

# A Stochastic Differential Equation SIS Epidemic Model

A. Gray<sup>1</sup>, D. Greenhalgh<sup>1</sup>, L. Hu<sup>2</sup>, X. Mao<sup>1,\*</sup>, J. Pan<sup>1</sup>

<sup>1</sup> Department of Mathematics and Statistics,  
University of Strathclyde, Glasgow G1 1XH, U.K.

<sup>2</sup> Department of Applied Mathematics,  
Donghua University, Shanghai 201620, China.

## Abstract

In this paper we extend the classical SIS epidemic model from a deterministic framework to a stochastic one, and formulate it as a stochastic differential equation (SDE) for the number of infectious individuals  $I(t)$ . We then prove that this SDE has a unique global positive solution  $I(t)$  and establish conditions for extinction and persistence of  $I(t)$ . We discuss perturbation by stochastic noise. In the case of persistence we show the existence of a stationary distribution and derive expressions for its mean and variance. The results are illustrated by computer simulations, including two examples based on real life diseases.

**Key words:** SIS model, Brownian motion, stochastic differential equations, extinction, persistence, basic reproduction number, stationary distribution, gonorrhoea, pneumococcus.

## 1 Introduction

Epidemics are commonly modelled by using deterministic compartmental models where the population amongst whom the disease is spreading is divided into several classes, such as susceptible, infected and removed individuals. The classical Kermack-McKendrick model [15] is sometimes used for modelling common childhood diseases where a typical individual starts off susceptible, at some stage catches the disease and after a short infectious period becomes permanently immune. This is sometimes called the SIR (susceptible-infected-removed) model.

However some diseases, in particular some sexually transmitted and bacterial diseases, do not have permanent immunity. For these diseases individuals start off susceptible, at some stage catch the disease and after a short infectious period become susceptible again. There is no protective immunity. For these diseases SIS (susceptible-infected-susceptible) models are appropriate [14]. If  $S(t)$  denotes the number of susceptibles and  $I(t)$  the

---

\*Corresponding author and e-mail: x.mao@strath.ac.uk

number of infecteds at time  $t$ , then the differential equations which describe the spread of the disease are:

$$\begin{cases} \frac{dS(t)}{dt} = \mu N - \beta S(t)I(t) + \gamma I(t) - \mu S(t), \\ \frac{dI(t)}{dt} = \beta S(t)I(t) - (\mu + \gamma)I(t), \end{cases} \quad (1.1)$$

with initial values  $S_0 + I_0 = N$ .  $N$  is the total size of the population amongst whom the disease is spreading. Here  $\mu$  is the per capita death rate, and  $\gamma$  is the rate at which infected individuals become cured, so  $1/\gamma$  is the average infectious period.  $\beta$  is the disease transmission coefficient, so that  $\beta = \lambda/N$  where  $\lambda$  is the per capita disease contact rate.  $\lambda$  is the average number of adequate contacts of an infective per day. An adequate contact is one which is sufficient for the transmission of an infection if it is between a susceptible and an infected individual. This is one of the simplest possible epidemic models and because it is so simple it and its variants are commonly studied. For example, SIS models are discussed by Brauer et al. [2]. There are many other examples of SIS epidemic models in the literature.

In their excellent monograph Hethcote and Yorke [14] outline several mathematical models for gonorrhoea with increasing levels of complexity. The simplest of these is the above model (1.1) with  $\mu$  set to zero (i.e. no demography). As  $S + I = N$ , the total population is constant, the two models with and without demographics are equivalent; just replace  $\gamma$  in the model with no demography by  $\mu + \gamma$  to get the model with demography.

This is a very simple model for gonorrhoea. It assumes that the population is homogeneous (so is more suitable for a homosexual than a heterosexual population) and mixing is homogeneous, whereas in practice sexual mixing is extremely heterogeneous. It also ignores the small but non-zero disease incubation period and assumes that contact rates remain constant and do not vary seasonally. Lajmanovich and Yorke [17] and Nold [28] discuss heterogeneously mixing SIS epidemic models for the spread of gonorrhoea. The model (1.1) also ignores screening. Hethcote and Yorke define the contact number to be  $\bar{\sigma} = \lambda/\gamma$ , so in our model  $\bar{\sigma} = \beta N/(\mu + \gamma)$ .

If  $\tilde{I} = I/N$  is the fraction of the population infected at time  $t$  then they show that in our notation the solution is

$$I(t) = \begin{cases} \left[ \frac{\beta}{\beta N - \mu - \gamma} (1 - e^{-(\beta N - \mu - \gamma)t}) + \frac{1}{I_0} e^{-(\beta N - \mu - \gamma)t} \right]^{-1}, & \text{if } \frac{\beta N}{\mu + \gamma} \neq 1, \\ \left[ \beta t + \frac{1}{I_0} \right]^{-1}, & \text{if } \frac{\beta N}{\mu + \gamma} = 1. \end{cases} \quad (1.2)$$

It is straightforward to show that  $\bar{\sigma}$  has the usual interpretation as the basic reproduction number  $R_0$ . This is the expected number of secondary cases produced by a single newly infected individual entering a disease-free population at equilibrium. In such a situation each newly infected individual remains infectious for time  $1/(\mu + \gamma)$  and during this period infects  $\beta N$  of the  $N$  susceptibles present. Hence

$$R_0 = \bar{\sigma} = \frac{\beta N}{\mu + \gamma}. \quad (1.3)$$

From now on we denote this  $R_0$  by  $R_0^D$  to emphasise that it is  $R_0$  for the deterministic model. It is a straightforward consequence of equations (1.2) that [14]:

- If  $R_0^D \leq 1$ ,  $\lim_{t \rightarrow \infty} I(t) = 0$ .
- If  $R_0^D > 1$ ,  $\lim_{t \rightarrow \infty} I(t) = N \left(1 - \frac{1}{R_0^D}\right)$ .

Another disease for which it is possible to use an SIS model is pneumococcal carriage. *Streptococcus pneumoniae* (*S.pneumoniae*) is a bacterium commonly found in the throat of young children. When an individual carries pneumococcus the infectious carriage normally lasts around 7 weeks [30] and at the end of this carriage period the individual is susceptible again.  $R_0$  for pneumococcal carriage and transmission is 1.8-2.2 [32]. Lipsitch [20] discusses mathematical models for the transmission of *S.pneumoniae* with multiple serotypes and vaccination. Lamb, Greenhalgh and Robertson [18] discuss a mathematical model for the transmission of a single serotype of *S.pneumoniae* with vaccination. If we have only a single serotype and no vaccination then the disease can be modelled by equations (1.1). Other bacterial diseases, for example tuberculosis, can be modelled by SIS models [9].

Allen [1] discusses a stochastic model of the above SIS epidemic model. However this is done by constructing a stochastic differential equation (SDE) approximation to the continuous time Markov Chain model. The latter is obtained by assuming that events occurring at a constant rate in the deterministic model occur according to a Poisson process with the same rate. McCormack and Allen [27] construct a similar SDE approximation to an SIS multihost epidemic model and explore the stochastic and deterministic models numerically. However this is not the only way to introduce stochasticity into the model, and in this paper we explore an alternative, namely parameter perturbation.

We have previously used the technique of parameter perturbation to examine the effect of environmental stochasticity in a model of AIDS and condom use [5]. We found that the introduction of stochastic noise changes the basic reproduction number of the disease and can stabilise an otherwise unstable system. Ding, Xu and Hu [6] apply a similar technique to a simpler model of HIV/AIDS transmission. Other previous work on parameter perturbation in epidemic models seems to have concentrated on the SIR model. Tornatore, Buccellato and Vetro [29] discuss an SDE SIR system with and without delay with a similar parameter perturbation as we shall discuss here. The system for the SDE SIR model with no delay is

$$\begin{cases} d\tilde{S}(t) = [\mu - \tilde{\beta}\tilde{S}(t)\tilde{I}(t) - \mu\tilde{S}(t)]dt - \tilde{\sigma}\tilde{S}(t)\tilde{I}(t)dB(t), \\ d\tilde{I}(t) = [\tilde{\beta}\tilde{S}(t)\tilde{I}(t) - (\mu + \gamma)\tilde{I}(t)]dt + \tilde{\sigma}\tilde{S}(t)\tilde{I}(t)dB(t), \\ d\tilde{R}(t) = \gamma\tilde{I}(t) - \mu\tilde{R}(t), \end{cases} \quad (1.4)$$

where  $B(t)$  is a Brownian motion. Here  $\tilde{S}$ ,  $\tilde{I}$  and  $\tilde{R}$  denote respectively the susceptible, infected and removed fractions of the population, rather than absolute numbers, so that  $\tilde{\beta}$  in this model corresponds to  $\beta N$  in (1.1). They study the stability of the disease-free equilibrium (DFE). They find that

$$0 < \tilde{\beta} < \min\left\{\gamma + \mu - \frac{\tilde{\sigma}^2}{2}, 2\mu\right\}$$

is a sufficient condition for the asymptotic stability of the DFE. Their computer simulations for the SDE SIR model agree well with the analytical results and show that the

introduction of noise into the system raises the threshold to  $\mu + \gamma + (\tilde{\sigma}^2/2)$ , so if

$$\min\left\{\mu + \gamma - \frac{\tilde{\sigma}^2}{2}, 2\mu\right\} < \tilde{\beta} < \gamma + \mu + \frac{\tilde{\sigma}^2}{2}$$

then the DFE  $E_0 = (S(0), I(0), R(0)) = (1, 0, 0)$  is stable and the disease does not occur, whereas if  $\tilde{\beta} > \gamma + \mu + (\tilde{\sigma}^2/2)$  then the DFE is unstable. These results are similar to those of [5]. Chen and Li [4] study an SDE version of the SIR model both with and without delay, but introduce stochastic noise in a different way than we and Tornatore, Buccellato and Vetro [29] do. Lu [22] studies an SIRS (susceptible-infected-removed-susceptible) model which generalises the model of [29] by allowing removed individuals to become susceptible again. He extends their results by including the possibility that immunity is only temporary and improving the analytical bound on the sufficient condition for the stability of the DFE to  $\beta < \mu + \gamma - (\tilde{\sigma}^2/2)$ .

In summary the deterministic SIS epidemic model is one of the simplest possible epidemic models, which has applications to transmission of real-life diseases, such as pneumococcus, gonorrhoea and tuberculosis. It is important to include the effect of uncertainty in estimation of parameters such as the disease transmission coefficient. Whilst several papers study the effect of stochastic parameter perturbation on SIR and SIRS epidemic models, we are not aware of any literature addressing this issue in SIS epidemic models. This paper is an attempt to fill this gap.

## 2 Stochastic Differential Equation SIS Model

Throughout this paper, we let  $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$  be a complete probability space with a filtration  $\{\mathcal{F}_t\}_{t \geq 0}$  satisfying the usual conditions (i.e. it is increasing and right continuous while  $\mathcal{F}_0$  contains all  $\mathbb{P}$ -null sets) and we let  $B(t)$  be a scalar Brownian motion defined on the probability space. We use  $a \vee b$  to denote  $\max(a, b)$ ,  $a \wedge b$  to denote  $\min(a, b)$ , and a.s. to mean almost surely. We will denote the indicator function of a set  $G$  by  $I_G$ .

Let us now consider the second equation of (1.1). To establish the stochastic differential equation (SDE) model, we naturally re-write this equation in the differential form

$$dI(t) = [\beta S(t)I(t) - (\mu + \gamma)I(t)]dt. \quad (2.1)$$

Here  $[t, t + dt)$  is a small time interval and we use the notation  $d \cdot$  for the small change in any quantity over this time interval when we intend to consider it as an infinitesimal change, for example  $dI(t) = I(t + dt) - I(t)$  and the change  $dI(t)$  is described by (2.1). Consider the disease transmission coefficient  $\beta$  in the deterministic model. This can be thought of as the rate at which each infectious individual makes potentially infectious contacts with each other individual, where a potentially infectious contact will transmit the disease if the contact is made by an infectious individual with a susceptible individual. Thus the total number of new infections in the small time interval  $[t, t + dt)$  is

$$\beta S(t)I(t)dt$$

and a single infected individual makes

$$\beta dt$$

potentially infectious contacts with each other individual in the small time interval  $[t, t + dt)$ .

Now suppose that some stochastic environmental factor acts simultaneously on each individual in the population. In this case  $\beta$  changes to a random variable  $\tilde{\beta}$ . More precisely each infected individual makes

$$\tilde{\beta} dt = \beta dt + \sigma dB(t)$$

potentially infectious contacts with each other individual in  $[t, t + dt)$ . Here  $dB(t) = B(t + dt) - B(t)$  is the increment of a standard Brownian motion. Thus the number of potentially infectious contacts that a single infected individual makes with another individual in  $[t, t + dt)$  is normally distributed with mean  $\beta dt$  and variance  $\sigma^2 dt$ . Hence  $E(\tilde{\beta} dt) = \beta dt$  and  $\text{var}(\tilde{\beta} dt) = \sigma^2 dt$ . As  $\text{var}(\tilde{\beta} dt) \rightarrow 0$  as  $dt \rightarrow 0$  this is a biologically reasonable model. Indeed this is a well-established way of introducing stochastic environmental noise into biologically realistic population dynamic models. See [7, 10, 11, 12, 19, 22, 29] and many other references.

To motivate our assumption we argue as follows. Suppose that the number of potentially infectious contacts between an infectious individual and another individual in successive time intervals  $[t, t + T), [t + T, t + 2T), \dots, [t + (n - 1)T, t + nT)$  are independent, identically distributed random variables and  $n$  is very large. Then by the Central Limit Theorem the total number of potentially infectious contacts made in  $[t, t + nT)$  has approximately a normal distribution with mean  $n\mu_0$  and variance  $n\sigma_0^2$ , where  $\mu_0$  and  $\sigma_0^2$  are respectively the mean and variance of the underlying distribution in each of the separate time intervals of length  $T$ . Thus it is reasonable to assume that the total number of potentially infectious contacts has a normal distribution whose mean and variance scale as the total length of the time interval as in our assumptions.

Therefore we replace  $\beta dt$  in equation (2.1) by  $\tilde{\beta} dt = \beta dt + \sigma dB(t)$  to get

$$dI(t) = S(t)I(t)(\beta dt + \sigma dB(t)) - (\mu + \gamma)I(t)dt. \quad (2.2)$$

Note that  $\beta dt$  now denotes the mean of the stochastic number of potentially infectious contacts that an infected individual makes with another individual in the infinitesimally small time interval  $[t, t + dt)$ . Similarly, the first equation of (1.1) becomes another SDE. That is, the deterministic SIS model (1.1) becomes the Itô SDE

$$\begin{cases} dS(t) = [\mu N - \beta S(t)I(t) + \gamma I(t) - \mu S(t)]dt - \sigma S(t)I(t)dB(t), \\ dI(t) = [\beta S(t)I(t) - (\mu + \gamma)I(t)]dt + \sigma S(t)I(t)dB(t). \end{cases} \quad (2.3)$$

This SDE is called an SDE SIS model.

Given that  $S(t) + I(t) = N$ , it is sufficient to study the SDE for  $I(t)$

$$dI(t) = I(t) \left( [\beta N - \mu - \gamma - \beta I(t)]dt + \sigma(N - I(t))dB(t) \right) \quad (2.4)$$

with initial value  $I(0) = I_0 \in (0, N)$ . In the following sections we will concentrate on this SDE only.

### 3 Existence of Unique Positive Solution

The SDE SIS model (2.4) is a special SDE. In order for the model to make sense, we need to show at least that this SDE SIS model does not only have a unique global solution but also the solution will remain within  $(0, N)$  whenever it starts from there. The existing general existence-and-uniqueness theorem on SDEs (see e.g. [23, 24, 25]) is not applicable to this special SDE in order to guarantee these properties. It is therefore necessary to establish such a new theory.

**Theorem 3.1** *For any given initial value  $I(0) = I_0 \in (0, N)$ , the SDE (2.4) has a unique global positive solution  $I(t) \in (0, N)$  for all  $t \geq 0$  with probability one, namely*

$$\mathbb{P}\{I(t) \in (0, N) \text{ for all } t \geq 0\} = 1.$$

*Proof.* Regarding equation (2.4) as an SDE on  $\mathbb{R}$ , we see that its coefficients are locally Lipschitz continuous. It is known (see e.g. [23, 24, 25]) that for any given initial value  $S_0 \in (0, N)$  there is a unique maximal local solution  $I(t)$  on  $t \in [0, \tau_e)$ , where  $\tau_e$  is the explosion time. Let  $k_0 > 0$  be sufficiently large for  $1/k_0 < I_0 < N - (1/k_0)$ . For each integer  $k \geq k_0$ , define the stopping time

$$\tau_k = \inf\{t \in [0, \tau_e) : I(t) \notin (1/k, N - (1/k))\},$$

where throughout this paper we set  $\inf \emptyset = \infty$  (as usual,  $\emptyset$  = the empty set). Clearly,  $\tau_k$  is increasing as  $k \rightarrow \infty$ . Set  $\tau_\infty = \lim_{k \rightarrow \infty} \tau_k$ , whence  $\tau_\infty \leq \tau_e$  a.s. If we can show that  $\tau_\infty = \infty$  a.s., then  $\tau_e = \infty$  a.s. and  $I(t) \in (0, N)$  a.s. for all  $t \geq 0$ . In other words, to complete the proof all we need to show is that  $\tau_\infty = \infty$  a.s. If this statement is false, then there is a pair of constants  $T > 0$  and  $\varepsilon \in (0, 1)$  such that

$$\mathbb{P}\{\tau_\infty \leq T\} > \varepsilon.$$

Hence there is an integer  $k_1 \geq k_0$  such that

$$\mathbb{P}\{\tau_k \leq T\} \geq \varepsilon \quad \text{for all } k \geq k_1. \quad (3.1)$$

Define a function  $V : (0, N) \rightarrow \mathbb{R}_+$  by

$$V(x) = \frac{1}{x} + \frac{1}{N - x}.$$

By the Itô formula (see e.g. [25]), we have, for any  $t \in [0, T]$  and  $k \geq k_1$ ,

$$\mathbb{E}V(I(t \wedge \tau_k)) = V(I_0) + \mathbb{E} \int_0^{t \wedge \tau_k} LV(I(s)) ds, \quad (3.2)$$

where  $LV : (0, N) \rightarrow \mathbb{R}$  is defined by

$$\begin{aligned} LV(x) &= x \left( -\frac{1}{x^2} + \frac{1}{(N-x)^2} \right) [\beta N - \mu - \gamma - \beta x] \\ &+ \sigma^2 x^2 (N-x)^2 \left( \frac{1}{x^3} + \frac{1}{(N-x)^3} \right). \end{aligned} \quad (3.3)$$

It is easy to show that

$$LV(x) \leq \frac{\mu + \gamma}{x} + \frac{\beta N}{N - x} + \sigma^2 N^2 \left( \frac{1}{x} + \frac{1}{N - x} \right) \leq CV(x), \quad (3.4)$$

where  $C = (\mu + \gamma) \vee (\beta N) + \sigma^2 N^2$ . Substituting this into (3.2) we get

$$\mathbb{E}V(I(t \wedge \tau_k)) \leq V(I_0) + \mathbb{E} \int_0^{t \wedge \tau_k} CV(I(s)) ds \leq V(I_0) + C \int_0^t \mathbb{E}V(I(s \wedge \tau_k)) ds.$$

The Gronwall inequality yields that

$$\mathbb{E}V(I(T \wedge \tau_k)) \leq V(I_0)e^{CT}. \quad (3.5)$$

Set  $\Omega_k = \{\tau_k \leq T\}$  for  $k \geq k_1$  and, by (3.1),  $\mathbb{P}(\Omega_k) \geq \varepsilon$ . Note that for every  $\omega \in \Omega_k$ ,  $I(\tau_k, \omega)$  equals either  $1/k$  or  $N - (1/k)$ , and hence

$$V(I(\tau_k, \omega)) \geq k.$$

It then follows from (3.5) that

$$V(I_0)e^{CT} \geq \mathbb{E} \left[ I_{\Omega_k}(\omega) V(I(\tau_k, \omega)) \right] \geq k \mathbb{P}(\Omega_k) \geq \varepsilon k.$$

Letting  $k \rightarrow \infty$  leads to the contradiction

$$\infty > V(I_0)e^{CT} = \infty,$$

so we must therefore have  $\tau_\infty = \infty$  a.s., whence the proof is complete.

## 4 Extinction

In the study of population systems, extinction and persistence are two of most important issues. We will discuss the extinction of the SDE SIS model (2.4) in this section but leave its persistence to the next section.

**Theorem 4.1** *If*

$$R_0^S := R_0^D - \frac{\sigma^2 N^2}{2(\mu + \gamma)} = \frac{\beta N}{\mu + \gamma} - \frac{\sigma^2 N^2}{2(\mu + \gamma)} < 1 \quad \text{and} \quad \sigma^2 \leq \frac{\beta}{N}, \quad (4.1)$$

*then for any given initial value  $I(0) = I_0 \in (0, N)$ , the solution of the SDE (2.4) obeys*

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \log(I(t)) \leq \beta N - \mu - \gamma - 0.5\sigma^2 N^2 < 0 \quad \text{a.s.}, \quad (4.2)$$

*namely,  $I(t)$  tends to zero exponentially almost surely. In other words, the disease dies out with probability one.*

*Proof.* By the Itô formula, we have

$$\log(I(t)) = \log(I_0) + \int_0^t f(I(s))ds + \int_0^t \sigma(N - I(s))dB(s), \quad (4.3)$$

where  $f : \mathbb{R} \rightarrow \mathbb{R}$  is defined by

$$f(x) = \beta N - \mu - \gamma - \beta x - 0.5\sigma^2(N - x)^2. \quad (4.4)$$

However, under condition (4.1), we have

$$\begin{aligned} f(I(s)) &= \beta N - \mu - \gamma - 0.5\sigma^2 N^2 - (\beta - \sigma^2 N)I(s) - 0.5\sigma^2 I^2(s), \\ &\leq \beta N - \mu - \gamma - 0.5\sigma^2 N^2, \end{aligned}$$

for  $I(s) \in (0, N)$ . It then follows from (4.3) that

$$\log(I(t)) \leq \log(I_0) + (\beta N - \mu - \gamma - 0.5\sigma^2 N^2)t + \int_0^t \sigma(N - I(s))dB(s). \quad (4.5)$$

This implies

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \log(I(t)) \leq \beta N - \mu - \gamma - 0.5\sigma^2 N^2 + \limsup_{t \rightarrow \infty} \frac{1}{t} \int_0^t \sigma(N - I(s))dB(s) \quad a.s. \quad (4.6)$$

But by the large number theorem for martingales (see e.g. [25]), we have

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \int_0^t \sigma(N - I(s))dB(s) = 0 \quad a.s.$$

We therefore obtain the desired assertion (4.2) from (4.6).

It is useful to observe that in the classical deterministic SIS model (1.1),  $I(t)$  tends to 0 if and only if  $R_0^D \leq 1$ ; while in the SDE SIS model (2.3),  $I(t)$  tends to 0 if  $R_0^S = R_0^D - 0.5\sigma^2 N^2 / (\mu + \gamma) < 1$  and  $\sigma^2 \leq \beta/N$ . In other words, the conditions for  $I(t)$  to become extinct in the SDE SIS model are weaker than in the classical deterministic SIS model. The following example illustrates this result more explicitly:

**Example 4.2** Throughout the paper we shall assume that the unit of time is one day and the population sizes are measured in units of 1 million, unless otherwise stated. With these units assume that the system parameters are given by

$$\beta = 0.5, \quad N = 100, \quad \mu = 20, \quad \gamma = 25, \quad \sigma = 0.035.$$

So the SDE SIS model (2.4) becomes

$$dI(t) = I(t) \left( [5 - 0.5I(t)]dt + 0.035(100 - I(t))dB(t) \right). \quad (4.7)$$

Noting that

$$R_0^S = \frac{\beta N}{\mu + \gamma} - \frac{\sigma^2 N^2}{2(\mu + \gamma)} = \frac{50}{45} - \frac{12.25}{90} = 1.111 - 0.136 < 1,$$

$$\text{and } \sigma^2 = 0.001225 \leq \frac{\beta}{N} = 0.005,$$



we can therefore conclude, by Theorem 4.1, that for any initial value  $I(0) = I_0 \in (0, 100)$ , the solution of (4.7) obeys

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \log(I(t)) \leq -1.125 \quad a.s.$$

That is,  $I(t)$  will tend to zero exponentially with probability one.

On the other hand, the corresponding deterministic SIS model (1.1) becomes

$$\frac{dI(t)}{dt} = I(t)(5 - 0.5I(t)). \quad (4.8)$$

For  $R_0^D > 1$ , it is known that, for any initial value  $I(0) = I_0 \in (0, 100)$ , this solution has the property

$$\lim_{t \rightarrow \infty} I(t) = N \left( 1 - \frac{1}{R_0^D} \right) = 10 \quad (\text{see section 1}).$$

The computer simulations in Figure 4.1, using the Euler Maruyama (EM) method (see e.g. [16, 25, 26]), support these results clearly, illustrating extinction of the disease.

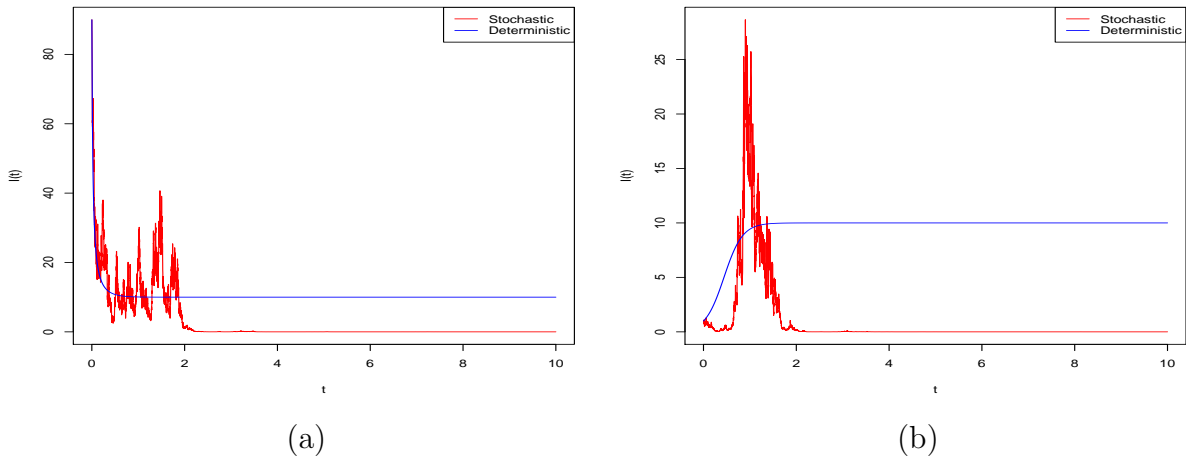


Figure 4.1: Computer simulation of the path  $I(t)$  for the SDE SIS model (4.7) and its corresponding deterministic SIS model (4.8), using the EM method with step size  $\Delta = 0.001$ , using initial values (a)  $I(0) = 90$  and (b)  $I(0) = 1$ .

In Theorem 4.1 we require the noise intensity  $\sigma^2 \leq \beta/N$ . The following theorem covers the case when  $\sigma^2 > \beta/N$ :

**Theorem 4.3** *If*

$$\sigma^2 > \frac{\beta}{N} \vee \frac{\beta^2}{2(\mu + \gamma)}, \quad (4.9)$$

*then for any given initial value  $I(0) = I_0 \in (0, N)$ , the solution of the SDE SIS model (2.4) obeys*

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \log(I(t)) \leq -\mu - \gamma + \frac{\beta^2}{2\sigma^2} < 0 \quad a.s., \quad (4.10)$$

namely,  $I(t)$  tends to zero exponentially almost surely. In other words, the disease dies out with probability one.

*Proof.* We use the same notation as in the proof of Theorem 4.1. It is easy to see that the quadratic function  $f : \mathbb{R} \rightarrow \mathbb{R}$  defined by (4.4) takes its maximum value  $f(\hat{x})$  at

$$x = \hat{x} := \frac{\sigma^2 N - \beta}{\sigma^2}.$$

By condition (4.9), it is easy to see that  $\hat{x} \in (0, N)$ . Compute

$$f(\hat{x}) = \beta N - \mu - \gamma - 0.5\sigma^2 N^2 + \frac{(\sigma^2 N - \beta)^2}{2\sigma^2} = -\mu - \gamma + \frac{\beta^2}{2\sigma^2},$$

which is negative by condition (4.9). It therefore follows from (4.3) that

$$\log(I(t)) \leq \log(I_0) + f(\hat{x})t + \int_0^t \sigma(N - I(s))dB(s).$$

This implies, in the same way as in the proof of Theorem 4.1, that

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \log(I(t)) \leq f(\hat{x}) \quad a.s.,$$

as required. The proof is hence complete.

Note that condition (4.9) implies that  $R_0^S \leq 1$ .

**Example 4.4** We keep the system parameters the same as in Example 4.2 but let  $\sigma = 0.08$ , so the SDE SIS model (2.4) becomes

$$dI(t) = I(t) \left( [5 - 0.5I(t)]dt + 0.08(100 - I(t))dB(t) \right). \quad (4.11)$$

It is easy to verify that the system parameters obey condition (4.9). We can therefore conclude, by Theorem 4.3, that for any initial value  $I(0) = I_0 \in (0, 100)$ , the solution of (4.11) obeys

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \log(I(t)) \leq -45 + \frac{0.5^2}{2 \times 0.08^2} = -25.4688 \quad a.s.$$

That is,  $I(t)$  will tend to zero exponentially with probability one. The computer simulations shown in Figure 4.2 support these results clearly.

## 5 Persistence

**Theorem 5.1** *If*

$$R_0^S := \frac{\beta N}{\mu + \gamma} - \frac{\sigma^2 N^2}{2(\mu + \gamma)} > 1 \quad (5.1)$$

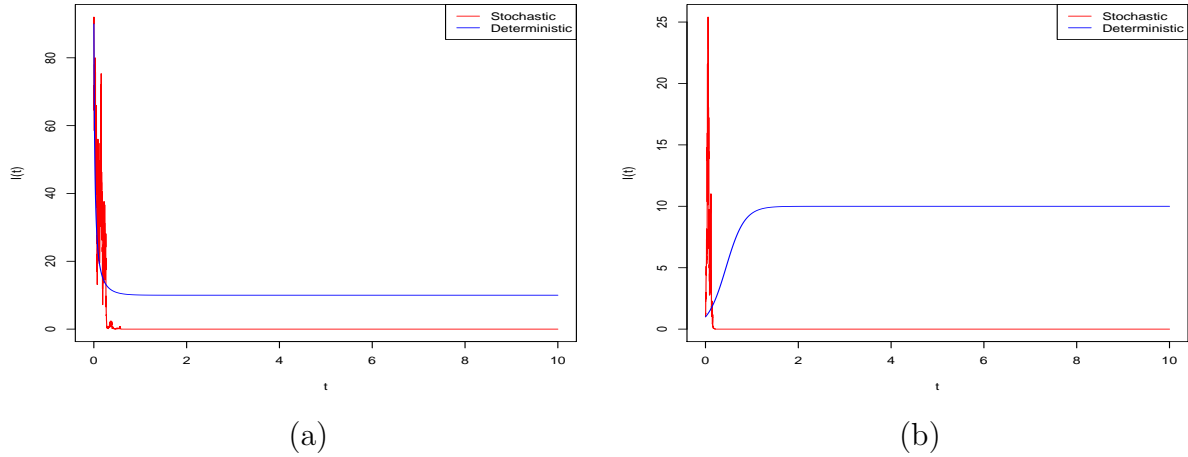


Figure 4.2: Computer simulation of the path  $I(t)$  for the SDE SIS model (4.11) and its corresponding deterministic SIS model (4.8), using the EM method with step size  $\Delta = 0.001$ , with initial values (a)  $I(0) = 90$  and (b)  $I(0) = 1$ .

then for any given initial value  $I(0) = I_0 \in (0, N)$ , the solution of the SDE SIS model (2.4) obeys

$$\limsup_{t \rightarrow \infty} I(t) \geq \xi \quad a.s. \quad (5.2)$$

and

$$\liminf_{t \rightarrow \infty} I(t) \leq \xi \quad a.s. \quad (5.3)$$

where

$$\xi = \frac{1}{\sigma^2} \left( \sqrt{\beta^2 - 2\sigma^2(\mu + \gamma)} - (\beta - \sigma^2 N) \right) \quad (5.4)$$

which is the unique root in  $(0, N)$  of

$$\beta N - \mu - \gamma - \beta \xi - 0.5\sigma^2(N - \xi)^2 = 0. \quad (5.5)$$

That is,  $I(t)$  will rise to or above the level  $\xi$  infinitely often with probability one.

*Proof.* Recall the definition (4.4) of function  $f : \mathbb{R} \rightarrow \mathbb{R}$ . By condition (5.1), it is easy to see that equation  $f(x) = 0$  has a positive root and a negative root. The positive one is

$$\begin{aligned} & \frac{1}{\sigma^2} \left( \sqrt{(\beta - \sigma^2 N)^2 + 2\sigma^2(\beta N - \mu - \gamma - 0.5\sigma^2 N^2)} - (\beta - \sigma^2 N) \right) \\ &= \frac{1}{\sigma^2} \left( \sqrt{\beta^2 - 2\sigma^2(\mu + \gamma)} - (\beta - \sigma^2 N) \right) = \xi. \end{aligned}$$

Noting that

$$f(0) = \beta N - \mu - \gamma - 0.5\sigma^2 N^2 > 0 \quad \text{and} \quad f(N) = -\mu - \gamma < 0,$$

we see that  $\xi \in (0, N)$  and

$$f(x) > 0 \quad \text{is strictly increasing on } x \in (0, 0 \vee \hat{x}), \quad (5.6)$$

$$f(x) > 0 \quad \text{is strictly decreasing on } x \in (0 \vee \hat{x}, \xi), \quad (5.7)$$

while

$$f(x) < 0 \quad \text{is strictly decreasing on } x \in (\xi, N). \quad (5.8)$$

We now begin to prove assertion (5.2). If it is not true, then there is a sufficiently small  $\varepsilon \in (0, 1)$  such that

$$\mathbb{P}(\Omega_1) > \varepsilon, \quad (5.9)$$

where  $\Omega_1 = \{\limsup_{t \rightarrow \infty} I(t) \leq \xi - 2\varepsilon\}$ . Hence, for every  $\omega \in \Omega_1$ , there is a  $T = T(\omega) > 0$  such that

$$I(t, \omega) \leq \xi - \varepsilon \quad \text{whenever } t \geq T(\omega). \quad (5.10)$$

Clearly we may choose  $\varepsilon$  so small (if necessary reduce it) that  $f(0) > f(\xi - \varepsilon)$ . It therefore follows from (5.6), (5.7) and (5.10) that

$$f(I(t, \omega)) \geq f(\xi - \varepsilon) \quad \text{whenever } t \geq T(\omega). \quad (5.11)$$

Moreover, by the large number theorem for martingales, there is a  $\Omega_2 \subset \Omega$  with  $\mathbb{P}(\Omega_2) = 1$  such that for every  $\omega \in \Omega_2$ ,

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t \sigma(N - I(s, \omega)) dB(s, \omega) = 0. \quad (5.12)$$

Now, fix any  $\omega \in \Omega_1 \cap \Omega_2$ . It then follows from (4.3) and (5.11) that, for  $t \geq T(\omega)$ ,

$$\begin{aligned} \log(I(t, \omega)) &\geq \log(I_0) + \int_0^{T(\omega)} f(I(s, \omega)) ds + f(\xi - \varepsilon)(t - T(\omega)) \\ &+ \int_0^t \sigma(N - I(s, \omega)) dB(s, \omega). \end{aligned} \quad (5.13)$$

This yields

$$\liminf_{t \rightarrow \infty} \frac{1}{t} \log(I(t, \omega)) \geq f(\xi - \varepsilon) > 0,$$

whence

$$\lim_{t \rightarrow \infty} I(t, \omega) = \infty.$$

But this contradicts (5.10). We therefore must have the desired assertion (5.2).

Let us now prove assertion (5.3). If it were not true, then there is a sufficiently small  $\delta \in (0, 1)$  such that

$$\mathbb{P}(\Omega_3) > \delta, \quad (5.14)$$

where  $\Omega_3 = \{\liminf_{t \rightarrow \infty} I(t) \geq \xi + 2\delta\}$ . Hence, for every  $\omega \in \Omega_3$ , there is a  $\tau = \tau(\omega) > 0$  such that

$$I(t, \omega) \geq \xi + \delta \quad \text{whenever } t \geq \tau(\omega). \quad (5.15)$$

Now, fix any  $\omega \in \Omega_3 \cap \Omega_2$ . It then follows from (4.3) and (5.8) that, for  $t \geq \tau(\omega)$ ,

$$\begin{aligned} \log(I(t, \omega)) &\leq \log(I_0) + \int_0^{\tau(\omega)} f(I(s, \omega)) ds + f(\xi + \delta)(t - \tau(\omega)) \\ &+ \int_0^t \sigma(N - I(s, \omega)) dB(s, \omega). \end{aligned} \quad (5.16)$$

This, together with (5.12), yields

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \log(I(t, \omega)) \leq f(\xi + \delta) < 0,$$

whence

$$\lim_{t \rightarrow \infty} I(t, \omega) = 0.$$

But this contradicts (5.15). We therefore must have the desired assertion (5.3). The proof is therefore complete.

**Example 5.2** Assume that the system parameters are given by

$$\beta = 0.5, \quad N = 100, \quad \mu = 20, \quad \gamma = 25, \quad \sigma = 0.03.$$

That is, we keep all the system parameters the same as in Example 4.2 except  $\sigma$  is reduced to 0.03 from 0.035. So the SDE SIS model (2.4) becomes

$$dI(t) = I(t) \left( [5 - 0.5I(t)]dt + 0.03(100 - I(t))dB(t) \right). \quad (5.17)$$

Noting that

$$R_0^S = \frac{\beta N}{\mu + \gamma} - \frac{\sigma^2 N^2}{2(\mu + \gamma)} = \frac{50}{45} - 0.1 > 1,$$

we compute

$$\xi = \frac{1}{\sigma^2} \left( \sqrt{\beta^2 - 2\sigma^2(\mu + \gamma)} - (\beta - \sigma^2 N) \right) = 1.2179.$$

We can therefore conclude, by Theorem 5.1, that for any initial value  $I(0) = I_0 \in (0, 100)$ , the solution of (5.17) obeys

$$\liminf_{t \rightarrow \infty} I(t) \leq 1.2179 \leq \limsup_{t \rightarrow \infty} I(t) \quad a.s.$$

In comparison, we recall that the solution of the corresponding deterministic SIS model (1.2) has the property

$$\lim_{t \rightarrow \infty} I(t) = N \left( 1 - \frac{1}{R_0^D} \right) = 10.$$

The computer simulations in Figure 5.1 support these results clearly, showing fluctuation around the level 1.2179.

**Example 5.3** To further illustrate the effect of the noise intensity  $\sigma$  on the SDE SIS model, we keep all the parameters in Example 5.2 unchanged but reduce  $\sigma$  to  $\sigma = 0.01$ , namely we have

$$\beta = 0.5, \quad N = 100, \quad \mu = 20, \quad \gamma = 25, \quad \sigma = 0.01.$$

So the SDE SIS model (2.4) now becomes

$$dI(t) = I(t) \left( [5 - 0.5I(t)]dt + 0.01(100 - I(t))dB(t) \right). \quad (5.18)$$

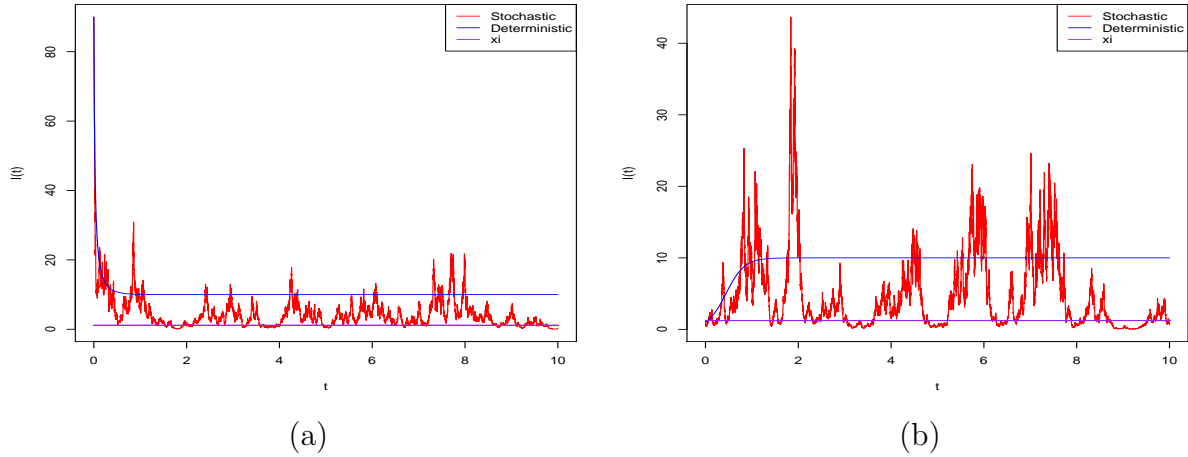


Figure 5.1: Computer simulation of the path  $I(t)$  for the SDE SIS model (5.17) and its corresponding deterministic SIS model (4.8), using the EM method with step size  $\Delta = 0.001$  and initial values (a)  $I(0) = 90$  and (b)  $I(0) = 1$ .

Noting that

$$R_0^S = \frac{\beta N}{\mu + \gamma} - \frac{\sigma^2 N^2}{2(\mu + \gamma)} = \frac{50}{45} - 0.011 > 1,$$

we compute

$$\xi = \frac{1}{\sigma^2} \left( \sqrt{\beta^2 - 2\sigma^2(\mu + \gamma)} - (\beta - \sigma^2 N) \right) = 9.1751.$$

We can therefore conclude, by Theorem 5.1, that for any initial value  $I(0) = I_0 \in (0, 100)$ , the solution of (5.18) obeys

$$\liminf_{t \rightarrow \infty} I(t) \leq 9.1751 \leq \limsup_{t \rightarrow \infty} I(t) \quad a.s.$$

The computer simulations in Figure 5.2 support these results clearly, illustrating persistence and that the effect of reducing the standard deviation  $\sigma$  is to increase the level  $\xi$ , which becomes closer to the limiting value of the corresponding deterministic SIS model.

These computer simulations indicate strongly that  $\xi$  will increase to  $N(1 - (1/R_0^D))$ , which is the equilibrium state of the deterministic SIS model (1.1), as the noise intensity  $\sigma$  decreases to zero. This is of course not surprising. The following proposition describes this situation rigorously:

**Proposition 5.4** *Assume that  $R_0^S > 1$  and regard  $\xi$  defined by (5.4) as a function of  $\sigma$  for*

$$0 < \sigma < \frac{\sqrt{2(\beta N - \mu - \gamma)}}{N} := \hat{\sigma}.$$

*Then  $\xi$  is strictly decreasing and*

$$\lim_{\sigma \rightarrow 0} \xi = N \left( 1 - \frac{1}{R_0^D} \right) \quad \text{and} \quad \lim_{\sigma \rightarrow \hat{\sigma}} \xi = \begin{cases} 0, & \text{if } 1 \leq R_0^D \leq 2, \\ N \left( \frac{R_0^D - 2}{R_0^D - 1} \right), & \text{if } R_0^D > 2. \end{cases}$$

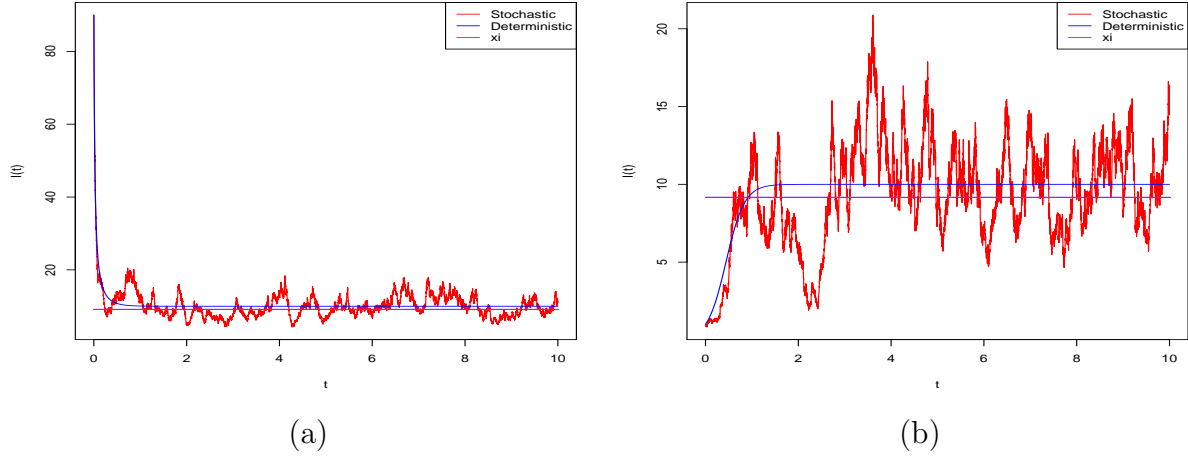


Figure 5.2: Computer simulation of the path  $I(t)$  for the SDE SIS model (5.18) and its corresponding deterministic SIS model (4.8), using the EM method with step size  $\Delta = 0.001$  and initial values (a)  $I(0) = 90$  and (b)  $I(0) = 1$ .

*Proof.* Compute

$$\begin{aligned}
\frac{d\xi}{d\sigma} &= \frac{1}{2} \left( \frac{\beta^2}{\sigma^4} - \frac{2(\mu + \gamma)}{\sigma^2} \right)^{-\frac{1}{2}} \left( -\frac{4\beta^2}{\sigma^5} + \frac{4(\mu + \gamma)}{\sigma^3} \right) + \frac{2\beta}{\sigma^3}, \\
&= \frac{-2\beta^2 + 2\sigma^2(\mu + \gamma) + 2\beta\sqrt{\beta^2 - 2\sigma^2(\mu + \gamma)}}{\sigma^3\sqrt{\beta^2 - 2\sigma^2(\mu + \gamma)}}, \\
&= \frac{-(\sqrt{\beta^2 - 2\sigma^2(\mu + \gamma)} - \beta)^2}{\sigma^3\sqrt{\beta^2 - 2\sigma^2(\mu + \gamma)}}.
\end{aligned}$$

Since  $\sigma > 0$  we have  $\sqrt{\beta^2 - 2\sigma^2(\mu + \gamma)} - \beta \neq 0$ . We therefore have that  $\frac{d\xi}{d\sigma} < 0$  which implies that  $\xi$  is strictly decreasing as  $\sigma$  increases. Moreover, by the well-known L'Hopital's rule,

$$\lim_{\sigma \rightarrow 0} \xi = \lim_{\sigma \rightarrow 0} [-(\beta^2 - 2\sigma^2(\mu + \gamma))^{-\frac{1}{2}}(\mu + \gamma) + N] = -\frac{\mu + \gamma}{\beta} + N = N \left( 1 - \frac{1}{R_0^D} \right)$$

as desired. Furthermore, it is obvious that

$$\lim_{\sigma \rightarrow \hat{\sigma}} \xi = \frac{\sqrt{\beta^2 - 2\hat{\sigma}^2(\mu + \gamma)} - \beta + \hat{\sigma}^2 N}{\hat{\sigma}^2}.$$

The numerator equals

$$\frac{|\beta N - 2(\mu + \gamma)|}{N} + \frac{\beta N - 2(\mu + \gamma)}{N},$$

so if  $1 \leq R_0^D \leq 2$  we have  $\lim_{\sigma \rightarrow \hat{\sigma}} \xi = 0$ , but if  $R_0^D > 2$  we have  $\lim_{\sigma \rightarrow \hat{\sigma}} \xi = N \left( \frac{R_0^D - 2}{R_0^D - 1} \right)$ . The proof is complete.

Note that Proposition 5.4 implies that for  $R_0^S > 1$ ,  $\xi$  lies between the deterministic equilibrium value (and limiting value)

$$N\left(1 - \frac{1}{R_0^D}\right)$$

for  $I(t)$  and

$$\max\left(0, N\left(1 - \frac{1}{R_0^D - 1}\right)\right).$$

If  $R_0^D$  is large then  $\xi$  will be close to but beneath the deterministic equilibrium value for  $I(t)$ .

**Example 5.5** The computer simulations of the solution to the SDE SIS model in the persistent case also suggest for higher  $\sigma$  that the distribution of the solution is skewed, as there are larger oscillations above  $\xi$  than below  $\xi$ , while for lower  $\sigma$  the oscillations about  $\xi$  appear to be more symmetrically distributed. This is confirmed by the histograms in Figure 5.3, showing the distribution of  $I(t)$  in the case of  $\beta = 0.5$ ,  $N = 100$ ,  $\mu = 20$ ,  $\gamma = 25$ , and  $\sigma = 0.03, 0.02, 0.01, 0.005$  and  $0.001$ , respectively. The simulations were run for 100,000 iterations with step size  $\Delta = 0.001$ , i.e. for 100 time steps, and the first 90,000 iterations were discarded to allow for  $I(t)$  to reach its recurrent level. The distribution is positively skewed for  $\sigma = 0.03$  and  $0.02$ , but as  $\sigma$  reduces it becomes more symmetric about  $\xi$ , so that the distribution appears closer to a normal distribution. The corresponding sample skewness coefficients are 4.8774, 0.9319, 0.1692, 0.1798, and  $-0.2106$ .

Testing these data for normality, all tests used were highly significant, conclusively rejecting normality in all cases. This is not surprising in view of the very large sample sizes (10,000), as even moderate deviations from the tested distribution will be significant, however the normal QQ plots in Figure 5.4 suggest that these data are not far from being normally distributed for smaller values of  $\sigma$ .

## 6 Stationary Distribution

In the previous section we showed that  $I(t)$  will fluctuate around the level  $\xi \in (0, N)$  with probability 1 when  $R_0^S > 1$ . The computer simulations also strongly indicate that the SDE SIS model (2.4) has a stationary distribution. To be more precise, let  $P_{I_0,t}(\cdot)$  denote the probability measure induced by  $I(t)$  with initial value  $I(0) = I_0$ , that is

$$P_{I_0,t}(A) = \mathbb{P}(I(t) \in A), \quad A \in \mathcal{B}(0, N),$$

where  $\mathcal{B}(0, N)$  is the  $\sigma$ -algebra of all the Borel sets  $A \subset (0, N)$ . If there is a probability measure  $P_\infty(\cdot)$  on the measurable space  $((0, N), \mathcal{B}(0, N))$  such that

$$P_{I_0,t}(\cdot) \rightarrow P_\infty(\cdot) \text{ in distribution for any } I_0 \in (0, N),$$

we then say that the SDE (2.4) has a stationary distribution  $P_\infty(\cdot)$  (see e.g. [13, 26]). To show the existence of a stationary distribution, let us first cite a known result from Has'minskii [13, pp.118–123] as a lemma.



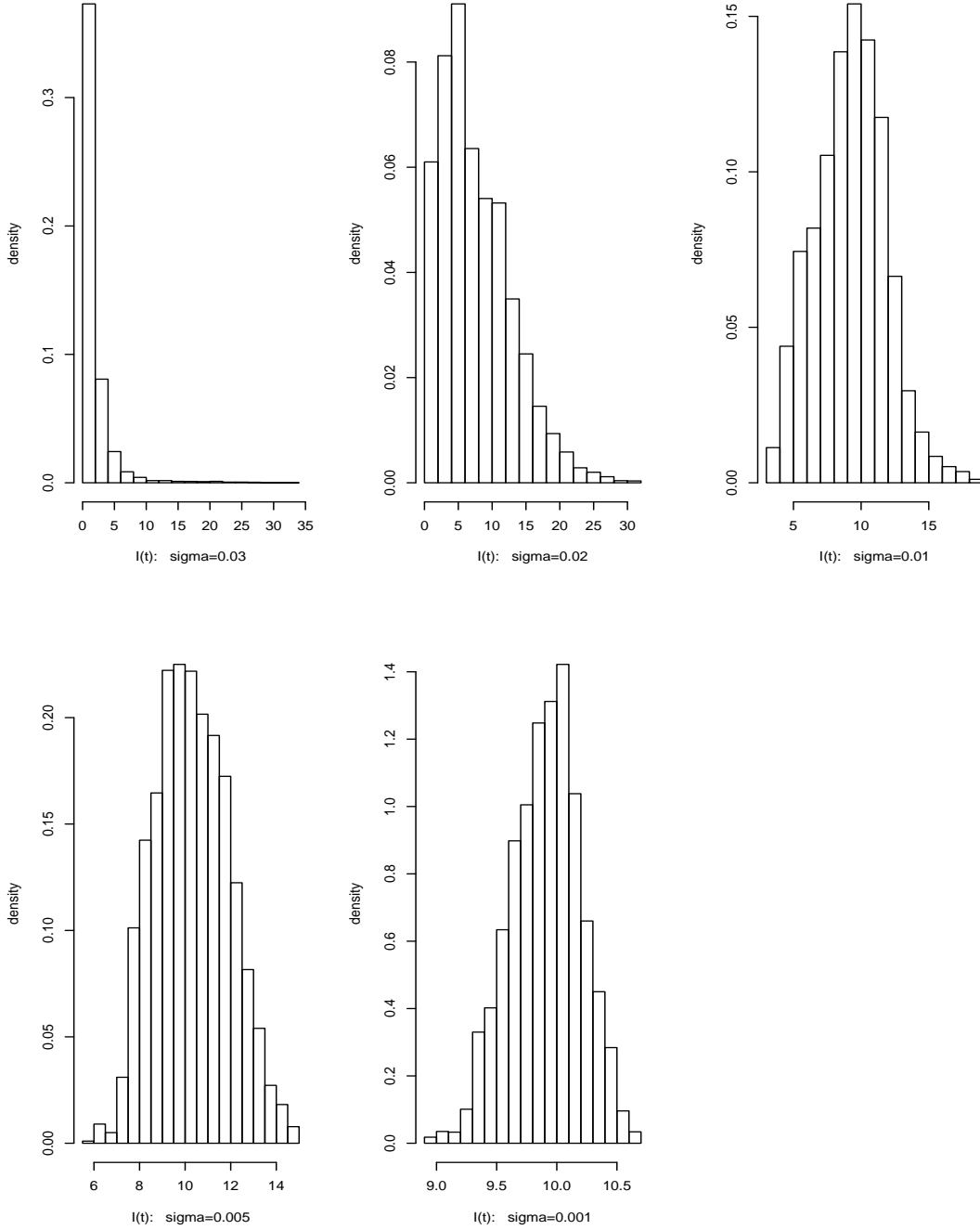


Figure 5.3: Histograms of the values of the path  $I(t)$  for the recurrent SDE SIS model (2.4), for parameter values  $\beta = 0.5$ ,  $N = 100$ ,  $\mu = 20$ ,  $\gamma = 25$ ,  $I(0) = 90$ , and differing values of  $\sigma = 0.03, 0.02, 0.01, 0.005$  and  $0.001$ .

**Lemma 6.1** *The SDE SIS model (2.4) has a unique stationary distribution if there is a strictly proper sub-interval  $(a, b)$  of  $(0, N)$  such that  $\mathbb{E}(\tau) < \infty$  for all  $I_0 \in (0, a] \cup [b, N)$ , where*

$$\tau = \inf\{t \geq 0 : I(t) \in (a, b)\},$$

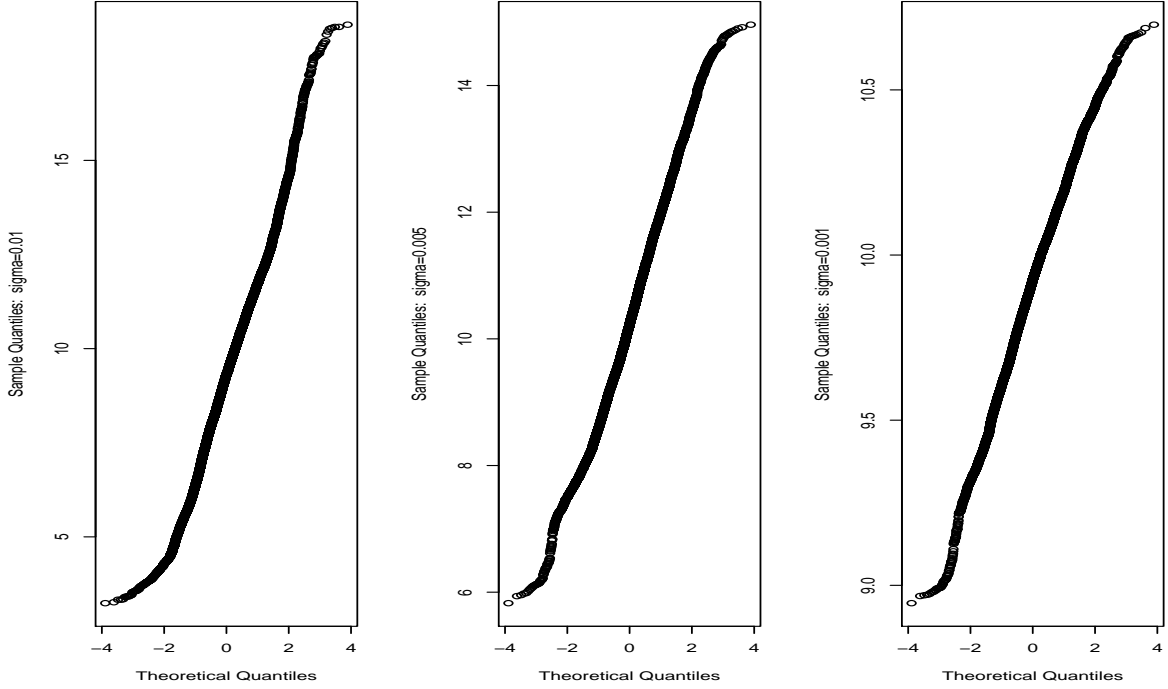


Figure 5.4: Normal quantile-quantile plots of the values of the path  $I(t)$  for the recurrent SDE SIS model (2.4), for parameter values  $\beta = 0.5$ ,  $N = 100$ ,  $\mu = 20$ ,  $\gamma = 25$ , and differing values of  $\sigma = 0.01$ ,  $0.005$  and  $0.001$ , corresponding to the last three histograms in Figure 5.3.

and, moreover,

$$\sup_{I_0 \in [\bar{a}, \bar{b}]} \mathbb{E}(\tau) < \infty \text{ for every interval } [\bar{a}, \bar{b}] \subset (0, N).$$

It should be pointed out that in the original Has'minskii theorem, there is one more condition which states that the square of the diffusion coefficient of the SDE (2.4), namely  $\sigma^2 I^2 (N - I)^2$ , is bounded away from zero for  $I \in (a, b)$ . But this is obvious for the SDE, hence there is no point to state.

**Theorem 6.2** *If  $R_0^S > 1$ , then the SDE SIS model (2.4) has a unique stationary distribution.*

*Proof.* We will use the same notation as used in the proofs of Theorems 4.1 and 5.1. Fix any  $0 < a < \xi < b < N$ . We observe from (5.6)–(5.8) that

$$f(x) \geq f(0) \wedge f(a) > 0 \text{ if } 0 < x \leq a \text{ and } f(x) \leq f(b) < 0 \text{ if } b \leq x < N. \quad (6.1)$$

Define  $\tau$  as in Lemma 6.1. For any  $I_0 \in (0, a)$ , it then follows from (4.3) and (6.1) that

$$\log(a) \geq \mathbb{E}(\log(I(\tau \wedge t))) \geq \log(I_0) + (f(0) \wedge f(a))\mathbb{E}(\tau \wedge t), \quad \forall t \geq 0.$$

Letting  $t \rightarrow \infty$  yields

$$\mathbb{E}(\tau) \leq \frac{\log(a/I_0)}{f(0) \wedge f(a)}, \quad \forall I_0 \in (0, a). \quad (6.2)$$

Similarly, for any  $I_0 \in (b, N)$ ,

$$\log(b) \leq \mathbb{E}(\log(I(\tau \wedge t))) \leq \log(I_0) - |f(b)|\mathbb{E}(\tau \wedge t), \quad \forall t \geq 0.$$

Letting  $t \rightarrow \infty$  yields

$$\mathbb{E}(\tau) \leq \frac{\log(N/b)}{|f(b)|}, \quad \forall I_0 \in (b, N). \quad (6.3)$$

The conditions in Lemma 6.1 follow clearly from (6.2) and (6.3). Hence the SDE SIS model (2.4) has a unique stationary distribution. The proof is complete.

The following theorem gives the mean and variance of the stationary distribution. Such explicit formulae are particularly useful in the test of computer simulations.

**Theorem 6.3** *Assume that  $R_0^S > 1$ . Let  $m$  and  $v$  denote the mean and variance of the stationary distribution of the SDE SIS model (2.4). Then*

$$m = \frac{2\beta(R_0^S - 1)(\mu + \gamma)}{2\beta(\beta - \sigma^2 N) + \sigma^2(\beta N - \mu - \gamma)} \quad (6.4)$$

and

$$v = \frac{m(\beta N - \mu - \gamma)}{\beta} - m^2. \quad (6.5)$$

*Proof.* Fix any  $I_0 \in (0, N)$ . It follows from (2.4) that

$$I(t) = I_0 + \int_0^t I(s)[\beta N - \mu - \gamma - \beta I(s)]ds + \int_0^t \sigma(N - I(s))dB(s).$$

Dividing both sides by  $t$ , letting  $t \rightarrow \infty$  and applying the ergodic property of the stationary distribution (see e.g. [13, 21]) and the large number theorem for martingales (see e.g. [25]), we obtain

$$0 = (\beta N - \mu - \gamma)m - \beta m_2, \quad (6.6)$$

where  $m_2$  denotes the second moment of the stationary distribution. Similarly, dividing both sides of (4.3) by  $t$  and letting  $t \rightarrow \infty$  we get

$$\lim_{t \rightarrow \infty} \frac{\log(I(t))}{t} = \beta N - \mu - \gamma - 0.5\sigma^2 N^2 - (\beta - \sigma^2 N)m - 0.5\sigma^2 m_2 \quad a.s. \quad (6.7)$$

This, together with Theorem 5.1, implies

$$\lim_{t \rightarrow \infty} \frac{\log(I(t))}{t} = 0 \quad a.s.$$

Writing  $\beta N - \mu - \gamma - 0.5\sigma^2 N^2 = (R_0^S - 1)(\mu + \gamma)$ , we then have

$$0 = (R_0^S - 1)(\mu + \gamma) - (\beta - \sigma^2 N)m - 0.5\sigma^2 m_2. \quad (6.8)$$

Substituting (6.6) into (6.8) yields

$$0 = (R_0^S - 1)(\mu + \gamma) - (\beta - \sigma^2 N)m - \frac{\sigma^2(\beta N - \mu - \gamma)m}{2\beta}.$$

This implies assertion (6.4). Moreover, it follows from (6.6) that

$$m_2 = \frac{m(\beta N - \mu - \gamma)}{\beta}.$$

Hence

$$v = m_2 - m^2 = \frac{m(\beta N - \mu - \gamma)}{\beta} - m^2,$$

which is the other assertion (6.5). The proof is therefore complete.

**Example 6.4** We now use the same parameter values  $\beta = 0.5$ ,  $N = 100$ ,  $\mu = 20$  and  $\gamma = 25$  for both  $\sigma = 0.02$  and  $0.001$  and show the results of running 1,000 simulations of the path  $I(t)$  for the recurrent SDE SIS model, for a longer run of 200,000 iterations with step size  $\Delta = 0.001$ , but storing only the last of these  $I(t)$  values in each case. Figure 6.1 shows the histogram of the last 10,000 samples from a single run of 200,000 iterations, beside the histogram of the last  $I(t)$  values from each of the 1,000 simulations, and also the corresponding two empirical cumulative distribution functions (ecdfs), for comparison, for both values of  $\sigma$ . In each of the two cases, the corresponding histograms are similar and the ecdfs are close to each other.

The similarity of these distributions in each case of  $\sigma$  may be taken as an illustration of the existence of the stationary distribution of  $I(t)$ , and that in the simulations the probability distribution of  $I(t)$  has more or less reached this stationary distribution. From (6.4) and (6.5), for these parameter values the mean and variance of the stationary distribution are 6.493506 and 22.76944 respectively when  $\sigma = 0.02$ , compared to the two sample means of 7.306251 and 6.557697 and the corresponding unbiased sample variances of 19.50794 and 23.18525, for the first and second histograms respectively. For  $\sigma = 0.001$ , the mean and variance of the stationary distribution are 9.991898 and 0.08094976, compared to the two sample means of 10.01330 and 9.98392 and unbiased sample variances of 0.06149254 and 0.07802619 respectively for the lower two histograms in Figure 6.1.

## 7 Two More Realistic Examples

As slightly more realistic examples to illustrate our theory, we suggest two SIS epidemic models with parameters estimated from actual disease situations. In this section the unit of time is still one day, but the population values are not scaled as previously:

**Model A** Gonorrhoea amongst homosexuals [14].

In this model, the parameters are given by  $N = 10,000$ ,  $R_0^D = 1.4$ ,  $\mu = (1/(40 \times 365.25))/\text{day} = 6.84463 \times 10^{-5}/\text{day}$  (average sexually active lifetime),  $\gamma = (1/55)/\text{day} = 0.018182/\text{day}$  (based on Yorke et al. [31]). Benenson [3] says that the infectious period

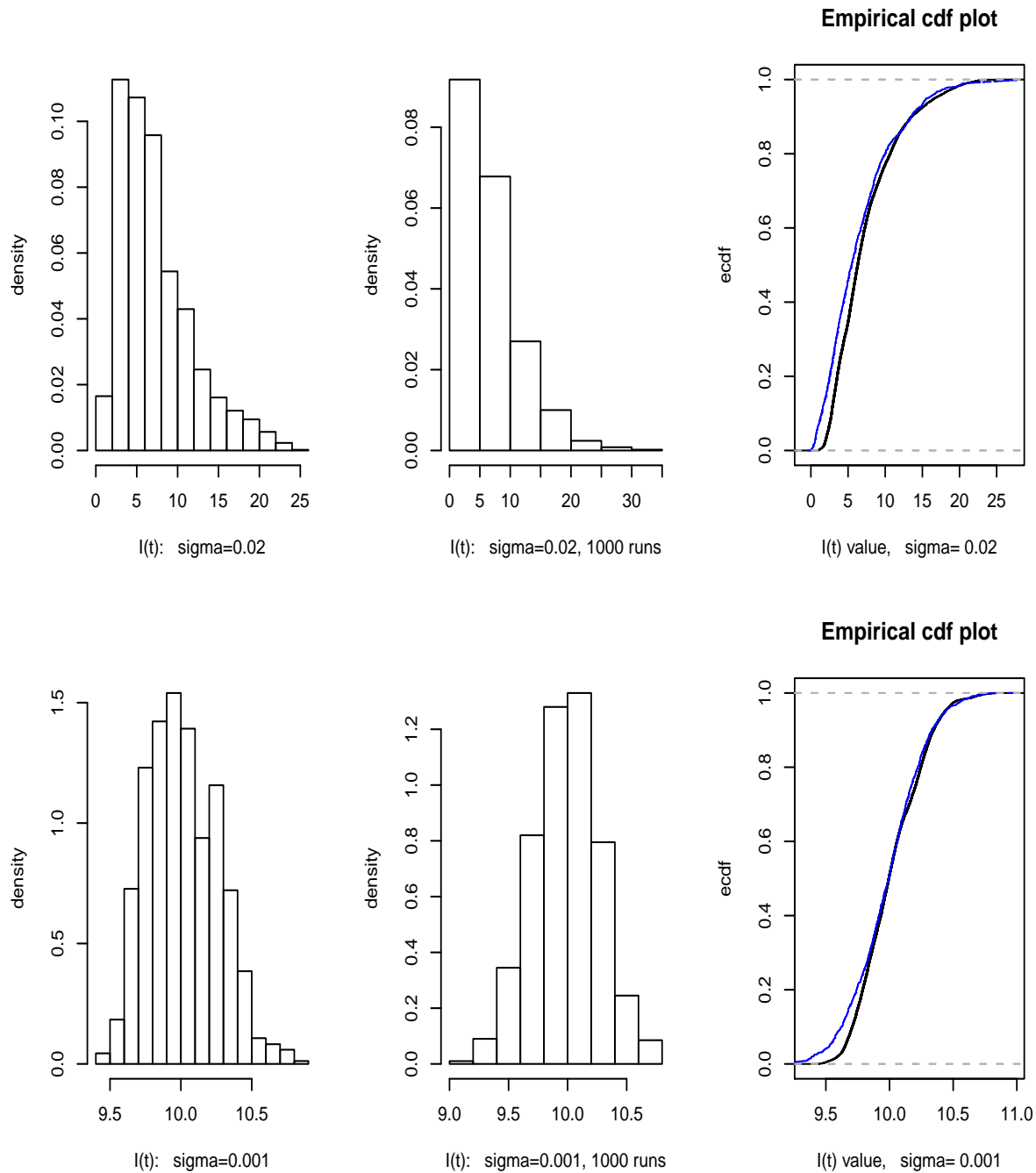


Figure 6.1: Histograms of the values of the path  $I(t)$  for the recurrent SDE SIS model (2.4) for the last 10,000 samples of a single run of 200,000 iterations (left plot in each row) and also for the last iteration from each of 1,000 such runs (middle plot in each row), and the empirical cumulative distribution plot of each of these (right plot in each row; the black line corresponds to the first histogram and the blue line to the second one), for parameter values  $\beta = 0.5$ ,  $N = 100$ ,  $\mu = 20$ ,  $\gamma = 25$ , and  $\sigma = 0.02$  (top row) and  $\sigma = 0.001$  (bottom row).

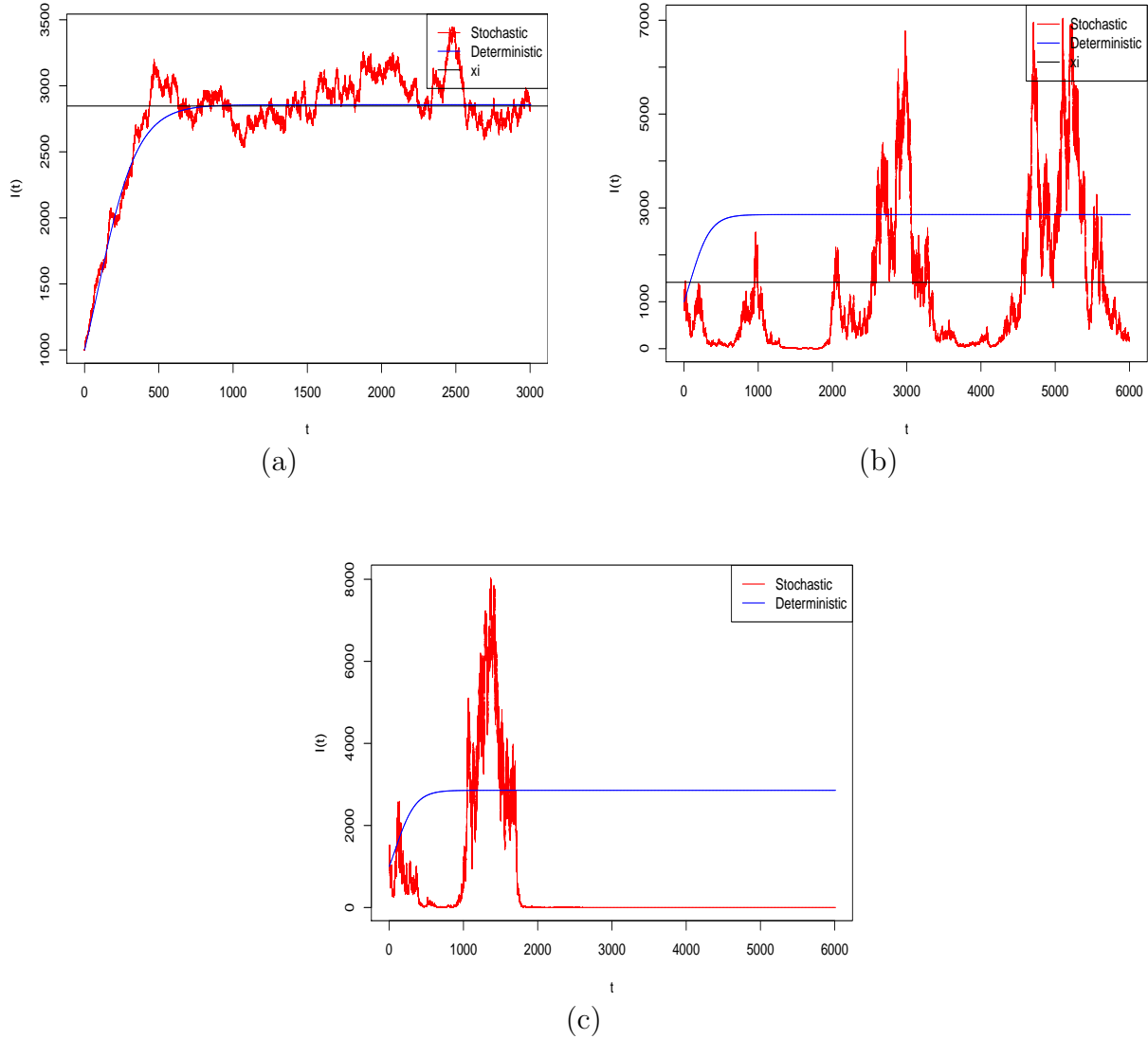


Figure 7.1: Computer simulation of the path  $I(t)$  for Model A using the EM method with step size  $\Delta = 0.001$  and initial value  $I(0) = 1,000$ . The deterministic case means  $\sigma = 0$ , while for the stochastic case we use (a)  $\sigma = 10^{-6}$ , (b)  $\sigma = 10^{-5}$ , (c)  $\sigma = 1.5 \times 10^{-5}$ .

is several months, but is not more precise, while Hethcote and Yorke ([14], Table 5.1) take  $1/\gamma = 20\text{-}40$  days for men and  $80\text{-}160$  days for women), and  $\beta = (\mu + \gamma)R_0^D/N = 2.55504 \times 10^{-6}/\text{day}$ .

Note that  $R_0^D = 1.4$ . Hence for the corresponding deterministic SIS model (1.1), we have

$$\lim_{t \rightarrow \infty} I(t) = 2,857$$

for any initial value  $I_0 \in (0, 10,000)$ . It is also easy to compute, for the SDE SIS model (2.4),

$$R_0^S = 1.4 - 2.739659 \times 10^9 \sigma^2.$$

To see the effect of the noise intensity, we consider three different values of  $\sigma$ :  $10^{-6}$ ,  $10^{-5}$  and  $1.5 \times 10^{-5}$ . The corresponding values of  $R_0^S$  are 1.397, 1.126 and 0.784, respectively.

By Theorem 5.1 we see that the SDE SIS model is persistent in the first two cases. However, in the last case, verifying  $\sigma^2 = 2.25 \times 10^{-10} < \beta/N = 2.55504 \times 10^{-10}$  we conclude by Theorem 4.1 that the SDE SIS model is extinctive. The computer simulations shown in Figure 7.1 support these results clearly. Figure 7.2 shows the level  $\xi$  and the value of the mean  $m$  in (6.4) as a function of  $\sigma$  in the range given by Proposition 5.4.

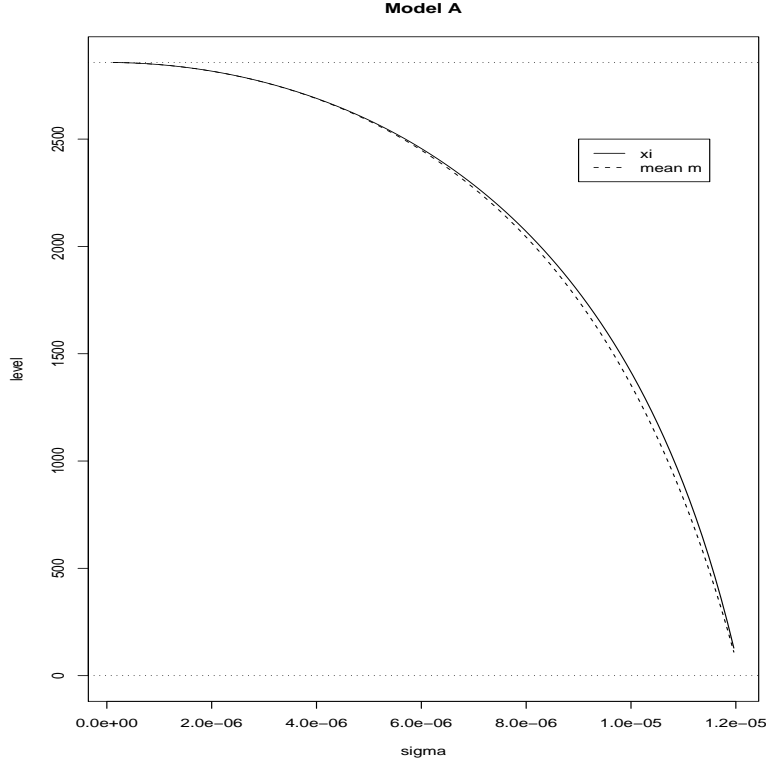


Figure 7.2: Plot of level of  $\xi$  (solid curve) and the mean  $m$  in (6.4) (dotted curve) against the value of  $\sigma$  in the range given in Proposition 5.4, for model A. The horizontal dotted lines show the levels 0 and  $N \left(1 - \frac{1}{R_0^D}\right)$  as limiting values for  $\xi$ .

Figure 7.3 shows histograms of the approximate stationary distribution of the two persistent cases (a) and (b) in Figure 7.1, resulting from the last 2 million iterations (last 2,000 days) for case (a) and the last 4 million iterations (4,000 days) for case (b). Both appear skewed to the right. The sample mean and unbiased sample variance are 2,901.894 and 31,045.18 respectively for case (a), compared to the theoretical mean and variance of the stationary distribution of 2,847.062 and 28,522.01 from (6.4) and (6.5). For case (b) the sample mean and unbiased sample variance are 1,051.77 and 2,125,489, compared to 1,354.592 and 2,035,258 from (6.4) and (6.5).

**Model B** Pneumococcus amongst children under 2 years in Scotland (Lamb, Greenhalgh and Robertson [18]).

In this model, the parameters are given by  $N = 150,000$ ,  $\gamma = 1/(7.1 \text{ wk}) = 0.02011/\text{day}$  (Weir [30]),  $\mu = 1/(104 \text{ wk}) = 1.3736 \times 10^{-3}/\text{day}$ , and  $\beta = 2.0055 \times 10^{-6}/\text{wk} = 2.8650 \times 10^{-7}/\text{day}$  (Zhang et al. [32]). (Farrington [8] has  $R_0^D = 1.5$ , which gives  $\beta = 2.1486 \times 10^{-7}/\text{day}$ .)

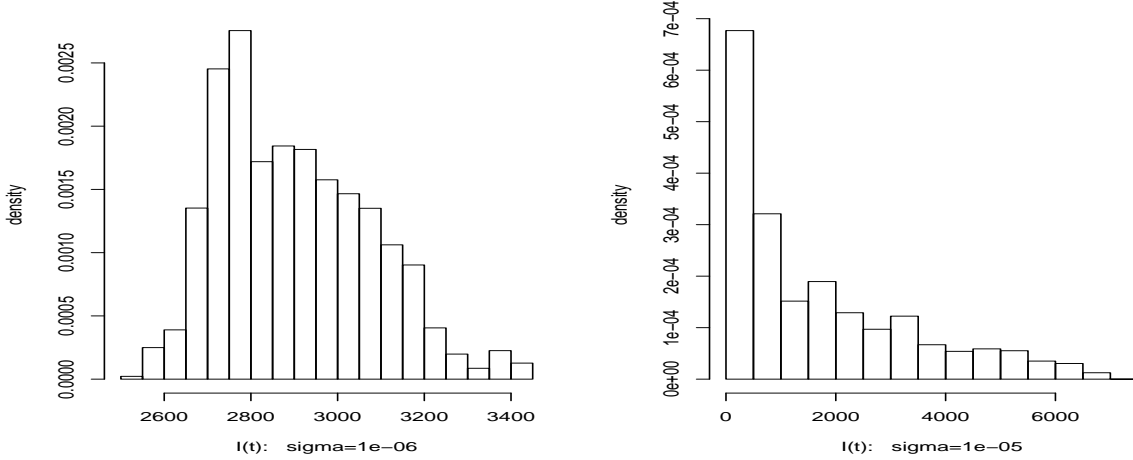


Figure 7.3: Histograms of the values of the path  $I(t)$  for model A for Figure 7.1(a) and 7.1(b), using the last 2 million iterations (2,000 days) for case (a) and the last 4 million iterations (4,000 days) for case (b).

It is easy to compute  $R_0^D = 2$ . Hence for the corresponding deterministic SIS model (1.1), we have

$$\lim_{t \rightarrow \infty} I(t) = 75,000$$

for any initial value  $I_0 \in (0, 150,000)$ . It is also easy to compute, for the SDE SIS model (2.4),

$$R_0^S = 2 - 5.23655 \times 10^{11} \sigma^2.$$

To see the effect of the noise intensity, we consider three different values of  $\sigma$ :  $10^{-6}$ ,  $1.3 \times 10^{-6}$  and  $1.5 \times 10^{-6}$ . The corresponding values of  $R_0^S$  are 1.476, 1.115 and 0.822, respectively. By Theorem 5.1 we see that the SDE SIS model is persistent in the first two cases. However, in the last case, verifying  $\sigma^2 = 2.25 \times 10^{-12} > (\beta/N) \vee (\beta^2/2(\mu + \gamma)) = 1.910347 \times 10^{-12}$  we conclude by Theorem 4.3 that the SDE SIS model is extinctive. The computer simulations shown in Figure 7.4 support these results clearly. In Figure 7.4(c) the deterministic simulation goes off the scale but is the same as in the other two simulations. Figure 7.5 shows the level  $\xi$  and the value of the mean  $m$  in (6.4) as a function of  $\sigma$  in the range given by Proposition 5.4.

Figure 7.6 shows histograms of the approximate stationary distribution of the two persistent cases (a) and (b) in Figure 7.4, resulting from the last 2 million iterations (the last 2,000 days). The first appears symmetric, the second positively skewed. The sample mean and unbiased sample variance are 61,448.49 and 647,526,916 respectively for case (a), compared to the theoretical mean and variance of the stationary distribution of 58,856.32 and 950,958,848 from (6.4) and (6.5). For case (b) the sample mean and unbiased sample variance are 45,633.17 and 1,494,311,814, compared to 25,718.33 and 1,267,792,373 from (6.4) and (6.5).



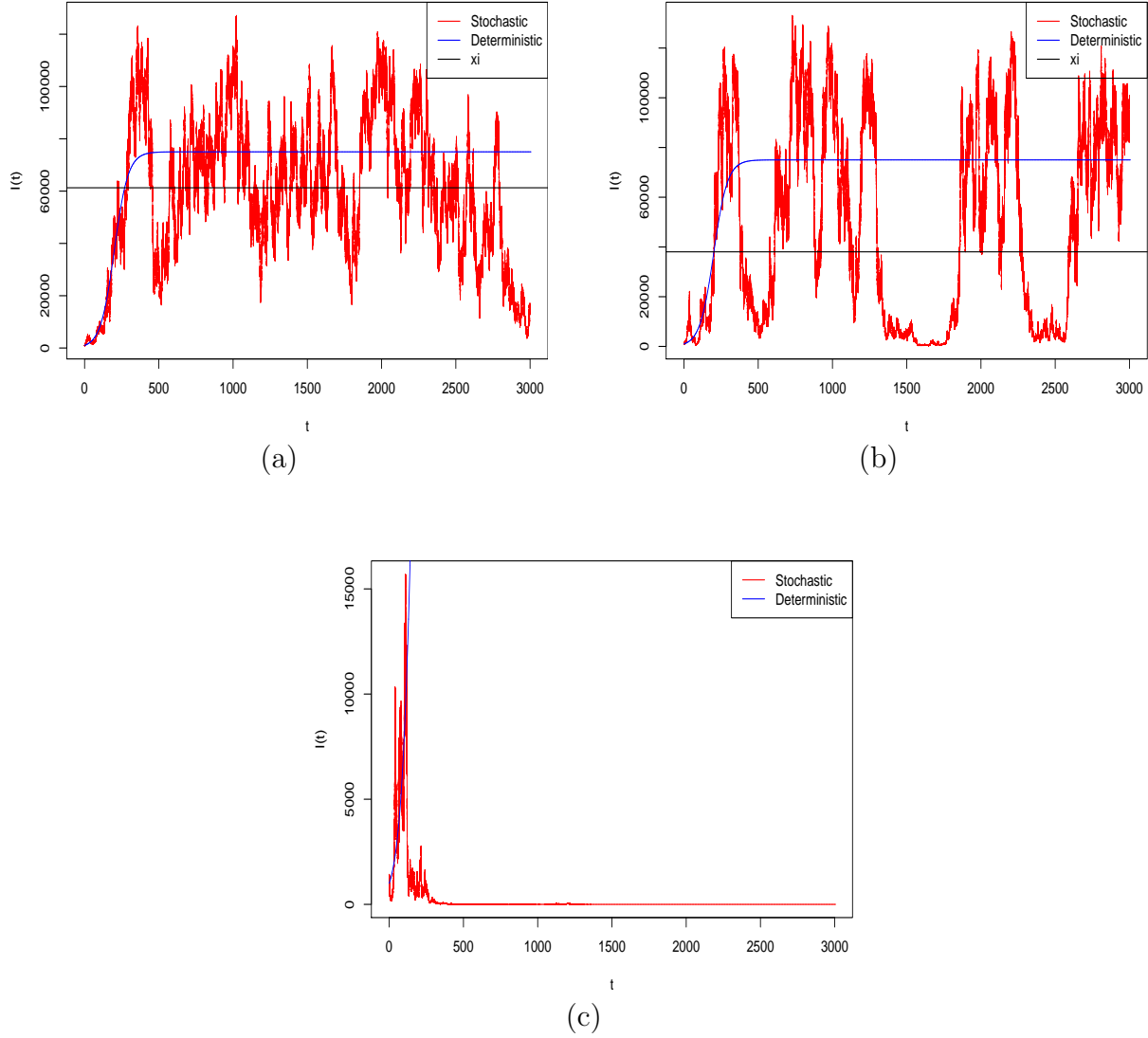


Figure 7.4: Computer simulation of the path  $I(t)$  for Model B using the EM method with step size  $\Delta = 0.001$  and initial value  $I(0) = 50,000$ . The deterministic case means  $\sigma = 0$ , while for the stochastic case we use (a)  $\sigma = 10^{-6}$ , (b)  $\sigma = 1.3 \times 10^{-6}$ , (c)  $\sigma = 1.5 \times 10^{-6}$ .

## 8 Discussion

Consider the stochastic SIS epidemic model in a neighbourhood of the DFE ( $I = 0$ ). Then equation (2.4) becomes approximately

$$dI = [\beta N - (\mu + \gamma)]I dt + \sigma N I dB(t),$$

with solution

$$I(t) = I_0 \exp \left[ \left( \beta N - (\mu + \gamma) - \frac{1}{2} \sigma^2 N^2 \right) t + \sigma N B(t) \right].$$

Hence as  $\lim_{t \rightarrow \infty} |B(t)|/t = 0$  [25] we expect that if

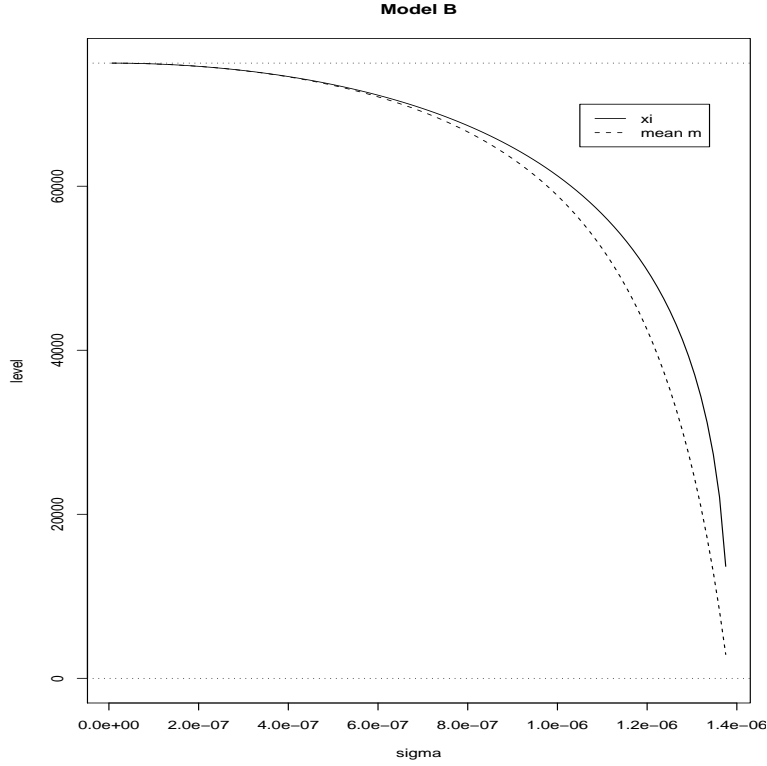


Figure 7.5: Plot of level of  $\xi$  (solid curve) and the mean  $m$  in (6.4) (dotted curve) against the value of  $\sigma$  in the range given in Proposition 5.4, for model B. The horizontal dotted lines show the levels 0 and  $N \left(1 - \frac{1}{R_0^D}\right)$  as limiting values for  $\xi$ .

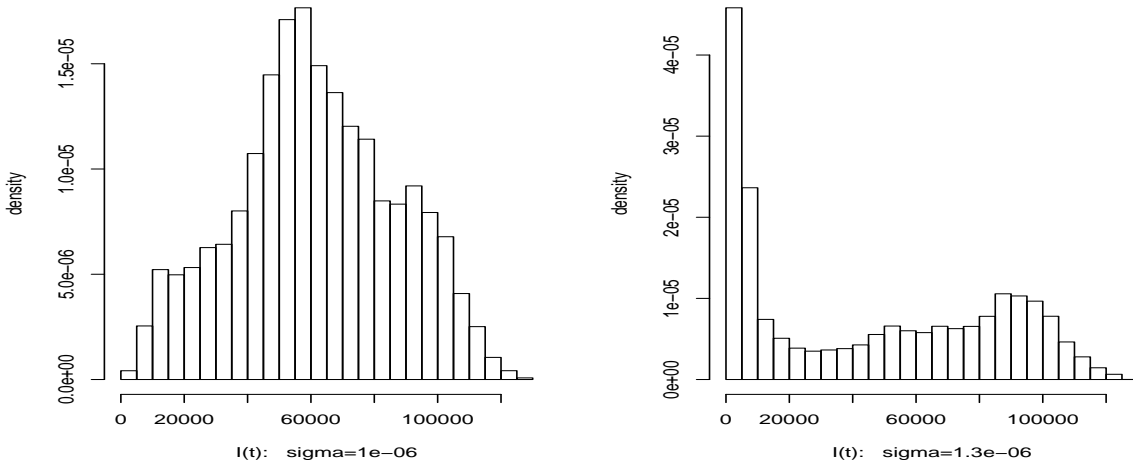


Figure 7.6: Histograms of the values of the path  $I(t)$  for model B for Figure 7.4(a) and 7.4(b), using the last 2 million iterations (2,000 days) in each case.

$$R_0^S = \frac{\beta N}{\mu + \gamma} - \frac{\sigma^2 N^2}{2(\mu + \gamma)} < 1,$$

then the approximate solution will die out, but if  $R_0^S > 1$  then the approximate solution will diverge from the DFE. Thus in this sense  $R_0^S$  is the natural interpretation of  $R_0$  in the SDE SIS model (2.4), although it is negative unless  $\sigma^2 < 2\beta/N$ .

This is almost what we have shown. Theorems 4.1 and 4.3 show that if either

$$(i) R_0^S < 1 \quad \text{and} \quad \sigma^2 \leq \frac{\beta}{N} \quad \text{or} \quad (ii) \sigma^2 > \frac{\beta}{N} \vee \frac{\beta^2}{2(\mu + \gamma)},$$

the disease will die out, whereas Theorem 5.1 shows that if  $R_0^S > 1$  then the disease will persist. It is natural to make the following conjecture:

**Conjecture 8.1** *If*

$$R_0^S < 1 \quad \text{and} \quad \frac{\beta^2}{2(\mu + \gamma)} \geq \sigma^2 > \frac{\beta}{N}, \quad (8.1)$$

*then the disease will die out with probability 1.*

While we have not so far been able to prove this, Example 8.2 provides an illustration of it.

**Example 8.2** We now use the system parameters

$$\beta = 0.5, \quad N = 100, \quad \mu = 10, \quad \gamma = 8,$$

and now let  $\sigma = 0.0825$ , so that condition (8.1) is satisfied, and so the SDE SIS model (2.4) becomes

$$dI(t) = I(t) \left( [32 - 0.5I(t)]dt + 0.0825(100 - I(t))dB(t) \right). \quad (8.2)$$

Figure 8.1 shows two simulations of the path  $I(t)$ , both becoming extinctive quickly. In Figure 8.2(b) the deterministic trajectory is as in Figure 8.2(a) although off the scale.

An alternative approach to including environmental stochasticity outlined by Allen [1] is to model the per capita disease transmission coefficient as a time dependent stochastic process  $\bar{\beta}(t)$ . The problem is discretised with a small timestep  $\Delta t$  so that  $\bar{\beta}(n\Delta t)$  follows a random walk with state-dependent transition probabilities. These transition probabilities include both a diffusion term which causes the random walk to diverge and a mean-reverting term which drives the process back to a given mean value, say  $\beta_0$ . Under these assumptions the limiting process as  $\Delta t \rightarrow 0$ ,  $\bar{\beta}(t)$ , follows an Ornstein-Uhlenbeck SDE

$$d\bar{\beta} = \gamma(\beta_0 - \bar{\beta})dt + \sigma dB(t).$$

This differential equation can be solved exactly and for large times the disease transmission coefficient is approximately normally distributed with mean  $\beta_0$  and variance  $\sigma^2/2\gamma$  [1].

In our method of including environmental stochasticity the total number of potentially infectious contacts between an infected individual and another individual in the infinitesimally small time interval  $[t, t + dt)$  is given by

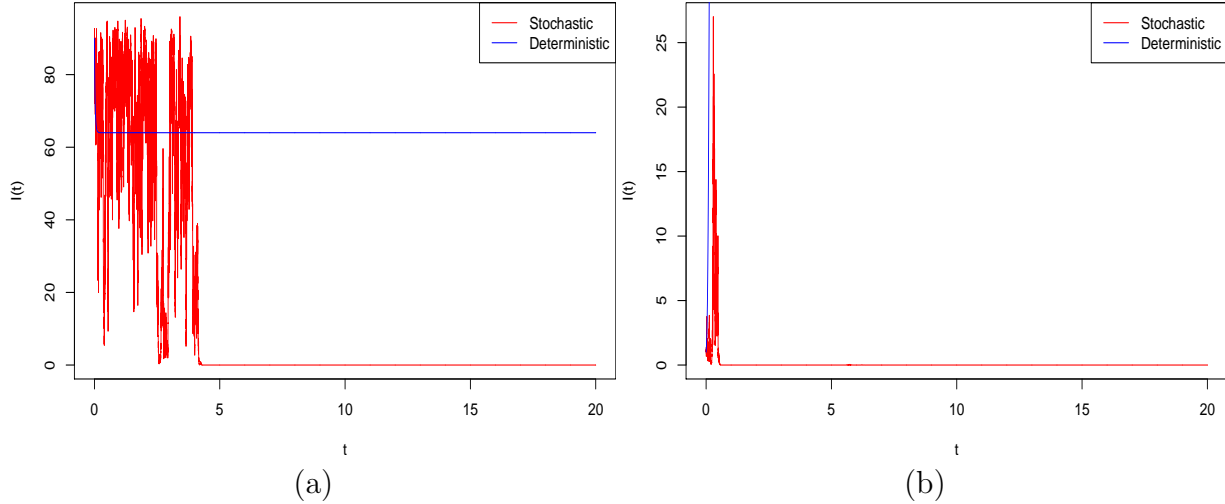


Figure 8.1: Computer simulation of the path  $I(t)$  for the SDE SIS model (8.2) and its corresponding deterministic SIS model (4.8), using the EM method with step size  $\Delta = 0.001$ , using initial values (a)  $I(0) = 90$  and (b)  $I(0) = 1$ .

$$\tilde{\beta}dt = \beta_0dt + \sigma dB(t)$$

(where  $\tilde{\beta}$  is as in Section 2 and  $\beta_0$  is the given value as above) which implies that

$$\int_0^t \tilde{\beta}dt = \beta_0t + \sigma B(t)$$

i.e. the total number of potentially infectious contacts between them in  $[0, t)$  has a normal distribution with mean  $\beta_0t$  and variance  $\sigma^2t$ . This is a well-established method [7, 10, 11, 12, 19, 22, 29], although both methods are biologically reasonable.

In this paper we have looked at an SDE version of the classical SIS epidemic model, with noise introduced in the disease transmission term. We showed that the SDE had a unique positive global solution and established conditions for extinction and persistence of disease. A key parameter was the basic reproduction number  $R_0^S$ , which was less than the corresponding deterministic version  $R_0^D$ . Theorems 4.1 and 4.3 show that if  $R_0^S \leq 1$ , under mild extra conditions the disease would die out. Theorem 5.1 shows that if  $R_0^S > 1$  then the disease will persist. We also showed (Theorem 6.2) that if  $R_0^S > 1$  then the model has a unique stationary distribution and derived expressions for its mean and variance (Theorem 6.3). We made a conjecture about the disease behaviour if  $R_0^S \leq 1$  and the conditions of Theorems 4.1 and 4.3 are not satisfied. Throughout the paper we have illustrated our theoretical results with computer simulations, including two sets with realistic parameter values for gonorrhoea amongst homosexuals and pneumococcus amongst young children.

## Acknowledgements

The authors would like to thank the referees and the editor for their very helpful comments and suggestions. The authors would also like to thank the Scottish Government, the British Council Shanghai and the Chinese Scholarship Council for their financial support.

## References

- [1] Allen, E., *Modelling with Itô Stochastic Differential Equations*, Springer-Verlag, 2007.
- [2] Brauer, F., Allen, L.J.S., Van den Driessche, P. and Wu, J. *Mathematical Epidemiology* Lecture Notes in Mathematics, No. 1945, Mathematical Biosciences Subseries, 2008.
- [3] Benenson, A.S., *Control of Communicable Diseases in Man*, Fifteenth Edition, American Public Health Association, 1990.
- [4] Chen, G. and Li, T., Stability of a stochastic delayed SIR model, *Stochastics and Dynamics*, 9(2), 231-252, 2009.
- [5] Dalal, N., Greenhalgh, D. and Mao, X., A stochastic model of AIDS and condom use, *Journal of Mathematical Analysis and Applications*, 325, 36-53, 2007.
- [6] Ding, Y., Xu, M. and Hu, L., Asymptotic behaviour and stability of a stochastic model for AIDS transmission, *Applied Mathematics and Computation*, 204, 99-108, 2008.
- [7] Engen, S. and Lande, R., Population dynamic models generating the lognormal species abundance distribution, *Mathematical Biosciences*, 132, 169-183, 1996.
- [8] Farrington, P., What is the reproduction number for pneumococcal infection, and does it matter? In: *4th International Symposium on Pneumococci and Pneumococcal Diseases, May 9-13, 2004 at Marina Congress Center, Helsinki, Finland*, 2004.
- [9] Feng, Z., Huang, W. and Castillo-Chavez, C., Global behaviour of a multi-group SIS epidemic model with age-structure, *Journal of Differential Equations*, 218(2), 292-324, 2005.
- [10] Foley, P., Predicting extinction times from environmental stochasticity and carrying capacity, *Conservation Biology*, 8(1), 124-137, 1994.
- [11] Gard, T.C., *Introduction to Stochastic Differential Equations*, Marcel Dekker Inc., 1988.
- [12] Grafton, R.Q., Kampas, T. and Lindenmayer, D., Marine reserves with ecological uncertainty, *Bulletin of Mathematical Biology*, 67, 957-971, 2005.
- [13] Has'minskii, R.Z., *Stochastic Stability of Differential Equations*, Sijthoff and Noordhoff, 1980.

- [14] Hethcote, H.W. and Yorke, J.A., *Gonorrhea Transmission Dynamics and Control*, Lecture Notes in Biomathematics 56, Springer-Verlag, 1994.
- [15] Kermack, W.O. and McKendrick, A.G. Contributions to the mathematical theory of epidemics. Part I. *Proceedings of the Royal Society Series A*, 115, 700-721, 1927.
- [16] Kloeden, P. E. and Platen, E., *Numerical Solution of Stochastic Differential Equations*, Springer, 1992.
- [17] Lajmanovich, A. and Yorke, J.A. A deterministic model for gonorrhea in a nonhomogeneous population, *Mathematical Biosciences*, 28, 221-236, 1976.
- [18] Lamb, K.E., Greenhalgh, D. and Robertson, C., A simple mathematical model for genetic effects in pneumococcal carriage and transmission, *Journal of Computational and Applied Mathematics*, 235, 1812-1818, 2011.
- [19] Lande, R., Engen, S. and Saether, B.-E., Spatial scale of population synchrony: Environmental correlation versus dispersal and density regulation, *The American Naturalist*, 154(3), 271-281, 1999.
- [20] Lipsitch, M., Vaccination against colonizing bacteria with multiple serotypes, *Proceedings of the National Academy of Sciences*, 94, 6571-6576, 1997.
- [21] Loève, M., *Probability Theory*, D. Van Nostrand Company Inc., 1963.
- [22] Lu, Q., Stability of SIRS system with random perturbations, *Physica A*, 388, 3677-3686, 2009.
- [23] Mao, X., *Stability of Stochastic Differential Equations with Respect to Semimartingales*, Longman Scientific and Technical, 1991.
- [24] Mao, X., *Exponential Stability of Stochastic Differential Equations*, Marcel Dekker, 1994.
- [25] Mao, X., *Stochastic Differential Equations and Applications*, 2nd Edition, Horwood, Chichester, UK, 2007.
- [26] Mao, X. and Yuan, C., *Stochastic Differential Equations with Markovian Switching*, Imperial College Press, 2006.
- [27] McCormack, R.K. and Allen, L.J.S., Stochastic SIS and SIR multihost epidemic models. *Proceedings of the Conference on Differential and Difference Equations and Applications*, R.P. Agarwal and K. Perera, Eds., Hindawi Publishing Corporation, pp.775-786, 2006.
- [28] Nold, A., Heterogeneity in disease transmission modelling, *Mathematical Biosciences*, 52, 227-240, 1980.
- [29] Tornatore, E., Buccellato, S.M. and Vetro, P., Stability of a stochastic SIR system, *Physica A*, 354, 111-126, 2005.

- [30] Weir, A. *Modelling the impact of vaccination and competition on pneumococcal carriage and disease in Scotland*, Unpublished Ph.D. Thesis, University of Strathclyde, Glasgow, Scotland, 2009.
- [31] Yorke, J.A., Hethcote, H.W. and Nold A. Dynamics and control of the transmission of gonorrhoea, *Sexually Transmitted Diseases*, 5, 51-56, 1978.
- [32] Zhang, Q., Arnaoutakis, K., Murdoch, C., Lakshman, R., Race, G., Burkinshaw, R. and Finn, A. Mucosal immune responses to capsular pneumococcal polysaccharides in immunized preschool children and controls with similar nasal pneumococcal colonization rates, *Pediatric Infectious Diseases Journal*, 23, 307-313, 2004.