Stochastic Model for Toxicity Assessment

by

Fei Zong

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Applied Mathematics

Department of Mathematical and Statistical Sciences
University of Alberta

Abstract

In this thesis, we study a mathematical model for the survival of a cells' population exposed to various chemical compounds with different concentrations. For experimental planning, it is important to find the threshold value for the initial concentrations at which the cells become extinct. First, we use a deterministic model to perform parameter estimations and model validations with the experimental data sets. We estimate the parameter sets for eight chemical compounds from different clusters. To account for parameter uncertainty, we derive a stochastic version of the model and perform the numerical analysis. The global Lipschitz condition is essential for strong convergence in most explicit methods; we relax this condition and prove convergence of the implicit Euler scheme with the one-sided Lipschitz conditions solely. Then we compare the Euler scheme, the tamed Euler scheme and the implicit Euler scheme by case studies in both stability and convergence analysis. Finally, the Monte Carlo simulations for three chemical compounds are presented to explore the distribution of thresholds.

Acknowledgements

I wish to express my great thanks to my supervisor Prof. Yaushu Wong and co-supervisor Prof. Cristina Anton, for their careful guidance and generous help for my master study and research. Meanwhile, I would like to thank Yaushu and Cristina for the help they gave me on my application to graduate school.

Meanwhile, I take this opportunity to thank all my course instructors and the supervisory committee members, for their advice and encouragement. In particular, I want to thank Dr. Hao Wang and Prof. Alexander Melnikov for the recommendations they gave me. Sincere thanks to the Department of Mathematical and Statistical Sciences which provides me many supports.

Finally, I want to express my special thanks to all my family members and friends who get me this far.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The steady increase of industrial activities around the world has led to a mass of chemical compounds being constantly released into the environment. Some of these compounds are classified as toxicants, because they are harmful to not only the environment but our health. Needless to say, there is an imperative demand to develop efficient and reliable methods for toxicity assessments. A traditional methodology for evaluating toxicity is by means of in vivo (Latin for "within the living") assays involving living organisms. Another approach is by means of in vitro (i.e., laboratory experiments) assays in which a real-time cell electronic sensing system is employed to perform cytotoxicity profiling [2]. The term cytotoxicity describes the process of putting cells into a toxicant environment and letting the toxicant environment affect the cells' morphology and adhesion. Due to efficiency reasons and moral/ethics related concerns, in vitro toxicity testing is naturally preferred.

Alberta Centre for Toxicology proposed a project for assessing toxicity by means of the xCELLigence Real-Time Cell Analysis High Throughput (RTCA HT) sys-

tem developed by ACEA Biosciences Inc. in San Diego, USA. The xCELLigence RTCA HT system records the response curve of the cell index dynamically [7]. More specifically, it consists of four 384 well E-Plates where the cell index (CI), as the representative for cell numbers, is recorded. In the experiments, the cells were seeded onto the E-Plate; the attachment was allowed to reach the exponential growth phase, which is approximately within 24 hours [4]. After the attachment phase, test substances were added to the lag cell growth phase during which the cells are the most viable. The compounds with 11 selected concentrations (1:3 serial dilution of the stock solution) were dissolved in the cell culture medium [1]. The monitor time was 75 hours and the cellular responses were recorded. The experiments were repeated 4 times to ensure the measurement reliability and data accuracy. The time-dependent concentration response curves (TCRCs) for each test substance in each cell line can be generated [9]. The cellular population is measured by the impedance signal generated by microelectrodes. Other responses, such as cell inhibitions and cell morphological changes, can also be generated by the impedance signal [4].

The dataset used in this study is from the cytotoxicity test on human hepatocellular carcinoma cell line HepG2 [10]. We have initially 63 chemical compounds which are divided into 10 clusters according to the mode of action (MOA), which involves functional and anatomical changes. The latter may include changes in cell population, morphology, and cellular functions [9]. A methodology involving a combination of clustering and machine learning models (artificial neural network (ANN) and support vector machine (SVM)) has been proposed in [3]. In the foregoing paper, the known MOAs in [9] are used for validations.

Fig. 1.1 shows an example of the TCRCs for a chemical compound in the test, which is BafilomycinA. We observe that the highest two concentrations of test substances used for BafilomycinA force all cells to go extinction, while the others are survivable for cells.

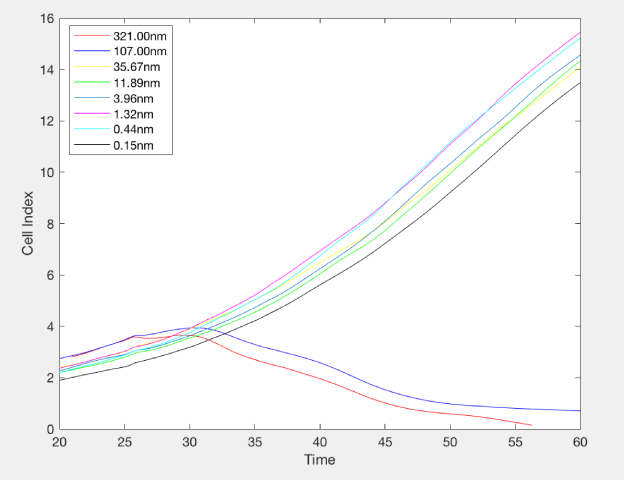


Figure 1.1: TCRC curves for BafilomycinA

A mathematical model consisting a system of three differential equations is developed in [1] to reproduce the TCRCs in the deterministic setting. For the purpose of parameter estimations, the Non-linear Least Square(NLS) method is first applied on the negative control data which are the data recorded in a non-toxic environment. Then, the Expectation-Maximization (EM) algorithm and the Unscented Filter(UF) are used for the model in the state-space form to determine the remaining parameters. One benefit of having a mathematical model is the ability to identify the critical point (by finding a small neighbourhood around it), where the concentrations change from persistence to extinction. We call such a critical point the threshold value. In presented Fig. 1.1, the threshold value is between 35.67nm-107nm for

BafilomycinA, which represents the critical point where the cells are headed towards extinction.

A global parameter sensitivity analysis is done in[1] to assess the effects of parameters uncertainty. To capture the effect of randomness/stochasticity, a stochastic model is proposed in this thesis. The most important contribution that we make is the numerical analysis of the proposed stochastic model. We compare the stochastic numerical schemes for several case studies from the perspectives of stability and convergence. One main difficulty that we encounter is the relaxation of the global Lipschitz condition on the coefficients of the stochastic differential equations (SDEs). It is widely known that the global Lipschitz condition is essential for strong convergence in most explicit methods; for instance, the Euler–Maruyama schemes proposed in [23] and the Milstein scheme[20]. Since it is challenging to prove the global Lipschitz condition for our model, we consider other schemes like the tamed Euler [25] and implicit Euler [28] schemes. Both of the schemes solely require one-sided Lipschitz conditions on the coefficients to converge. We employ Monte Carlo simulations to determine the stochastic thresholds and we also explore the distribution of the thresholds.

This thesis is organized in the following way. In Chapter 2, we present the mathematical model in the deterministic setting. Following the presentation of the model, we discuss methods used for parameter estimations: the Non-linear Least Square method and EM algorithm. We perform model validation by comparing the TCRCs with the predicted values for the cell index. To conclude Chapter 2, we present separation graphs using Runge-Kutta-Fehlberg Method (RKF45) applied to the deterministic model. Based on the global sensitivity analysis done in [1], we derive

the stochastic model in Chapter 3, with two parameters being random variables. Next we present Euler and Milstein schemes by using stochastic Taylor expansions and explore sufficient conditions for convergence. The introduction of tamed and implicit Euler schemes follows immediately. At the end of Chapter 3 we include conditions for convergence of the implicit Euler scheme. Chapter 5 mainly consists of the final numerical results of various schemes for different noise levels. Monte Carlo simulations based on the selected schemes are performed to explore the distribution of the stochastic thresholds. Chapter 6 contains the conclusion and some future works.

Chapter 2

Model Evaluation

In this chapter, we first present our model and perform data analysis for parameter estimations and model validations. Some preliminary tools like expectation maximization algorithm and unscented filter will also be provided. An application of the model, separation graph, will be shown at the end of this chapter.

2.1 Model Introduction

The model we choose is a three-dimensional model for an acute dose of test subsenses. The variables set are the cell index $n(t)$, the concentration of toxicant in the cell $CI(t)$, and the concentration of toxicant in the environment $CE(t)$. We focus on modeling the phase when cells actively proliferate and have an exponential increase in the cell density[1].

First, for the non-toxic environment, the negative control data measures the cell index in the whole process without no testing substances. To model the negative

control curve, we use the famous logistic model.

$$\frac{dn(t)}{dt} = \beta n(t)(1 - \frac{n(t)}{K}) \tag{2.1}$$

symbol	definition
$n(t)$	cell index at time t
β	cell growth rate
K	capacity volume

Table 2.1: Parameters for model (2.1)

For a toxic environment, we suppose the death rate of the cell is linearly dependent on the concentration of internal toxicants $CI(t)$ with coefficient $\alpha > 0$. Since only small-molecular compounds are tested in our experiments, carrier-mediated transportation is ignored, and linear kinetic is only considered. It is assumed that the change rate of the amount of toxicant in the environment as the sum of the input rates and losses from the environment[1]. Consider following model:

$$\begin{aligned}
\frac{dn(t)}{dt} &= \beta n(t)(1 - \frac{n(t)}{K}) - \alpha CI(t)n(t) \\
\frac{dCI(t)}{dt} &= \lambda_1^2 CE(t) - \eta_1^2 CI(t) \\
\frac{dCE(t)}{dt} &= \lambda_2^2 CI(t)n(t) - \eta_2^2 CE(t)n(t)
\end{aligned} \tag{2.2}$$

where $CI(t)$ and $CE(t)$ denote the intracellular and the extracellular concentrations of toxicant at time t, respectively.

symbol	definition
α	coefficient of toxicant on the cell's growth
λ_1^2	the uptake rate of the toxicant from environment
λ_2^2	the toxicant uptake rate from cells
η_1^2	the toxicant input rate to the environment
η_2^2	the losses rate of toxicant absorbed by cells

Table 2.2: Parameters for model (2.2)

2.2 Parameter Estimation

Parameter estimation plays an essential role for model validation and implementation. To use the model (2.2), we need to determine the parameters in the model first. The parameter set $\Theta = \{\beta, K, \alpha, \lambda_1, \lambda_2, \eta_1, \eta_2\}$ is estimated using the information from the real experimental data. We use two different methods to perform parameter estimation: Nonlinear Least Square(NLS) method and Expectation Maximization(EM) algorithm with the Unscented Filter(UF).

2.2.1 Nonlinear Least Square Method

A Nonlinear Least Square method is first employed to find β and K with negative control data (no toxicity) based on the analytic solution of logistic model. We do this by fitting the model with real experiemntal data from 20 to 90 hours using R programming. From the figure below we can see that the curve shows the logistic solution while the dots are experimental data.

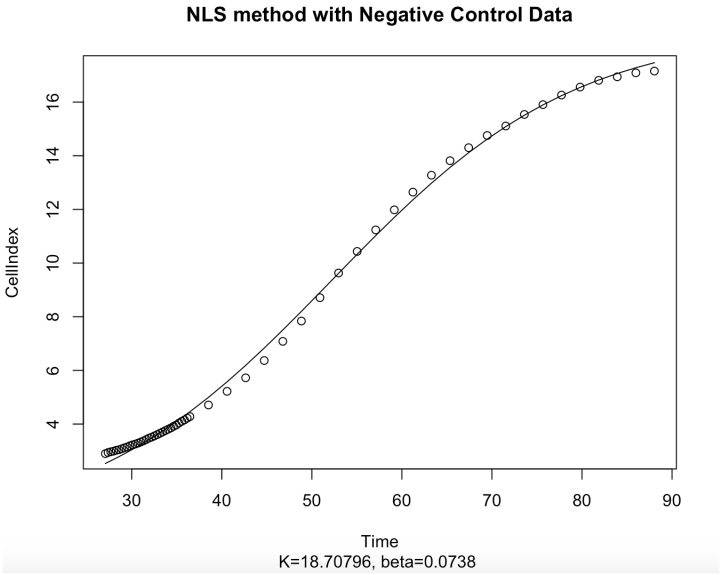


Figure 2.1: Parameter estimation using negative control data

2.2.2 State-space System

Since the non-linear system cannot be solved analytically, we apply the state-space model using the Euler integration scheme with time step h:

$$x_{k+1} = x_k + h \begin{bmatrix} \beta x_k[1](1 - \frac{x_k[1]}{K}) - \alpha x_k[2]x_k[1] \\ \lambda_1^2 x_k[3] - \eta_1^2 x_k[2] \\ x_k[1](\lambda_2^2 x_k[2] - \eta_2^2 x_k[3]) \end{bmatrix} + v_{k+1} \tag{2.3}$$

$$y_{k+1} = Cx_{k+1} + w_{k+1} \tag{2.4}$$

where

$$x_k = \begin{bmatrix} n \\ CI \\ CE \end{bmatrix}, C = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}^T$$

v_k and w_k are Gaussian white noise vectors, $v_k \sim N(0;Q)$ and $w_k \sim N(0;R)$ with The 3×3 diagonal covariance matrix Q, the variance $R > 0$.

2.2.3 The Unscented Filter and Expectation Maximization algorithm

We use the Expectation Maximization algorithm based on the Unscented Filter to estimate the rest of the parameters$\{\alpha, \lambda_1, \lambda_2, \eta_1, \eta_2\}$. Julier and Uhlmann introduced the Unscented Filter for non-linear state-space models in [5]. Eight diluted concentrations of the compound are used (largest 8 of 11 values of initial concentrations).

For the state-spaced model, we define

$$\text{filtered values:} \overline{x_i} = E[x_i | y_1, \cdots, y_i]$$
$$\text{predicted values:} \hat{x}_{i+1} = E[x_{i+1} | y_1, \cdots, y_i]$$

So we have the conditional covariances:

$$\overline{P_i} = E[(x_i - \overline{x_i})(x_i - \overline{x_i})^T | y_1, \cdots, y_i]$$
$$\hat{P}_{i+1} = E\left[(x_{i+1} - \hat{x}_{i+1})(x_{i+1} - \hat{x}_{i+1})^T | y_1, \cdots, y_i\right].$$

And smoothed values

$$x_{i|N} = E[x_i | y_1, \cdots, y_N]$$
$$P_{i|N} = E[(x_i - x_i^N)(x_i - x_i^N)^T | y_1, \cdots, y_N]$$

Then, based on the state-space system, we define

$$f(x) = x + h \begin{bmatrix} \beta x_k[1](1 - \frac{x_k[1]}{K}) - \alpha x_k[2]x_k[1] \\ \lambda_1^2 x_k[3] - \eta_1^2 x_k[2] \\ x_k[1](\lambda_2^2 x_k[2] - \eta_2^2 x_k[3]) \end{bmatrix}, x \in R^3 \tag{2.5}$$

$$x_{k+1} = f(\bar{x}_k) + \frac{\partial f}{\partial x}(\bar{x}_k)(x_k - \bar{x}_k) + v_{k+1} \tag{2.6}$$

$$y_k = Cx_k + w_k \tag{2.7}$$

Where $x_1[1] = n(0)$, the initial cell numbers, $x_1[2] = CI(0) = 0$ and $x_1[3] = CE(0)$, the initial concentration of chemical compound. Then we have

$$\overline{P}_{x_k x_{k+1}} = E[(x_k - \bar{x}_k)(x_{k+1} - \hat{x}_{k+1})^T | x_1, \cdots, x_k] = \overline{P}_k \frac{\partial f^T}{\partial x}(\bar{x}_k) \tag{2.8}$$

We can use the Unscented Filter algorithm [5] to compute $\bar{x}_{k+1}, \overline{P}_{k+1}, \hat{x}_{k+1}, \hat{P}_{k+1}$ and $\overline{P}_{x_k x_{k+1}}$.

$$\begin{aligned} J_k &= \overline{P}_{x_k x_{k+1}} \hat{P}_{k+1}^{-1} \\ x_{k-1|N} &= \bar{x}_{k-1} + J_{k-1}(x_{k|N} - \hat{x}_k) \\ P_{k-1|N} &= \overline{P}_{k-1} + J_{k-1}(P_{k|N} - \hat{P}_k) J_{k-1}^T \end{aligned} \tag{2.9}$$

The smoothed values can be computed using 2.9 starting with $k = N$.

The EM algorithm: The EM algorithm is an iterative method for estimating parameters of linear models. The iteration is combined by an expectation step, which creates a function for the expectation of the log-likelihood evaluated using the current estimate for the parameters, and a maximization step, which computes

parameters maximizing the expected log-likelihood found on the E step[11]. Since our system is non-linear, we approximate the likelihood based on the linearization, and the UF is used for filtering and smoothing. We start with an initial guess θ_0 and get the estimation θ^* by iterations, the following implementation regards (2.3) are from [1].

1.The expectation-step:

$$Q(\theta|\theta_n) = E[\log\ p(\theta|Y_{aug})|Y_{obs}, \theta_n]$$

2.The Maximization-step :

$$\theta_{n+1} = arg\ \max_{\theta}\ Q(\theta|\theta_n) :$$

3.The complete log-likelihood:

$$\begin{aligned}\log(L) &= \log P(x_1, \cdots, x_N, y_1, \cdots, y_N) \\ &= \log P(y_N|y_{N-1}, \cdots, y_1, x_N, \cdots, x_1) + \cdots + \log P(y_1|x_N, \cdots, x_1) \\ &+ \log P(x_N|x_{N-1}, \cdots, x_1) + \cdots + \log P(x_1)\end{aligned}$$

Process for the EM algorithm

- Initialize the model parameters $\Theta = \{Q, R, \alpha, \lambda_1, \lambda_2, \eta_1, \eta_2\}$
- Repeat the following two steps until convergence:
  - For $k = 1, \cdots, N$, compute $\bar{x}_{k+1}, \overline{P}_{k+1}, \hat{x}_{k+1}, \hat{P}_{k+1}$ and $\overline{P}_{x_k x_{k+1}}$; For $k = N, \cdots, 1$ Calculate the smoothed values $x_{k|N}$, and $P_{k|N}$ using (2.9).
  - Get new values of the parameters Θ to maximize the E step.

2.2.4 Model Validation

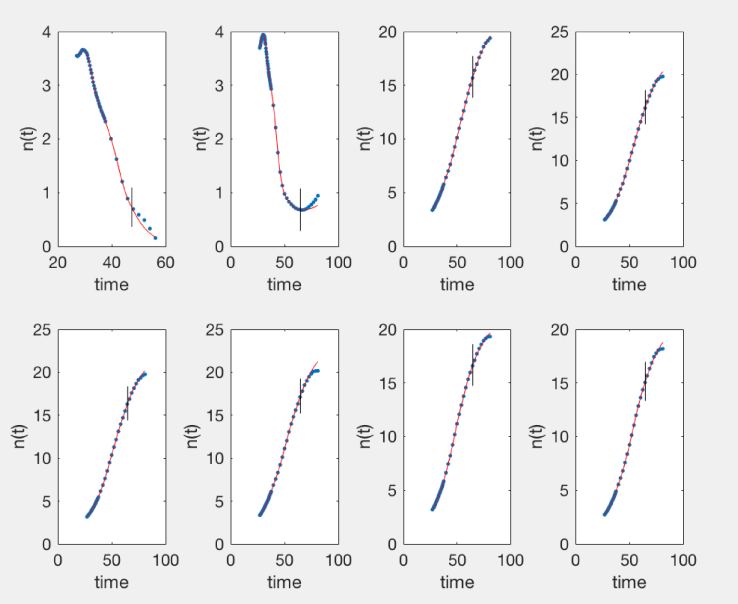
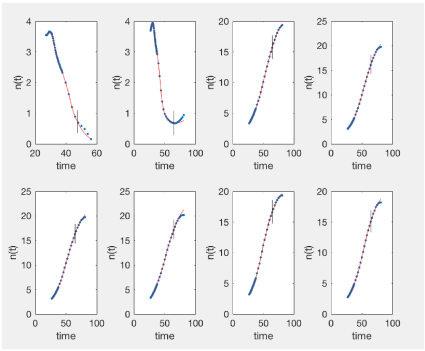
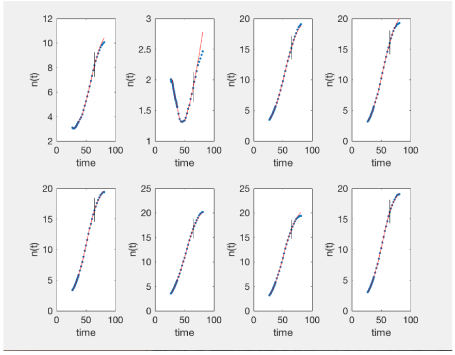


Figure 2.2: Parameter estimation for BafilomycinA1

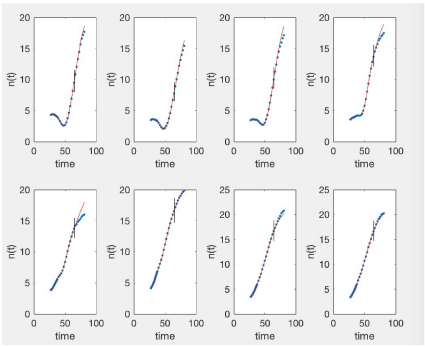
Fig 2.2 shows one of the compounds, BafilomycinA1, that we run parameter estimation using EM algorithm and Unscented Filter with. All eight pictures show the cell index under eight diluted toxic environments. We can roughly see that for the highest concentration, the curve is downward, which indicates that cells go extinction under that concentration. And the cells survive under a concentration less or equal to the one used in the third graph. The observations before the vertical line represent the training set (70% of the data), which are used for parameter estimation. The test set is formed with the data after the vertical line. And we use the parameters we get to do the prediction and validate the model by comparing prediction results to the experimental data. The concentration reduces by a third from top to bottom, and from left to right[1]. Here we show the model validation graphs of 8 chemical compounds in various clusters and the parameter table we get:



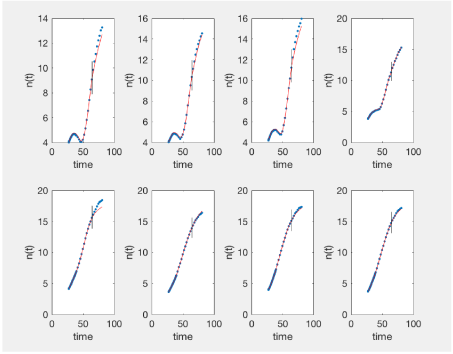
(a) BafilomycinA1



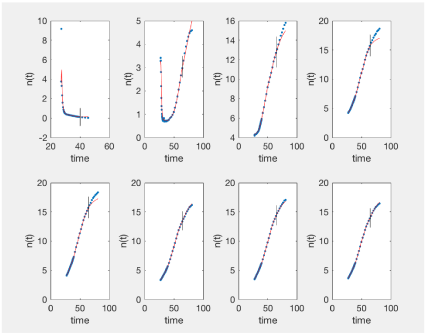
(b) ML7hydrochloride



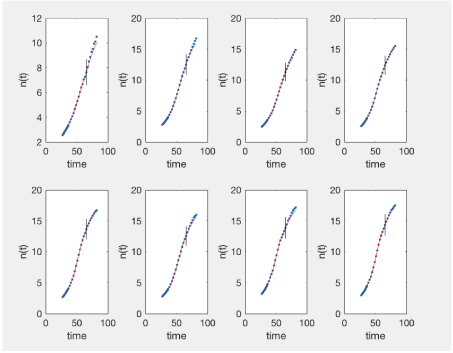
(c) Docetaxel



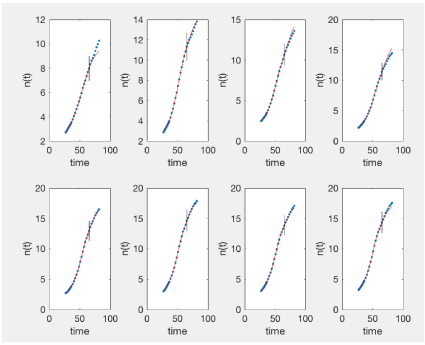
(d) Dimethylenenastrone



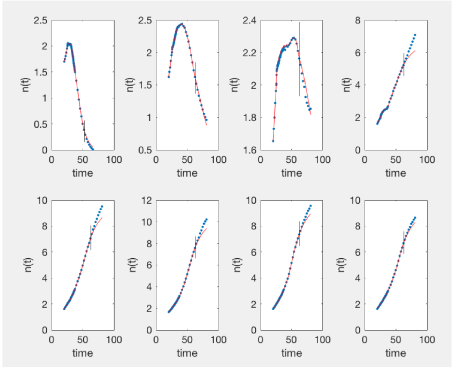
(e) W7 HCl



(f) NU7026



(g) CRT0044876



(h) CCCP

Figure 2.3: Model validations with chemical compounds

Toxicant	Cluster	β	K	λ_1	η_1	λ_2	η_2	α
BafilomycinA1	2	0.0766	21.9126	0.1488	0.9064	0.0412	0.2576	0.1793
ML7hydrochloride	10	0.0766	21.9126	0.0786	0.0064	0	0.0113	0.3051
Docetaxel	4	0.0766	21.9126	0.1758	0.2320	0	0.2265	0.6169
Dimethylenenastrone	10	0.074	18.17	0.0911	0.0985	0	0.0078	6.1813e-04
W7 HCl	8	0.074	18.17	0.2037	0.4947	0	0.1000	1.5769e-04
NU7026	1	0.0738	18.70796	0.2617	0.0131	0	0.1177	0.9535
CRT0044876	1	0.0738	18.70796	0.1693	0.0719	0.0043	0.5000	0.0351
CCCP	6	0.0658	5.8458	0	0.0421	0.0343	0.1072	0.0311

Table 2.3: Parameters for chemical compounds

2.3 Seperation Graph

We can get the numerical solution using Runge-Kutta-Fehlberg Method (RKF45) for the deterministic model (2.2). Here are two pictures of BafilomycinA1 with different initial concentrations of $CE(0)$. From direct observation from Fig 2.4, $CE(t)$ and $CI(t)$ tend to go to zero as t increases while $n(t)$ approaches to its capacity K for the survival case. If all cells extinct at the end Fig 2.5, $CE(t)$ tends to stay at a particular positive value.

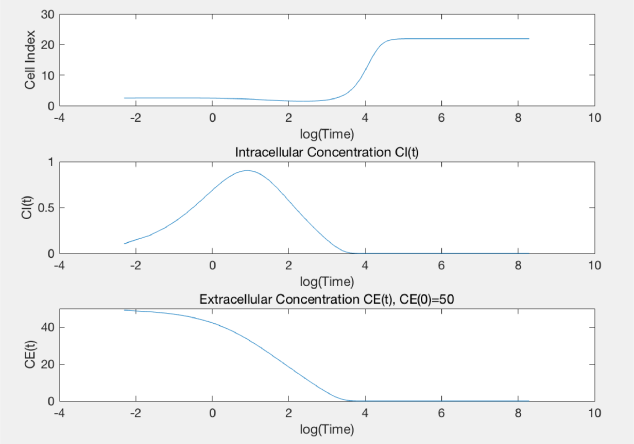


Figure 2.4: Survival case

To determine if cells will survive at the end, we choose t to be big enough (around 1000 hours in the simulation). If $n(t)$ is not zero (bigger than 10^{-3} in

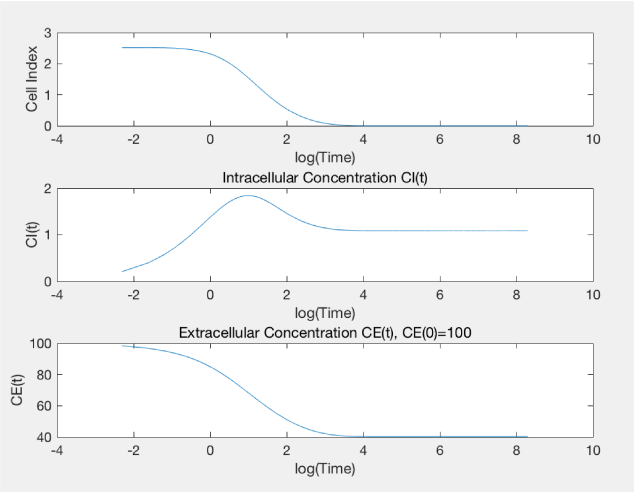


Figure 2.5: Extinction case

numerical sense), the cells survive under that initial concentration of the test substance.

After running tests for a specific compound for several times, we can narrow the range of the turning point for cells to go from survival to extinction. The threshold value is the lowest concentration of toxicants that will make cells go extinction, which can be determined using Runge-Kutta-Fehlberg Method (RKF45). And the threshold for a specific condition (fixed starting cell index $n(0)$) can be tested under the requirement of accuracy. Fig 2.6 shows two examples of seperation graphs. Every point shows an outcome for a combination with $n(0)$ and $CE(0)$. In the separation graph, a blue circle means that cells go extinction after a period while a red asterisk indicates that the cells survive in the end. In this sense, the threshold value is the first blue circle in any verticle line with different initial cell index $n(0)$.

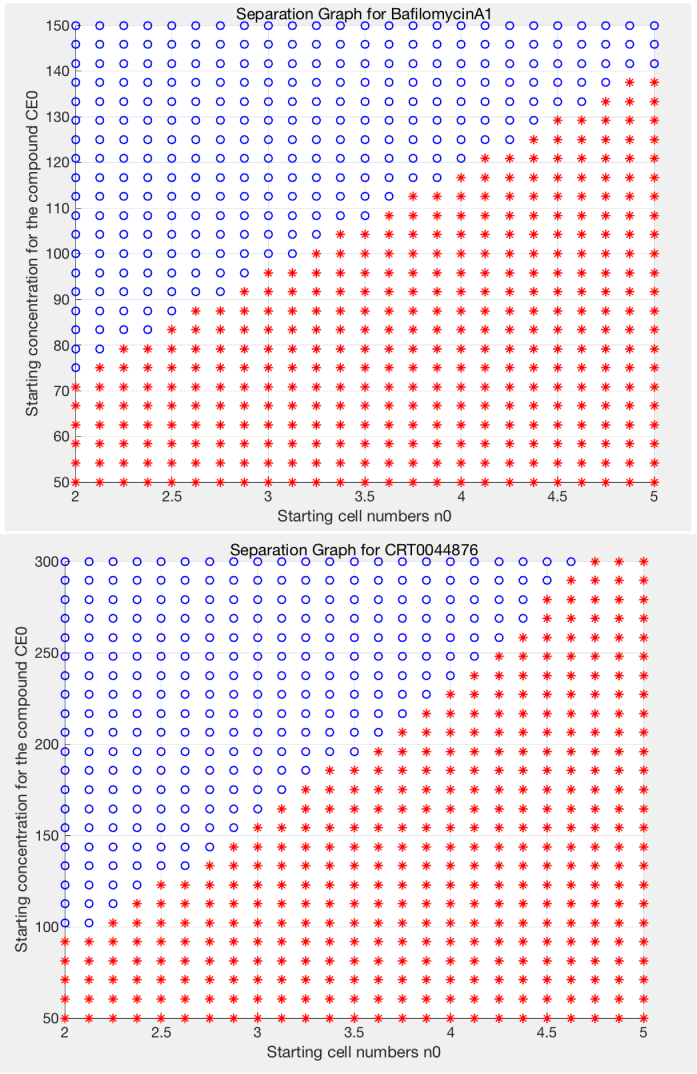


Figure 2.6: Separation graphs

Chapter 3

Convergence Analysis

In this chapter, a stochastic model is derived starting from the deterministic model (2.2). To solve the system numerically, we consider multiple dimensional Taylor Expansions. After that, the convergent conditions for numerical schemes are explored.

3.1 Stochastic Model

A global parameter sensitivity analysis in [1] shows that the parameters K and β are the most sensitive parameters of the model. We consider parameters β and $\gamma = \frac{\beta}{K}$. In practice we usually estimate a parameter by an average value plus an error term. So the parameters β and γ are replaced by random variables :

$$\tilde{\beta} = \beta + error_1$$

$$\tilde{\gamma} = \gamma + error_2$$

By the central limit theorem, the error terms may be approximated by a normal distribution with zero mean.The deterministic model is now replaced by a system of stochastic differential equations:

$$\begin{aligned} dn(t) &= n(t)(\beta - \gamma n(t) - \alpha CI(t))dt + \sigma_1 n(t)dB_1(t) - \sigma_2 n^2(t)dB_2(t) \\ dCI(t) &= (\lambda_1^2 CE(t) - \eta_1^2 CI(t))dt \\ dCE(t) &= (\lambda_2^2 CI(t)n(t) - \eta_2^2 CE(t)n(t))dt \end{aligned} \tag{3.1}$$

where $\sigma_i, i = 1,2$ are noise intensities and $B_i(t), i = 1,2$ are independent standard Brownian motions.

We can express the model (3.1) as

$$dX_t = f(X_t)dt + \begin{pmatrix} \sum_{j=1}^{2} \sigma_j(X_t)dB_j(t) \\ 0 \\ 0 \end{pmatrix} \tag{3.2}$$

where X_t is a three-dimensional variable

$$X_t = \begin{pmatrix} n_t \\ CI_t \\ CE_t \end{pmatrix}$$

and the drift part is

$$f(X_t) = \begin{pmatrix} n_t(\beta - \gamma n_t - \alpha CI_t) \\ \lambda_1^2 CE_t - \eta_1^2 CI_t \\ (\lambda_2^2 CI_t - \eta_2^2 CE_t)n_t \end{pmatrix}$$

The diffusion term has two components: a linear part $\sigma_1 n(t)dB_1(t)$ and a non-linear part $\sigma_2 n^2(t)dB_2(t)$.

3.2 Stochastic Taylor Expansion

In this section, the multi-dimension of Ito-Taylor expansion and the multi-dimension of Stratonovich-Taylor expansion are presented. Two explicit numerical schemes, Euler scheme and Milstein scheme are derived according to the expansions.

3.2.1 Multi-Dimensional Ito-Taylor Expansion

Suppose we have stochastic differential equation [20]:

$$dX_t^k = f_k(X_t)dt + \sum_{j=1}^{n} \sigma_{kj}(X_t)dB_t^j (k = 1,2,\cdots d) \tag{3.3}$$

and the corresponding Ito processes:

$$X_t^k = X_{t_0}^k + \int_{t_0}^{t} f_k(X_s)ds + \sum_{j=1}^{n} \int_{t_0}^{t} \sigma_{kj}(X_s)dB_s^j$$

Using Ito's Lemma we get

$$\begin{aligned} X_t^k = X_{t_0}^k &+ \int_{t_0}^{t} (f_k(X_{t_0}) + \int_{t_0}^{s} L^0 f_k(X_\tau)d\tau + \sum_{j=1}^{n} \int_{t_0}^{s} L^j f_k(X_\tau)dB_\tau^j)ds \\ &+ \sum_{j=1}^{n} \int_{t_0}^{t} (\sigma_{kj}(X_{t_0}) + \int_{t_0}^{s} L^0 \sigma_{kj}(X_\tau)d\tau + \sum_{j=1}^{n} \int_{t_0}^{s} L^j \sigma_{kj}(X_\tau)dB_\tau^l)dB_s^j \end{aligned}$$

where the linear operators L^0 and L^j are defined by

$$L^0 := \frac{\partial}{\partial t} + \sum_{k=1}^{d} f_k \frac{\partial}{\partial x_k} + \frac{1}{2} \sum_{k,m=1}^{d} \sum_{j=1}^{n} \sigma_{kj}\sigma_{mj} \frac{\partial^2}{\partial x^k \partial x^m}$$

$$L^j := \sum_{k=1}^{d} \sigma_{kj} \frac{\partial}{\partial x^k} (j = 1, 2, \cdots, n)$$

From the formulas above we can derive the Ito-Taylor expansion for our system, where we have

$$f_1 = n(t)(\beta - \gamma n(t) - \alpha CI(t))$$

$$f_2 = \lambda_1^2 CE(t) - \eta_1^2 CI(t)$$

$$f_3 = \lambda_2^2 CI(t)n(t) - \eta_2^2 CE(t)n(t)$$

$$\sigma_{11} = \sigma_1 n(t)$$

$$\sigma_{12} = -\sigma_2 n^2(t)$$

$$\sigma_{21} = \sigma_{22} = \sigma_{31} = \sigma_{32} = 0$$

$$L^0 f_1 = f_1 \frac{\partial f_1}{\partial n} + f_2 \frac{\partial f_1}{\partial CI} + \frac{1}{2}(\sigma_{11}^2 \frac{\partial f_1^2}{\partial^2 n} + \sigma_{12}^2 \frac{\partial f_1^2}{\partial^2 n})$$

$$= f_1(\beta - 2\gamma n(t) - \alpha CI(t)) - f_2 \alpha n(t) - 2\gamma((\sigma_1 n(t))^2 + (\sigma_2 n(t)^2)^2)$$

$$L^0 f_2 = \lambda_1^2 f_3 - \eta_1^2 f_2$$

$$L^0 f_3 = f_1(\lambda_2^2 CI(t) - \eta_2^2 CE(t)) + \lambda_2^2 n(t) f_2 - \eta_2^2 n(t) f_3$$

$$L^1 f_1 = \sigma_1 n(t)(\beta - 2\gamma n(t) - \alpha CI(t))$$

$$L^1 f_2 = L^2 f_2 = 0$$

$$L^1 f_3 = \sigma_1 n(t)(\lambda_2^2 CI(t) - \eta_2^2 CE(t))$$

$$L^2 f_1 = -\sigma_2 n(t)^2(\beta - 2\gamma n(t) - \alpha CI(t))$$

$$L^2 f_3 = -\sigma_2 n(t)^2(\lambda_2^2 CI(t) - \eta_2^2 CE(t))$$

$$L^0 \sigma_{11} = f_1 \sigma_1$$

$$L^1 \sigma_{11} = \sigma_1^2 n(t)$$

$$L^2 \sigma_{11} = -\sigma_1 \sigma_2 n(t)^2$$

$$L^0 \sigma_{12} = -2 f_1 \sigma_2 n(t) - \sigma_2(\sigma_1^2 + \sigma_2^2)$$

$$L^1 \sigma_{12} = -2\sigma_1 \sigma_2 n(t)^2$$

$$L^2 \sigma_{12} = 2\sigma_2^2 n(t)^3$$

3.2.2 Multi-Dimensional Stratonovich-Taylor Expansion

The Stratonovich representation [20] of the stochastic differential equation (3.3) is

$$dX_t^k = \overline{f_k}(X_t)dt + \sum_{j=1}^{n} \sigma_{kj}(X_t) \circ dB_t^j (k = 1, 2, \cdots d) \tag{3.4}$$

And

$$X_t^k = X_{t_0}^k + \int_{t_0}^{t} \overline{f_k}(X_s)ds + \sum_{j=1}^{n} \int_{t_0}^{t} \sigma_{kj}(X_s) \circ dB_s^j$$

$$\overline{f_i} = f_i - \frac{1}{2} \sum_{k=1}^{d} \sum_{j=1}^{n} \sigma_{kj}(X_t) \frac{\partial \sigma_{ij}}{\partial x_k}$$

From the Stratonovich form of Ito's Lemma we get

$$X_t^k = X_{t_0}^k + \int_{t_0}^t (\overline{f_k}(X_{t_0}) + \int_{t_0}^s \overline{L^0}\overline{f_k}(X_\tau)d\tau + \sum_{j=1}^n \int_{t_0}^s \overline{L^j}\overline{f_k}(X_\tau)dB_\tau^j)ds$$
$$+ \sum_{j=1}^n \int_{t_0}^t (\sigma_{kj}(X_{t_0}) + \int_{t_0}^s \overline{L^0}\sigma_{kj}(X_\tau)d\tau + \sum_{l=1}^n \int_{t_0}^s \overline{L^l}\sigma_{kj}(X_\tau)dB_\tau^l)dB_s^j,$$

where the linear operators $\overline{L^0}$ and $\overline{L^j}$ are defined by

$$\overline{L^0} := \frac{\partial}{\partial t} + \sum_{k=1}^d \overline{f_k}\frac{\partial}{\partial x_k}$$
$$\overline{L^j} := \sum_{k=1}^d \sigma_{kj}\frac{\partial}{\partial x^k}(j = 1,2,\cdots,n)$$

Using formulas above we can derive the Stratonovich-Taylor expansion for our system.

$$\begin{aligned}
\overline{f_1}(X_t) &= f_1 - \frac{1}{2}(\sigma_{11}\frac{\partial\sigma_{11}}{\partial n} + \sigma_{12}\frac{\partial\sigma_{12}}{\partial n}) \\
&= f_1 - \frac{1}{2}(\sigma_1^2 n(t) + 2\sigma_2^2 n(t)^3) \\
&= n(t)(\beta - \gamma n(t) - \alpha CI(t) - \frac{1}{2}\sigma_1^2 - \sigma_2^2 n(t)^2) \\
\overline{f_2}(X_t) &= f_2 = \lambda_1^2 CE(t) - \eta_1^2 CI(t) \\
\overline{f_3}(X_t) &= f_3 = \lambda_2^2 CI(t)n(t) - \eta_2^2 CE(t)n(t) \\
\overline{L^0}\overline{f_1} &= \overline{f_1}\frac{\partial\overline{f_1}}{\partial n} + \overline{f_2}\frac{\partial\overline{f_1}}{\partial CI} + \overline{f_3}\frac{\partial\overline{f_1}}{\partial CE} \\
&= \overline{f_1}(\beta - 2\gamma n(t) - \alpha C_0(t) - \frac{1}{2}\sigma_1^2 - 3\sigma_2^2 n(t)^2) - \overline{f_2}\alpha n(t) \\
\overline{L^0}\overline{f_2} &= -\overline{f_2}\eta_1^2 + \overline{f_3}\lambda_1^2 \\
\overline{L^0}\overline{f_3} &= \overline{f_1}(\lambda_2^2 CI(t) - \eta_2^2 CE(t)) + \overline{f_2}\lambda_2^2 n(t) - \overline{f_3}\eta_2^2 n(t)
\end{aligned}$$

$$\overline{L^1}\overline{f_1} = \sigma_{11}\frac{\partial \overline{f_1}}{\partial n} = \sigma_1 n(t)(\beta - 2\gamma n(t) - \alpha CI(t) - \frac{1}{2}\sigma_1^2 - 3\sigma_2^2 n(t)^2)$$

$$\overline{L^1}\overline{f_2} = \sigma_{11}\frac{\partial \overline{f_2}}{\partial n} = 0$$

$$\overline{L^1}\overline{f_3} = \sigma_{11}\frac{\partial \overline{f_3}}{\partial n} = \sigma_1 n(t)(\lambda_2^2 CI(t) - \eta_2^2 CE(t))$$

$$\overline{L^2}\overline{f_1} = \sigma_{12}\frac{\partial \overline{f_1}}{\partial n} = -\sigma_2 n^2(t)(\beta - 2\gamma n(t) - \alpha CI(t) - \frac{1}{2}\sigma_1^2 - 3\sigma_2^2 n(t)^2)$$

$$\overline{L^2}\overline{f_2} = \sigma_{12}\frac{\partial \overline{f_2}}{\partial n} = 0$$

$$\overline{L^2}\overline{f_3} = \sigma_{12}\frac{\partial \overline{f_3}}{\partial n} = -\sigma_2 n^2(t)(\lambda_2^2 CI(t) - \eta_2^2 CE(t))$$

$$\overline{L^0}\sigma_{11} = \sigma_1\overline{f_1} = \sigma_1 n(t)(\beta - \gamma n(t) - \alpha CI(t) - \frac{1}{2}\sigma_1^2 - \sigma_2^2 n(t)^2)$$

$$\overline{L^0}\sigma_{12} = -2\sigma_2 n(t)\overline{f_1} = -2\sigma_2 n(t)(\beta - \gamma n(t) - \alpha CI(t) - \frac{1}{2}\sigma_1^2 - \sigma_2^2 n(t)^2)$$

$$\overline{L^1}\sigma_{11} = \sigma_{11}\frac{\partial \sigma_{11}}{\partial n} = \sigma_1^2 n(t) \tag{3.5}$$

$$\overline{L^1}\sigma_{12} = \sigma_{11}\frac{\partial \sigma_{12}}{\partial n} = -2\sigma_1\sigma_2 n(t)^2 \tag{3.6}$$

$$\overline{L^2}\sigma_{11} = \sigma_{12}\frac{\partial \sigma_{11}}{\partial n} = -\sigma_1\sigma_2 n(t)^2 \tag{3.7}$$

$$\overline{L^2}\sigma_{12} = \sigma_{12}\frac{\partial \sigma_{12}}{\partial n} = 2\sigma_2^2 n(t)^3 \tag{3.8}$$

3.2.3 Euler Method(order 0.5)

The $k-th$ component of the Euler scheme is given by ([20], pp 341).

$$Y_{i+1}^k = Y_i^k + f_i^k\Delta + \sum_{j=1}^{n}\sigma_j^k\Delta B^j \tag{3.9}$$

for $k = 1, 2, \cdots, d$ where the drift and diffusion coefficients are d-dimensional vectors, and

$$\Delta B^j = B^j_{\tau_{i+1}} - B^j_{\tau_i} = I_{(j)} = J_{(j)}$$

is the $N(0;\Delta)$ distributed increment of the $j-th$ component of the m-dimensional standard Wiener process on $[i, i+1]$, and B^{j_1} and B^{j_2} are independent for $j_1 \neq j_2$. Here $I_{(j)}$ denotes the multiple Ito integral while $J_{(j)}$ denotes the multiple Stratonovich integral.

For our stochastic system we have,

$$\begin{aligned} n_{i+1} &= n_i + n_i(\beta - \gamma n_i - \alpha CI_i)\Delta t + \sigma_1 n_i I_{(1)} - \sigma_2 n_i^2 I_{(2)} \\ CI_{i+1} &= CI_i + (\lambda_1^2 CE_i - \eta_1^2 CI_i)\Delta t \\ CE_{i+1} &= CE_i + (\lambda_2^2 CI_i n_i - \eta_2^2 CE_i n_i)\Delta t \end{aligned}$$

where $I_{(1)}$ and $I_{(2)}$ are independent Gaussian random variables $I_{(1)} \sim N(0;\Delta), I_{(2)} \sim N(0;\Delta)$.

3.2.4 Milstein Method(order 1.0)

In the general multi-dimensional case with $k = 1, 2, \cdots, d$ the $k-th$ component of the Milstein scheme is given by ([20] pp 346)

$$Y^k_{i+1} = Y^k_i + \overline{f^k_i}\Delta + \sum_{j=1}^{n} \sigma^k_j \Delta B^j + \sum_{j_1,j_2=1}^{n} \overline{L^{j_1}}\sigma^k_{j_2} J_{(j_1,j_2)} \tag{3.10}$$

So the only difference compared to the Euler scheme is on the first equation of our system, from (3.5)(3.6)(3.7)(3.8) and from [20] we obtain:

$$
\begin{aligned}
n_{i+1} &= n_i + n_i(\beta - \gamma n_i - \alpha C I_i - \frac{1}{2}\sigma_1^2 - \sigma_2^2 n_i^2)\Delta t + \sigma_1 n_i I_{(1)} - \sigma_2 n_i^2 I_{(2)} + \\
&\quad \overline{L^1} h_{11} J_{(11)} + \overline{L^1} b_{12} J_{(12)} + \overline{L^2} b_{11} J_{(21)} + \overline{L^2} b_{12} J_{(22)} \\
J_{(j,j)} &= \frac{1}{2}(\Delta B^j)^2 \\
J^p_{(j_1,j_2)} &= \Delta(\frac{1}{2}\xi_{j_1}\xi_{j_2} + \sqrt{\rho_p}(\mu_{j1,p}\xi_{j_2} - \mu_{j_2,p}\xi_{j_1})) + \frac{\Delta}{2\pi}\sum_{r=1}^{p}\frac{1}{r}(\xi_{j_1,r}(\sqrt{2}\xi_{j_2} + \eta_{j_2,r}) - \\
&\quad \xi_{j_2,r}(\sqrt{2}\xi_{j_1} + \eta_{j_1,r})) \\
\rho_p &= \frac{1}{12} - \frac{1}{2\pi^2}\sum_{r=1}^{p}\frac{1}{r^2} \\
J^3_{(12)} &= \Delta(\frac{1}{2}\xi_1\xi_2 + \sqrt{\rho_3}(\mu_{(1,3)}\xi_2 - \mu_{(2,3)}\xi_1)) + \frac{\Delta}{2\pi}\sum_{r=1}^{3}\frac{1}{r}(\xi_{(1,r)}(\sqrt{2}\xi_2 + \eta_{(2,r)}) - \\
&\quad \xi_{(2,r)}(\sqrt{2}\xi_1 + \eta_{(1,r)})) \\
J^3_{(21)} &= \Delta(\frac{1}{2}\xi_1\xi_2 + \sqrt{\rho_3}(\mu_{(2,3)}\xi_1 - \mu_{(1,3)}\xi_2)) + \frac{\Delta}{2\pi}\sum_{r=1}^{3}\frac{1}{r}(\xi_{(2,r)}(\sqrt{2}\xi_1 + \eta_{(1,r)}) - \\
&\quad \xi_{(1,r)}(\sqrt{2}\xi_2 + \eta_{(2,r)})) \\
J_{(1,1)} &= \frac{1}{2}(\xi_1)^2 h \\
J_{(2,2)} &= \frac{1}{2}(\xi_2)^2 h \\
\rho_3 &= \frac{1}{12} - \frac{49}{72\pi^2}
\end{aligned}
$$

The rest of the scheme stays the same

$$
\begin{aligned}
CI_{i+1} &= CI_i + (\lambda_1^2 CE_i - \eta_1^2 CI_i)\Delta t \\
CE_{i+1} &= CE_i + (\lambda_2^2 CI_i n_i - \eta_2^2 CE_i n_i)\Delta t
\end{aligned}
$$

3.3 Convergent Conditions

Before doing numerical analysis on our non-linear stochastic system, we need to consider which numerical scheme should be used. Since different schemes have various conditions for convergence, we need to make sure that our system satisfies certain conditions. Besides, a good scheme should also be efficient and accurate. Form SDE (3.1), we have

$$f(X) = \begin{pmatrix} n\beta - \gamma n^2 - \alpha nCI \\ \lambda_1^2 CE - \eta_1^2 CI \\ \lambda_2^2 CIn - \eta_2^2 CEn \end{pmatrix}$$

$$f(X_1) - f(X_2) = \begin{pmatrix} (n_1 - n_2)\beta - \gamma(n_1^2 - n_2^2) - \alpha(n_1 CI_1 - n_2 CI_2) \\ \lambda_1^2(CE_1 - CE_2) - \eta_1^2(CI_1 - CI_2) \\ \lambda_2^2(CI_1 n_1 - CI_2 n_2) - \eta_2^2(CE_1 n_1 - CE_2 n_2) \end{pmatrix}$$

$$\begin{aligned} |f(X_1) - f(X_2)|^2 = {} & (n_1 - n_2)^2\beta^2 + \gamma^2(n_1 + n_2)^2(n_1 - n_2)^2 + \alpha^2(n_1 CI_1 - n_2 CI_2)^2 \\ & - 2\gamma\beta(n_1 + n_2)(n_1 - n_2)^2 - 2\alpha\beta(n_1 CI_1 - n_2 CI_2)(n_1 - n_2) \\ & + 2\gamma\alpha(n_1 CI_1 - n_2 CI_2)(n_1 + n_2)(n_1 - n_2) \\ & + \lambda_1^4(CE_1 - CE_2)^2 + \eta_1^4(CI_1 - CI_2)^2 - 2\lambda_1^2\eta_1^2(CE_1 - CE_2)(CI_1 - CI_2) \\ & + \lambda_2^4(CI_1 n_1 - CI_2 n_2)^2 + \eta_2^4(CE_1 n_1 - CE_2 n_2)^2 \\ & - 2\lambda_2^2\eta_2^2(CI_1 n_1 - CI_2 n_2)(CE_1 n_1 - CE_2 n_2) \end{aligned}$$

Since we can not garantee $n(t)$ will be bounded the whole time, we need to relax the global Lipschitz condition for $f(X)$.

We want to explore a numerical scheme which relaxes the globally Lipschitz condition for the coefficient fuction. In [22], it is shown that methods with global Lipschitz condition can also work well with nonglobally Lipschitz coefficients if a small number of bad-behavior trajectories are discarded. Higham and Mao prove in [23] that the Euler scheme is convergent if the SDE is locally Lipschitz and the p-th moments of the exact and numerical solution are bounded. In [24] it is proven that if with the assumptions of the coefficients are one-sided Lipschitz and the moments of X_k are bounded and the scheme is convergent with order q in the global Lipschitz case then the scheme has the same convergence order q in the considered nonglobal Lipschitz case. In [24], the boundedness of moments for a balanced method is proved. But in general, the boundness on moments of a method X_k is often rather difficult to prove. For Implicit scheme, Mao and Szpruch show in [27] that under a dissipative condition on the drift coefficient and superlinear growth condition on the diffusion coefficient the Backward Euler-Maruyama scheme converges with strong order of $\frac{1}{2}$. The paper [25] proves that for a large class of SDEs with non-globally Lipschitz continuous coefficients, the Euler scheme converges neither in the strong mean-square sense nor in the numerically weak sense to the exact solution at a finite time point. In [25] Martin Hutzenthaler proposes a "tamed" version of the explicit Euler scheme. The article shows that tamed Euler scheme converges strongly with convergence order $\frac{1}{2}$ to the exact solution of the SDE with the drift coefficient function being globally one-sided Lipschitz continuous and at most polynomially growing derivative. Also, the diffusion coefficient should be globally Lipschitz continuous [25]. Sotirios Sabanis explores the tamed Euler scheme further in [29].

3.4 The Tamed Euler Scheme

In [25], the tamed Euler scheme is presented, and it is shown that it converges strongly with the standard convergence order $\frac{1}{2}$ to the exact solution of the SDE:

$$dX_t = f(X_t)dt + \sigma(X_t)dB_t \tag{3.11}$$

The tamed Euler scheme

$$X_{t+1} = X_t + \frac{T/N \cdot f(X_t)}{1 + \|T/N \cdot f(X_t)\|} + \sigma(X_t)(B_{(t+1)\cdot T/N} - B_{t\cdot T/N}) \tag{3.12}$$

Convergent Conditions:

- $f(X) : R^d \to R^d$ be a continuously differentiable and globally one-sided Lipschitz continuous function whose derivative grows at most polynomially. i.e. $L, c \in (0, \infty)$, for all $X, Y \in R^d$

$$\langle X - Y, f(X) - f(Y)\rangle \leq L|X - Y|^2 \tag{3.13}$$

$$|f'(X)| \leq c(1 + |X|^c) \tag{3.14}$$

- $\sigma(X)$ be a globally Lipschitz continuous function.

For our model (3.2), notice that the tamed Euler scheme requires the diffusion term $\sigma_1 n(t)dB_1(t) - \sigma_2 n^2(t)dB_2(t)$ to be globally Lipschitz continuous. This condition can be satisfied only when $\sigma_2 = 0$.

3.5 The Implicit Euler method

The primary object of this section is to present the implicit Euler scheme and prove its convergence. The implicit Euler scheme is:

$$X_{t+1} = X_t + f(X_{t+1}) \cdot T/N + \sigma(X_t)(B_{(t+1)\cdot T/N} - B_{t\cdot T/N}) \tag{3.15}$$

We apply the implicit Euler scheme 3.15 to our system

$$\begin{pmatrix} n_{t+1} \\ CI_{t+1} \\ CE_{t+1} \end{pmatrix} = \begin{pmatrix} n_t \\ CI_t \\ CE_t \end{pmatrix} + h \begin{pmatrix} n_{t+1}\beta - \gamma n_{t+1}^2 - \alpha n_{t+1} CI_{t+1} \\ \lambda_1^2 CE_{t+1} - \eta_1^2 CI_{t+1} \\ \lambda_2^2 CI_{t+1} n_{t+1} - \eta_2^2 CE_{t+1} n_{t+1} \end{pmatrix} + \begin{pmatrix} \sigma_1 n_t \Delta B_t - \sigma_2 n_t \Delta B_t \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} n_{t+1} \\ CI_{t+1} \\ CE_{t+1} \end{pmatrix} - h \begin{pmatrix} n_{t+1}\beta - \gamma n_{t+1}^2 - \alpha n_{t+1} CI_{t+1} \\ \lambda_1^2 CE_{t+1} - \eta_1^2 CI_{t+1} \\ \lambda_2^2 CI_{t+1} n_{t+1} - \eta_2^2 CE_{t+1} n_{t+1} \end{pmatrix} = \begin{pmatrix} n_t \\ CI_t \\ CE_t \end{pmatrix} + \begin{pmatrix} \sigma_1 n_t \Delta B_t - \sigma_2 n_t \Delta B_t \\ 0 \\ 0 \end{pmatrix}$$

To make the equations easier to read, we denote variables $a = n_t, b = CI_t, c = CE_t, x = n_{t+1}, y = CI_{t+1}, z = CE_{t+1}, B = \sigma_1 n_t \Delta B_t - \sigma_2 n_t^2 \Delta B_t$. The new system is presented below

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} - h \begin{pmatrix} x(\beta - \gamma x - \alpha y) \\ \lambda_1^2 z - \eta_1^2 y \\ x(\lambda_2^2 y - \eta_2^2 z) \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} + \begin{pmatrix} B \\ 0 \\ 0 \end{pmatrix}$$

$$
\begin{aligned}
x - hx(\beta - \gamma x - \alpha y) &= a + B \\
y - h(\lambda_1^2 z - \eta_1^2 y) &= b \\
z - hx(\lambda_2^2 y - \eta_2^2 z) &= c
\end{aligned} \tag{3.16}
$$

Mathematica is used to solve y and z with respect to x and combine the system into one equation

```
Solve[{y - h (g^2 z - k^2 y) == b, z - h x (a^2 y - d^2 z) == c}, {y, z}]
```

$$\left\{\left\{y \to -\frac{b + c\,g^2\,h + b\,d^2\,h\,x}{-1 - h\,k^2 - d^2\,h\,x + a^2\,g^2\,h^2\,x - d^2\,h^2\,k^2\,x},\ z \to -\frac{c + c\,h\,k^2 + a^2\,b\,h\,x}{-1 - h\,k^2 - d^2\,h\,x + a^2\,g^2\,h^2\,x - d^2\,h^2\,k^2\,x}\right\}\right\}$$

```
A (b + c g^2 h) = G
1 + k^2 h = K
a + W = W
d^2 h + d^2 k^2 h^2 - a^2 g^2 h^2 = T
A (b d^2 h) = R
```

$$\left[x - h\,x\left(B - Y\,x - \frac{G + R\,x}{K + T\,x}\right) == W,\ x\right]$$

The new equations for y and z with respect to x are

$$CI_{t+1} = y = \frac{b + c\lambda_1^2 h + b\eta_2^2 hx}{1 + h\eta_1^2 + \eta_2^2 hx + \delta h^2 x} \tag{3.17}$$

$$CE_{t+1} = z = \frac{c + c\eta_1^2 h + b\lambda_2^2 hx}{1 + h\eta_1^2 + \eta_2^2 hx + \delta h^2 x} \tag{3.18}$$

Using Mathematica, (3.17) and (3.18) are put back into (3.16); the cubic equation are solved with respect to x. Since only the real root should be used for our numerical analysis, we can solve the implicit Euler using the only real root we get.

3.5.1 Proof of Convergence

For the convergence of the implicit Euler method, we assume that $n_t \geq 0$ for all $t \geq 0$ and we make the following assumption:

Assumption 1. *Suppose* $\delta = \eta_1^2\eta_2^2 - \lambda_1^2\lambda_2^2 > 0$ *and* $\eta_1^2 > \lambda_1^2, \eta_2^2 > \lambda_2^2$ *in (3.1).*

Lemma 3.5.1. *For the implicit Euler method (3.16), we have*

i. $0 \leq CI_t \leq CE_0$ *and* $0 \leq CE_t \leq CE_0$ *for all* $t \geq 0$.

ii. $\eta_2^2 CE_t > \lambda_2^2 CI_t$.

iii. If $n_{t_1} > n_{t_2}$, *then* $CI_{t_1} < CI_{t_2}$, $CE_{t_1} < CE_{t_2}$, $n_{t_1}CI_{t_1} > n_{t_2}CI_{t_2}$.

Proof. i. By induction we assume that $0 \leq b = CI_t \leq CE_0$ and $0 \leq c = CE_t \leq CE_0$. From (3.17) and (3.18), with the Assumption1, since $x = n_{t+1} \geq 0$,

$$0 \leq CI_{t+1} = y = \frac{b + c\lambda_1^2 h + b\eta_2^2 hx}{1 + h\eta_1^2 + \eta_2^2 hx + \delta h^2 x} < max(b,c)$$
$$0 \leq CE_{t+1} = z = \frac{c + c\eta_1^2 h + b\lambda_2^2 hx}{1 + h\eta_1^2 + \eta_2^2 hx + \delta h^2 x} < max(b,c)$$

So $0 \leq CI \leq CE_0$ and $0 \leq CE \leq CE_0$ if $\eta_1^2 > \lambda_1^2, \eta_2^2 > \lambda_2^2$. That proves 1.

ii. This part is proved by induction:

Step 1: For the initial concentration $CE_0 \neq 0$ and $CI_0 = 0$, it is true that $\eta_2^2 CE_0 > \lambda_2^2 CI_0$.

Step 2: Suppose it is true for the $k-th$ component b and c so that $\eta_2^2 c > \lambda_2^2 b$.

Step 3: For the $k+1-th$ component

$$
\begin{aligned}
CI_{k+1} &= \frac{b+c\lambda_1^2 h+b\eta_2^2 hx}{1+h\eta_1^2+\eta_2^2 hx+\delta h^2 x} \\
CE_{k+1} &= \frac{c+c\eta_1^2 h+b\lambda_2^2 hx}{1+h\eta_1^2+\eta_2^2 hx+\delta h^2 x} \\
\lambda_2^2 CI_{k+1} &= \frac{\lambda_2^2 b+c\lambda_1^2\lambda_2^2 h+b\lambda_2^2\eta_2^2 hx}{1+h\eta_1^2+\eta_2^2 hx+\delta h^2 x} \\
\eta_2^2 CE_{k+1} &= \frac{\eta_2^2 c+c\eta_1^2\eta_2^2 h+b\lambda_2^2\eta_2^2 hx}{1+h\eta_1^2+\eta_2^2 hx+\delta h^2 x}
\end{aligned}
$$

So $\eta_2^2 CE_t > \lambda_2^2 CI_t$ is true by induction.

iii. Calculate the derivatives of CI and CE using (3.17) and (3.18) with respect to x

$$
\begin{aligned}
CI &= \frac{b+c\lambda_1^2 h+b\eta_2^2 hx}{1+h\eta_1^2+\eta_2^2 hx+\delta h^2 x} \\
\frac{\partial CI}{\partial x} &= \frac{b\eta_2^2 h(1+h\eta_1^2+\eta_2^2 hx+\delta h^2 x)-(\eta_2^2 h+\delta h^2)(b+c\lambda_1^2 h+b\eta_2^2 hx)}{(1+h\eta_1^2+\eta_2^2 hx+\delta h^2 x)^2} \\
&= \frac{bh^2\lambda_1^2\lambda_2^2-ch^2\lambda_1^2\eta_2^2-c\delta h^3\lambda_1^2}{(1+h\eta_1^2+\eta_2^2 hx+\delta h^2 x)^2} < 0 \text{ by (2)} \\
CE &= \frac{c+c\eta_1^2 h+b\lambda_2^2 hx}{1+h\eta_1^2+\eta_2^2 hx+\delta h^2 x} \\
\frac{\partial CE}{\partial x} &= \frac{b\lambda_2^2 h(1+h\eta_1^2+\eta_2^2 hx+\delta h^2 x)-(\eta_2^2 h+\delta h^2)(c+c\eta_1^2 h+b\lambda_2^2 hx)}{(1+h\eta_1^2+\eta_2^2 hx+\delta h^2 x)^2} \\
&= \frac{bh\lambda_2^2-ch\eta_2^2+bh^2\lambda_2^2\eta_1^2-ch^2\eta_2^2\eta_1^2-ch^2\delta-c\delta h^3\eta_1^2}{(1+h\eta_1^2+\eta_2^2 hx+\delta h^2 x)^2} < 0 \text{ by (2)} \\
nCI &= \frac{bx+cx\lambda_1^2 h+bx^2\eta_2^2 h}{1+h\eta_1^2+\eta_2^2 hx+\delta h^2 x} \\
\frac{\partial nCI}{\partial x} &= \frac{(b+c\lambda_1^2 h+2bx\eta_2^2 h)(1+h\eta_1^2+\eta_2^2 hx+\delta h^2 x)-(\eta_2^2 h+\delta h^2)(bx+cx\lambda_1^2 h+bx^2\eta_2^2 h)}{(1+h\eta_1^2+\eta_2^2 hx+\delta h^2 x)^2} > 0
\end{aligned}
$$

So If $n_{t_1} > n_{t_2}$, then $CI_{t_1} < CI_{t_2}, CE_{t_1} < CE_{t_2}$ and $n_{t_1} CI_{t_1} > n_{t_2} CI_{t_2}$. □

Convergent Conditions:

From [28] we know the convergent conditions for the implicit Euler (3.15)

1 Coefficients $f(X_t)$ and $\sigma(X_t)$ are locally Lipschitz continuous.

2 For some $\rho \geq 1$, and $r \in N$, $r \geq 1$, there exist positive constants $\alpha_0, \alpha_1, \beta, \beta_0, \beta_1 > 0$, such that for all $X_t \in R^n$

$$-\beta_1 |X_t|^{r+1} - \beta_0 \leq \langle X_t, f(X_t) \rangle \leq \alpha_0 - \alpha_1 |X_t|^{r+1} \tag{3.19}$$

$$\sigma(X_t) \leq \beta(1 + |X_t|^{\rho}) \tag{3.20}$$

3 (3.19)(3.20) must obey $r + 1 > 2\rho$.

4 There exists a constant $L > 0$, such that

$$\langle X - Y, f(X) - f(Y) \rangle \leq L|X - Y|^2 \tag{3.21}$$

5 The coefficients of Equation 3.15 satisfy the polynomial growth condition, there exist positive constants H and h, $h \geq 1$, such that

$$|f(X_t)| \vee |\sigma(X_t)| \leq H(1 + |X_t|^h), \forall X_t \in R^n$$

Condition 1, 5 are fullfilled so we start by proving condition 2 and 3 first then condition 4. Note that we only focus on the linear diffusion term here.

Proof.

$$f(X_t) = \begin{pmatrix} n(\beta - \gamma n - \alpha CI) \\ \lambda_1^2 CE - \eta_1^2 CI \\ (\lambda_2^2 CI - \eta_2^2 CE)n \end{pmatrix}$$

$$X_t = \begin{pmatrix} n \\ CI \\ CE \end{pmatrix} \sigma(X_t) = \sigma_1 n$$

To fullfill (3.20), $\rho \geq 1$, we choose $\rho = 1$, from condition 3, $r > 1$, we choose $r = 2$. From (3.19) we need to find positive constants $\alpha_0, \alpha_1, \beta_0, \beta_1$ so that

$$-\beta_1 |X_t|^3 - \beta_0 \leq \langle X_t, f(X_t) \rangle \leq \alpha_0 - \alpha_1 |X_t|^3$$

We start with the right half $\langle X_t, f(X_t) \rangle \leq \alpha_0 - \alpha_1 |X_t|^3$. From preliminary results 1 we know that CE_t, CI_t are positive and upper bounded by CE_0. n_t is also positive but we don't know if it is bounded.

$$\begin{aligned} \langle X_t, f(X_t) \rangle &= n^2(\beta - \gamma n - \alpha CI) + \lambda_1^2 CI \cdot CE - \eta_1^2 CI^2 + (\lambda_2^2 CI \cdot CE - \eta_2^2 CE^2)n \\ &= -\gamma n^3 + (\beta - \alpha CI)n^2 + (\lambda_2^2 CI \cdot CE - \eta_2^2 CE^2)n + \lambda_1^2 CI \cdot CE - \eta_1^2 CI^2 \\ &\leq -\gamma n^3 + (\beta - \alpha CI)n^2 + \lambda_2^2 CI \cdot CE \cdot n + \lambda_1^2 CI \cdot CE \\ &\leq -\gamma n^3 + \beta n^2 + \lambda_2^2 CE_0^2 \cdot n + \lambda_1^2 CE_0^2 \end{aligned}$$

So we transfer the problem into proving

$$-\gamma n^3 + \beta n^2 + \lambda_2^2 CE_0^2 \cdot n + \lambda_1^2 CE_0^2 \leq \alpha_0 - \alpha_1 |X_t|^3 \tag{3.22}$$

Since

$$\begin{aligned}
|X_t|^3 &= (n^2+CE^2+CI^2)\sqrt{n^2+CE^2+CI^2} \\
&\leq (n^2+CE^2+CI^2)\sqrt{n^2+2CE_0^2} \\
&\leq (n^2+CE^2+CI^2)(n+\sqrt{2}CE_0) \\
&= n^3+\sqrt{2}CE_0n^2+(CE^2+CI^2)n+\sqrt{2}CE_0(CE^2+CI^2) \\
&\leq n^3+\sqrt{2}CE_0n^2+2CE_0^2n+2\sqrt{2}CE_0^3
\end{aligned}$$

So (3.22) becomes

$$\begin{aligned}
-\gamma n^3+\beta n^2+\lambda_2^2CE_0^2\cdot n+\lambda_1^2CE_0^2 &\leq \alpha_0-\alpha_1(n^3+\sqrt{2}CE_0n^2+2CE_0^2n+2\sqrt{2}CE_0^3) \\
0 \leq \alpha_0+\underbrace{(\gamma-\alpha_1)n^3-(\alpha_1\sqrt{2}CE_0+\beta)n^2-(2\alpha_1CE_0^2+\lambda_2^2CE_0^2)n-2\sqrt{2}\alpha_1CE_0^3-\lambda_1^2CE_0^2}_{PolonomialFunction}
\end{aligned} \tag{3.23}$$

If $\gamma > \alpha_1$, we can find the mimimum of the polonomial function for $n \geq 0$, so the value of α_0 can be determined such that (3.23) holds.

We focus on the left half of (3.19)

$$\begin{aligned}
-\beta_1|X_t|^3-\beta_0 &\leq \langle X_t, f(X_t)\rangle \\
\langle X_t, f(X_t)\rangle &= n^2(\beta-\gamma n-\alpha CI)+\lambda_1^2CI\cdot CE-\eta_1^2CI^2+(\lambda_2^2CI\cdot CE-\eta_2^2CE^2)n \\
&= -\gamma n^3+(\beta-\alpha CI)n^2+(\lambda_2^2CI\cdot CE-\eta_2^2CE^2)n+\lambda_1^2CI\cdot CE-\eta_1^2CI^2 \\
&\geq -\gamma n^3-\alpha CIn^2-\eta_2^2CE^2n-\eta_1^2CI^2 \\
&\geq -\gamma n^3-\alpha CIn^2-\eta_2^2CE_0^2n-\eta_1^2CI^2 \\
&\geq -\gamma n^3-\alpha CE_0n^2-\eta_2^2CE_0^2n-\eta_1^2CE_0^2
\end{aligned}$$

$$
\begin{aligned}
|X_t|^3 &= (n^2 + CE^2 + CI^2)\sqrt{n^2 + CE^2 + CI^2} \\
&\geq (n^2 + CE^2 + CI^2)n \\
&\geq n^3
\end{aligned}
$$

We want to prove that

$$
\begin{aligned}
&-\beta_1 n^3 - \beta_0 \leq -\gamma n^3 - \alpha CE_0 n^2 - \eta_2^2 CE_0^2 n - \eta_1^2 CE_0^2 \\
&\underbrace{-(\beta_1 - \gamma)n^3 + \alpha CE_0 n^2 + \eta_2^2 CE_0^2 n + \eta_1^2 CE_0^2}_{PolonomialFunction} - \beta_0 \leq 0 \qquad (3.24)
\end{aligned}
$$

So if we choose $\beta_1 > \gamma$, we can find the maximum of the polonomial function for $n \geq 0$ and the value of β_0 can be determined such the inequality (3.24) holds thus the convergent conditions 2 and 3 are fullfilled. □

We prove the one sided Lipschitz condition (3.21)

$$
\langle X - Y, f(X) - f(Y) \rangle \leq L|X - Y|^2
$$

Let

$$
X = \begin{pmatrix} n_1 \\ CI_1 \\ CE_1 \end{pmatrix} Y = \begin{pmatrix} n_2 \\ CI_2 \\ CE_2 \end{pmatrix}
$$

$$
f(X) = \begin{pmatrix} n_1(\beta - \gamma n_1 - \alpha CI_1) \\ \lambda_1^2 CE_1 - \eta_1^2 CI_1 \\ (\lambda_2^2 CI_1 - \eta_2^2 CE_1)n_1 \end{pmatrix}
$$

$$f(Y) = \begin{pmatrix} n_2(\beta - \gamma n_2 - \alpha CI_2) \\ \lambda_1^2 CE_2 - \eta_1^2 CI_2 \\ (\lambda_2^2 CI_2 - \eta_2^2 CE_2) n_2 \end{pmatrix}$$

$$f(X) - f(Y) = \begin{pmatrix} (n_1 - n_2)\beta - \gamma(n_1^2 - n_2^2) - \alpha(n_1 CI_1 - n_2 CI_2) \\ \lambda_1^2(CE_1 - CE_2) - \eta_1^2(CI_1 - CI_2) \\ \lambda_2^2(CI_1 n_1 - CI_2 n_2) - \eta_2^2(CE_1 n_1 - CE_2 n_2) \end{pmatrix}$$

According to Condition (3.21)

$$\begin{aligned} \langle X - Y, f(X) - f(Y) \rangle &= (n_1 - n_2)^2 \beta - \gamma(n_1 + n_2)(n_1 - n_2)^2 && (3.25) \\ &\quad - \alpha(n_1 CI_1 - n_2 CI_2)(n_1 - n_2) && (3.26) \\ &\quad + \lambda_1^2(CE_1 - CE_2)(CI_1 - CI_2) && (3.27) \\ &\quad - \eta_1^2(CI_1 - CI_2)^2 && (3.28) \\ &\quad + \lambda_2^2(CI_1 n_1 - CI_2 n_2)(CE_1 - CE_2) && (3.29) \\ &\quad - \eta_2^2(CE_1 n_1 - CE_2 n_2)(CE_1 - CE_2) && (3.30) \\ |X - Y|^2 &= (n_1 - n_2)^2 + (CI_1 - CI_2)^2 + (CE_1 - CE_2)^2 \end{aligned}$$

Consider (3.25) first, so $L \geq \beta$.

Then we deal with 3.27.

$$\lambda_1^2(CE_1 - CE_2)(CI_1 - CI_2) \leq \lambda_1^2(|CE_1 - CE_2|^2 + |CI_1 - CI_2|^2)$$

so we choose $L \geq \lambda_1^2$.

We only need to consider (3.26)(3.29)(3.30). If we assume Lemma 3.5.1 (iii) hold, $n_1 > n_2$ so $CI_1 < CI_2, CE_1 < CE_2, n_1 CI_1 > n_2 CI_2$.

Since

$$-\alpha \underbrace{(n_1 CI_1 - n_2 CI_2)}_{>0} \underbrace{(n_1 - n_2)}_{>0} < 0$$

$$\lambda_2^2 \underbrace{(CI_1 n_1 - CI_2 n_2)}_{>0} \underbrace{(CE_1 - CE_2)}_{<0} < 0$$

only $-\eta_2^2(CE_1 n_1 - CE_2 n_2)(CE_1 - CE_2)$ can be positive.

Since $CE_1 < CE_2$, consider $(CE_1 n_1 - CE_2 n_2) > 0$, so $\frac{n_1}{n_2} > \frac{CE_2}{CE_1} > 1$.

$$\begin{aligned}(CE_1 n_1 - CE_2 n_2)^2 &= CE_1^2 n_1^2 + CE_2^2 n_2^2 - 2 n_1 n_2 CE_1 CE_2 \\ &= CE_1^2 (n_1 - n_2)^2 + (CE_2^2 - CE_1^2) n_2^2 + 2 n_1 n_2 (CE_1^2 - CE_1 CE_2)\end{aligned}$$

Since $1 < \frac{(CE_1 + CE_2)}{2CE_1} < \frac{CE_2}{CE_1} < \frac{n_1}{n_2}$, so $(CE_2^2 - CE_1^2) n_2^2 + 2 n_1 n_2 (CE_1^2 - CE_1 CE_2) < 0, (CE_1 n_1 - CE_2 n_2)^2 < CE_1^2 (n_1 - n_2)^2$.

$$\begin{aligned}&-\eta_2^2 (CE_1 n_1 - CE_2 n_2)(CE_1 - CE_2) \\ &\quad \le \eta_2^2 [(CE_1 n_1 - CE_2 n_2)^2 + (CE_1 - CE_2)^2] \\ &\quad \le \eta_2^2 [CE_1^2 (n_1 - n_2)^2 + (CE_1 - CE_2)^2] \\ &\quad \le \eta_2^2 [CE_0 (n_1 - n_2)^2 + (CE_1 - CE_2)^2]\end{aligned}$$

so we choose $L \ge \eta_2^2 \max(CE_0, 1)$. The function $f(x)$ is one-sided Lispchitz continuous.

Chapter 4

Numerical Result

In this chapter, we compare results obtained with various numerical schemes. Monte Carlo simulations are performed to determine the threshold of which cells go from surviving to extinction depending on values of the initial concentrations of the chemical compounds.

4.1 Comparison of Numerical Schemes

Euler and implicit Euler

Figs 4.1 and 4.2 illustrate the results obtained with the Euler scheme and implicit Euler scheme. Fig 4.1 shows that both schemes converge with a step size of 0.001. Fig 4.2 indicates that with the same level of noise, but a larger step size, the Euler scheme results in negative values, which is impossible since the y-axis stands for the cell number. For this reason, the Euler should not be used, unless we choose a very small time step.

Tamed Euler and implicit Euler

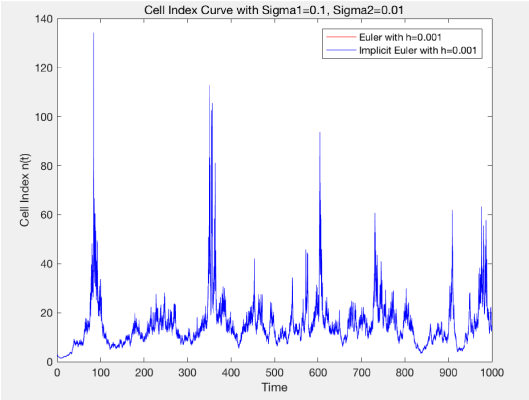


Figure 4.1: Case study $\sigma_1 = 0.1, \sigma_2 = 0.01, h = 0.001$

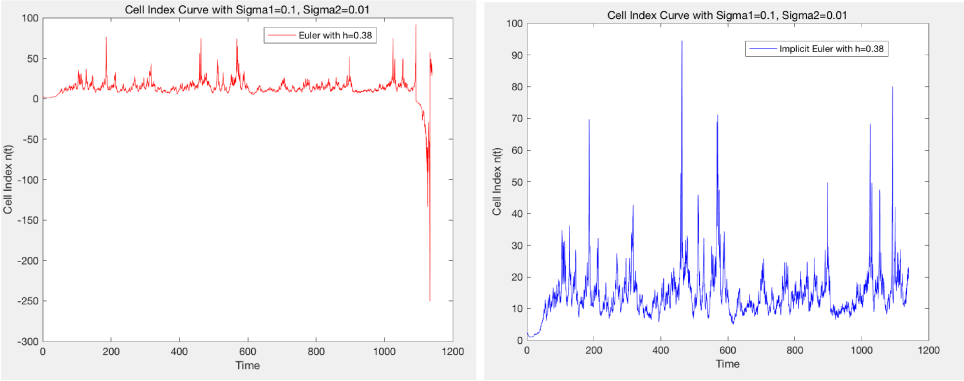


Figure 4.2: Case study $\sigma_1 = 0.1, \sigma_2 = 0.01, h = 0.38$

Next, the tamed Euler and the implicit Euler are compared. Figs 4.3 and 4.4 are the numerical results of the same case, but with step sizes of 0.01 and 0.8. Fig 4.3 illustrates that with a step size of 0.01, both tamed Euler and implicit Euler are convergent. In Fig 4.4, the tamed Euler shows larger fluctuations compared to the implicit Euler. It can be concluded that the Implicit Euler is more stable than tamed Euler.

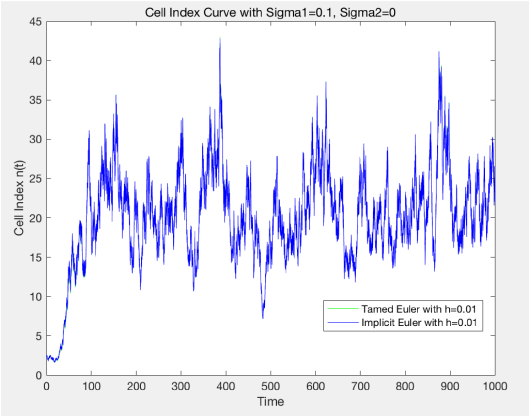


Figure 4.3: Case study $\sigma_1 = 0.1, \sigma_2 = 0, h = 0.01$

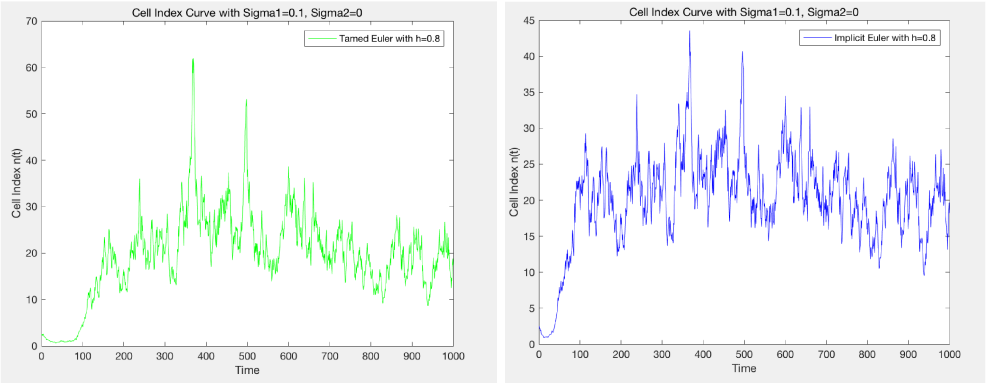


Figure 4.4: Case study $\sigma_1 = 0.1, \sigma_2 = 0, h = 0.8$

4.2 Monte Carlo Simulations

In this section, the impact of the initial concentrations of the chemical compounds on cell survival will be discussed. The deterministic thresholds for chemical compounds can be retrieved by using Runge-Kutta-Fehlberg Method (RKF45) when $n_0 = 2.5$. We denote the threshold value as the lowest concentration of the compounds that will make all cells extinct. It is the lowest blue circle in the vertical direction for a specific n_0 Fig 2.6.

In the stochastic case, since different random samples will come out with different

thresholds, Monte Carlo simulations were explored to evaluate the thresholds distributions. After 1000 simulations, a histogram of the threshold values is plotted.

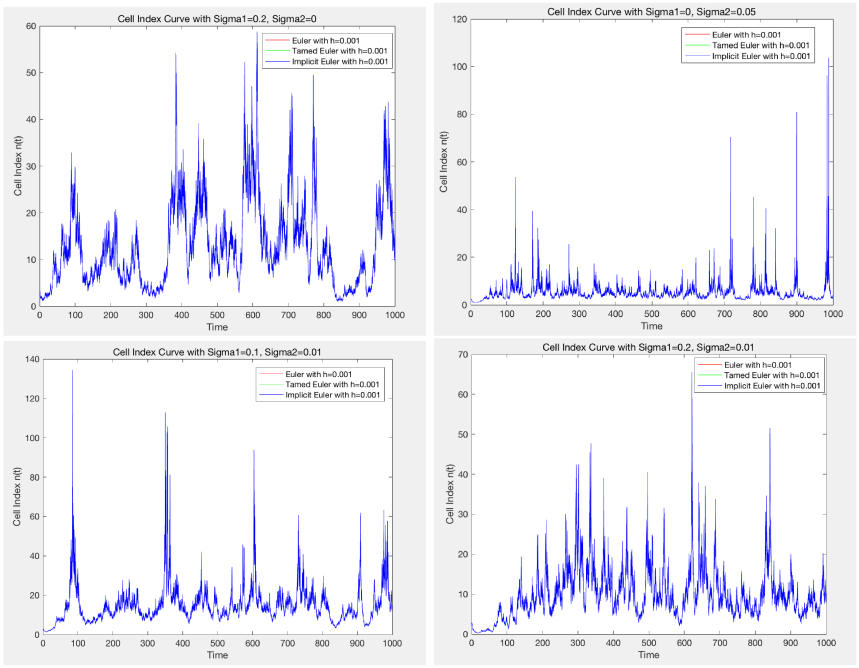


Figure 4.5: Case study with step size $h = 0.001$

Using an iterative algorithm we determine the threshold within a small range according to the accuracy required. The bisection algorithm is used to reduce the

range of the changing point of the initial concentration.

Result: Narrow the range of the threshold within a limit $r = 10^{-3}$.

Initials: range $[a_0, b_0]$.

while $|b_t - a_t| > r$ **do**

  Test n_t when $CE_0 = \frac{b_t + a_t}{2}$

  if $n_t \to 0$ **then**

    $a_{t+1} = a_t; b_{t+1} = \frac{b_t + a_t}{2}$

  else

    $a_{t+1} = \frac{b_t + a_t}{2}; b_{t+1} = b_t$

  end

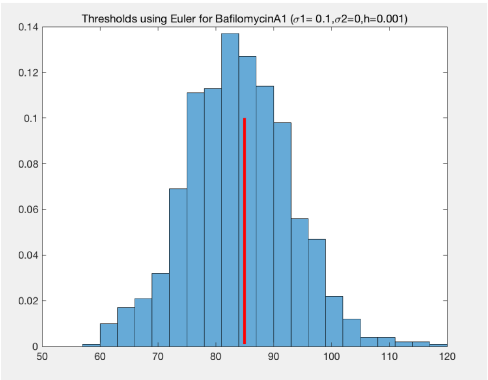
end

Algorithm 1: Bisection algorithm

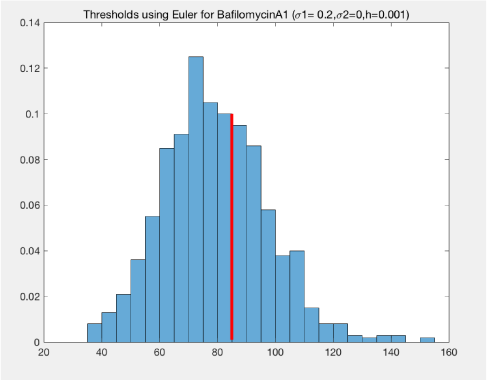
Case studies are presented in Fig 4.5 with step size $h = 0.001$. It shows from the graphs that Euler scheme, tamed Euler scheme and implicit Euler scheme converge with step size 0.001. Because it is faster, we choose Euler scheme with step size 0.001 for simulations and plot the histogram after 1000 simulations. The red line in each histogram represents the threshold for each chemical compound in the deterministic case, $C = 86$ for BafilomycinA1, $C = 129$ for CRT0044876 and $C = 160$ for Dimethylenenastrone.

Compound and Noise	Cluster	$\sigma_1 = 0.1$	$\sigma_1 = 0.2$	$\sigma_1 = 0.3$	$\sigma_1 = 0.4$	$\sigma_2 = 0.05$	$\sigma_2 = 0.1$	$\sigma_2 = 0.2$	$\sigma_2 = 0.3$
BafilomycinA1	2	83.9354	79.3604	69.9481	32.1170	68.9428	63.9447	55.3176	48.8668
CRT0044876	1	122.1212	113.2431	86.0886	42.4520	114.8955	104.4070	80.5948	60.3685
Dimethylenenastrone	10	155.8048	121.1223	68.9385	55.1117	157.6048	152.5725	148.4866	143.1971

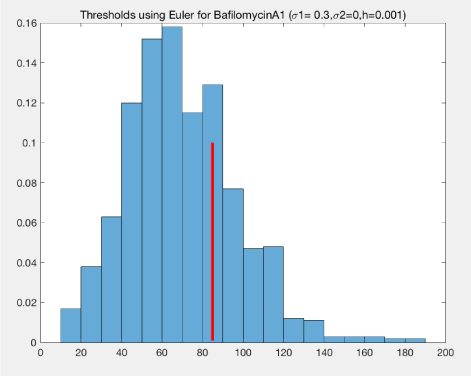
Table 4.1: Average of thresholds



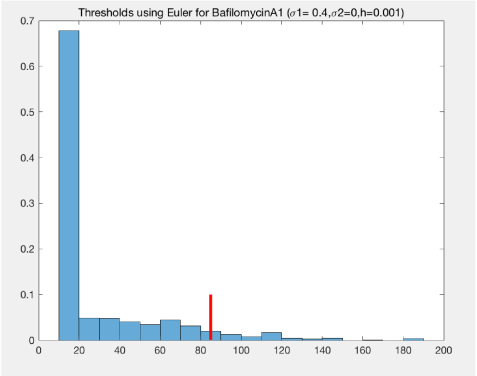
(a) $\sigma_1 = 0.1$



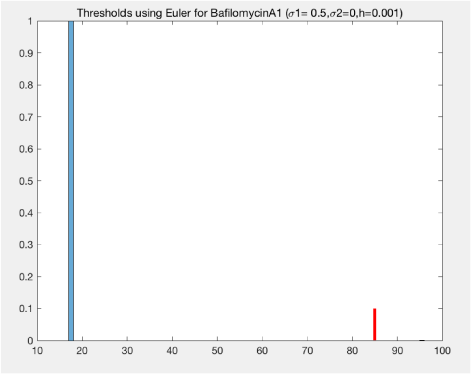
(b) $\sigma_1 = 0.2$



(c) $\sigma_1 = 0.3$

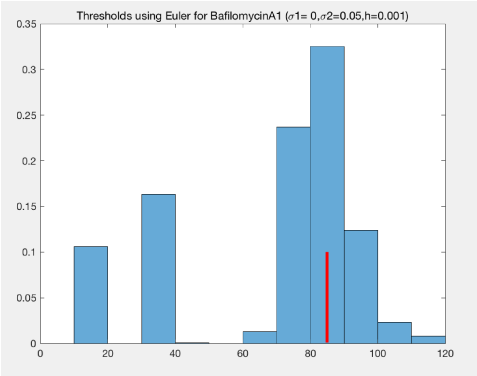


(d) $\sigma_1 = 0.4$

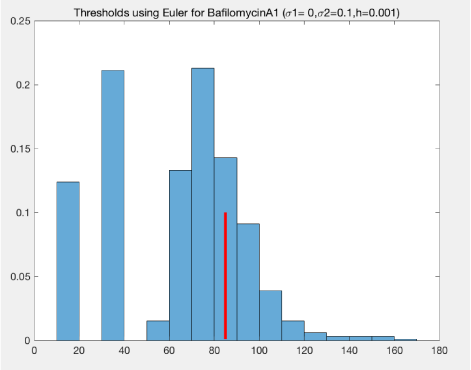


(e) $\sigma_1 = 0.5$

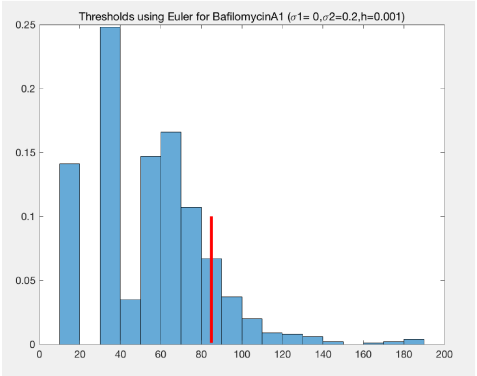
Figure 4.6: Simulations for BafilomycinA1 with $\sigma_1 \neq 0, \sigma_2 = 0$



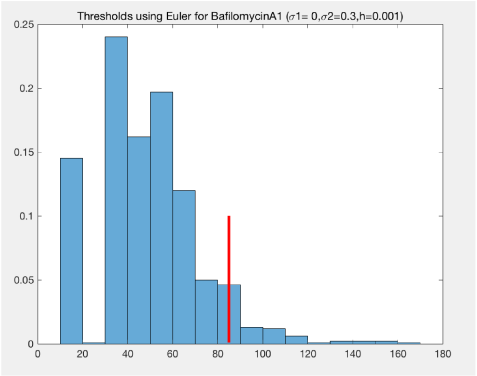
(a) $\sigma_2 = 0.05$



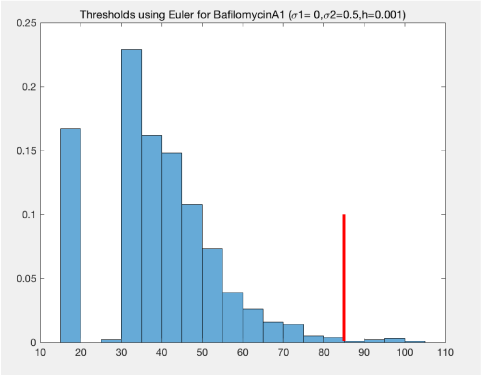
(b) $\sigma_2 = 0.1$



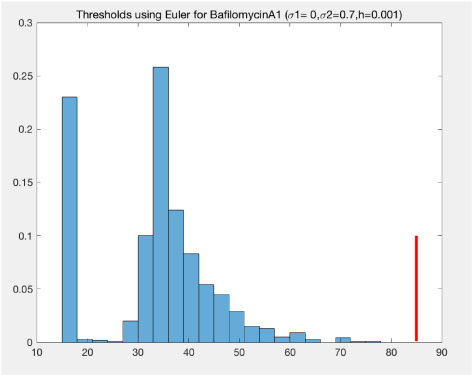
(c) $\sigma_2 = 0.2$



(d) $\sigma_2 = 0.3$

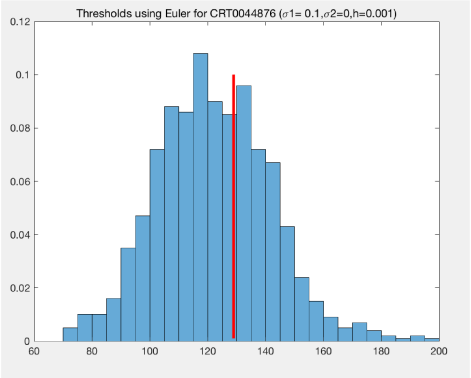


(e) $\sigma_2 = 0.5$

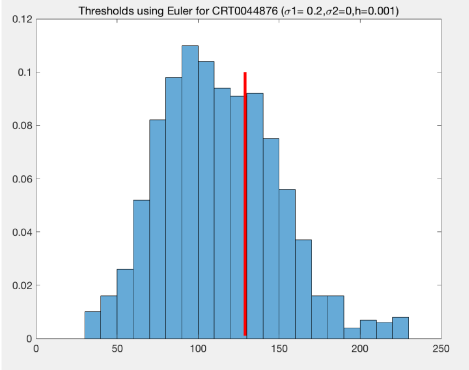


(f) $\sigma_2 = 0.7$

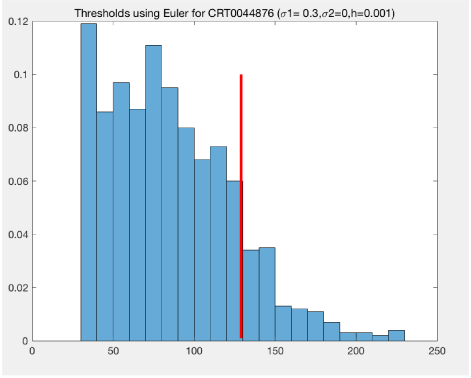
Figure 4.7: Simulations for BafilomycinA1 with $\sigma_1 = 0, \sigma_2 \neq 0$



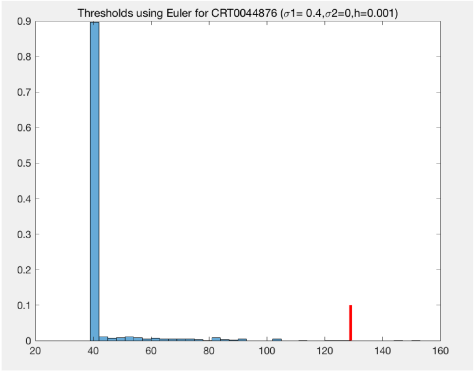
(a) $\sigma_1 = 0.1$



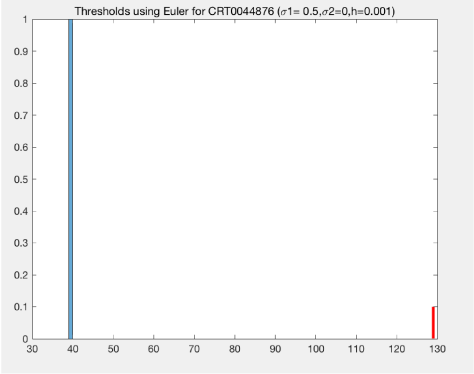
(b) $\sigma_1 = 0.2$



(c) $\sigma_1 = 0.3$

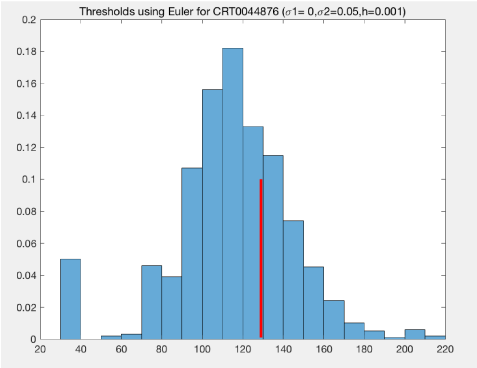


(d) $\sigma_1 = 0.4$

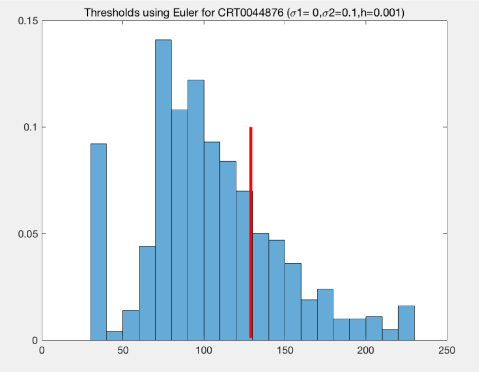


(e) $\sigma_1 = 0.5$

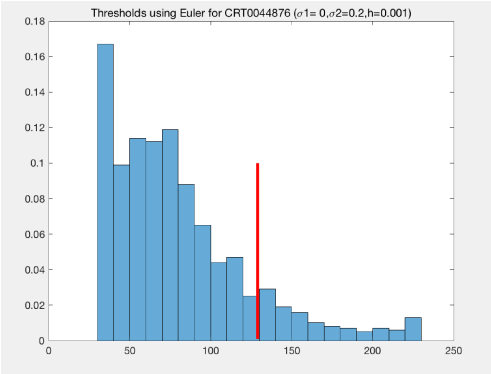
Figure 4.8: Simulations for CRT0044876 with $\sigma_1 \neq 0, \sigma_2 = 0$



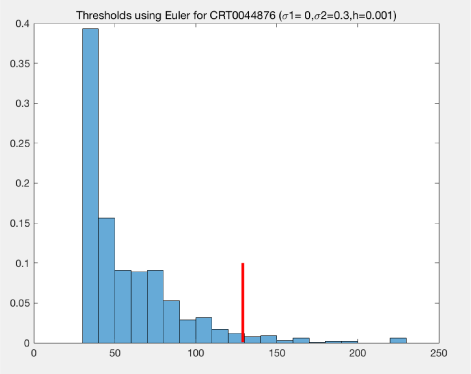
(a) $\sigma_2 = 0.05$



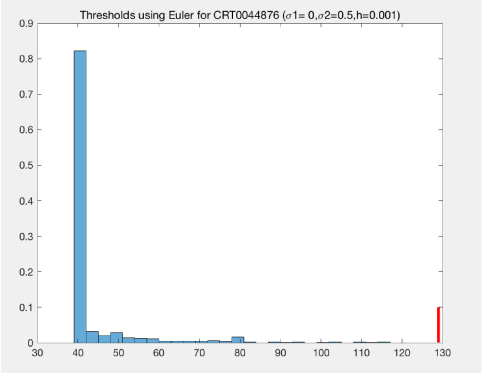
(b) $\sigma_2 = 0.1$



(c) $\sigma_2 = 0.2$

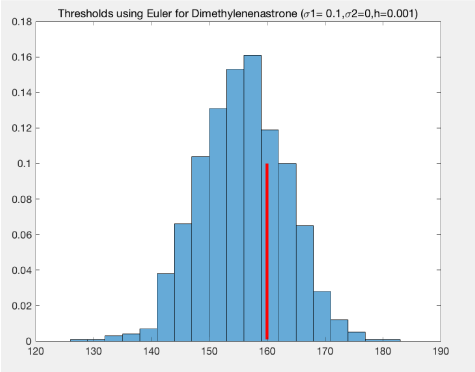


(d) $\sigma_2 = 0.3$

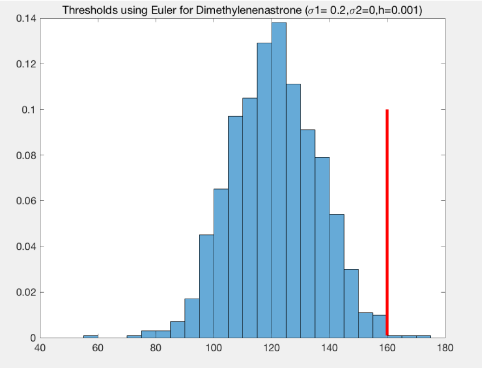


(e) $\sigma_2 = 0.5$

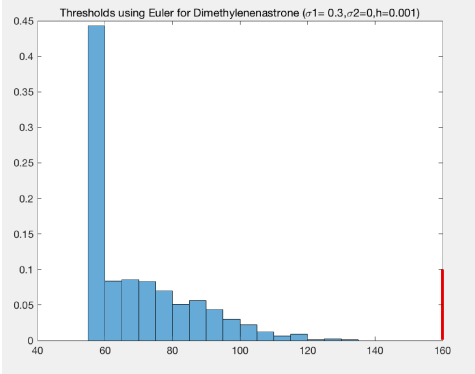
Figure 4.9: Simulations for CRT0044876 with $\sigma_1 = 0, \sigma_2 \neq 0$



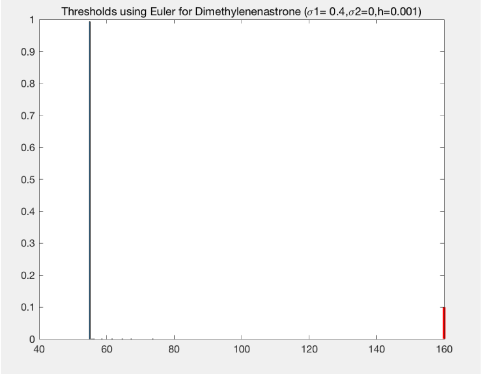
(a) $\sigma_1 = 0.1$



(b) $\sigma_1 = 0.2$

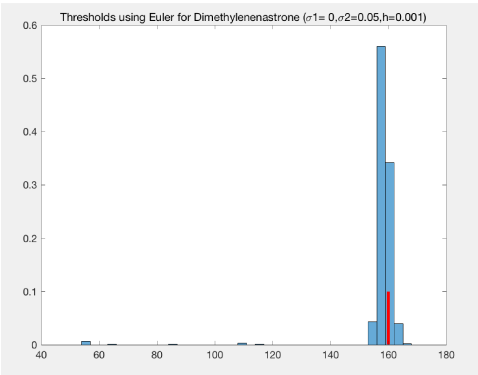


(c) $\sigma_1 = 0.3$

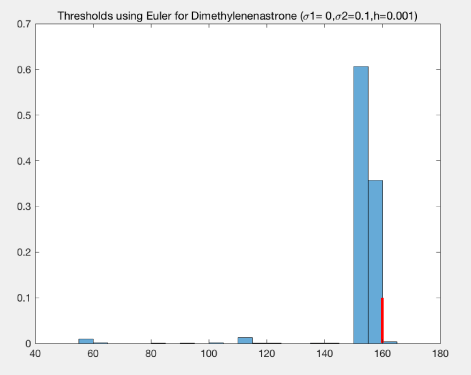


(d) $\sigma_1 = 0.4$

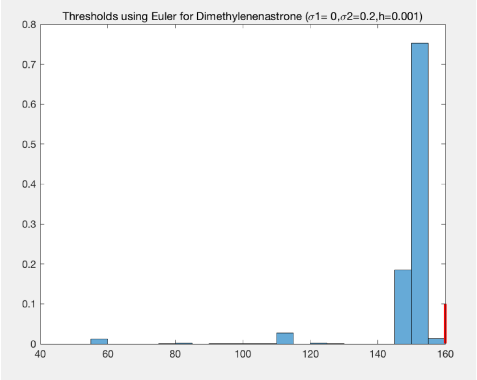
Figure 4.10: Simulations for Dimethylenenastrone with $\sigma_1 \neq 0, \sigma_2 = 0$



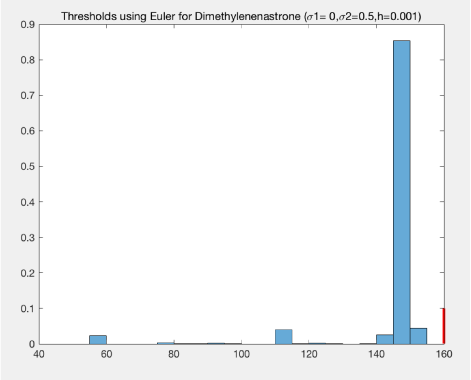
(a) $\sigma_2 = 0.05$



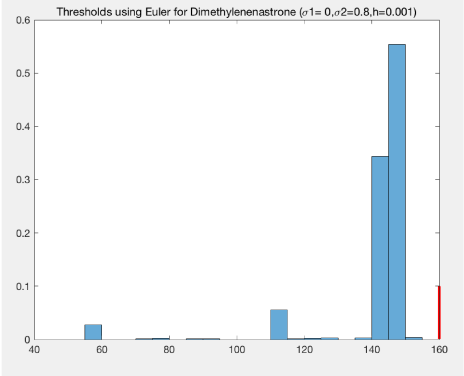
(b) $\sigma_2 = 0.1$



(c) $\sigma_2 = 0.2$



(d) $\sigma_2 = 0.5$



(e) $\sigma_2 = 0.8$

Figure 4.11: Simulations for Dimethylenenastrone with $\sigma_1 = 0, \sigma_2 \neq 0$

Chapter 5

Conclusion and Discussion

We study the effects of parameters uncertainty for toxicity assessment and consider a mathematical model which represents our system with stochastic differential equations. We solve this system numerically and run Monte Carlo simulations to find the distributions of the threshold values of initial concentration of the chemical compounds are to which the cells become extinct. Since the drift coefficient is not a globally Lipschitz function, we explore the convergence of several finite difference schemes for stochastic differential equations: the explicit and implicit Euler schemes and the tamed Euler scheme.

We choose three chemical compounds, CRT0044876 from cluster 1, BafilomycinA1 from cluster 2 and Dimethylenenastrone from cluster 10, to perform Monte Carlo simulations and show the result in Table 4.1. Each chemical compound is simulated with $\sigma_1 \neq 0$ and $\sigma_2 \neq 0$. From the histograms of the thresholds, we conclude that all cells tend to go to extinction if noises σ_1 or σ_2 is large enough. As a result, the threshold values for the initial concentration $CE(0)$ shift to the left when noise intensities increase. In other words, when planning the in vitro experiments, a con-

servative estimation for the threshold should be used.

In the future, we would like to focus on the following aspects of our project:

- In our numerical result, we know that the Euler scheme is convergent with small step sizes. We want to find a condition for the convergence of Euler scheme for our model.

- All schemes we have explored till now are for the linear diffusion term, which is a limitation of our model. We need to consider a better method which can be used with non-linear diffusion term.

- The effects of the non-linear term is complex and it requires further study. In the future, we will explore the impact that the non-linear diffusion terms have on the stochastic model.

Bibliography

[1] Cristina Anton, Jian Deng and Yau Shu Wong and Yile Zhang. *Modeling And Simulation For Toxicity Assessment*. Mathematical Bioscience And Engineering, Volume 14, Number 3, June 2017

[2] J. Xing, L. Zhu, S. Gabos and L. Xie. *Microelectronic cell sensor assay for detection of cytotoxicity and prediction of acute toxicity*. Toxicology in Vitro, 20 (2006), 995-1004

[3] Yile Zhang, Yau Shu Wong, Jian Deng, Cristina Anton, Stephan Gabos, Weiping Zhang, Dorothy Yu Huang and Can Jin *Machine learning algorithms for mode-of-action classification in toxicity assessment*. BioData Mining (2016) 9:19 DOI 10.1186/s13040-016-0098-0

[4] T. Pan, B. Huang, W. Zhang, S. Gabos, D. Huang and V. Devendran. *Cytotoxicity assessment based on the AUC50 using multi-concentration time-dependent cellular response curves*. Anal. Chim. Acta, 764 (2013), 44-52.

[5] S. Julier, J. Uhlmann and H. Durrant-White. *A new approach for filtering nonlinear systems*. American Control Conference, 21-23 June 1995

[6] S. Julier, J. Uhlmann and H. Durrant-White, *A new method for the nonlinear transformation of means and covariances in filters and estimators*. IEEE Trans. Aut. Control, 45 (2000), 477-482.

[7] T. Pan, S. Khare, F. Ackah, B. Huang, W. Zhang, S. Gabos, C. Jin and M. Stamp. *In vitro cytotoxicity assessment based on KC50 with real-time cell analyzer (RTCA) assay*. Comp. Biol. Chem., 47 (2013), 113-120.

[8] C. F. Jeff Wu *On the Convergence Properties of the EM Algorithm*. The Annals of Statistics, 11 (1983), 95-103.

[9] Z. Xi, S. Khare, A. Cheung, B. Huang, T. Pan, W. Zhang, F. Ibrahim, C. Jin and S. Gabos *Mode of action classification of chemicals using multi-concentration time-dependent cellular response profiles*. Comp. Biol. Chem., 49 (2014), 23-35.

[10] T. Pan, B. Huang, W. Zhang, S. Gabos, D. Huang and V. Devendran *Cytotoxicity assessment based on the AUC50 using multi-concentration time-dependent cellular response curves*. Anal. Chim. Acta, 764 (2013), 44-52.

[11] *Expectation–maximization algorithm*. In Wikipedia, The Free Encyclopedia. Retrieved 20:01, January 15, 2018, from https://en.wikipedia.org/w/index.php?title=Expectation

[12] Z. Ghahramani and S. Roweis, *Learning nonlinear dynamical systems using an EM algorithm*. in Advances in Neural Information Processing Systems (eds. M. Kearns, S. Solla and C. D.A.), MIT Press, 1999, 599-605.

[13] X. Meng and D. Van Dyk, *The EM algorithm - an old folk-song to a fast new tune*. J.R. Statist. Soc.B, 59 (1997), 511-567.

[14] R. Neal and G. Hinton, *A view of the EM algorithm that justifies incremental, sparse, an other variants*. in Learning in Graphical Models (ed. M. Jordan), 89 (1998), 355-368.

[15] M. Liu and K. Wang, *Survival analysis of stochastic single-species population models in polluted environments*. Ecological Modeling, 220 (2009), 1347-1357.

[16] J. He and K. Wang, *The survival analysis for a population in a polluted environment*. Nonlinear Analysis: Real World Applications 10 (2009), 1555-1571.

[17] F. Ibrahim, B. Huang, J. Xing and S. Gabos, *Early determination of toxicant concentration in water supply using MHE*. Water Research, 44 (2010), 3252-3260.

[18] A. Kiparissides, S. Kucherenko, A. Mantalaris and E. N. Pistikopoulos, *Global sensitivity analysis challenges in biological systems modeling*. Industrial and Engineering Chemistry Research, 48 (2009), 7168-7180.

[19] I. M. Sobol, *Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates*. Mathematics and computers in simulation, 55 (2001), 271-280.

[20] P. Kloeden and E. Platen. *Numerical solutions of stochastic differential equations*. Springer Verlag, Berlin, 1992.

[21] Sheldon M. Ross *Stochstic Processes*. Second Edition John Wiley and Sons, INC.

[22] G. N. Milstein, M. V.Tretyakov. *Numerical integration of stochastic differential equations with nonglobally Lipschitz coefficients*. SIAM J. Numer. Anal., 43(3), pp.1139-1154, 2005.

[23] Desmond J. Higham, Xuerong Mao, And Andrew M. Stuart. *Strong Convergence Of Euler-Type Methods For Nonlinear Stochastic Differential Equations*. Siam J. Numer. Anal. Vol. 40, No. 3, pp. 1041–1063

[24] M.V. Tretyakov and Z. Zhang. *A fundamental mean-square convergence theorem for SDEs with locally Lipschitz coefficients and its applications*. SIAM J. Numer. Anal. 51 (2013) 3135–3162

[25] Martin Hutzenthaler, Arnulf Jentzen and Peter E. Kloeden. *Strong Convergence Of An Explicit Numerical Method For SDES With Nonglobally Lipschitz Continuous Coefficents*. The Annals of Applied Probability 2012, Vol. 22, No. 4, 1611-1641.

[26] Adam Andersson, Raphael Kruse *Mean-square convergence of the BDF2-Maruyama and backward Euler schemes for SDE satisfying a global monotonicity condition*. BIT Numerical Mathematics 57 (1), 21-53, 2017.

[27] Xuerong Mao, Lukasz Szpruch *Strong convergence and stability of implicit numerical methods for stochastic differential equations with non-globally Lipschitz continuous coefficients*.

[28] Xuerong Mao, Lukasz Szpruch. *Strong convergence rates for backward Euler-Maruyama method for non-linear dissipative-type stochastic differential equations with super-linear diffusion coefficients*. Stochastics: An International

Journal of Probability and Stochastic Processes Vol. 85, No. 1, February 2013, 144-171.

[29] Sotirios Sabanis. *A note on tamed Euler approximations*. Electron. Commun. Probab. 2013, Vol. 18, no. 47.

[30] Lukasz Szpruch, Xuerong Mao, Desmond J. Higham, Jiazhu Pan *Numerical simulation of a strongly nonlinear Ait-Sahalia-type interest rate model*. BIT Numer Math (2011) 51:405–425