First Exit Time Analysis for the Stochastic Reaction Diffusion Process in a One Dimensional Domain

Daodao Zhou

Thesis submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master of Science
in
Computer Science and Applications

Yang Cao, Chair
Adrian Sandu
Alexey Onufriev

September 24, 2025
Blacksburg, Virginia

Keywords: stochastic simulation,reaction-diffusion processes, first exit time.

First Exit Time Analysis for the Stochastic Reaction Diffusion Process in a One Dimensional Domain

Daodao Zhou

(ABSTRACT)

Recent advances in modeling stochastic reaction–diffusion (RD) process have focused on particle-based and master equation formulations. While these models offer strong theoretical foundation, a practical challenge remains: how does the choice of spatial discretization affect the accuracy and computational efficiency of simulation results, particularly when estimating first exit times.

This thesis addresses this research gap by investigating the accuracy of first exit time estimates in one-dimensional stochastic RD systems. We design and analyze three simplified models using stochastic simulations: (1) model 1: pure diffusion, (2) model 2: diffusion with monomolecular reaction, and (3) model 3: diffusion with bimolecular reaction. We conduct theoretical study for the mean first exit times and evaluate them based on these models.

Our results show that strictly following the Gillespie SSA is not necessary to obtain accurate results under certain conditions and a moderate discretizations size (e.g., $K \geq 5$)already provides highly accurate estimates for first exit times. Our results can guide efficient and accurate simulation of RD systems.

First Exit Time Analysis for the Stochastic Reaction Diffusion Process in a One Dimensional Domain

Daodao Zhou

(GENERAL AUDIENCE ABSTRACT)

Many natural processes can be described using reaction–diffusion systems. These models often rely on computer simulations to predict when and where particles move or react. One practical question is that: for the computer simulation program, how does the way we divide space or how fine or coarse the grid is affecting the accuracy of the results and the time it takes to compute them? In this thesis, we focus on three simplified models that represent different situations to help answer the question: one with pure diffusion, one where particles can react individually, and one where two types of particles can interact with each other. The results show that simulations do not always need to use the most detailed possible grid to be accurate. Even moderately detailed setups can produce reliable predictions while saving time and computational resources. These findings provide practical guidance for scientists who use computer models to study diffusion and reaction processes efficiently and accurately.

Dedication

Dedicated to Jian.

Acknowledgments

Finally, I graduated. I want to thank many people who helped me reach this milestone.

I would like to express my gratitude to my advisor, Dr. Cao Yang, for his continuous support and guidance which helped me move forward throughout my graduate studies. I am also grateful to my committee members, Dr. Adrian Sandu and Dr. Alexey Onufriev, for their valuable comments and suggestions on my thesis.

I want to thank my parents for giving me the opportunity to pursue higher education in the United States. I am also thankful to those whose support made me feel the warmth of family in a place far away from home, especially during the holiday seasons.

Finally, I would like to thank my partner, Jian Liu, for his love, care, and understanding in everyday life, and for always being there whenever I needed to talk. I also thank my cat, Maomao, for keeping me company and bringing me so much joy.

I could not have finished this journey without all of you. Now, it is time to move on.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Reaction–diffusion (RD) processes are mathematical models used to describe how particles move and interact in space. They are widely applied in biology, chemistry, and physics, where they help explain processes such as molecular transport inside cells, chemical reactions in solution, and heat or material transfer in engineering systems. At small scales where particle numbers are low, randomness plays a major role and fluctuations can no longer be ignored. Therefore stochastic models are needed to capture the true behavior [8].

One of the most widely used methods for stochastic modeling is Gillespie's stochastic simulation algorithm (SSA), which generates trajectories of chemical reactions at each time step [8, 10]. Over the years, many extensions have been developed to improve efficiency or handle spatial effects. For example, Tau-leaping allows multiple reactions to be grouped into a single step [9]. Other work has reformulated the SSA to make it more efficient [2],and introduced the slow-scale SSA to address multiscale systems [1]. Although these methods differ in detail, they all face the same challenge: balancing accuracy with computational cost [5]. Finer spatial discretization improves accuracy but makes simulations slower. Coarse discretization speeds up simulations but risks losing important dynamics.

The first passage time (FPT) problem has long been used to study diffusion processes [13], yet its connection to stochastic simulations under different discretization schemes has not been fully examined. FPT studies the time it takes for a diffusive molecule to reach or exit a specified region. There has been a long history of study on the FPT analysis in

different areas. Some important results have been presented in Gillespie's book [11], in which theoretical foundation for FPT analysis for general one-dimensional diffusion systems has been well established.

This thesis addresses the gap by combining FPT theory with stochastic simulations of RD processes in discrete one-dimensional domains. The goal is to use FPT analysis to quantitatively study the impact of different discretization sizes on simulation accuracy. Three representative models are studied: pure diffusion, diffusion with a monomolecular reaction, and diffusion with a bimolecular reaction. For each case, analytical results for mean first exit times are derived and compared with simulations across different discretization schemes.

The main contributions of this thesis are as follows:

1. Provide a systematic framework for studying first exit times with three simplified yet representative RD models: pure diffusion, diffusion with a monomolecular reaction, and diffusion with a bimolecular reaction.

2. The analytical mean first exit times were derived from FPT theory and validated through stochastic simulations.

3. Quantified the impact of spatial discretization on simulation accuracy and the results show that moderate discretization is often sufficient for reliable exit time estimates.

4. It was demonstrated that strict adherence to Gillespie's exponential waiting-time rule is not always required. Alternative timing schemes can still produce accurate results, though with differences in convergence speed.

5. For the bimolecular case, an explicit cost formula for estimating simulation effort was derived with respect to the discretization size, reaction rates, and diffusion coefficients.

6. Overall, the findings show that accurate and efficient simulations can be achieved without excessively fine discretization or rigid adherence to classical SSA assumptions, offering general guidelines for researchers working with stochastic RD systems.

The structure of this thesis is organized as follows. Chapter 2 introduces Foundations of Reaction–Diffusion Modeling including Overview of Reaction and Diffusion Processes, the first passage time problem and and stochastic simulation algorithm, and Self-Distance in Stochastic Simulation; Chapter 3 develops the first model, which focuses on pure diffusion in a discrete one-dimensional domain. Theoretical analysis based on FPT is presented and compared with stochastic simulations to examine the effect of discretization size, diffusion rate, and particle population on exit times. Chapter 4 extends the framework by incorporating a monomolecular reaction into the diffusion process. This second model combines theory and simulation to analyze how reaction and diffusion interact to influence the mean first exit time and event-time distribution. Chapter 5 further generalizes the approach to a bimolecular reaction system, where two diffusing particle types interact. This chapter derives an analytical formula for the mean reaction time, quantifies the effects of spatial discretization and reaction rate, and evaluates the computational cost of different simulation resolutions; Chapter 6 concludes with a summary of findings and discusses future work.

Chapter 2

Foundations of Reaction–Diffusion Modeling

2.1 Overview of Reaction and Diffusion Processes

Reaction and diffusion systems are often used to mathematically model particle movement or cell behavior in complex scientific fields such as chemistry, biology, and physics[12, 14]. Diffusion processes is the random movement of particles due to thermal energy and reaction processes describe how particles undergo state changes due to chemical interactions. At a large scale, diffusion is often formulated using Fick's laws [6], which describe the rate of mass transfer in a medium. At a microscopic scale, diffusion is described by random walks [4]. Reaction might involve a single particle undergoing a transformation (monomolecular), or two or more particles colliding and reacting (bimolecular or higher order). Reactions can occur independently or be influenced by diffusion, leading to a reaction-diffusion system. Such systems capture the interplay between particle random motion and chemical transformations, and it is particularly important at microscopic scales where noise and fluctuations become significant.

Simulation for RD systems involves a trade-off: Higher accuracy often comes at the cost of greater computational cost. Therefore in the context of this thesis, RD systems are not only physical models but also computational challenges. Although stochastic simulation provides

a accurate representation, it requires dividing the spatial domain into discrete compartments. Understanding how discretization choices influence simulation results is the main focus of the work presented here.

2.2 First Passage Time Problem

Problem Definition

The first passage time (FPT) problem studies the first time a solute molecule that diffuses within a domain first reaches or exits the boundary of the region. In a one-dimensional scenario, the domain will be an interval $[a,\ b]$, and the boundary will be the corresponding points at $x = a$, or b. In the last chapter of the book *Simple Brownian Diffusion* [11], Gillespie et. al. examine the FPT problem in a simple form. The problem considers a solute molecule undergoing Einstein diffusion on the x-axis with diffusion rate D, and assumes that at time $t = 0$ the molecule is located at a specified point x_0 inside a specific interval $[a, b]$. Gillespie defined the problem using the random variable as:

$$T^{(a,a)}(x_0; a, b) \equiv \text{the time at which the solute molecule, if released at time } 0 \text{ from } x_0 \in [a, b],\ \text{will first reach either } x = a \text{ or } x = b. \tag{2.1}$$

Here the superscript (a, a) indicates that the upper and lower boundaries of the interval $[a, b]$ are both absorbing boundaries. A different scenario would be reflective boundary, and the corresponding superscript will be noted as (r, r). In the following, if we don't specify, the boundary will be both absorbing and we will not show the superscripts.

We would like to find the probability density function (PDF) of the random variable

$$T(x_0; a, b)$$

, as well as its cumulative distribution function (CDF), defined as the cumulative integral of the PDF. With the equations of PDF/CDF, we can also compute the mean and variance of $T(x_0; a, b)$. This analysis will give us a comprehensive statistical understanding of $T(x_0; a, b)$ and can be very useful for theoretical study of other systems.

Analytical Solution through Fokker–Planck equation

To analyze the first-passage time $T(x_0; a, b)$ in the context of Einstein's diffusion model, a classic method [11] focuses on solving the corresponding Fokker–Planck equation for $t_0 = 0$, given by:

$$\frac{\partial P_X(x, t|x_0, 0)}{\partial t} = D\frac{\partial^2 P_X(x, t|x_0, 0)}{\partial x^2}, \tag{2.2}$$

where $P_X(x, t|x_0, 0)$ represents the probability density function for finding the particle in position x_0 at time t, and D is the diffusion coefficient. With the initial condition

$$P_X(x, 0|x_0, 0) = \delta(x - x_0) \quad (x_0 \in [a, b]), \tag{2.3}$$

and with the boundary conditions for the respective definitions

$$P_X^{(a,a)}(a, t|x_0, 0) = 0 \quad (\forall t \geq 0), \tag{2.4}$$

Solving the Fokker–Planck equation allows us to determine the probability that the particle remains in the interval $[a, b]$ and, consequently, to compute every existing moment of the first-passage time $T(x_0; a, b)$.

While the direct approach using the forward Fokker–Planck equation requires solving for the full probability distribution function

$$P(x, t \mid x_0, 0), \tag{2.5}$$

an alternative approach, using the backward Fokker–Planck equation, simplifies the computation of first-passage time statistics to one step task. The backward Fokker–Planck equation is given by:

$$\frac{\partial}{\partial t} P(x, t \mid x_0, 0) \;=\; D \frac{\partial^2}{\partial x_0^2} P(x, t \mid x_0, 0). \tag{2.6}$$

Here, differentiation is performed with respect to the *initial* position x_0, rather than the current position x. This approach directly provides the probability function

$$G(x_0, t; a, b) \;=\; \int_a^b P(x, t \mid x_0, 0)\, dx, \tag{2.7}$$

which represents the probability that the particle, starting from $x(0) = x_0 \in (a,\ b)$, has not yet reached the boundary ($x = a$ or b) by time t. In other words,

$$G(x_0, t; a, b) = P\left(T(x_0; a, b) > t\right). \tag{2.8}$$

Using this function, the cumulative distribution function of the first-passage time can be determined as:

$$1 - G(x_0, t; a, b) \;=\; P\big(T(x_0; a, b) \leq t\big). \tag{2.9}$$

Thus, the probability density function (PDF) of the FPT is given by

$$-\frac{\partial}{\partial t}G(x_0,t;a,b)$$

This formulation allows for a single-step computation of first-passage time probabilities without explicitly solving for $P(x,t \mid x_0,0)$.

Mean First-Passage Time

(2.6) and (2.7) lead to an equation for G as following,

$$\frac{\partial G(x_0,t;a,b)}{\partial t} = D\frac{\partial^2 G(x_0,t;a,b)}{\partial x_0^2}, \tag{2.10}$$

With (2.10), the function $G(x_0,t;a,b)$ can be used to compute statistical moments of the first-passage time distribution, including the mean first-passage time, formulated as

$$\langle T(x_0;a,b)\rangle = \int_0^\infty t\Big(-\frac{\partial}{\partial t}G(x_0,t;a,b)\Big)\,dt, \tag{2.11}$$

which is also

$$\langle T(x_0;a,b)\rangle = \int_0^\infty G(x_0,t;a,b)\,dt. \tag{2.12}$$

Consider

$$\int_0^\infty \frac{\partial G(x_0,t;a,b)}{\partial t}dt = G(x_0,\infty;a,b) - G(x_0,;a,b) = -1, \tag{2.13}$$

Coupling (2.13) with the backward equation (2.10) and (2.12), we have

$$D\frac{\partial^2}{\partial x_0^2}\langle T(x_0;a,b)\rangle = -1 \tag{2.14}$$

Solving (2.14) with boundary conditions that $\langle T(a;a,b)\rangle = \langle T(b;a,b)\rangle = 0$, a straightforward formula for the mean first-passage time can be derived as:

$$\langle T(x_0;a,b)\rangle \;=\; \frac{(x_0-a)\,(b-x_0)}{2\,D}. \tag{2.15}$$

This formula gives the average time required for a particle, initially located at x_0, to reach either absorbing boundary at a or b. The result is quadratic in x_0, reflecting the spatial dependence of diffusion processes. In our theoretical analysis, we will use (2.15) to derive formulas for mean reaction time or exit time under different scenario.

2.3 Stochastic Simulation Algorithm

Traditional modeling of dynamical systems relies on deterministic reaction rate equations, typically based on ordinary differential equations. While this approach is sufficient for large-scale systems with high concentrations, it fails to capture the noise and fluctuations in small scales especially at the molecular or particle level. Stochastic models not only address this limitation but also provide a more accurate and detailed explanation of the reaction-diffusion system. One of the most widely used stochastic simulation methods is Gillespie's Stochastic Simulation Algorithm (SSA) [7].

Gillespie's SSA offers an exact method for simulating the time evolution of well-stirred chemical system. The key concept is to formulate a *propensity function* so that $\alpha_i(t)$ product with dt defines the probability that any given chemical reaction occurs within an infinitesimally small time interval$[t, t+dt)$.

To implement Gillespie SSA, consider a system of q chemical reactions. For each reaction channel i, the propensity function $\alpha_i(t)$ determines the probability that the reaction occurs

at time t. Then the algorithm consists of the following four steps:

(1) Random Number Generation:

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

(2) Compute Propensities:

Compute the propensity function $\alpha_i(t)$ of each reaction. Then compute the total propensity:

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

(3) Determine Time of Next Reaction:

Using the first random number $r1$ to compute time when the next chemical reaction takes place as $t + \tau$ where τ is given by

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

.

(4) Select and Execute a Reaction:

Using the second random number $r2$ to compute which reaction occurs at time $t + \tau$. Find j such that

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

Then the j-th reaction takes place, and the system state is updated accordingly.

After completion of these steps, the numerical simulation advances with time τ, and the process repeats from step (1). This loop continues until a specified simulation time or ending condition is reached.

2.4 Self-Distance in Stochastic Simulation

Stochastic simulation typically involves generating independent sets of samples from the same underlying probability distribution. To quantitatively measure the error between these sample sets, the concept of distribution distance [3] is often used. However, as a Monte Carlo simulation, different random number sequences may result in different sample sets. Thus, we have to distinguish whether the distribution distance is due to numerical error or purely from inherent randomness in Monte Carlo simulations.

Self-distance plays a similar role as the unit round-off error in classical numerical analysis. When the distribution distance of the two samples is less than the self-distance of one sample, one cannot tell the difference is from numerical error or inherent randomness. Thus self-distance is the maximum distance that we can measure between different Monte Carlo simulations. In the following, Theorem 2.1 gives the method to compute the mean and variance of the histogram self-distance.

Theorem 2.1. *For sufficiently large N and M, the mean of the histogram self distance is bounded by*

$$\sqrt{\frac{2K}{\pi}\left(\frac{1}{N}+\frac{1}{M}\right)}, \tag{2.16}$$

and the variance of the self distance is bounded by

$$\frac{(\pi-2)K}{\pi}\left(\frac{1}{N}+\frac{1}{M}\right), \tag{2.17}$$

where N and M are the sample sizes of two independent Monte Carlo simulations, and K is the number of bins that are used to generate the histogram.

Typically, we set $K = 50$, and $N = M = 10,000$, in that case, the histogram self distance is around 0.080, if $N = M = 100,000$, the histogram self distance is around 0.025.

Chapter 3

Model 1: Particle Diffusion in a Discrete One-dimensional Space

3.1 Model Description

In this chapter we present a simple toy model named model 1, which models the random movement of particles, initially located at $x = 0$, diffusing in a one dimensional domain bounded by $[-L, L]$. We will present the theoretical study for this system and then apply Gillespie's SSA, as well as other similar simulation methods, to this model, then report simulation results and compare them with the theoretical study. For the purpose of simulation, this simplified model focuses on a single type of particle jumping between neighboring bins.

At the start of each simulation run, all particles are initialized at the center of the domain (i.e., at location $x = 0$). The computational domain [-L, L] is discretized into 2^*K compartments, where K represents the number of bins on one side of the domain (i.e., from 0 to L). As a result, the total number of bins spans the entire interval from $[-L, L]$ (Figure 3.1) and the length of each bins h is given by:

$$h = \frac{L}{K}. \tag{3.1}$$

During numerical simulation, each particle randomly jumps to its neighboring one bin either to the left or to the right with equal probabilities. In a discrete setting, the basic jump rate

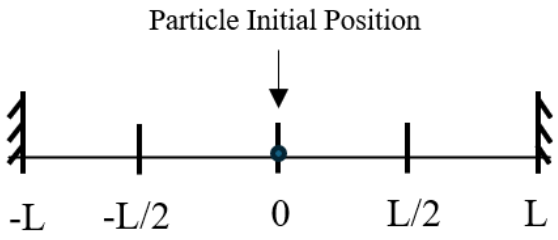

Figure 3.1: Schematic representation of the one-dimensional bounded domain used in the simulation,with particles initialized at the center and allowed to jump to adjacent compartments.

is derived from the diffusion rate D as: $\frac{D}{h^2}$.

Since each particle can move in two directions (left or right), the total jump rate used in the simulation must account for both possibilities. Therefore, the total jump rate is given by:

$$\text{Total jump rate} = d = \frac{2D}{h^2} \tag{3.2}$$

The total jump rate is then used to compute the total jumping propensity α_0 by:

$$\alpha_0 = N_p \cdot d, \tag{3.3}$$

where N_p is the number of particles and d is the total jump rate given in (3.2).

In the classical Gillespie SSA, the time τ at which the next event takes place follows an exponential distribution with the total propensity of all possible reactions in the system α_0. To investigate how different distributions of τ affect the dynamics of particle diffusion and the exit time behavior, we experimented with several different approaches including exponential distribution, uniform distribution, normal (Gaussian) distribution, and fixed mean values.

The simulation ends when a particle exits the domain. The time of first particle exit is

recorded for each simulation for analysis.

3.2 Theoretical Discussion for the First Exit Time in Model 1

The exit time analysis for this simple model has been presented in the classic book of Gillespie et. al. [11] as a theoretical approach in the analysis of the first passage time (see Section 2.2). Here we will follow their idea and focus on the error analysis of Gillespie's algorithm as well as the impact of different discretization sizes and different probability distribution for jumps.

To start, let us consider model 1 with only one particle starting from position 0 at time $t = 0$ in figure 3.1, Suppose that there are K bins on each side, and the particle will jump to its two neighbor bins with a combined jump rate

$$d = \frac{2D}{h^2}, \tag{3.4}$$

where $h = \frac{L}{K}$.

In each simulation step, Gillespie's algorithm will simulate the particle jump time by an exponential distribution $E(\lambda)$, where $\lambda = \frac{1}{d}$, and there is a 50% probability for the particle to jump to its left or right neighbor bin. If we follow the index for the location of the particle, its location can be represented by its index i, where $x_i = i * h$ is the location of the particle. Thus, initially the particle starts with index 0, and the simulation ends when $i = \pm K$.

As the particle jumps on the uniform lattice of $x_i, i = -(K-1), \ldots, K-1$, in each step, the jump process is regulated by two random variables: One is related to the jump time τ, which follows an unknown distribution $F(\tau)$. In the Gillespie algorithm, F will be exponential, but

we will show that this assumption is not necessary. The other random variable is related to the jump direction: an unbiased jump $\Omega(\cdot)$ will have the following probabilities:

$$P(\omega) = \begin{cases} \frac{1}{2}, & \omega = \pm 1, \\ 0, & else. \end{cases} \tag{3.5}$$

Thus, we have $E(\omega) = 0$ and $D(\omega) = 1$.

For a large K, there should be at least K jumps before the particle can exit the interval. Actually if we denote N as the number of steps for the particle to exit the interval, based on the mean exit time analysis, we have the following:

$$\langle N \rangle * \frac{h^2}{2D} = \frac{L^2}{2D}, \tag{3.6}$$

and thus,

$$\langle N \rangle = \frac{L^2}{h^2} = K^2. \tag{3.7}$$

For $K = 5$, $\langle N \rangle = 25$, and for $K = 10$, $\langle N \rangle = 100$. Meanwhile, following a series of jump events $j = 1, 2, \ldots, N$, the index of the particle will be at $l = \sum_j^N \omega_j$, and the time will be $t = \sum_j^N \tau_j$, where for each j, ω_j follows the random distribution $\Omega(\cdot)$ and τ_j follows the event time distribution F. Because ω_j and τ_j each follows a distribution with a fixed mean and variance, based on the law of large numbers, before $|l| = K$, l and t follows a normal random distribution. For l, its distribution can be approximated by a normal distribution: $\mathbb{N}(0, N)$. If we consider the location of the particle $x_l = l * h$, then x_l follows a distribution $\mathbb{N}(0, Nh^2)$. For the accumulated time $t = \sum_j^N \tau_j$, it also follows a normal distribution $\mathbb{N}\left(\frac{Nh^2}{2D}, \frac{Nh^2}{2D}\right)$.

Since x_l depends on two distributions $\Omega(\cdot)$ and $F(\cdot)$, which is more important? As $\Omega(\cdot)$ follows a binomial distribution and approximates a normal distribution, for any jump process $F(\cdot)$, as long as its mean remains the same and there is a bounded variance, the distribution of t

still follows the original diffusion distribution given in Section 2.2. Thus, in the numerical experiment below, we will compare the distribution of the exit time corresponding to different choices of $F(\cdot)$.

On the other hand, although the different choices of $F(\cdot)$ all lead to the same distribution, the convergence rate is different. Apparently, the exponential distribution and normal distribution result in faster convergence than the fixed value distribution. Errors in the distribution of the first exit time happen mainly at the region of small values. Because of that, the errors may become more important when the total population in model 1 becomes large. If there are initially N_p particles in model 1, let T_j represents the exit time of the j-th particle's, then the exit time for the first particle to exit the interval $[-L, L]$ is given by

$$T = \min_{j}^{N_p} T_j, \tag{3.8}$$

where each T_j follows the same exit time distribution. Because T takes the minimum of all T_j, the long tail of T_j become less important, and the accuracy depends mainly on the region of the small values. So potentially the errors in the small region will become more obvious when N_p increases. In the following experiment part, we will compare how the errors change with different N_p.

3.3 Simulation Results

3.3.1 Effect of Discretization Sizes on Exit Time

Below we present the results of the numerical experiment for model 1. The base parameters for this experiment are given in table 3.1.

Table 3.1: The parameter set for experiment 1

Parameter	Value
L	1
D	1
N_p	1

We consider three different sampling schemes for the reaction interval τ: exponential distribution, normal Distribution and a fix mean value.

In the first experiment, τ is sampled from an exponential distribution following the classical Gillespie formulation:

$$\tau = \frac{1}{\alpha_0} \ln \left[\frac{1}{r_1} \right]. \tag{3.9}$$

where r_1 is a random number uniformly distributed in (0,1) and α_0 is the total propensities. Figure 3.2 shows the plots for the exit time distribution for the first experiment.

In the second experiment, τ is sampled from a normal distribution:

$$\tau = \frac{r_1}{\alpha_0} \tag{3.10}$$

where r_1 is drawn from a normal distribution with mean 1.0 and standard deviation 0.5. Figure 3.3 shows the plots for the exit time distribution for the second experiment.

Finally, τ is fixed as the mean value:

$$\tau = \frac{1}{\alpha_0}. \tag{3.11}$$

where α_0 is the total propensities. Figure 3.4 shows the plots for the distribution of the exit time.

We can see from these numerical simulation plots that for all three sampling schemes, the exit time distributions begin to converge consistently after $K = 5$. By $K = 20$, the distributions

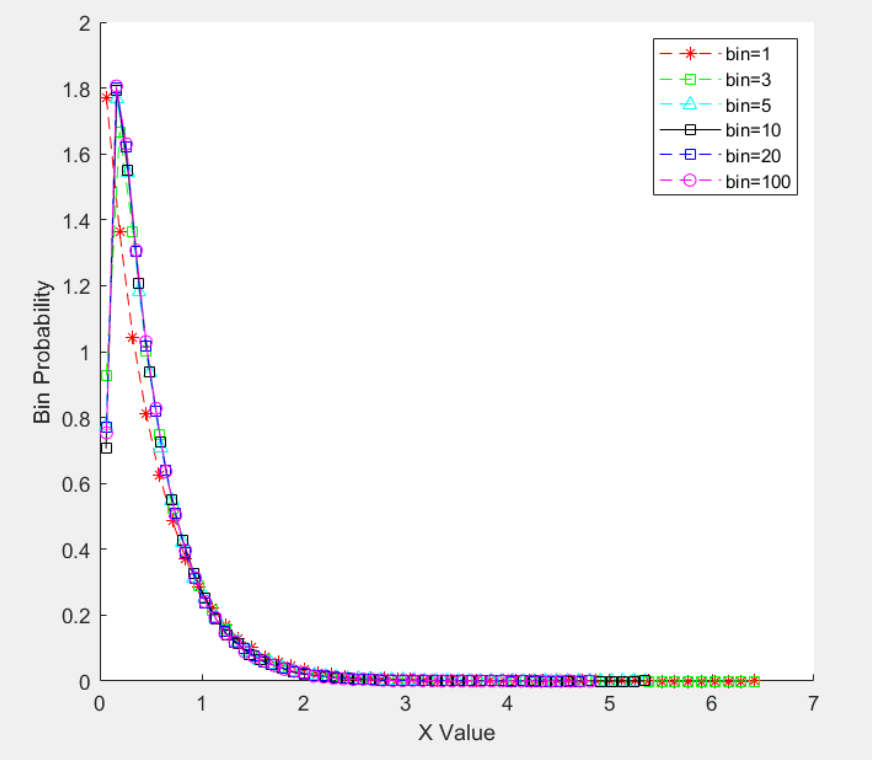

Figure 3.2: Exit time distribution for varying discretization sizes with $\tau \sim$ Exponential(α_0) The plot is generated from 100,000 runs of stochastic simulation.

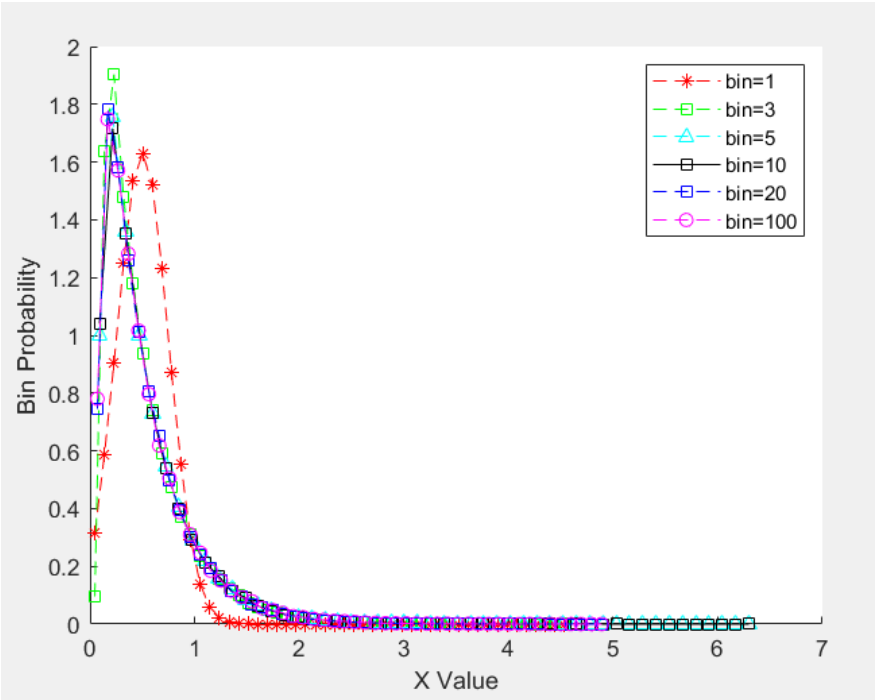

Figure 3.3: Exit time distribution for varying discretization sizes with $\tau = \frac{r_1}{\alpha_0}$, where $r_1 \sim \mathcal{N}(1, 0.5)$.The plot is generated from 100,000 runs of stochastic simulation.

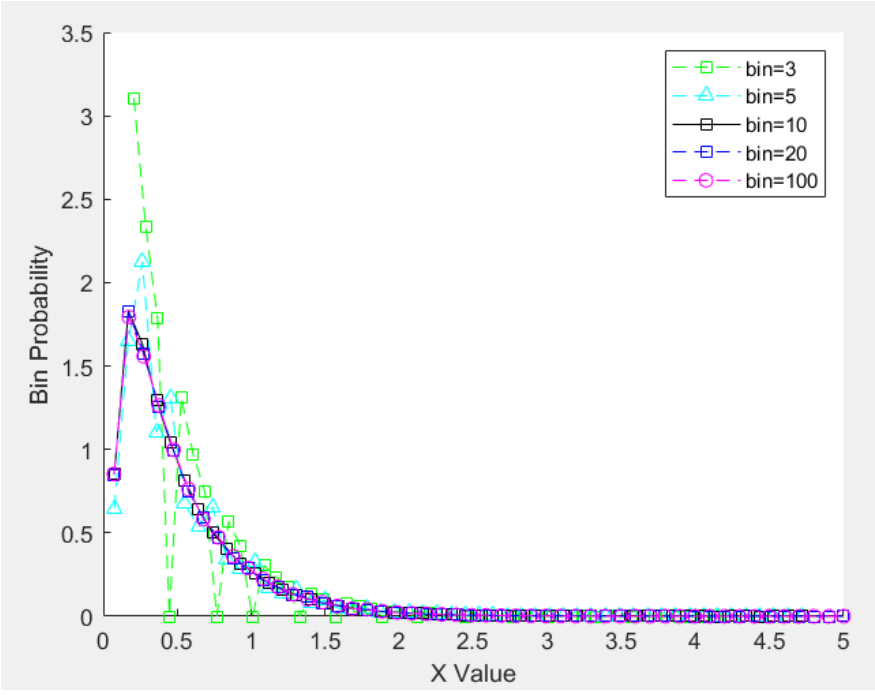

Figure 3.4: Exit time distribution for varying discretization sizes with $\tau = \frac{1}{\alpha_0}$ (fixed value).The plot is generated from 100,000 runs of stochastic simulation

across all methods are nearly indistinguishable as shown in figure 3.5. Therefore, we consider simulations with $K = 20$ to serve as the ground truth for evaluating accuracy and efficiency.

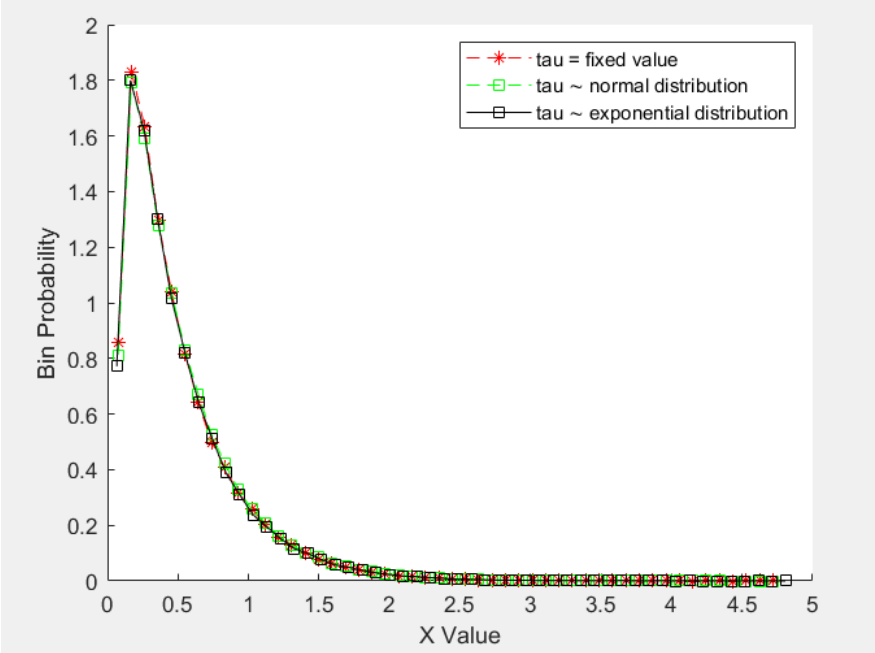

Figure 3.5: Exit time distribution for K =20 with τ in different distribution.The plot is generated from 100,000 runs of stochastic simulation

Table 3.2 summarize self distance value(Refer to 2.4) for different bin numbers compare with 20 and different τ distributions.

Table 3.2: Distribution Difference values for different bin numbers and τ distributions

Bin Number	**1**	**3**	**5**	**8**	**10**
Exponential	0.3082	0.0563	0.0267	**0.0129**	0.01502
Normal	0.572	0.0865	0.0291	**0.014**	0.0144
Fixed Value	1.8061	0.2913	0.2124	0.0599	**0.0146**

An important observation from this experiment is that strictly following the Gillespie SSA is not required to obtain accurate results in this simulation setting. Our results demonstrate that when the spatial discretization resolution is greater than 5 ($K > 5$), the exit time distributions closely match those produced by the classical SSA.

3.3.2 Effect of Diffusion Rate and Domain Size on Exit Time

Below we present the numerical results for the experiment 2. The parameters for this experiment are given in table 3.3.

Table 3.3: The parameter set for experiment 2

Parameter	Value
N_p	1
D	1
K	20

Figure 3.6 shows the relationship between the square of the domain size L^2 and the average exit time. Figure 3.7 shows the relationship between the inverse of particle population $1/N_p$ and the average exit time.

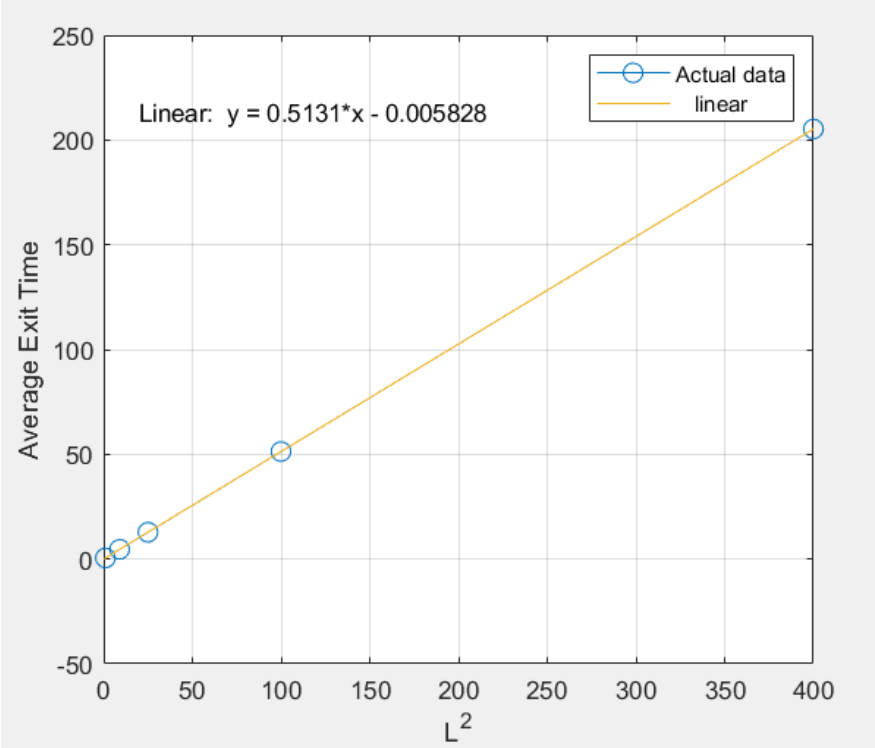

Figure 3.6: Relationship between the square of the domain size L^2 and the average exit time.

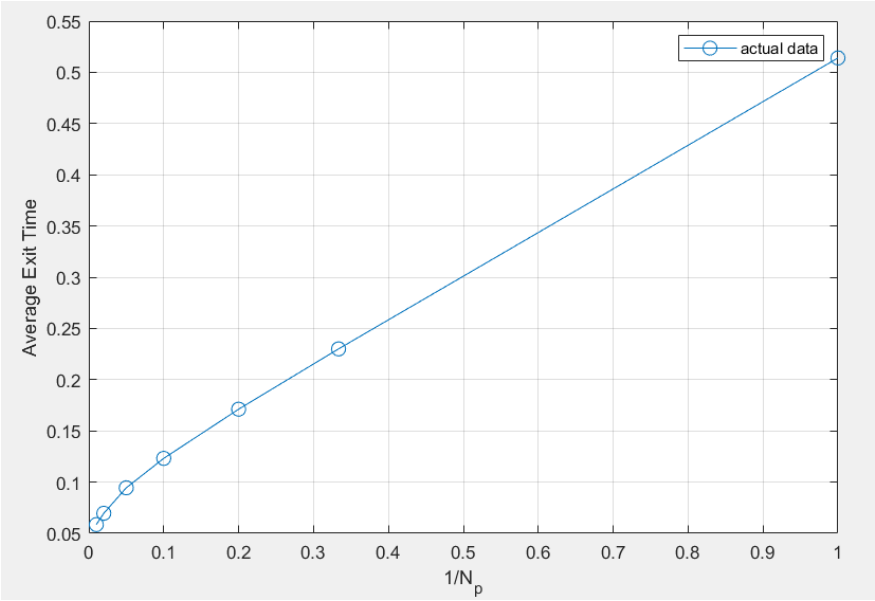

Figure 3.7: Relationship between the inverse of particle population $1/N_p$ and the average exit time.

Interestingly, although Figure 3.7 looks like the first exit time is linear with $1/N_p$, when $N_p > 100$, this linearity disappeared. However, the first exit time is still linear with $\frac{L^2}{D}$.

3.3.3 Effect of Particle Population on Exit Time

In the following, we present the results of the numerical experiment for experiment 3. The parameters for this experiment are given in table 3.4. We consider the sampling schemes

Table 3.4: The parameter set for experiment 3

Parameter	Value
L	1
D	1
N_p	2,3,10,50,100

for the reaction interval τ to be an exponential distribution. Figure 3.8 shows the plots for the exit time distribution for the initial population = 2. Figure 3.9 shows the plots for the exit time distribution for the initial population = 3. Figure 3.10 shows the plots for the

exit time distribution for the initial population = 10. Table 3.5 summarized the Distribution Difference values for different bin numbers and N_p values compared to $K = 20$. It is apparent in Table 3.5 that as the number of particles increases, the error also increases, requiring a higher value of K for the results to converge.

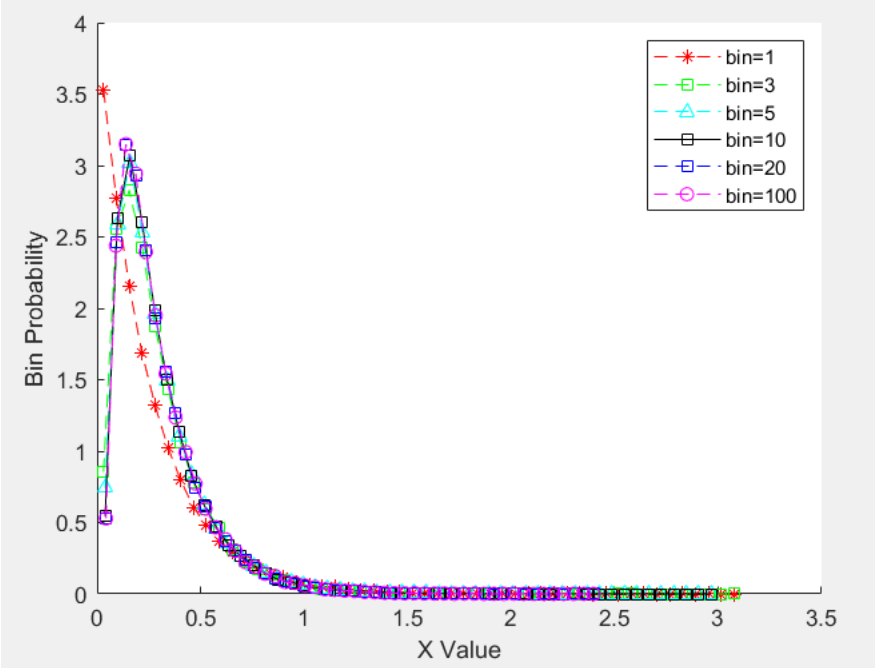

Figure 3.8: Exit time distribution for varying population $N_p = 2$.The plot is generated from 100,000 runs of stochastic simulation

Table 3.5: Distribution Difference values for different bin numbers and N_p values

	K =1	**3**	**5**	**8**	**10**	**12**	**16**	**17**
N_p **= 2**	0.4681	0.0857	0.0389	**0.0202**				
3	0.6196	0.1233	0.0493	**0.0233**				
10	1.2327	0.3063	0.1261	0.0499	0.0295	**0.0243**		
50	1.8983	0.8062	0.3663	0.2184	0.0839	0.0505	**0.0233**	
100	1.9794	1.0763	0.5257	0.3502	0.1315	0.0788	0.0315	**0.024**

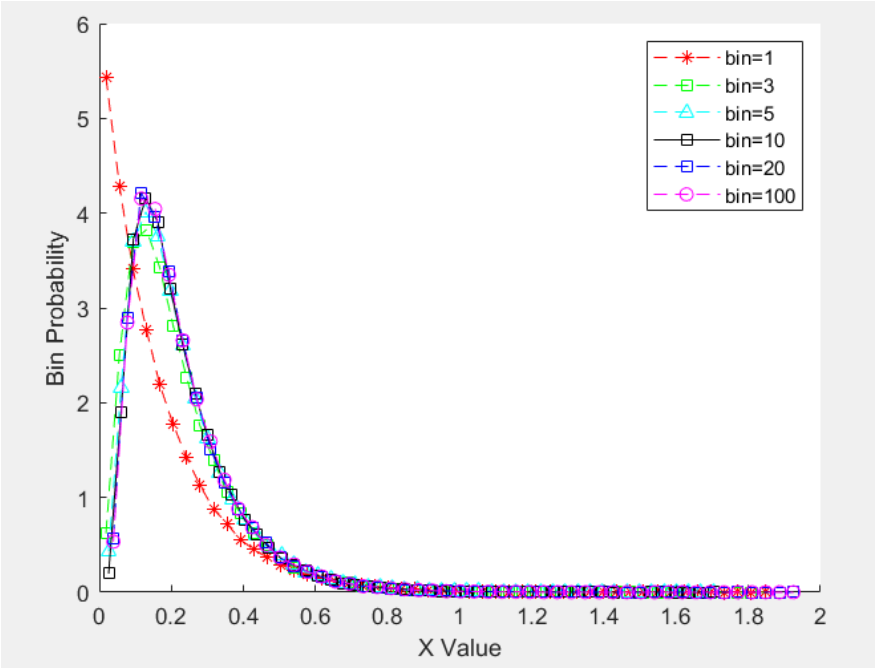

Figure 3.9: Exit time distribution for varying population $N_p = 3$. The plot is generated from 100,000 runs of stochastic simulation

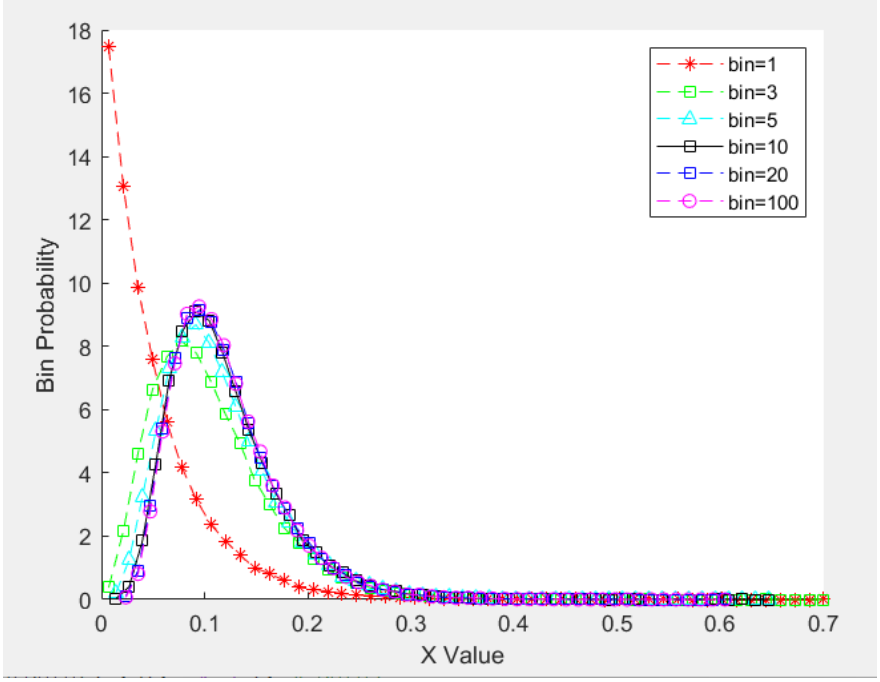

Figure 3.10: Exit time distribution for varying population $N_p = 10$. The plot is generated from 100,000 runs of stochastic simulation

3.4 Conclusion for Model 1

In this chapter we perform numerical experiments to study the first exit time of one or multiple particles in a one-dimensional domain. We demonstrate that Gillespie's exponential distribution is not necessary for modeling the diffusion jumps, but it may affects the convergence rates related to discretization size h. We also demonstrate that the distribution of the first exit time when $K \geq 5$ is already very close to the true distribution. From a pure diffusion point of view, we do not need very fine grids for the purpose of simulating particle diffusion. When there are multiple particles, the first exit time may show greater error with the same K values. But they are stil very close to the converged distribution when $K \geq 20$.

Chapter 4

Model 2: One Dimensional Diffusion and Monomolecular Reaction

4.1 Model Description

In this chapter, we present another model named model 2, where a mono-molecular reaction is added to the system. We will present the simulation results along with the corresponding theoretical study. Model 2 extends the framework of Model 1 by incorporating both diffusion and monomolecular reaction in a one-dimensional domain bounded by $[-L, L]$. Species A is the only reactive species in this model. When that reaction fires, A becomes a product B. Species A diffuse within the one-dimensional domain, while B stays at the location where it is generated. The reaction can be expressed as follows for this simple model:

$$A \xrightarrow{r} B \tag{4.1}$$

where r is the reaction rate. Initially, species A are in location $x = 0$. If the one-dimensional domain is discretized, we assume that each side is discretized into K bins. species A jump to neighboring bins with a jump rate d similar to (5.2) in Model 1. In each time step of the simulation, a particle of species A can either diffuse (jump left or right) or react to form species B. The jump process for species A follows the same dynamics in Model 1, where the

propensity function for the reaction diffusion system can be formulated as:

$$\alpha_0 = N_p \cdot (d + r) \tag{4.2}$$

where N_p denotes the population of A in the domain.

The simulation ends when all particles of species A have either crossed the boundary or reacted to become species B. The event time of boundary exit or reaction is recorded in each simulation for analysis.

4.2 Theoretical Study

When both reactions and diffusion jumps are involved in the system dynamics, the analysis will be similar to that of Model 1. Before we proceed with our analysis, we define the following:

$$\begin{aligned} T_0(x_0; a, b) = &\text{ the time at which a molecule of type A, starting} \\ &\text{at time0 at } x_0 \in [a, b], \text{ will either react and becomes} \\ &\text{type B, or first reach either } x = a \text{ or } x = b. \end{aligned} \tag{4.3}$$

$$\begin{aligned} T_1(x_0; a, b) = &\text{ the time at which a molecule of type A, starting} \\ &\text{at time0 at } x_0 \in [a, b], \text{ will react and becomes} \\ &\text{type B, before it reaches either } x = a \text{ or } x = b. \end{aligned} \tag{4.4}$$

$$\begin{aligned} T_2(x_0; a, b) = &\text{ the time at which a molecule of type A, starting} \\ &\text{at time0 at } x_0 \in [a, b], \text{ will first reach either} \\ &x = a \text{ or } x = b \text{ and will not have any reaction.} \end{aligned} \tag{4.5}$$

Specifically for model 2, where $a = -L$, $b = L$, and $x_0 = 0$, we will denote these three variables as T_0, T_1 and T_2 to simplify the presentation.

Similarly to the definition (2.7), we can consider two probability distributions together. One is the probability P_1 that particle A, starting at x_0 at time 0, reacts at time t before it reaches the boundary either at $x = a$ or $x = b$. The other is the probability P_2 that particle A, starting at x_0 at time 0, reaches the boundary at time t before reacting. We know that if there is no boundary limitation, the probability that there is no reaction before t for the reaction is given by: e^{-rt}; Similarly, when there is no reaction present, the probability that the particle has not reached either boundary is given by function G as in (2.7). Thus if we consider the probability that by time t, particle A, starting at x_0 at time 0, has not reacted nor reached a boundary, it should be given by

$$Prob(T_0 > t) = e^{-rt} G(x_0, t). \tag{4.6}$$

Thus, in order to calculate the mean of T_0 (denoted by $\langle T_0 \rangle$), we have

$$\begin{aligned} \langle T_0 \rangle &= -\int_0^\infty t \; d\left(e^{-rt} G(x_0, t)\right) \\ &= \int_0^\infty \left(e^{-rt} G(x_0, t)\right) dt \\ &= -\tfrac{1}{r}\left[\int_0^\infty G(x_0, t) d\left(e^{-rt}\right)\right] \\ &= \tfrac{1}{r} + \tfrac{1}{r}\int_0^\infty e^{-rt} \tfrac{\partial}{\partial t} G(x_0, t) dt. \end{aligned} \tag{4.7}$$

In the above equation, we used the corresponding results of $G(x_0, 0) = 1$ and $G(x_0, \infty) = 0$, which have already been provided in Gillespie's classic book [**?**].

Also as a result from the book, the function G satisfies the equation

$$\frac{\partial}{\partial t}G(x_0, t) = D\frac{\partial^2}{\partial x_0^2}G(x_0, t). \tag{4.8}$$

Thus, (4.7) leads to

$$\langle T_0 \rangle = \frac{1}{r} + \frac{D}{r}\frac{\partial^2}{\partial x_0^2}\langle T_0 \rangle . \tag{4.9}$$

(4.9) is a fundamental equation that we can solve for the mean event time T_0.

Denote $k = \sqrt{\frac{r}{D}}$ and $S(x_0) = \langle T_0 \rangle - \frac{1}{r}$. (4.9) leads to

$$\frac{\partial^2}{\partial x_0^2}S = k^2 S. \tag{4.10}$$

Note that if x_0 starts from the boundary $-L$ or L, then $T_0 = 0$ and $S(-L) = S(L) = -\frac{1}{r}$.

Solving equation (4.10) with the above boundary conditions, we have

$$S(x_0) = -\frac{e^{kx_0} + e^{-kx_0}}{r\left(e^{kL} + e^{-kL}\right)}. \tag{4.11}$$

Thus,

$$\langle T_0 \rangle = \frac{1}{r}\left(1 - \frac{e^{kx_0} + e^{-kx_0}}{e^{kL} + e^{-kL}}\right). \tag{4.12}$$

When $x_0 = 0$, then we have

$$\langle T_0 \rangle = \frac{1}{r}\left(1 - \frac{2}{e^{kL} + e^{-kL}}\right) = \frac{1}{r}\frac{(e^{kL} - 1)^2}{(e^{2kL} + 1)}. \tag{4.13}$$

We can analyze the formula (4.13) in two familiar situation: when there is no reaction; and

when there is no diffusion. When there is no diffusion, $D = 0$. Since $k = \sqrt{r/D}$, when $D \to 0$, the particle cannot diffuse, we have $k \to \infty$ and

$$\lim_{D\to 0} \langle T_0 \rangle = \frac{1}{r},$$

which is the mean reaction time when there is no diffusion involved. When there is no reaction, $r = 0$. If $r \to 0$, $k \to 0$, and

$$\lim_{r\to 0} \langle T_0 \rangle = \frac{L^2}{2D},$$

which is the mean exit time when there is no reaction.

On the other hand, since $e^{-rt}G(x_0, t)$ is the probability that no reaction fires and the particle does not reach the boundary, the probability density function for T_1 can be represented by

$$p_{T_1}(t) = re^{-rt}G(x_0, t). \tag{4.14}$$

Thus,

$$\langle T_1 \rangle = \int_0^\infty tre^{-rt}G(x_0, t)dt, \tag{4.15}$$

and

$$\begin{aligned} \langle T_0 \rangle &= -\int_0^\infty t \; d\left(e^{-rt}G(x_0, t)\right) \\ &= \int_0^\infty tre^{-rt}G(x_0, t)dt - \int_0^\infty te^{-rt}dG(x_0, t) \\ &= \langle T_1 \rangle - \int_0^\infty te^{-rt}\frac{\partial G(x_0,t)}{\partial t}dt. \end{aligned} \tag{4.16}$$

Note that $G(x_0, t)$ decreases as t increases. Thus $\frac{\partial G(x_0,t)}{\partial t} < 0$ and we have

$$\langle T_0 \rangle > \langle T_1 \rangle .$$

If we consider the probability p that the particle A will react and become B before it reaches the boundary. Then we have

$$\langle T_0 \rangle = p \langle T_1 \rangle + (1-p) \langle T_2 \rangle , \tag{4.17}$$

and thus

$$p = \frac{\langle T_2 \rangle - \langle T_0 \rangle}{\langle T_2 \rangle - \langle T_1 \rangle}, \tag{4.18}$$

which shows that $\langle T_2 \rangle > \langle T_0 \rangle > \langle T_1 \rangle$. Although we do not have an analytic solution for $\langle T_1 \rangle$ and $\langle T_2 \rangle$. This inequality gives us some relative connections between these three mean values.

4.3 Simulation Results

4.3.1 Numerical Comparison for T_0

To examine our formula for (4.13), we conduct a simple numerical experiment based on model 2.

Figure 4.1 shows a comparison between the mean values of measured T_0 from Model 2 simulations and the theoretical results based on Equation (4.13), with varying values of the reaction rate r. Figure 4.2 shows a similar comparison, but for varying values of the diffusion rate D. In both cases, the simulation results closely follow the theoretical curve. Table 4.1 and Table 4.2 lists the detailed measured values and the theoretical results of T_0 for each tested value of the diffusion rate D and reaction rate r.

Figure 4.3 shows the plot of T_0 for a fixed set of r, D, and L. The figure also plot an exponential function with the same mean value as the measured T_0 for comparison. It is

interesting to see that although the tail of the the measured T_0 follows the exponential curve, it shows a slower initial rise, suggesting that T_0 is not purely exponentially distributed at short times.

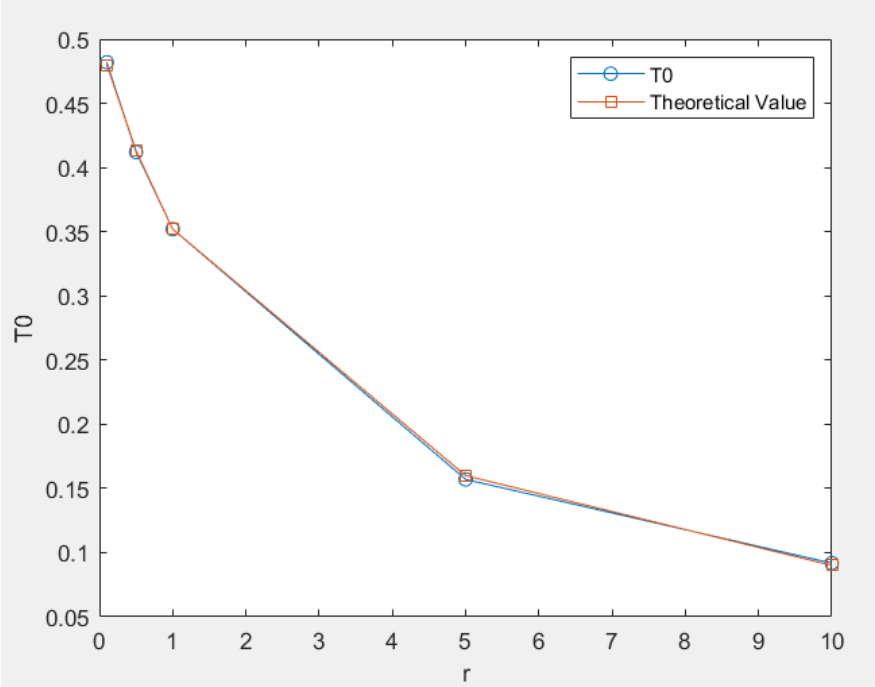

Figure 4.1: Comparison of mean values of measured T_0 and the theoretical value. The parameters are: total length L = 1.0, and D =1. The plot is generated from 1 run of stochastic simulation with 100,000 initial particles

r	0.1	0.5	1	5	10
T0	0.482	0.412	0.352	0.157	0.0918
Theoretical Value	0.48	0.413	0.352	0.16	0.09

Table 4.1: Comparison of T0 and theoretical values for different values of r

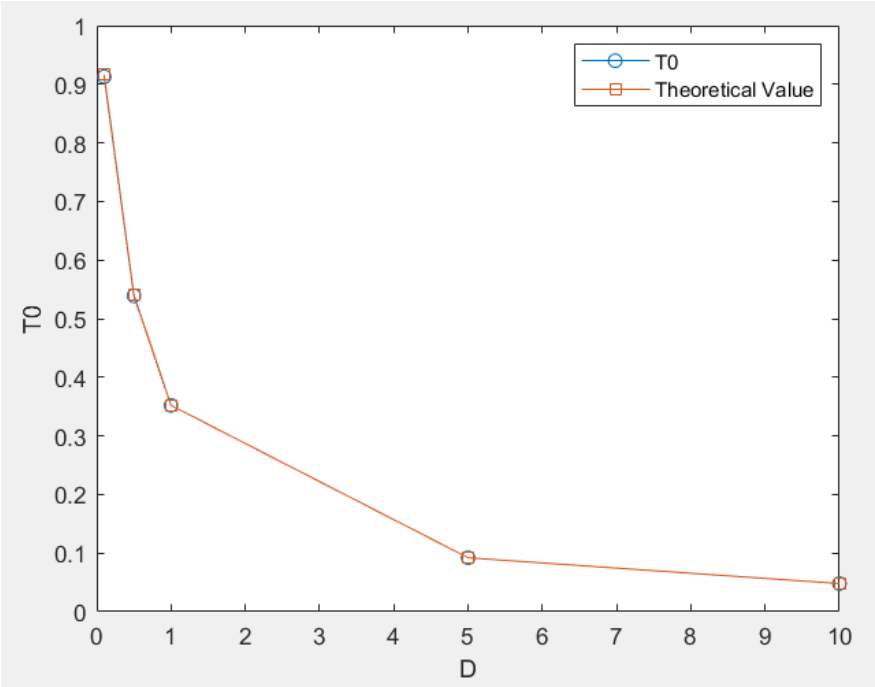

Figure 4.2: Comparsion of the mean values of measured T_0 and the theoretical value. The other parameters are set as: total length L = 1.0, and r =1. The plot is generated from 1 run of stochastic simulation with 100000 initial particles

D	0.1	0.5	1	5	10
T0	0.913	0.539	0.352	0.092	0.048
Theoretical Value	0.916	0.541	0.352	0.092	0.048

Table 4.2: Comparison of T0 and theoretical values for different values of D

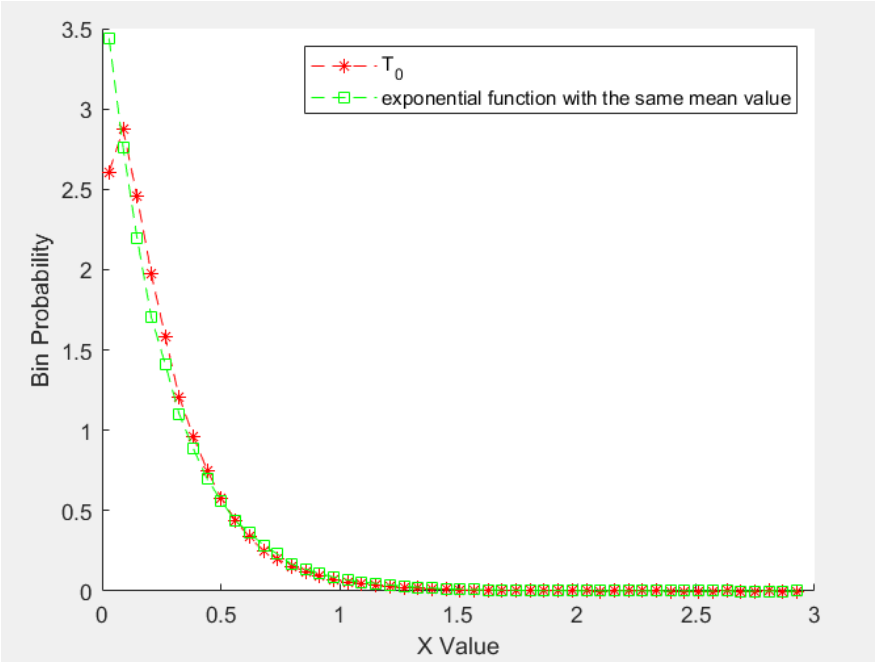

Figure 4.3: Plot of T_0 and the exponential function with the same mean value. The parameters are: total length L = 1.0, r = 2, and D = 1. The plot is generated from 1 run of stochastic simulation with 100000 initial particles

4.3.2 Event Time

The event time is defined as the reaction time and jump time recorded before a reaction fires. Table 4.2 shows the distribution difference values between the event time distributions in different bin numbers, using $BINNUM = 20$ as reference. Figure 4.4 shows the distribution

Table 4.3: Distribution Difference values for different D values and bin numbers

D	**Bin = 1**	**2**	**3**	**4**	**5**
0.1	0.048218	0.046	0.0348	0.0307	**0.0203**
0.5	0.059	**0.0205**	–	–	–
1	0.0823	**0.022056**	–	–	–
5	0.1175	0.0359	**0.0153**	–	–
10	0.1226	0.0324	**0.0128**	–	–

of the event time T_0 for $BINNUM = 2$ and the diffusion rate $D = 1$.

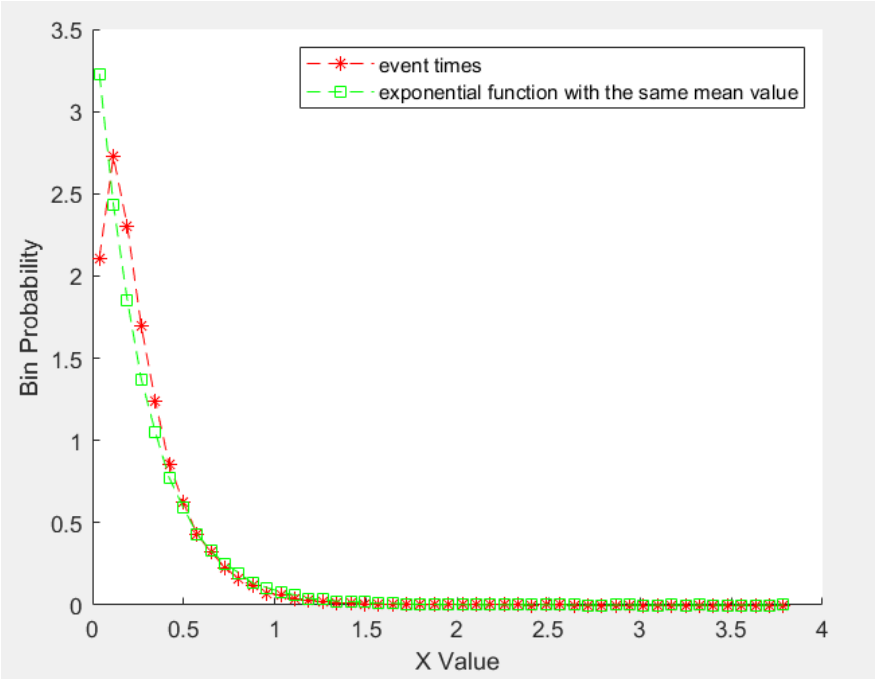

Figure 4.4: Distribution of event times and the exponential function with the same mean value. The other parameters are set as: total length L = 1.0, r = 2, and D = 1. The plot is generated from 1 run of stochastic simulation with 100000 initial particles

4.4 Conclusion for Model 2

In this chapter, we derived an analytical formula for the first exit time and reaction time of diffusion and monomolecular reaction in a one-dimensional domain and performed numerical experiments to evaluate the result. We demonstrate that our simulation result match with the theoretical result pretty well.

Chapter 5

Model 3: Bimolecular Reaction and One Dimensional Diffusion

5.1 Model Description

In this chapter, we will study Model 3, in which the mono-molecular reaction in Model 2 is replaced with a bi-molecular reaction. We will analyze the reaction time and then use numerical simulation results to compare with corresponding theoretical study. Model 3 contains two types of diffusive molecules, A and B, which diffuse in a one-dimensional domain and undergo a bimolecular reaction:

$$A + B \xrightarrow{r} C$$

where r is the reaction rate and C is the product that we do not track in our model.

Similar to Model 1 and Model 2, the domain is defined over the interval $[-L, L]$ and discretized into a finite number of compartments. At the start of each simulation run, particle A is initialized at position $-L/2$, and particle B is initialized at position $L/2$(Figure 5.1). Both particles$A$ and B can jump randomly to neighboring bins with diffusion coefficients d_A and d_B.

$$d_A = \frac{2D_A}{h^2} \tag{5.1}$$

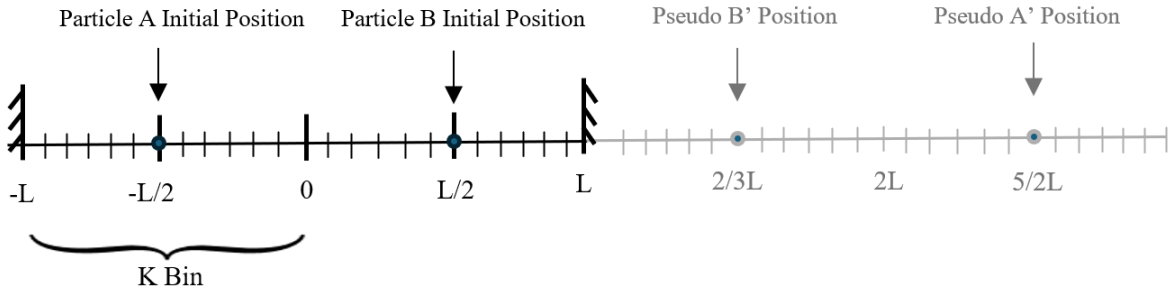

Figure 5.1: Schematic representation of the one-dimensional bounded domain used in the simulation, with particles A initialized at position $-L/2$ and particles B at $+L/2$, allowed to jump to adjacent compartments.

$$d_B = \frac{2D_B}{h^2} \tag{5.2}$$

Because the domain is one-dimensional, the reaction rate between one A particle and one B particle is given by:

$$r' = \frac{r}{h} \tag{5.3}$$

where h is given in 3.1. Note that in the current setting, A and B can only react when they are in the same bin.

At each time step, one of the following events occurs: particle A jumps left or right (with rate d_A), particle B jumps left or right (with rate d_B), or particles A and B react (with rate r') if they are in the same bin. Therefore, the total propensity is given by:

$$\alpha_0 = d_A + d_B + r' \cdot \mathbf{1}_{\{A=B\}},$$

where $\mathbf{1}_{\{A=B\}}$ is a Boolean function that equals 1 if A and B are both in the same bin and 0 otherwise.

The simulation ends when a reaction occurs. The time of first reaction is recorded for each

simulation for analysis.

5.2 Theoretical Study

In order to analyze the reaction time under the setting of Model 3, we will consider the whole process to be two parts. One is that they have to jump into the same bin. In this part, it will only be involved with diffusion jumps, and we can analyze it very similar to Model 1. Once there are in the same bin, there could be either a jump event or a reaction.

Let us start with the diffusion part first. Both A and B particles can diffuse. To simplify the analysis, we can fix one of the particles, assuming that it is A, and let only B diffuse, but with a diffusion rate $D = D_A + D_B$. In this case, the diffusion part can be viewed as A is fixed, and B jumps within the interval $[-\frac{L}{2}, \frac{5L}{2}]$, because we can extend the interval from the location L, up to $3L$. The interval $[L, 3L]$ can be considered as a reflection of interval [-L, L] with the point $x = L$. Thus, a symmetric A particle can be considered to be located at $x = \frac{5L}{2}$. Thus, the first time A and B jump into the same bin can be viewed as the exit time for B to jump in the interval $[-\frac{L}{2}, \frac{5L}{2}]$. In this case, initially B has a distance L to the left boundary $-\frac{L}{2}$ and a distance $2L$ to the right boundary $\frac{5L}{2}$. With the analysis in Chapter 3, we know that the mean time for that exit time can be given by:

$$T_0 = \frac{L * (2L)}{2D} = \frac{L^2}{D}, \tag{5.4}$$

where $D = D_A + D_B$.

When both A and B are in the same bin, diffusion jumps and bi-molecular reaction are both

possible. The probability of a reaction is given by:

$$p = \frac{\frac{r}{h}}{\frac{r}{h} + \frac{2D}{h^2}}, \tag{5.5}$$

where $\frac{2D}{h^2}$ represents diffusive jumps towards both directions. With the other probability $1 - p$, B jumps to the neighboring bin, and the bimolecular reaction may fire again only when B jumps back to A. If the B particle jumps to the right bin, the mean time for it to come back to A is given by

$$T_1 = \frac{h * (3L - h)}{2D}. \tag{5.6}$$

But if the B particle jumps to the left bin, it will face a fixed A at $x = -\frac{L}{2}$ and a reflected pseudo A at $x = -\frac{3L}{2}$. Thus, the total distance between the two absorbing boundaries is given by L, and the corresponding mean exit time is given by

$$T_2 = \frac{h * (L - h)}{2D}. \tag{5.7}$$

After B jumps back to A, it returns to the same situation that A and B land in the same bin. Thus, if we denote T_3 as the mean time that, when A and B are in the same bin, how long it takes for the bimolecular reaction to fire. We have

$$T_3 = p\frac{1}{\frac{r}{h} + \frac{2D}{h^2}} + (1 - p) \cdot \left[\frac{1}{2}(T_1 + T_2) + T_3\right]. \tag{5.8}$$

Solving (5.8) for T_3, we obtain the following.

$$T_3 = \frac{h^2}{2D + rh} + \frac{2L - h}{r}. \tag{5.9}$$

When the number of bins increases, $h \to 0$, and then the above equation leads to

$$\lim_{h\to 0} T_3 = \frac{2L}{r}, \tag{5.10}$$

and the overall mean time for the bimolecular reaction to fire in Model 3 is given by

$$T = \frac{L^2}{D} + \frac{2L}{r}. \tag{5.11}$$

Interestingly, this mean time contains exactly two parts: the first term $\frac{L^2}{D}$ represents the time it takes A and B to jump to the same bin, while the second term $\frac{2L}{r}$ is exactly the average time for the reaction if the interval $[-L,\ L]$ is not discretized at all. Thus, if the first term is much smaller than the second term, numerically we will obtain the same mean reaction time even without spatial discretization. This criteria can be formulated as

$$\frac{L^2}{D} \ll \frac{2L}{r}, \tag{5.12}$$

or equivalently

$$\frac{r}{D} \ll \frac{2}{L}. \tag{5.13}$$

It indicates that if the reaction is very slow, or if the combined diffusion rate is high, numerically one can treat the system as a well-mixed one and directly apply Gillespie's SSA to it.

The mean reaction time itself is independent of the discretization size h. Based on the above analysis and our results in Chapter 3, T_0 accuracy can be viewed as independent of h as long as $K \geq 5$. On the other hand, (5.9) shows that there is an error term related to h that can be formulated as:

$$E_{abs} = \frac{h}{r}. \tag{5.14}$$

If we only consider relative error, the relative error can be formulated as

$$E_{rel} = \frac{h/r}{\frac{L^2}{D} + \frac{2L}{r}} = \frac{h}{2L + \frac{r}{D}L^2}. \tag{5.15}$$

If L is small, or if r is small compared to D, this error term may be relatively large; then the corresponding stochastic simulation requires greater K and this increases the computational cost. In order to develop a formula for a reasonable K, we can set up a parameter $0 < \epsilon \ll 1$, and requires that

$$E_{rel} \leq \epsilon.$$

Thus, (5.15) leads to

$$\frac{h}{2L + \frac{r}{D}L^2} \leq \epsilon, \tag{5.16}$$

as $K = \frac{L}{h}$, the above inequality leads to

$$K \geq \frac{1}{\epsilon\left(2 + \frac{r}{D}L\right)} \tag{5.17}$$

So what is the computational cost corresponding to discretization size h or bin number K? We can simply consider the mean reaction time and the mean jump time in each step. Because in Model 3, the system dynamics stops when A and B react, we can deduct the reaction time in the final step from the mean reaction time. Then all the rest time is spent on jumps, given that the average jump time for each step is given by $\frac{h^2}{2D}$, we can estimate the mean number of jumps before the reaction fires, and that is also the computational cost. From this analysis, we formulate the computational cost as follow:

$$Cost = \frac{\frac{L^2}{D} + \frac{2L}{r} - \frac{h}{r}}{\frac{h^2}{2D}} \approx \frac{2rL^2 + 4DL}{rh^2}. \tag{5.18}$$

Note that $h = \frac{L}{K}$, thus this cost is also

$$Cost = \left(2 + \frac{4D}{rL}\right) K^2. \tag{5.19}$$

5.3 Simulation Results

5.3.1 Convergence of bimolecular reaction time

To examine the formula for the bimolecular reaction (5.11), a simple numerical experiment in model 3 is presented here. Figure 5.2 shows the distributions of the bimolecular reaction time between different bins. It is apparent that as the number of bins increases, the distributions converge.

Table 5.1 summarized the mean reaction time, absolute error (compared to the theoretical value), and relative errors for each bin size. Based on equation (5.15), when L, r, and D are fixed, the relative error should decrease as h decreases. In our simulation results, as the number of bins increases, h becomes smaller and the relative error also decreases. Therefore, the results are consistent with the theoretical expectations. It is notable that for K(Bin) = 32, the relative error drops below 1% and effectively converges to the theoretical value.

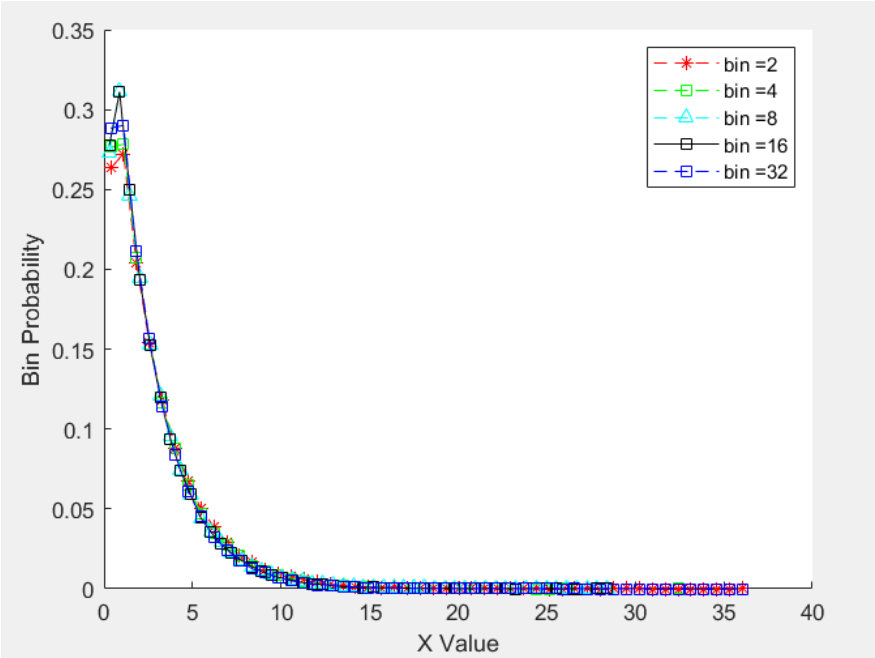

Figure 5.2: Distribution of bimolecular reaction times. The parameters are: total length L = 1.0, and D = D_A + D_B = 2, r = 1. The plot is generated from 100000 run of stochastic simulation

Bin=	**2**	**4**	**8**	**16**	**32**
Mean T	2.77488	2.64787	2.57693	2.55129	2.5219
Error	0.27488	0.14787	0.07693	0.05129	0.0219
Relative Error	0.109952	0.059148	0.030772	0.020516	0.00876

Table 5.1: Mean value of bimolecular reaction time T, error, and relative error across different bin sizes. The theoretical T is 2.5

5.3.2 Effect of r on convergence of bimolecular reaction time

In this experiment, the goal is to examine how the reaction rate r affects the accuracy of the reaction time distribution in Model 3. Table 5.2 shows the distribution difference values for different bin numbers between Bin = 32. It is interesting to see that as r increases, fewer bins are required for the distribution to converge.

Bin =	**2**	**4**	**8**
r = 0.1	0.09800	0.04026	**0.02350**
r= 1	0.07392	0.03934	**0.01952**
r = 10	0.03142	**0.01036**	
r = 100	0.0529	**0.0168**	

Table 5.2: Distribution Difference values for different bin numbers between Bin = 32. The parameters are: L = 1, D = 1.The result is generated from 100000 run of stochastic simulation

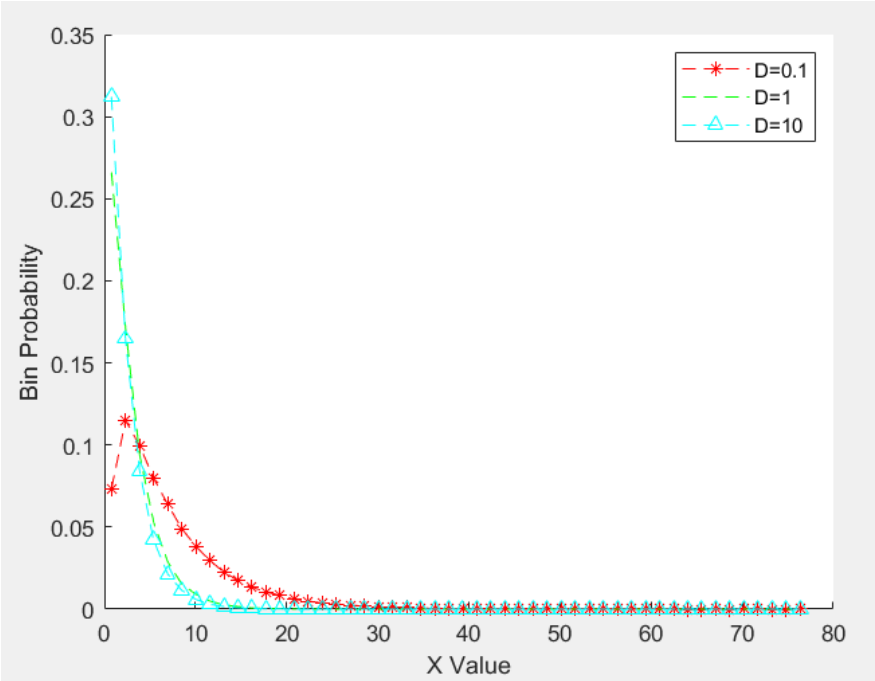

Figure 5.3: Distribution of bimolecular reaction times for K=2

5.3.3 Computational Cost

To examine the formula for computational cost 5.19, we performed experiments using a range of bin sizes. The results are summarized in Table 5.3 and plotted in 5.4. It is apparent that the measured number of jumps closely matches the theoretical values calculated from the

formula given in 5.19.

Bin =	2	4	8	16	32
Average Number of Jump	44.3467	168.474	662.968	2621.62	10403
Theoretical Result	40	160	640	2560	10240

Table 5.3: Average number of jumps for different bin

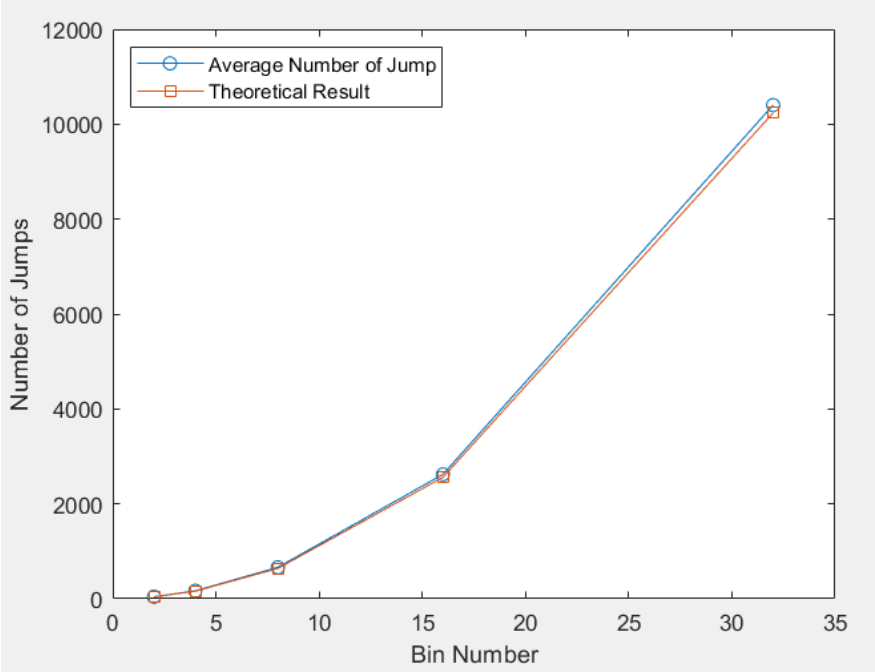

Figure 5.4: Relationship between average number of jump and theoretical Result for different bin

5.4 Conclusion for Model 3

In this chapter, we derived an analytical formula for the first exit time, reaction time and computational cost of diffusion and bimolecular reaction in a one-dimensional domain and performed numerical experiments to evaluate the result. We demonstrate that our simulation

result match with the theoretical result pretty well. We also show that when a bimolecular reaction is involved, increasing the reaction rate r reduces the number of spatial bins K required to achieve accurate simulation results.

Chapter 6

Conclusions

In this thesis three toy models related to stochastic reaction diffusion processes and the corresponding stochastic simulation are studied. The theoretical analysis method based on the first passage time (or first exit time) analysis is adopted to analyze the three toy models. The mean values of the first passage time for these three different models were analytically presented. Numerical results also verified the accuracy of these formulas. Meanwhile, we use these three models to study the simulation accuracy and efficiency related to different discretization sizes. Based on our analysis, when reaction is not involved, a discretization scheme of $K \geq 5$ is already quite accurate for the first passage time simulation. When reaction is involved, it also seems that we do not need a very high number of bins to achieve good accuracy for monomolecular reaction and bimolecular reaction.

We believe the theoretical and numerical results present here point to a good direction to theoretically study the numerical scheme of stochastic simulation for reaction diffusion systems. Although the three toy models cannot cover all possible cases in one-dimensional RD systems, the method present here can be applied to all of them. Of course, the real challenges lie in the two or three dimensional domain. We look forward to extending our analysis to those cases in following studies.

Bibliography

[1] Yang Cao, Daniel T. Gillespie, and Linda R. Petzold. The slow-scale stochastic simulation algorithm. *The Journal of Chemical Physics*, 122(1):014116, 2005.

[2] Yang Cao, Hong Li, and Linda R. Petzold. Efficient formulation of the stochastic simulation algorithm for chemically reacting systems. *The Journal of Chemical Physics*, 121(9):4059–4067, 2004.

[3] Yang Cao and Linda Petzold. Accuracy limitations and the measurement of errors in the stochastic simulation of chemically reacting systems. *Journal of Computational Physics*, 212:6–24, 2005.

[4] Einstein, A. *Investigations on the Theory of Brownian Movement*. New York: Dover. ISBN 978-0-486-60304-9. , 1956.

[5] Radek Erban and S. Jonathan Chapman. Stochastic modelling of reaction–diffusion processes: algorithms for bimolecular reactions. *Physical Biology*, 6(4):046001, 2009.

[6] Adolph Fick. V. on liquid diffusion. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 10(63):30–39, 1855.

[7] Daniel T Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics*, 22(4):403–434, 1976.

[8] Daniel T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, 81(25):2340–2361, 1977.

[9] Daniel T. Gillespie. Approximate accelerated stochastic simulation of chemically reacting systems. *The Journal of Chemical Physics*, 115(4):1716–1733, 2001.

[10] Daniel T. Gillespie. Stochastic simulation of chemical kinetics. *Annual Review of Physical Chemistry*, 58(1):35–55, 2007.

[11] Daniel Thomas Gillespie and Effrosyni Seitaridou. *Simple Brownian Diffusion: An Introduction to the Standard Theoretical Models*. Oxford University Press, 1st edition, 2012.

[12] Bartosz A. Grzybowski. *Chemistry in Motion: Reaction-Diffusion Systems for Micro- and Nanotechnology*. John Wiley & Sons, Hoboken, NJ, 2009.

[13] Sidney Redner. *A Guide to First-Passage Processes*. Cambridge University Press, 2001.

[14] Siowling Soh, Marta Byrska, Kristiana Kandere-Grzybowska, and Bartosz A. Grzybowski. Reaction-diffusion systems in intracellular molecular transport and control. *Angewandte Chemie International Edition*, 49(23):4170–4198, 2010.