\boldsymbol{\xi}, \ t \in [0, T],$$

Figure 3.5: Three sample paths of the linearized stochastic Lotka-Volterra I on the time interval $[0, 20]$, corresponding to $n = 4$ (cyan), $n = 7$ (green) and $n = 10$ (magenta). The time-step is 2^{-8}, the initial condition is $\mathbf{Z}_0 = (20, -10)$ and the values of the remaining parameters are $a = 1$, $b = 0.02$, $c = 1$, $d = 0.01$. The blue line represents the solution of the corresponding deterministic linearized model with the same time-step and the same parameters' values.

with

$$\mathbf{\Phi}(t) = exp\left[\left(\mathbf{A} - \frac{1}{2}\mathbf{B}^2\right)t + \mathbf{B}W_t\right], \ t \in [0, T]. \tag{3.8}$$

Three sample paths of $\mathbf{Z}$ are presented in figure 3.5.

Finally, notice that the random effect introduced to obtain the models (3.2), (3.5) and (3.8) results from environmental randomness and not from demographic randomness, i.e., the randomness associated to the number of births and deaths, which exists even when the birth and death rates are constant. This case is considered in the next section.

3.2. Demographic randomness

The phenomena studied in population dynamics are intrinsically discrete and are often modeled by Markov chains. However, as the discretization becomes finer, Markov chains might be well approximated by diffusion processes corresponding to

solutions of stochastic differential equations, therefore facilitating the mathematical analysis (see [3], [10], [12], [13], [11] and [28]).We will then make the assumption that the biological phenomenon is well described by a vector-valued time-homogeneous Markov chain $\mathbf{X} = (\mathbf{X}_t^{(n)}, t \in [0, T])$, with $\mathbf{X}_t^{(n)} = (X_{t,1}^{(n)}, ..., X_{t,d}^{(n)})$, $d \in \mathbb{N}$, with state space $\frac{1}{n}\mathbb{Z}$, where $X_t^{(n)}$ represents its population density at time t, depending on the parameter n, which represents the order of magnitude of the populationsŠ size or of the area of the habitat.

The procedure followed for the model deduction comprises the following steps:

- define $\mathbf{X}_0^{(n)} = \mathbf{x}_0$ and the transition probabilities of the Markov chain $\mathbf{X}^{(n)}$,

- for large n, we approximate $\mathbf{X}^{(n)}$ by a diffusion process $\mathbf{Y}^{(n)}$ with coefficients $\boldsymbol{\rho}$ and $\mathbf{Q}$,

- $\mathbf{Y}^{(n)}$ is a solution of the matrix SDE

$$d\mathbf{Y}_t^{(n)} = \boldsymbol{\rho}(\mathbf{Y}^{(n)}, t)dt + \frac{1}{2\sqrt{n}}\boldsymbol{\sigma}(\mathbf{Y}^{(n)}, t)d\mathbf{W}_t, \quad \mathbf{Y}_0^{(n)} = \mathbf{x}_0$$

with $\boldsymbol{\sigma}\boldsymbol{\sigma}^T = \mathbf{Q}$.

We begin by briefly presenting the discrete model for one population. An heuristic deduction of the continuous model shall then be described and analysed in order to draw some conclusions about the underlying biological phenomenon. Consider one species in a certain habitat at time $t > 0$. For a small positive constant δt let $\Delta X^{(n)} := X_{t+\Delta t}^{(n)} - X_t^{(n)}$ denote the variation of the density occurring within $[t, t+\Delta t]$ and D_n the support of $\Delta X^{(n)}$, which is defined by

$$D_n = \left\{ \frac{l}{n} : l \in D \right\},$$

where $D := \mathbb{Z} \cap E$ and E is a limited interval. Assume that the transition probabilities of $X^{(n)}$,

$$p_{t,k/n}\left(t + \Delta t, \frac{k+l}{n}\right) := P\left(\Delta X^{(n)} = \frac{l}{n} / X_t^{(n)} = \frac{k}{n}\right),$$

take the general form

$$p_{t,k/n}\left(t + \Delta t, \frac{k+l}{n}\right) = n\beta_l\left(t, \frac{k}{n}\right)\Delta t + o(\Delta t), \tag{3.9}$$

for each $l \in D \setminus \{0\}$ and $\beta_l : [0, T] \times \mathbb{R} \to \mathbb{R}_0^+$, which represents the transition change rate of the Markov chain. Since $p_{t,k/n}$ is a probability, then for $l = 0$ the probability of the variation to be null is naturally given by

$$p_{t,k/n}\left(t + \Delta t, \frac{k}{n}\right) = 1 - n\Delta t \sum_{l \in D \setminus \{0\}} \beta_l\left(t, \frac{k}{n}\right) + o(\Delta t). \tag{3.10}$$

The Markov chain is, thus, completely determined by the transition probabilities (3.9) and (3.10) and the distribution of $X_0^{(n)}$.

The deduction of the diffusion approximation undergoes now through the heuristic construction of a relation involving the transition probability functions $p_{t,x}(s,y)$, for $x, y \in \frac{1}{n}\mathbb{Z}$ and $0 \le t < s \le T$, based on the forward Kolmogorov equations. We begin by writing down the Chapman-Kolmogorov equation for time-continuous Markov chains,

$$p_{t,x}(s + \Delta s, y) = \sum_{l \in D} p_{t,x}(s, y - l/n) p_{s,y-l/n}(s + \Delta s, y).$$

with $0 \le t < s < s + \Delta s \le T$. Using the expressions in (3.9) and (3.10), the previous summation can be rewritten as

$$p_{t,x}(s + \Delta s, y) = p_{t,x}(s, y) \left(1 - n\Delta s \sum_{l \in D \setminus \{0\}} \beta_l(s, y) + o(\Delta s) \right)$$
$$+ n\Delta s \sum_{l \in D \setminus \{0\}} p_{t,x}(s, y - l/n)\beta_l(s, y - l/n) + o(\Delta s).$$

Dividing by Δs and making Δs to approach 0 we derive a differential equation for the transition probability function $p_{t,x}$. From (3.11) we thus obtain,

$$\frac{\partial p_{t,x}}{\partial s}(s, y) = -n \sum_{l \in D \setminus \{0\}} p_{t,x}(s, y)\beta_l(s, y) \tag{3.11}$$
$$+ n \sum_{l \in D \setminus \{0\}} p_{t,x}(s, y - l/n)\beta_l(s, y - l/n)$$

Furthermore, for l/n sufficiently small, which is easily verified if n is large, we can find a local approximation for $p_{t,x}(s, y - l/n)\beta_l(s, y - l/n)$, by using the second order Taylor polynomial at y. Under adequate regularity conditions, it follows from (3.12) that

$$\frac{\partial p_{t,x}}{\partial s} \approx -n \sum_{l \in D \setminus \{0\}} p_{t,x}\beta_l + n \sum_{l \in D \setminus \{0\}} \left[p_{t,x}\beta_l - \frac{l}{n}\frac{\partial(p_{t,x}\beta_l)}{\partial y} + \frac{1}{2}\frac{l^2}{n^2}\frac{\partial^2(p_{t,x}\beta_l)}{\partial y^2} \right] \tag{3.12}$$

evaluated at (s, y). Finally, simplifying this equation, we conclude that the transition probability, $p_{t,x}$, satisfies approximately the second-order partial differential equation

$$\frac{\partial f_{t,x}}{\partial s} = \frac{\partial(f_{t,x}\rho)}{\partial y} + \frac{1}{2n}\frac{\partial^2(f_{t,x}Q)}{\partial y^2} \qquad \text{at } (s, y), \tag{3.13}$$

with

$$\rho := \sum_{l \in D \setminus \{0\}} l\beta_l \text{ and } Q := \sum_{l \in D \setminus \{0\}} l^2\beta_l.$$

This gives us an idea about the diffusion process $Y^{(n)}$ we may use as an approximation for the Markov chain $X^{(n)}$. Whenever ρ and Q satisfy the conditions described in theorem 2, (3.13) coincides with the Kolmogorov forward equation of the diffusion process with coefficients ρ and $\frac{Q}{n}$, whose the unique solution, $f_{t,x}$, is a transition density function of the process $Y^{(n)}$. Indeed, it can be shown that, for n large enough, $Y^{(n)}$ represents a good approximation of $X^{(n)}$ (see [10]).

We finally arrived to the SDE's stochastic model for one species given by

$$dY_t^{(n)} = \rho(Y_t^{(n)}, t)dt + \frac{1}{2\sqrt{n}}\sigma(Y_t^{(n)}, t)dW_t, \; Y_0^{(n)} = x_0 \; a.s. \tag{3.14}$$

with $\sqrt{Q} = \sigma$.

The previous deduction can easily be generalized for the case of d interacting species. In this case, define $D := \mathbb{Z}^d \cap E$, where E a limited d-dimensional rectangle, and

$$\rho := \sum_{\mathbf{l} \in D \setminus \{0\}} \mathbf{l}\beta_{\mathbf{l}} \text{ and } Q := \sum_{\mathbf{l} \in D \setminus \{0\}} \mathbf{l}\mathbf{l}^T \beta_{\mathbf{l}},$$

where $\beta_{\mathbf{l}}$ is a real function defined in D_n The resulting model is given by

$$d\mathbf{Y}_t^{(n)} = \rho(\mathbf{Y}_t^{(n)}, t)dt + \frac{1}{2\sqrt{n}}\sigma(\mathbf{Y}_t^{(n)}, t)d\mathbf{W}_t, \; \mathbf{Y}_0^{(n)} = \mathbf{x}_0 \; a.s. \tag{3.15}$$

such that $\mathbf{Q} = \boldsymbol{\sigma}\boldsymbol{\sigma}^T$.

3.2.1. Stochastic logistic growth model II

The functions β_l can be obtained from the deterministic logistic growth model, which is given by $\frac{dy}{dt} = y\left(r - r\frac{y}{K}\right)$. To establish the stochastic model, we need some more information related to the population in consideration, namely, for instance, the birth and death rates, which are represented by the parameters β_1 and β_{-1}, respectively, and must satisfy $\beta_1 - \beta_{-1} = y\left(r - r\frac{y}{K}\right)$. In order to present latter numerical examples we will consider $\beta_1(t, y) = ry$, $\beta_{-1}(t, y) = r\frac{y^2}{k}$ and $r_l = 0$ for $l \neq -1, 0, 1$. Consequently, $m(s, y) = y\left(r - r\frac{y}{K}\right)$ and $Q(s, y) = y\left(r - r\frac{y}{K}\right)$. A stochastic logistic growth model for the population density, whose size of the habitat has a order of magnitude of n, assumed constant, is thus given by the following SDE

$$dY_t^{(n)} = \left(rY_t^{(n)} - \frac{r}{K}\left(Y_t^{(n)}\right)^2\right)dt + \frac{1}{2\sqrt{n}}\sqrt{\left|rY_t^{(n)} + \frac{r}{K}\left(Y_t^{(n)}\right)^2\right|}dW_t.$$

Some sample paths of the solution are plotted in figure 3.6. Similarly to the stochastic logistic model I studied in section 3.1.2, we observe that the sample paths remain in the vicinity of the solution of the deterministic model. Increasing the parameter

n	Number of sample paths	Error
	10	7.8640
1	100	3.3491
	1000	1.9547
	10	2.5097
100	100	0.7685
	1000	0.2748

Table 3.2: Euclidian norm of the difference between the vectors corresponding to the approximation of the mean function and the solution of the deterministic logistic model within the time interval $[0, 20]$, for different values of n and $r = 1$ and $k = 5$.

n we notice a decreasing in the variability, which confirms the theoretical reasoning described in the deduction of the model, in the beginning of this section.

Contrarily to the stochastic version of the logistic model considered on the previous section, this model allows the sample paths to be negative (see figure 3.6 b)), which is not realistic in this context. In [21], this problem is overcame by an adaptation to the model, giving rise to a mixed distribution. One hundred sample paths were generated for different time intervals and for the parameters $X_0 = 0.5$, $n = 1$, $k = 5$ and $r = 1$, resulting in 5 extinctions within $[1, 20]$, 3 within $[0, 40]$ and 4 within $[0, 200]$. Apparently, we could not find any relation between the length of the time interval and the probability of extinction. In fact, by analyzing the graph of the sample paths, it is possible to realize that the extinction occurs just until a certain instant. This observation allows us to expect that if the population resisted until a certain moment, then it will probably survive at least in the medium term. New simulations were made for the time interval $[0, 20]$ in order to obtain the times to extinction. Four sample paths led to extinction at times 1.3203, 1.9258, 0.8047 and 2.7695. A rough estimate of the mean time to extinction (when it occurs) is then about 1.7051.

Finally, we present an approximation for the mean function of the solution by applying the Monte Carlo simulations as before. The mean function was obtained for two different values of n from 10, 100 and 1000 sample paths. The comparison between the estimates and the deterministic solution are made by using the euclidian norm. The results, which are presented in table 3.2, are, once again, in accordance with the theoretical reasoning: as n increases, the distance to the deterministic
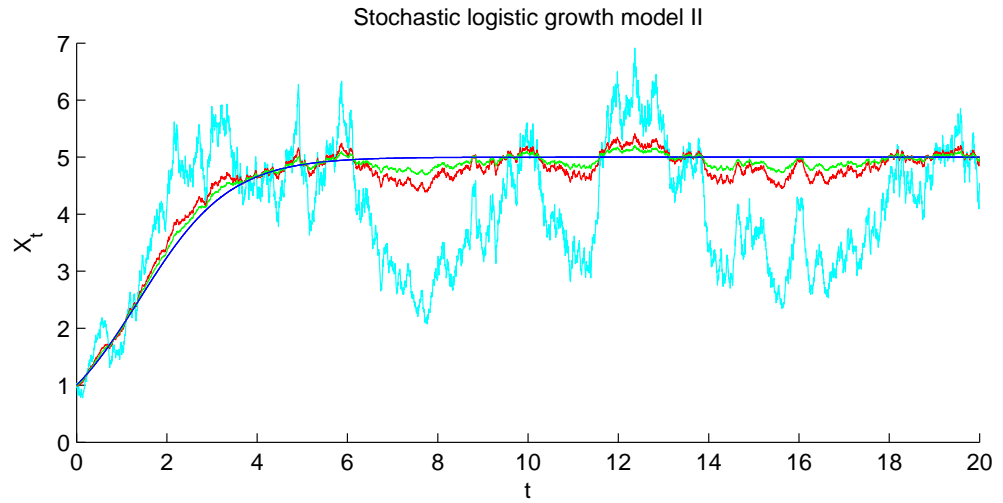
(a)
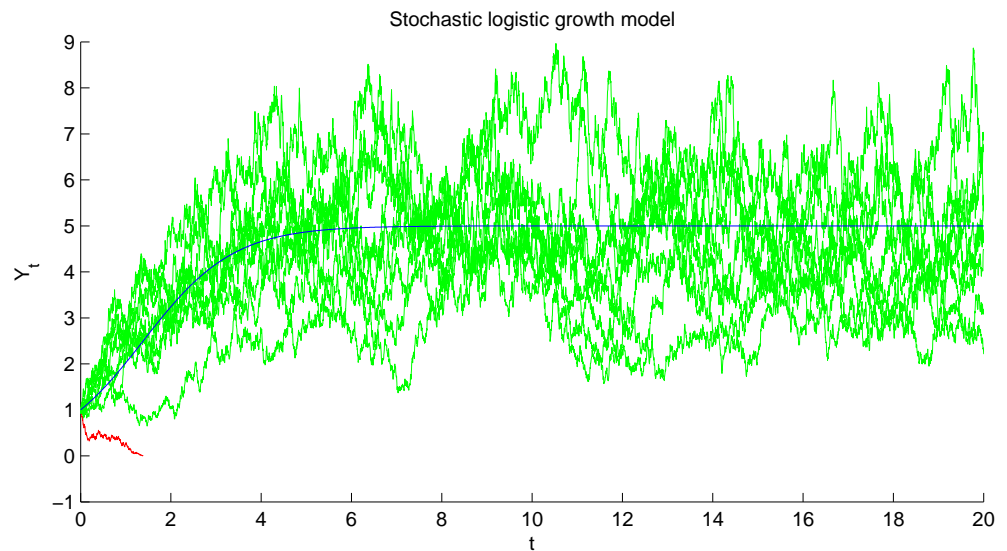
(b)

Figure 3.6: In a): three sample paths of the stochastic logistic growth model II on the time interval $[0, 20]$ with the same underlying Wiener process, corresponding to $n = 1$ (cyan), $n = 25$ (red) and $n = 100$ (green). The time-step is 2^{-8}, the initial condition is $Y_0 = 1$ and the values of the remaining parameters are $r = 1$, $k = 5$. The blue line represents the solution of the corresponding deterministic model with the same time-step and the same parameters' values. In b): ten sample paths of the same model with the same parameters and $n = 0.8$. The red line corresponds to a sample path for which extinction occurs.
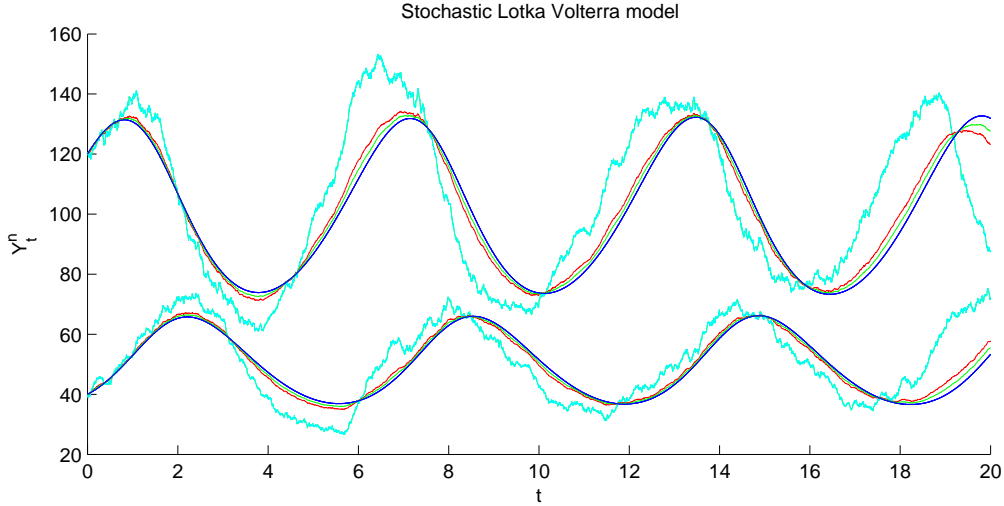
Figure 3.7: Three sample paths of the stochastic Lotka-Volterra model on the time interval $[0, 20]$ with the same underlying Wiener process, corresponding to $n = 1$ (cyan), $n = 25$ (red) and $n = 100$ (green). The time-step is 2^{-8}, the initial condition is $\mathbf{Z}_0 = (120, 40)$ and the values of the remaining parameters are $a = 1$, $b = 0.02$, $c = 1$, $d = 0.01$. The blue line represents the solution of the corresponding deterministic model with the same time-step and the same value of the parameters.

model decreases. This fact suggests that the empirical mean function converges to the solution of the deterministic model.

3.2.2. Stochastic predator-prey model II

Consider the deterministic Lotka-Volterra model (3.6). Define $\mathbf{z} := (x, y)$ and $D := \{(0,0), (1,0), (-1,0), (0,1), (0,-1)\}$, where $(1,0)$ and $(-1,0)$ represent one birth and one death, respectively, in the prey population and $(0,1)$ and $(0,-1)$ represent one birth and one death in the predator population. The birth and death rates in each population are given by $\beta_{(1,0)} = ax$, $\beta_{(0,1)} = dxy$, $\beta_{(-1,0)} = bxy$, $\beta_{(0,-1)} = cy$. Consequently, we have

$$\mathbf{m}(t, y) = \sum_{\mathbf{l} \in D} \mathbf{l}\beta_{\mathbf{l}} = \begin{bmatrix} ax - bxy \\ dxy - cy \end{bmatrix}$$

and

$$\mathbf{Q}(t, y) = \sum_{\mathbf{l} \in D} \mathbf{l}\mathbf{l}^T \beta_{\mathbf{l}} = \begin{bmatrix} ax + bxy & 0 \\ 0 & dxy + cy \end{bmatrix}.$$

38

Prey			Predator		
n	No. of sample paths	Error	n	No. of sample paths	Error
	10	698.7552		10	387.3996
1	100	341.6737	1	100	164.3148
	1000	286.4792		1000	107.8320
	10	59.9920		10	34.1499
100	100	26.2427	100	100	12.6041
	1000	12.2385		1000	6.0136

Table 3.3: Euclidian norm of the difference between the vectors corresponding to the approximation of the mean function and the solution of the deterministic logistic model within the time interval $[0, 20]$, for $a = 1$, $b = 0.02$, $c = 1$, $d = 0.01$ and for different values of n.

Once $\mathbf{Q}$ is a diagonal matrix, the diffusion matrix is given by

$$\boldsymbol{\sigma}(t, y) = \begin{bmatrix} \sqrt{|ax + bxy|} & 0 \\ 0 & \sqrt{|dxy + cy|} \end{bmatrix}.$$

If the habitat area has a order of magnitude of n, then a stochastic predator-prey model is given by the system

$$\begin{cases} dX_t &= (aX_t - bX_tY_t)dt + \frac{1}{2\sqrt{n}}\sqrt{|aX_t + bX_tY_t|}dW_t^{(1)} \\ dY_t &= (dX_tY_t - cY_t)dt + \frac{1}{2\sqrt{n}}\sqrt{|dX_tY_t + cY_t|}dW_t^{(2)}. \end{cases}$$

For simplicity, we suppress the index n on the random variables X_t and Y_t. Recall that the Wiener processes $(W_t^{(1)}, t \in [0, T])$ and $(W_t^{(2)}, t \in [0, T])$ are independent. Some sample paths are plotted in figure 3.7 corresponding to different values of the parameter n. We see that the general dynamics of the stochastic model is consistent with the deterministic model: when the prey population increases (decreases), the density of the predators increases (decreases) too, with a delay. Once again, as n increases, the variability decreases and the sample paths get apparently closer to the solution of the deterministic model. The results outlined in table 3.3 were obtained through Monte Carlo simulations and they confirm the previous conclusions.

3.2.3. Stochastic SIR epidemic model

An epidemic model pretends to represent the dynamic associated to the transmission of a disease within a population. In this section, we begin by considering the most

basic deterministic epidemic model, the SIR model, which describes the number of susceptible (S), infected (I) and recovered (R) individuals of a population. The following assumptions are made:

1. the recovered individuals are not able to be infected again or to transmit the infection to any susceptible individual,

2. the size of the population is considered to be fixed, i.e., $S(t) + I(t) + R(t) = N$,

3. all individuals of the population have the same probability of contracting the disease at a rate proportional to the number of meetings between infected and susceptible, aSI,

4. infected individuals recover at a rate per capita b

Summarizing the above description, the evolution of each class of individuals is governed by the following rules

$$S + I \xrightarrow{a} 2I, \quad I \xrightarrow{b} R.$$
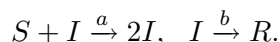
By the Law of Mass Action, the deterministic model is then given by

$$\begin{cases} \frac{ds}{dt} &= -asj \\ \frac{dj}{dt} &= asj - bj, \end{cases},$$

with $s = S/N$, $j = I/N$. Obviously, the percentage of recovered individuals follows directly from the second assumption: $r = R/N = 1 - s - j$. In the setting of figure 3.8, at $t = 0$ the fraction of infected individuals corresponds to 1%, while the remaining 99% are susceptible. The epidemic increases until reaching its maximum 0.1586 at time $t = 6.25$. At this time, the susceptible population falls below $c = b/a = 0.5$, as we can see in figure 3.8. The parameter c is called *basic parameter number* and has a crucial importance in the model, since it determines whether the disease will dye out $(c < 1)$ or whether it may become epidemic $(c > 1)$. The stochastic model is capable of giving a more realistic description of this value, as we will see next.

Proceeding as before, we define $\mathbf{z} := (s, j)$ and $D := \{(0,0), (-1,1), (0,-1)\}$, where $(-1,1)$ represents the infection of one individual and $(0,-1)$ represents one recovery. The rates of change of each class of individuals is given by $\beta_{(0,-1)} = bj$ and $\beta_{(-1,1)} = asj$. Consequently, we have
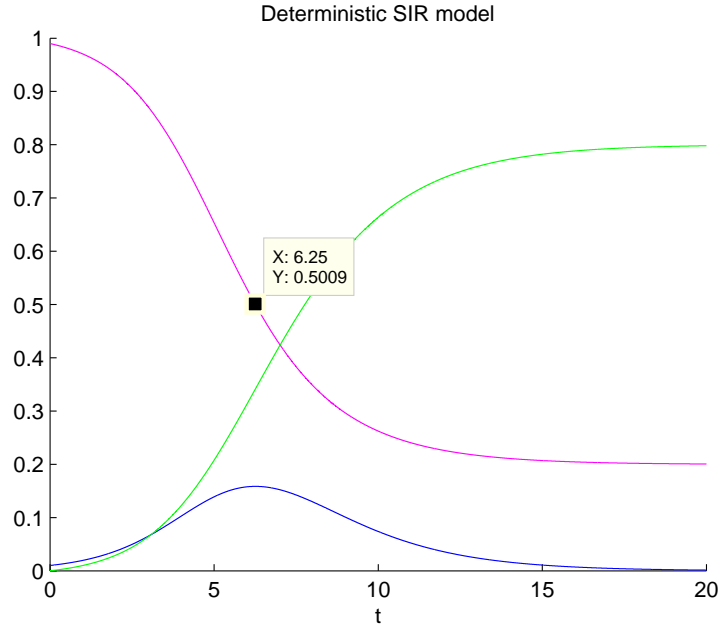
Figure 3.8: Evolution of the percentages of susceptible (magenta), infected (blue) and recovered (green) individuals within the time interval $[0, 20]$ in the deterministic SIR model (3.16). The time-step is 2^{-8}, the initial condition is $x_0 = (0.99, 0.01, 0)$ and the values of the parameters are $a = 1.4$, $b = 0.7$.

$$\mathbf{m}(t, \mathbf{z}) = \begin{bmatrix} -asj \\ asj - bj \end{bmatrix} \text{ and } \mathbf{Q}(t, \mathbf{z}) = \begin{bmatrix} asj & -asj \\ -asj & asj + bj \end{bmatrix}. \text{ By using the Cholesky}$$

decomposition, we obtain $\mathbf{Q} = \boldsymbol{\sigma}^T \boldsymbol{\sigma}$, with

$$\boldsymbol{\sigma} = \begin{bmatrix} \sqrt{|asj|} & 0 \\ -\sqrt{|asj|} & \sqrt{|bj|} \end{bmatrix}.$$

Notice that the decomposition $\mathbf{Q} = \boldsymbol{\sigma}^T \boldsymbol{\sigma}$ is not unique. Another decomposition for the matrix $\mathbf{Q}$ satisfying the previous relation is given by

$$\boldsymbol{\sigma}'(t, \mathbf{z}) = \frac{1}{d} \begin{bmatrix} asj + g & -asj \\ -asj & asj + bj + g \end{bmatrix},$$

where $g = \sqrt{det(\mathbf{Q})}$ and $d = \sqrt{|2asj + bj + 2g|}$. This ambiguity allows that different SDE models can be used to construct the same diffusion process, which seems a contradiction. In fact, once both systems are associated to the same Kolmogorov forward equation, the solutions of each system have the same distribution. Moreover, under the conditions of theorem 3, it can be proved that a sample path solution of one system is also a sample path solution of the other system and vice versa (see [13]). The systems may be considered as thus equivalent.
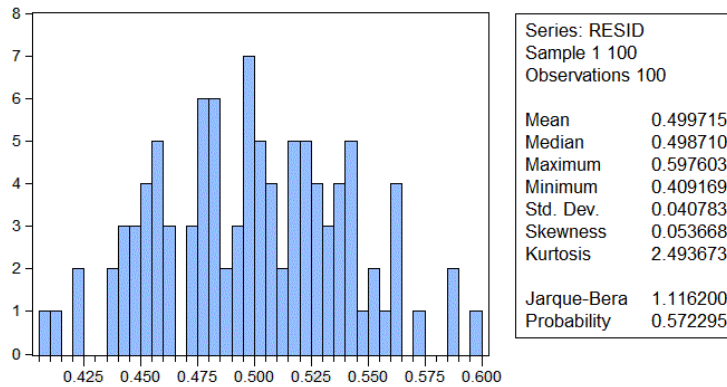
Figure 3.9: Histogram of W obtained from 100 sample paths that were generated from the model (3.16) within the time interval $[0, 20]$ for $n = 1000$ and the remaining parameters as in figure 3.10. The software used was *Eviews*.
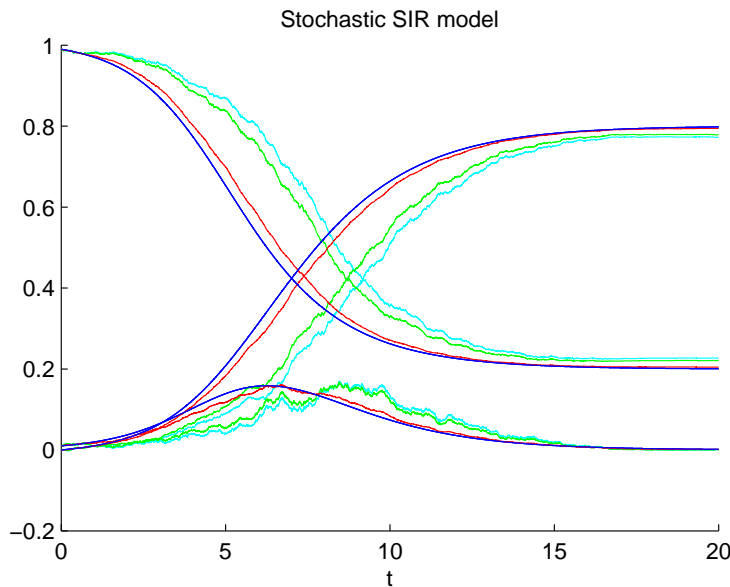

Figure 3.10: Three sample paths of the stochastic SIR model (with the same underlying Wiener sample path) describing the percentages of susceptible and infected individuals within the time interval $[0, 20]$ for $n = 80$ (cyan), $n = 100$ (green) and $n = 1000$ (red). The time-step is 2^{-8}, the initial condition is $(S_0, I_0, R_0) = (0.99, 0.01, 0)$ and the values of the parameters are $a = 1.4$, $b = 0.7$.

For a big enough population, and considering the matrix $\boldsymbol{\sigma}$, a stochastic SIR model is given by the system

$$\begin{cases} dS_t &= -aS_tI_tdt + \frac{1}{2\sqrt{n}}\sqrt{|aS_tI_t|}dW_t^{(1)} \\ dI_t &= (aS_tI_t - bI_t)dt - \frac{1}{2\sqrt{n}}\sqrt{|aS_tI_t|}dW_t^{(2)} + \frac{1}{2\sqrt{n}}\sqrt{|bI_t|}dW_{t,2} \end{cases}.$$

For simplicity, we suppress the index n on the random variables I_t and S_t. Some sample paths are plotted in figure 3.10 corresponding to different values of n. Due to the importance of the parameter $c = b/a$ referred before, we will explore the behavior of the correspondent random variable C in the stochastic model. From 100 sample paths we obtained the histogram of C presented in figure 3.9 and confirmed by Jarque-Béra test that C follows a gaussian distribution with mean 0.4997 and standard deviation 0.0408. It follows that about 99.7% of the values of C drawn from this model are within three standard deviations from the mean value, i.e., $[0.3273, 0.5721]$. Consequently, in this case, the disease will be eradicated, which is in accordance with the results illustrated in figure 3.10.

Chapter 4

Conclusion

As we could see, from the perspective of modellation, SDEs can provide a rich variety of models in the context of population dynamics. One can settle different stochastic models corresponding to the same phenomenon, depending on the assumption and on the random aspect we want to model. In this thesis, two different random effects were considered: the environmental and the demographic randomness.

In order to build SDE models, one can add a random *ad hoc* perturbation in the classic ODE model, assuming that the dynamics are partly driven by noise, or consider a diffusion approximation of a continuous-time Markov chain, frequently a simple birth and death process (see [10], [29] and [30]). The last method seems more natural as it is based on the relationship between diffusion processes and SDE solutions, while the first one is more vague, because of the wide variety of different environmental aspects existing in Nature. A combination of both would naturally be the best option from the modellation point of view.

In what concerns the analysis of the established models, different technics were used. When the model was linear, it was possible to find the exact solution. However, realistic models in population dynamics are usually nonlinear. In this case, two different approaches were followed: in the first one, we proceeded to the linearization of the deterministic model and introduced then the random environmental perturbation, which led to a linear stochastic differential system for describing the prey-predator interaction; in the second one, the Euler-Maruyama method was implemented in order to give an approximation for the sample paths solution of the original non-linear model.

The dynamic of one population following a logistic growth as well as the predator-prey interaction were analysed in the context of the two construction methods mentioned above. It was thus possible to obtain two different stochastic models for the same phenomenon. Lastly, a stochastic SIR epidemic model was established and the distribution of the basic parameter number was estimated by Monte-Carlo simula-

tions. Curiously, this parameter has apparently a normal distribution.

During the preparation of this thesis, we could realize that stochastic differential equations constitute a fundamental modelling tool in science and an extensive field of mathematical research. During the last decades, the subject has been developing in many exciting ways and in several fields (for instance in image denoising, see [31]), both theoretical and application specific. In the theoretical framework, we stress the study of the assymptotic behavior of their solutions and extinction times, which are particularly important in biology, parameters estimation (see [30] for an introduction), as well as the properties of numerical approximations of the solutions (see [32]) such as error estimation, consistency, stability and convergence. More recently, we also find in the literature new and more realistic modeling approaches that comprise stochastic partial differential equation systems, in order to obtain also a spatial distribution of the populations, or stochastic differential equations driven by Wiener and Poisson processes, known as SDEs with jumps, where the jump component can capture event-driven uncertainties.

Appendix A

Matlab code

A.1. Euler-Maruyama method for stochastic logistic growth model

```matlab
1  %Método de Euler−Maruyama
2  %Generating 100 sample paths of stochastic logistic growth
       model
3  %dX=mudt+sigmadW
4  clear all
5  delta=2^(−8);
6  T=20;
7  n=T/delta;
8  X0=1;
9  t0=0;
10 num_trajetorys=100;
11 Y=zeros(num_trajetorys,n);
12 Ye=zeros(1,n);
13 Yt=X0;
14 Yte=X0;
15 meantime=zeros(n,10);
16 extinction=0;
17 N=0.8;
18 k=5;
19 r=1;
20 hold on
21 for i=1:num_trajetorys
22     dW=sqrt(delta)*randn(1,n);
23     Yt=X0;
24         for j=1:n
```

```
25        mu=r*Yt*(1-Yt/k);
26          sigma=sqrt(abs(r*Yt*(1+Yt/k)))/(2*sqrt(N));
27          Yt=Yt+mu*delta+sigma*dW(j);%random solution
28          Y(i,j)=Yt;
29      end
30      bool=Y(i,:)>0;
31      if (sum(bool)==n)
32          plot([0:delta:T],[X0,Y(i,:)],'g');
33      else
34          meantime(:,extinction+1)=1-bool;
35          extinction=extinction+1;
36          time=find(1-bool, 1 );
37          plot([0:delta:time*delta],[X0,Y(i,1:time)],'r');
38      end
39  end
40  for j=1:n
41      mu=r*Yte*(1-Yte/k);
42      Yte=Yte+mu*delta;
43      Ye(j)=Yte;%deterministic solution
44  end
45  plot([0:delta:T],[X0,Ye]);
46  hold off
47  MeanTimeExtinction=find(meantime(:,1),1)*delta;
```

A.2. Euler-Maruyama method for stochastic SIR model

```
1  %Euler-Maruyama method for systems of SDEs
2  %Generating 100 sample paths of stochastic SIR model
3  %dX=mudt+sigmadW
4  clear all
5  delta=2^(-8);
6  T=20;
7  n=T/delta;
8  X0=[0.99,0.01]';
9  m=length(X0);
```

```matlab
10   t0=0;
11   num_trajetorys=100;
12   Y=zeros(num_trajetorys,n,m);
13   Ye=zeros(m,n);
14   C=zeros(1,num_trajetorys);
15   Yte=X0;
16   a=1.4;
17   b=0.7;
18   N=1000;
19   hold on
20   for i=1:num_trajetorys
21       dW=sqrt(delta)*randn(m,n);
22       Yt=X0;
23       for j=1:n
24           mu=[-a*Yt(1)*Yt(2);a*Yt(1)*Yt(2)-b*Yt(2)];
25           sigma=[sqrt(abs(a*Yt(1)*Yt(2))),0;-sqrt(abs(a*Yt(1)*
                   Yt(2))),sqrt(abs(b*Yt(2)))];
26           Yt=Yt+mu*delta+sigma/sqrt(N)/2*dW(:,j);%random
                   solution
27           Y(i,j,1)=Yt(1);
28           Y(i,j,2)=Yt(2);
29       end
30       [maximo,indice]=max(Y(i,:,2));
31       C(i)=Y(i,indice,1);
32       plot([0:delta:T],[X0(1),Y(i,:,1)],'c');%S
33       plot([0:delta:T],[X0(2),Y(i,:,2)],'c');%I
34       plot([0:delta:T],[1-X0(1)-X0(2),1-Y(i,:,1)-Y(i,:,2)],'c'
                   );%R
35   end
36   for j=1:n
37       mu=[-a*Yte(1)*Yte(2);a*Yte(1)*Yte(2)-b*Yte(2)];
38       Yte=Yte+mu*delta;
39       Ye(1,j)=Yte(1);%deterministic solution
40       Ye(2,j)=Yte(2);
```

```matlab
41  end
42  plot([0:delta:T],[X0(1),Ye(1,:)]);%S
43  plot([0:delta:T],[X0(2),Ye(2,:)]);%I
44  plot([0:delta:T],[1-X0(1)-X0(2),1-Ye(2,:)-Ye(1,:)]);%R
45  hold off
```

Bibliography

[1] R. Jarrow and P. Protter. *A Short History of Stochastic Integration and Mathematical Finance the Early Years, 1880-1970.* IMS Lecture Notes Monograph, Vol. 45, 1-17, 2004.

[2] L. Arnold. *Stochastic Differential Equations: Theory and Applications.* John Wiley and Sons, Inc., United States of America, 1974.

[3] P. E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations.* Springer-Verlag, Berlin, 1992.

[4] D. Bakstein V. Capasso. *An Introduction to Continuous-Time Stochastic Processes: Theory, Models and Applications to Finance, Biology and Medicine.* Birkhäser Boston, United States of America, 2nd edition, 2005.

[5] F. C. Klebaner. *Introduction to Stochastic Calculus with Applications.* Imperial College Press, Singapore, 2nd edition, 2005.

[6] D. Heath E. Platen. *A Benchmark Approach to Quantitative Finance.* Springer-Verlag, Berlin, 2006.

[7] C. A. Braumann. *Uma Introdução às Equações Diferenciais Estocásticas e Aplicações.* Sociedade Portuguesa de Matemática, Lisboa, 2005.

[8] B. Oksendal. *Stochastic Differential Equations: an introduction with applications.* Springer-Verlag, Berlin Heidelberg.

[9] M. Grigoriu. *Stochastic Calculus: applications in science and engineering.* Springer Science+Business Media, LLC, Birkhiiuser Boston.

[10] S. N. Ethier and T. Kurtz. *Markov Processes.* John Wiley & Sons, Inc., United States of America.

[11] P. K. Pollett. *Diffusion Approximations for Ecological Models.* Proc. International Congress on Modelling and Simulation society of Australia and New Zeland, 843-848, 2000.

[12] E. Allen L. Allen. *A comparison of three different stochastic population models with regard to persistence time.* Theoretical Population Biology, Vol. 64, 439-449, 2003.

[13] A. Arciniega E. Allen, L. Allen and P. Greenwood. *Construction of Equivalent Stochastic Differential Equation Models.* Stochastic Analisys and Applications, Vol. 26, pp. 274-297, United States of America, 2008.

[14] H. Kuo. *Introduction to Stochastic Integration.* Springer, United States of America, 2006.

[15] M. F. Gardner R. B. Ash. *Topics in Stochastic Processes.* Academic Press, 1975.

[16] Z. Schuss. *Theory and Applications of Stochastic Differential Equations.* John Wiley & Sons, New York, 1980.

[17] T. Wai-Yuan. *Stochastic Models with Applications to Genetics, AIDS and other Biomedical Systems.* World Scientific Publishing Co. Pte. Ltd., Singapore, 2002.

[18] S. Karlin and H. M. Taylor. *A Second Course in Stochastic Processes.* Academic Press, Inc., London, 1981.

[19] A. Friedman. *Partial Differential Equations of Parabolic Type.* Prentice-Hall, Inc., United States of America, 1964.

[20] L. B. Koralov and Y. G. Sinai. *Theory of Probability and Random Processes.* Springer, Berlin, 2007.

[21] M. Joannides F. Campillo and I. L. Valverde. *Estimation of the Parameters of a Stochastic Logistic Growth Model.* Cornell University Library, asXiv:1307.2217[math.ST], United States of America, 2013.

[22] D. F. Alves. *Uma Introdução à Resolução de Equações Diferenciais Estocásticas: Soluções Exatas e Aproximações Numéricas.* Tese de Mestrado, Universidade de Coimbra, Coimbra, 2012.

[23] A. Friedman. *Stochastic Differential Equations and Applications.* Academic Press, Inc., United Kingdom, 1975.

[24] T. Gard. *Introduction to Stochastic Diferential Equations*. Marcel Dekker, Inc., New York and Basel.

[25] T. Mikosch. *Elementary Stochastic Calculus with Finance in View*. World Scientific Publishing, 1998.

[26] F. Black and M. Scholes. *The Pricing of Options and Corporate Liabilities*. The Journal of Political Economy, Vol. 81, No. 3, pp. 637-654, United States of America, 1973.

[27] J.-M. Marin W. S. Kendall and C. P. Robert. *Confidence Bands for Brownian Motion and Applications to Monte Carlo Simulation*. Journal Statistics and Computing, V.7, No. 1, pp.1-10, United States of America, 2007.

[28] C. Dargatz. *A Diffusion Approximation for an Epidemic Model*. Technical report, Ludwig-Maximilian Universität München, 2007.

[29] L. Allen. *An Introduction to Stochastic Processes with Applications to Biology*. Taylor and Francis Group, LLC, United States of America, 2nd edition, 2011.

[30] E. Allen. *Modeling with Itô Stochastic Differential Equations*. Springer, United States of America, 2007.

[31] E. Zhizhina X. Descombes. *Image Denoising using Stochastic Differential Equations*. INRIA, No. 4814, 2003.

[32] D. Higham. *An Algorithmic Introduction to Numerical Simulation of Stochastic Differential Equations*. SIAM Review, Vol. 43, No.3, 525-546, Birkhiiuser Boston.