MICROSCOPIC LATTICE METHOD FOR STUDYING INTRACELLULAR REACTION-DIFFUSION KINETICS

CHEW WEI XIANG

FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR

2019

MICROSCOPIC LATTICE METHOD FOR STUDYING INTRACELLULAR REACTION-DIFFUSION KINETICS

CHEW WEI XIANG

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF PHYSICS
FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR

2019

UNIVERSITI MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: **CHEW WEI XIANG**

Registration/Matric No.:**SHC140082**

Name of Degree:**DOCTOR OF PHILOSOPHY**

Title of Project Paper/Research Report/Dissertation/Thesis ("this Work"):

MICROSCOPIC LATTICE METHOD FOR STUDYING INTRACELLULAR REACTION-DIFFUSION KINETICS

Field of Study:**THEORETICAL PHYSICS**

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date:

Subscribed and solemnly declared before,

Witness's Signature Date:

Name:
Designation:

MICROSCOPIC LATTICE METHOD FOR STUDYING INTRACELLULAR REACTION-DIFFUSION KINETICS

ABSTRACT

Reaction and diffusion plays fundamental roles in many intracellular biochemical process. Bottom-up modelling approaches provide a way to connect the microscopic reaction-diffusion influenced by intracellular conditions to the macroscopic cellular behavior. However, microscopic computational method that diffuses the molecule in the continuum, such as the Brownian dynamics simulation, require long duration to simulate reaction-diffusion up to the subcellular scale. In contrast, simulation scheme based on microscopic lattice method (MLM) that represent each molecule explicitly as a random walker on lattice voxels has a relatively low computational cost. Yet, the theory and consistency of MLM have not been clarified in detail. In this work, we performed in-depth analysis on MLM based on the Spatiocyte simulation scheme in solving diffusion-influenced reaction. We construct the theoretical framework for MLM based on the Smoluchowski-Collins-Kimball model for bimolecular reaction and the random walk theory. By matching the MLM theory with the continuum-based theory, we obtain the expressions for determining the simulation parameter. The MLM theory and the simulation result is then validated with the continuum-based theory and simulation method. We also show that the run time of MLM is an order of magnitude faster than continuum-based method for simulating macromolecules diffusion at typical intracellular concentrations. We have performed several case studies on reaction-diffusion process often encountered in the cell. Finally, we discuss the applicability and limitation of MLM.

Keywords: reaction-diffusion, Monte Carlo simulation, random walk, spatial stochastic simulation method, lattice-based method

KAEDAH KEKISI MIKROSKOPIK UNTUK KAJIAN TINDAKBALAS-RESAPAN DALAM SEL

ABSTRAK

Tindakbalas-resapan berperanan penting dalam banyak proses biokimia dalam sel. Pendekatan bottom-up model membenarkan penyambungan tindakbalas-resapan dari skala mikroskopik yang dipengaruhi oleh keadaan intra-sel ke skala sel. Walau bagaimanapun, kaedah kiraan mikroskopik konvensional yang mengerakkan molekul di kontinum, seperti simulasi dinamik Brownian, memerlukan masa yang lama untuk mensimulasikan sehingga skala subselular. Sebaliknya, skim simulasi berdasarkan kaedah kekisi mikroskopik (KKM), dimana molekul digerakkan secara rawak dalam kisi mikroskopik, mempunyai kos pengiraan yang agak rendah. Walau bagaimanapun, teori dan konsistensi KKM belum dijelaskan dan dianalisasi secara terperinci. Dalam kerja ini, kami menganalisasi kebolehan KKM untuk mensimulasi tindakbalas yang dipengaruhi oleh resapan dengan memakai skim Spatiocyte. Kami membina kerangka teori tindakbalas bimolekul untuk KKM mengikuti model Smoluchowski-Collins-Kimball dan teori gerakan rawak. Dengan memadankan teori kinetik tindakbalas KKM dengan teori berasaskan kontinum, kita memperoleh ekspresi untuk menentukan parameter KKM untuk simulasi. Teori dan simulasi KKM kemudian disahkan dengan teori dan simulasi kontinum. Kami menunjukkan bahawa masa yang diperlukan oleh KKM untuk simulasi adalah lebih cepat daripada kaedah konvensional apabila kepekatan yang tipikal dalam sel diguna. Kami juga membentangkan kes kajian mengenai beberapa sistem tindakbalas-resapan yang ditemui dalam sel. Akhirnya, kami membincangkan applikasi dan had KKM.

Kata kunci: tindakbalas-resapan, Monte Carlo simulasi, gerakan rawak, kaedah simulasi stokastik ruangan, kaedah kekisi

ACKNOWLEDGEMENTS

This research would not have been possible without the support from many individuals. First, I would like to extend my gratitude to my supervisors: Prof. Sithi V. Muniandy (Department of Physics, UM) and Dr. Satya N. V. Arjunan (RIKEN). My PhD journey would not have started without the introduction to Dr Satya's work by Prof. Sithi. Dr. Satya's constant guidance has helped my transition into a new research field. Prof Sithi's resourcefulness and creativity are the source of inspiration for me in tackling new research problems. I have inherited not only their technical skills, but also their research attitudes and ethics. I am very fortunate to have the freedom for exploring research topics, and received help during challenging times.

I would also like to thank my host in RIKEN Quantitative Biological Center, Dr. Koichi Takahashi, for providing a great working environment and research facility. I thank Dr. Kazunari Kaizu for his help in the theoretical analysis and numerical implementation; Dr. Kozo Nishida for his technical support on computing aspects; Dr. Masaki Watabe for his critical inputs and feedbacks; Suguru Kato and Toru Nina for their help in software development and debugging; and Kaoru Ikegami for her generous assistance.

I appreciate the comments from the thesis examiners: Prof. Per Lötstedt, Prof. Masanori Tachiya and Dr. Chew Khian Hooi. I acknowledge the financial support provided by RIKEN through the International Program Associate fellowship, University of Malaya under the Skim Biasiswa Universiti Malaya fellowship and the Frontier Research Grant (FG032-17AFR).

Finally, I thank my family members for always believing in me. Special thanks to my wife, for her unconditional support and encouragement. Last but not least, my baby girl, for her cooperation throughout the thesis writing.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

P'_a : actual reaction probability

s_1 : an element of the set of immediate neighbor voxels of s_0

s_2 : an element of the set of second nearest neighbor voxels of s_0

S : area of 2D plane (L^2)

β_1 : average reaction rate

β_2 : average hopping rate

λ : chemical kinetics relaxation rate constant

α : coefficient of reaction probability and diffusion time step

k_D : collision rate for three-dimensional reaction ($L^3\ T^{-1}$)

K : compartment length in number of voxels

$K(x)$: complete elliptic integral of the first kind

$[x]$: concentration of species x

$G_{n+1}(t)$: continuous rebinding-time probability density after $(n + 1)$ steps

D_x : diffusion coefficient for species x ($L^2\ T^{-1}$)

$\widetilde{F}(s)$: diffusion factor function

n : diffusion step

t_d : diffusion time step (T)

d : dimension

k_{eff3D} : effective reaction rate constant for three-dimensional reaction ($L^3\ T^{-1}$)

R_{eff} : effective radius (L)

K_{eq} : equilibrium constant

P_e : escape probability

γ : Eular constant

$F_n(s_a|s_b)$: first-passage time distribution for a random walker to walk from voxel s_b to s_a

$\overline{F}_n(s_1|s_2)$: first-passage time distribution from s_2 to s_1 at step n

f_{2D} : fraction of equilibrium two-dimensional rate

$\psi_h(t)$: hopping time distribution/ survival probability of the hopping events

c_0 : initial molecule concentration

δ : infinitesimal value

k_a : intrinsic rate constant

k'_a : initial lattice rate constant

k_{a1D} : intrinsic rate constant for one-dimensional reaction ($\mathrm{L\ T^{-1}}$)

k_{a2D} : intrinsic rate constant for two-dimensional reaction ($\mathrm{L^2\ T^{-1}}$)

k_{a3D} : intrinsic rate constant for three-dimensional reaction ($\mathrm{L^3\ T^{-1}}$)

k_{sa} : intrinsic adsorption rate constant ($\mathrm{L\ T^{-1}}$)

k'_{sa} : intrinsic lattice adsorption rate constant ($\mathrm{L\ T^{-1}}$)

k_{d2D} : intrinsic dissociation rate constant for two-dimensional reaction ($\mathrm{T^{-1}}$)

k_{d3D} : intrinsic dissociation rate constant for three-dimensional reaction ($\mathrm{T^{-1}}$)

k_{sd} : intrinsic desorption rate constant ($\mathrm{T^{-1}}$)

$\hat{G}(s)$: Laplace form of $G_{n+1}(t)$

$\hat{S}_n(s_1|s_1;s)$: Laplace form of $S_n(s_1|s_1)$

$k(s)$: Laplace form of time-dependent rate coefficient

$k'(s)$: Laplace form of on-lattice time-dependent rate coefficient

$\hat{S}_{irr,x}(s)$: Laplace form of the irreversible reaction survival probability for species x

ϕ_p : lattice percolation threshold

ϕ : lattice occupancy

k_{on2D} : macroscopic rate constant for two-dimensional association reaction ($L^2\ T^{-1}$)

k_{on3D} : macroscopic rate constant for three-dimensional association reaction ($L^3\ T^{-1}$)

τ_m : mean life-time (T)

$[x]'$: modified concentration of species x

$K_\nu(z)$: modified Bessel functions of the second kind

r_x : molecular radii of species x (L)

$\emptyset$: non-existence

$\mathbb{N}$: non-negative integers

N_x : number of molecules of species x

N_v : number of volumetric voxels

N_{sv} : number of surface voxels

N_o : number of obstacles

Z : number of collision

k'_D : on-lattice collision rate for three-dimensional reaction ($L^3\ T^{-1}$)

k'_{eff3D} : on-lattice effective reaction rate constant for three-dimensional reaction ($L^3\ T^{-1}$)

$k'_{sa}(t)$: on-lattice time-dependent rate adsorption rate coefficient

$k'_i(t)$: on-lattice time-dependent rate coefficient for i-dimensional reaction

$O(x)$: order of expression

ϕ_l : packing fraction for lattice

$\mathbb{Z}^+$: positive integers

t_r : predefined run time (T)

$F(s_a|s_b; z)$: probability generating function for $F_n(s_a|s_b)$

$F_n^j(s_0|s_0)$: probability to reach the origin for the jth time at the nth step

$H(s_0|s_1;z)$: probability generating function for $H_n(s_0|s_1)$

$P_1(s_0|s_1)$: probability to select s_1 given that the molecule is in s_0

$P(s_a|s_b;z)$: probability generating function of $P_n(s_a|s_b)$

$P_h(t=m)$: probability of successful hopping after m trials

$S_n(s_1|s_1)$: probability that a particle is in contact after n steps

$S(s_1|s_1;z)$: probability generating function of $S_n(s_1|s_1)$

$p(r,t;r_0,0)$: probability that the particle is at position r at time t given that it was at r_0 at time t_0

$G(s_0|s_1;z)$: probability generating function of $G_n(s_0|s_1)$

f : ratio of voxel to molecule size

κ : ratio of k_a and D

$P_a,\ P_r$: reaction probability

P_{ai} : reaction probability for i-dimensional association reaction

P_{axy} : reaction probability for reactive pair of species x and y

$\psi_r(t)$: reaction time distribution/ survival probability of the hopping events

τ_{rel} : reactivation time (T)

$p_{reb}(R,t|R,0)$: rebinding-time probability distribution for a reactive particle-pair separated by a distance R at time t.

$H_n(s_0|s_1)$: rebinding-time probability distribution for the activation-limited scheme in diffusion step n

$G_n(s_0|s_1)$: rebinding-time probability distribution for the diffusion-influenced scheme in diffusion step n

D : relative/total diffusion coefficient

D' : reduced long-time diffusion coefficient ($L^2\ T^{-1}$)

t_s : simulation run time (T)

m : simulation step

P_w : step acceptance probability

k_{ss3D} : steady-state reaction rate constant for three-dimensional reaction ($L^3\ T^{-1}$)

k_{ss2D} : steady-state reaction rate constant for two-dimensional reaction ($L^3\ T^{-1}$)

k_m : step-dependent rate coefficient at simulation step m

k_m^i : step-dependent rate coefficient for i-dimensional reaction

R : sum of the radii of two reacting molecules (L)

$P_h(t > m)$: survival probability (no hopping) until mth trial

$S_n(s_1|s_1;t)$: survival-time probability density of a particle that started and ended at s_1 on the nth step

$S^i_{irr,x}(t)$: survival probability of species x in a i-dimensional irreversible reaction

$S_{rev,x}(t)$: survival probability of species x in a reversible reaction

$S(t;R)$: survival probability of an isolated reactant pair at time t, given that they were initially in contact

t : time (T)

$k(t)$: time-dependent rate coefficient

$k_i(t)$: time-dependent rate coefficient for i-dimensional reaction

$k_{sa}(t)$: time-dependent rate adsorption rate coefficient

G_{reb} : total rebinding probability for diffusion-influenced scheme

H_{reb} : total rebinding probability for activation-limited scheme

P_{reb} : total rebinding probability

$p(s_0 \rightarrow s_1)$: transition probability from s_0 to s_1

V : volume (L^3)

s_0 : voxel at the origin

$P_n(s_a|s_b)$: voxel occupation probability from s_b to s_a

l : voxel size (L)

1D : one dimensional or one dimension

2D : two dimensional or two dimension

3D : three dimensional or three dimension

EGFRD : enhanced Green's function reaction dynamics

HCP : hexagonal closed packed

K : MAPK

KK : reactivated MAPK kinase

KK* : inactivated MAPK kinase

MPK1 : multiparticle kernel theory

MLM : microscopic lattice method

MAPK : mitogen-activated protein kinase

MAPKK : MAPK kinase

ODE : ordinary differential equation

P : reactivated phosphatase

P* : inactivated phosphatase

RD : reaction-diffusion

RDME : reaction-diffusion master equation

RDMEm : modified reaction-diffusion master equation

SVTA : small-voxel tracking algorithm

SCK : Smoluchowski-Collins-Kimball

CHAPTER 1: INTRODUCTION

1.1 Introduction and Motivation

Reaction and diffusion play fundamental roles in many intracellular biochemical processes, such as cell polarity establishment (Halatek et al., 2018), symmetrical cell division (Wettmann & Kruse, 2018), cell migration (Devreotes et al., 2017) and signal transduction (Zhou & Hancock, 2018). Although the key elements underlying these biochemical processes can be identified using conventional biological techniques, it is less clear how individual element interacts dynamically to produce the observable macroscopic behaviors. A model which either consist of mathematical formula or computational calculation, provides a means to explain the experimental data quantitatively, and to formulate and test hypotheses or to predict the behavior. Because of the broad spatio-temporal scales in intracellular biochemical system, it is important to determine the scales of interest and to neglect unnecessary details of the system in order to construct a practical model (see Figures 1.1 and 1.2) (Phillips et al., 2012; Milo & Phillips, 2015). In intracellular reaction-diffusion (RD) process that forms subcellular or cellular scale spatio-temporal pattern, macromolecules, such as protein, are often the smallest unit of interest (Halatek et al., 2018). In such cases, faster dynamics such as DNA synthesis and movement of smaller molecules is typically not represented explicitly in the model.

Reaction and diffusion in the intracellular environment differ from in-vitro system in several aspects. Due to the small number of some macromolecule ensemble, the number fluctuation has non-negligible influence on the overall cellular behavior (Rao et al., 2002; Bar-Even et al., 2006). For such cases, it is therefore inappropriate and insufficient to use the notion of concentration together with deterministic description of the reaction kinetics. Intracellular space is also spatially organized (Ryan & Shapiro, 2003) and

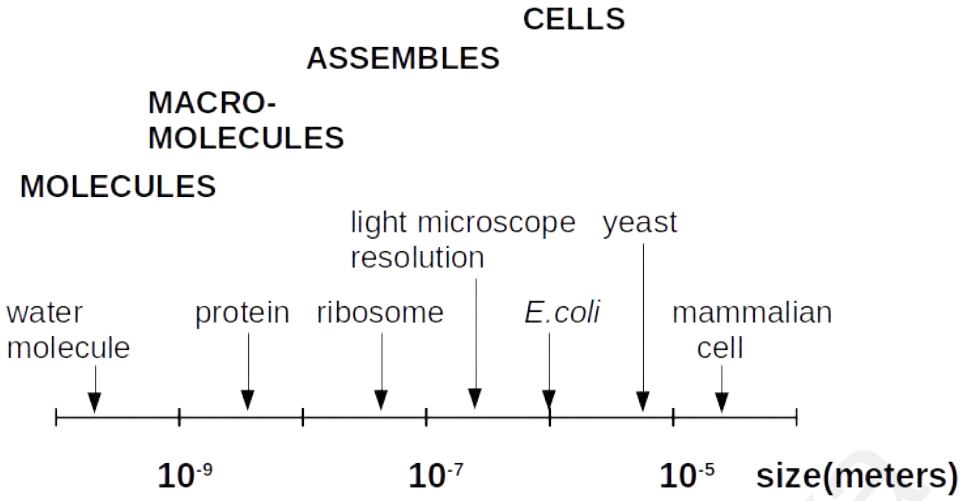


Figure 1.1: Characteristic size of some biological entities in a cell. Adapted from (Milo & Phillips, 2015).

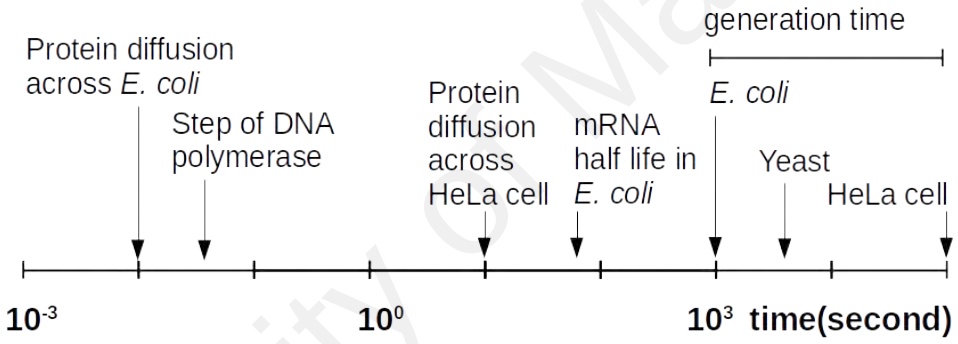


Figure 1.2: Characteristic time scale of some intracellular processes in a cell. Adapted from (Milo & Phillips, 2015).

crowded (Ellis, 2001; Ellis & Minton, 2003). The geometrical and volume exclusion effects has been reported to alter the transport behavior (Nicolau Jr et al., 2007; Dix & Verkman, 2008; Höfling & Franosch, 2013) and the reaction kinetics of macromolecules (Saxton & Jacobson, 1997; Melo & Martins, 2006; Kalay, 2012; Zhdanov & Höök, 2015; Kerketta et al., 2016). Therefore, conventional description of the reaction and diffusion that assume homogeneous environment is insufficient for intracellular system.

Taken all these into consideration, a spatial and stochastic model is necessary to reflect the realistic intracellular environment into the RD process. A natural approach to achieve this is by adopting the particle-based computational model. In the particle-based approach,

molecules and the intracellular compartments are represented explicitly. Due to the highly viscous intracellular environment, long-range interactions between macromolecules such as hydrodynamical forces can be neglected to some extent. To probe the dynamics at subcellular or cellular scale, reaction dynamics at smaller spatial scale and faster time scale can be effectively integrated out. These eventually lead to the hard-sphere representation of the macromolecule, where molecule size is described by the hydrodynamical radius, whereas reaction is described by the intrinsic rate constant and the molecular binding radius. These hard-sphere molecules undergo normal diffusion explicitly in the cellular compartments that mainly consist of the cytosolic volume and the membrane surface. Molecules undergo the unimolecular reaction (first order reaction), where species changes to another distinct species with rate constant of dimension (1/T), or the bimolecular reaction (second order reaction), where two reactants react upon collision with each other. Bimolecular reaction can be classified into (i) the volume-volume reaction, where cytosolic molecules react with each other; (ii) surface-surface reactions, where membrane-bound molecules react with each other; and (iii) volume-surface reactions, where cytosolic molecules react with the membrane or membrane-bound molecules.

Particle-based simulation method such as Brownian dynamics is as accurate as the underlying physical model. To ensure accuracy in simulating diffusion, small time step is required to resolve the molecule collision. However, the computational cost involved increases with the number of molecules and the details incorporated. Brownian dynamics is feasible for simulating subcellular scale phenomena but demanding for whole-cell simulation. In this study, we focus on an alternative microscopic simulation method that is computationally less costly. In this approach, molecules are propagated on uniform-sized lattice instead of in the continuum. We refer this approach as the microscopic lattice method (MLM), which is also known as the Monte Carlo lattice gas automata in some literature

(Saxton, 2007). MLM also retain the important features of particle-based approach in simulating RD process, while being computationally tractable.

1.2 Problem Statement and Objectives

MLM has some advantage in the computational cost of simulating RD process. However, realistic biological space does not arranged as fixed lattice structures. It is uncertain whether the diffusion behavior on lattice can recapitulate the diffusion behavior in the continuum. Furthermore, diffusion and reaction kinetics can be influenced by the lattice arrangement as shown previously(Hughes, 1995; Saxton, 1987; Meinecke & Eriksson, 2016). Thus the accuracy of MLM compared to continuum particle-based methods requires detailed examination. Notably, a consistent approach is needed to determine MLM parameters such as voxel size and reaction probability that can replicate the reaction kinetics in the continuum.

This research focuses on the study of reaction-diffusion processes in complex biological system with the following objectives: (i) to determine the accuracy and consistency of microscopic lattice method as a reaction-diffusion simulation method through theoretical and numerical analysis; (ii) to derive the on-lattice reaction rate coefficient based on the Smoluchowski-Collins-Kimball model for reaction-diffusion in 1D, 2D, and 3D spatial dimensions; and (iii) to benchmark the performance and accuracy of Spatiocyte simulation scheme before applying it to selected reaction-diffusion processes (Arjunan & Tomita, 2010).

1.3 Thesis Organization

This thesis consists of six chapters. In Chapter 2, we address the need for spatial stochastic modeling approach to study the RD process in the cell. The continuum-based and lattice-based spatial stochastic simulation methods are reviewed. We introduce the

Smoluchowski-Collins-Kimball theory for the bimolecular reaction in different spatial dimensions and the important concepts in random walk theory. We also discuss the concept of anomalous diffusion and its occurrence in the cell.

Chapter 3 gives the description of the Spatiocyte scheme, which is the basis for constructing the MLM theoretical framework. The main contribution of this thesis is the lattice theory derived for MLM, which appears in the form of rebinding probability and reaction rate coefficient. The lattice theory are derived separately for the activation-limited and diffusion-influenced cases in 1D, 2D and 3D spatial dimensions. From the lattice theory, we compare and match the MLM and continuum reaction-diffusion behaviors, which allow us to determine the MLM parameters for accurate simulation.

In Chapter 4, we validate the accuracy of lattice theory with numerical simulation and continuum-based theory. The accuracy of Spatiocyte in simulating irreversible and reversible diffusion-influenced reaction is verified in 1D, 2D and 3D. The performance of MLM in comparison to other well-known particle-based methods such as Smoldyn and eGFRD is shown.

In Chapter 5, we apply MLM to study several biological reaction-diffusion process that include the production-degradation process and the dual phosphorylation-dephosphorylation cycle. We show how diffusion influences the steady-state protein concentration and the macroscopic response behavior of the MAPK cascade. We also study the reaction-diffusion behavior in the presence of obstacles and discuss the validity of Smoluchowski-Collins-Kimball model in such situation. Finally, we investigate the contribution of 2D reaction pathway in surface reaction with typical intracellular condition.

In Chapter 6, we summarize this thesis and discuss the current limitations of MLM with suggestions for future work.

Novel contributions from this thesis are published in (Chew et al., 2018) for the 3D reaction and in (Chew et al., 2019) for the 2D and 1D reactions. Some texts and figures in this thesis are reproduced and adapted from these publications, with the permission of AIP Publishing. Source and implementation codes used in this thesis can be found in the online repository mentioned in the relevant chapters.

CHAPTER 2: LITERATURE REVIEW AND BASIC CONCEPTS

We begin with the introduction of spatial stochastic modeling approach for intracellular reaction-diffusion (RD) process and review the associated simulation methods. The simulation methods are described based on their spatial representation, either in the continuum or on lattice space. The MLM is highlighted and compared with other lattice-based methods. Then, we introduce the bimolecular reaction theory based on the well-known Smoluchowski-Collins-Kimball (SCK) model. Detailed expression of rate coefficient according to the SCK model are given for 3D, 2D and 1D reaction. Useful expressions of the continuum-based theory such as the survival probability are also defined. We then introduce the basic terminologies and concepts of random walk theory required to derive the lattice reaction theory. Finally, we introduce anomalous transport and its origin in intracellular environment.

2.1 Spatial Stochastic Modeling Approach

Macroscopic modeling approaches that apply the rate (Edelstein-Keshet, 1988) or the reaction-diffusion equation (Edelstein-Keshet, 1988; Murray, 2001) has been successful in modeling many biochemical processes, when the molecules copies are abundant and are well-mixed in the compartment. However, these approaches have several limitations. For example, they cannot explicitly capture the effects of space at the microscopic scale that arise from the interactions of finite-sized molecules (Almeida & Vaz, 1995; Melo & Martins, 2006; Guigas & Weiss, 2016), the fast rebinding of reactants (Lagerholm & Thompson, 1998; Takahashi et al., 2010; Loose et al., 2011; Mugler et al., 2012) and the inhomogeneity in the medium such as lipid domains and membrane-associated cytoskeletal structures (Almeida & Vaz, 1995; Kalay et al., 2012; Kalay, 2012; Zhdanov & Höök, 2015; Kerketta et al., 2016).

The spatial effects can alter the diffusion behavior and the reaction kinetics, leading to different physiological outcomes. For example, clustering of membrane receptors changes the response of signaling network (Mugler et al., 2012); fluctuation in protein copy number promotes cell polarization in the absence of spatial cue (Pablo et al., 2018); rapid protein rebinding affects spatiotemporal patterns (Loose et al., 2011) and amplifies noise during ligand interactions (Kaizu et al., 2014). Therefore, a realistic RD model should consider the spatial and stochastic aspects of the RD process (Howard & Rutenberg, 2003; Klann & Koeppl, 2012; Engblom et al., 2017; Li & Holmes, 2018; Earnest et al., 2018; Smith & Grima, 2018).

2.2 Continuum-based and Lattice-based Simulation Methods

Generally, in spatial stochastic simulation methods, space is represented by the continuum or as a lattice. In the continuum particle-based method, each molecule is represented as a point or a hard-body sphere that propagates in continuous space with Gaussian distributed displacement (Stiles et al., 1998; Plimpton & Slepoy, 2005; van Zon & Wolde, 2005; Opplestrup et al., 2006; Sanford et al., 2006; Ridgway et al., 2008; Byrne et al., 2010; Gruenert et al., 2010; Takahashi et al., 2010; Klann et al., 2012; Schöneberg & Noé, 2013; Andrews, 2016; Michalski & Loew, 2016; Lehnert & Figge, 2017; Bittig & Uhrmacher, 2017; Bashardanesh & Lötstedt, 2017; Donev et al., 2018; Sayyidmousavi & Rohlf, 2018). These molecules react with each other via bimolecular reactions according to the Smoluchowski-Collins-Kimball (SCK) model (Smoluchowski, 1917; Collins & Kimball, 1949) or the Doi physical model (Doi, 1976; Erban & Chapman, 2009; Agbanusi & Isaacson, 2014). In the SCK model, the reaction occurs either immediately or with a probability of reflection when the distance between reactants equals to a predefined reaction radius, whereas in the Doi model, the reaction occurs with a fixed probability per unit time when the reactants are closer than the radius. Continuum-based SCK methods that support

reactions on volume and surface compartments include Smoldyn (Andrews, 2009, 2016), CDS (Byrne et al., 2010) and eGFRD (Sokolowski et al., 2019), while the Doi model is adopted by MCell (Kerr et al., 2008), ReaDDy (Schöneberg & Noé, 2013; Hoffmann et al., 2019) and SpringSaLaD (Michalski & Loew, 2016). Smoldyn also recently included the option to support the Doi model. All of these methods except MCell can simulate volume occupying molecules.

On the other hand, in lattice approaches, the average diffusion behavior is adopted and the reactions follow either the simple first-order process, or the second-order process when two reactive molecules meet on the same lattice voxel. Such approaches reduce the computational cost even in crowded space and provide an efficient way to simulate large numbers of molecules and reactions. Within lattice approaches, variation exists depending on how each molecule is represented and reaction is modeled. In the well-established reaction-diffusion master equation (RDME) models (Baras & Mansour, 1996; Hattne et al., 2005; Drawert et al., 2012; Cowan et al., 2012; Hepburn et al., 2012; Roberts et al., 2013; Drawert et al., 2016), space is discretized into lattice voxels called subvolumes. In each subvolume, point-like molecules are assumed to be dilute and well-mixed. To obey the well-mixed condition, there is a limit to the size of the subvolume (Hellander et al., 2012; Hattne et al., 2005; Hellander et al., 2015), which in turn imposes a limit to the spatial resolution. Diffusion of molecules across subvolumes is modeled as a first-order reaction with a concentration dependent rate. Unimolecular and bimolecular reactions only occur within each subvolume with a rate defined by the propensity function (Gillespie, 1976). Compared to continuum particle-based schemes, RDME models RD from the mesoscopic to the macroscopic scale but not at the microscopic scale. However, there have been several efforts to overcome the well-mixed limit in RDME models and to bridge mesoscopic and microscopic scales (Fange et al., 2010; Hellander et al., 2012, 2015; Isaacson, 2013;

Gillespie et al., 2013).

Apart from the RDME lattice models, there is another class of schemes, which is the MLM that represents molecules at single particle resolution (Torney & McConnell, 1983; Montalenti & Ferrando, 2000; Saxton, 2002; Berry, 2002; Schnell & Turner, 2004; Grima & Schnell, 2006; Boulianne et al., 2008; Schmit et al., 2009; Arjunan & Tomita, 2010; Gillespie et al., 2014; Pitulice et al., 2014; Gomez & Klumpp, 2015). In most of these schemes (Torney & McConnell, 1983; Saxton, 2002; Berry, 2002; Schnell & Turner, 2004; Grima & Schnell, 2006; Boulianne et al., 2008; Arjunan & Tomita, 2010; Gomez & Klumpp, 2015), the size of the voxel follows the molecule size, whereas in the small-voxel tracking algorithm (SVTA) (Gillespie et al., 2014), a particle can occupy multiple voxels, providing greater spatial resolution at the cost of higher computational complexity. In MLM, a molecule hops into a neighbor voxel at a constant rate such that normal diffusion is satisfied. Excluded volume arises naturally since the size of molecule is directly reflected by the voxel size and occupancy in the lattice. Similar to RDME models, unimolecular reaction is modeled as a first-order process. Bimolecular reactions are coupled to molecular collisions in all of these schemes except GridCell (Boulianne et al., 2008). In the collision-based reaction schemes, the steady-state reaction rate follows the macroscopic effective reaction rate when the reaction is activation-limited.

Schemes based on the MLM attempt to reduce the cost of resolving molecular collisions by discretizing the space into fine molecule-sized voxels. In the Spatiocyte simulation scheme, a molecule only checks its destination voxel for occupancy before performing a bimolecular reaction with the occupying molecule or moving into it if it is vacant (Arjunan & Tomita, 2010). The reduced computational cost and the simplicity of MLM implementation have promoted its applications, for example, in biological (Berry, 2002; Schnell & Turner, 2004; Saxton, 2008; Arjunan & Tomita, 2010; Tsourkas & Raychaudhuri,

2011; Shimo et al., 2015) and chemical (Lombardo & Bell, 1991; Imbihl & Ertl, 1995; Lukkien et al., 1998) surface reactions.

2.3 Smoluchowski-Collins-Kimball Theory for Bimolecular Reaction

In this section, we review bimolecular reaction model based on the Smoluchowski-Collins-Kimball (SCK) theory. Consider a many-body bimolecular reaction,

$$A + B \rightarrow B, \tag{2.1}$$

with A and B having radii r_A and r_B, and diffusion coefficients D_A and D_B, respectively. Let us denote the fraction of A remaining in the system by the survival probability, $S_{irr,A}(t) = [A(t)]/[A(0)]$, where [ ] denotes the concentration. The survival probability of A is provided in the rate equation (Szabo, 1989):

$$\frac{dS_{irr,A}(t)}{dt} = -[B]k(t)S_{irr,A}(t), \tag{2.2}$$

where $k(t)$ represents the time-dependent rate coefficient. The solution for the survival probability requires the integration of the rate coefficient (Szabo, 1989):

$$S_{irr,A}(t,[B]) = \exp\left[-[B]\int_0^t k(\tau)d\tau\right]. \tag{2.3}$$

Smoluchowski derived the rate coefficient by incorporating molecular diffusion into the reaction flux (Smoluchowski, 1917). In Smoluchowski model, A is made up of an immobile molecule and is surrounded by multiple diffusing B molecules in which $[B] \gg [A]$. The reaction is modeled by absorbing boundary condition, that is, A and B will react once the distance between the pair is $R = r_A + r_B$. While the Smoluchowski model assumption of perfect adsorption can describe the diffusion-limited reaction. Collins and Kimball (Collins

& Kimball, 1949) extended the Smoluchowski theory to model both activation-limited and diffusion-influenced reaction by assuming partial adsorption. This is achieved by using a radiation boundary condition, in which a reaction will occur with an intrinsic rate k_a whenever the distance between a reactant pair is $R = r_A + r_B$. In this thesis, the term, Smoluchowski-Collins-Kimball (SCK) model, refer to the model with the radiation boundary condition.

According to the particle-pair formalism of SCK model (Noyes, 1954; Berg, 1978; Naqvi et al., 1980; Tachiya, 1983), the time-dependent rate coefficient can be expressed as the product of k_a and $S(t;R)$:

$$k(t) = k_a S(t;R). \tag{2.4}$$

Here, $S(t;R)$ is the survival probability of an isolated reactant pair at time t given that they were initially in contact. Additionally, let $p_{reb}(R,t|R,0)$ denote the rebinding-time probability distribution for a reactive particle-pair separated by distance R at time t, given that the pair were initially in contact.

Note that the survival probability $S(t;R)$ is the same as the probability that the first rebinding event between an initially in-contact pair has not yet occurred at time t. Hence we can rewrite Equation (2.4) as

$$k(t) = k_a \left[1 - \int_0^t p_{reb}(R,\tau|R,0) d\tau \right]. \tag{2.5}$$

The specific functional form of the rate coefficient depends on the spatial dimension and will be treated separately in the following subsections.

2.3.1 3D Reaction

In this section, we describe the rebinding probability distribution of a reactant pair and the time-dependent rate coefficient of an irreversible reaction in 3D space.

2.3.1.1 Rebinding Probability Distribution

Let k_{a3D} be the 3D intrinsic rate constant with dimensions $\mathrm{L}^3\mathrm{T}^{-1}$ and $p(r,t;r_0,0)$ be the probability that the particle is at position r at time t given that it was at r_0 at time $t_0 = 0$. The rebinding probability distribution for a reactive particle-pair separated by distance R at time t, given that the pair were initially in contact is defined as (Equation (3.10) in (Naqvi et al., 1980) and Equation (S27) in (Takahashi et al., 2010))

$$p_{reb}(R,t;R,0) = k_{a3D}p(R,t;R,0), \tag{2.6}$$

where $p(r,t;r_0,0)$ is the Green's function in the diffusion equation:

$$\frac{\partial p(r,t;r_0,0)}{\partial t} = D\nabla^2 p(r,t;r_0,0), \tag{2.7}$$

subjected to initial condition $p(r,0) = \delta(r-r_0)/4\pi r^2$ and boundary conditions such that $p(r,t) \to 0$ as $r \to \infty$ and

$$4\pi R^2 D \left.\frac{\partial p(r,t;r_0,0)}{\partial r}\right|_{r=R} = k_{a3D}p(R,t;r_0,0). \tag{2.8}$$

The latter condition is known as the radiation boundary condition. The Green's function $p(r,t;r_0,0)$ has been solved in (pp. 368 in (Carslaw & Jaeger, 1959)) to be

$$\begin{aligned} p(r,t;r_0,0) = \frac{1}{8\pi r r_0}\frac{1}{\sqrt{\pi Dt}}\Big[&\exp\left[-(r-r_0)^2/4Dt\right] + \exp\left[-(r+r_0-2R)^2/4Dt\right] \\ &-2B\sqrt{\pi Dt}\exp\left[B^2Dt + B(r+r_0-2R)\right]\mathrm{erfc}\left(B\sqrt{Dt}\right)\Big], \end{aligned} \tag{2.9}$$

where $B = (1 + k_{a3D}/4\pi RD)/R$. For $r = r_0 = R$, we thus have

$$p(R,t;R,0) = \frac{1}{4\pi R^2}\frac{1}{\sqrt{\pi Dt}}\left[1 - B\sqrt{\pi Dt}\exp\left(B^2Dt\right)\text{erfc}\left(B\sqrt{Dt}\right)\right]. \tag{2.10}$$

Finally by substituting Equation (2.10) into Equation (2.6), we obtain the rebinding probability distribution

$$p_{reb}(R,t;R,0) = \left(\frac{k_{a3D}}{4\pi R^3}\right)\left(\frac{k_{a3D}}{4\pi RD} + 1\right)\left(\frac{1}{\sqrt{\pi\tau}} - \exp(\tau)\text{erfc}(\sqrt{\tau})\right), \tag{2.11}$$

where $\tau = tD(1 + k_{a3D}/4\pi RD)^2/R^2$.

2.3.1.2 Time-dependent Rate Coefficient

The time-dependent rate coefficient in 3D space is defined as (Collins & Kimball, 1949):

$$k_{3D}(t) = \frac{k_D k_{a3D}}{k_D + k_{a3D}}\left[1 + \frac{k_{a3D}}{k_D}\Phi\left(\frac{k_{a3D}}{k_D}\sqrt{\frac{t}{\tau'}}\right)\right]. \tag{2.12}$$

Here, $k_D = 4\pi RD$ is the collision rate, $D = D_A + D_B$, $\Phi(x) = \exp(x^2)\text{erfc}(x)$ and $\tau' = (1/D)(k_{a3D}R/(k_{a3D} + k_D))^2$. The rate coefficient in Equation (2.12) starts ($t = 0$) at k_{a3D} but decays rapidly to (Agmon & Szabo, 1990)

$$k_{3D}(t) \simeq \frac{k_D k_{a3D}}{k_D + k_{a3D}}\left[1 + \frac{k_{a3D}}{k_{a3D} + k_D}\frac{R}{\sqrt{\pi Dt}}\right], \tag{2.13}$$

at long-time. Following Equation (2.13), the steady-state or the effective reaction rate constant k_{eff3D} is defined as (Noyes, 1961):

$$k_{eff3D} \coloneqq k_{3D}(t \to \infty) = \frac{k_{a3D}k_D}{k_{a3D} + k_D}. \tag{2.14}$$

Alternatively, by taking the long-time limit of Equation (2.5), we have

$$k_{eff3D} \coloneqq k_{3D}(t \to \infty) = k_{a3D}\left[1 - \int_0^\infty p_{reb}(R, \tau|R, 0)d\tau\right], \tag{2.15}$$

where the integrated term gives the total rebinding probability:

$$P_{reb} = \int_0^\infty p_{reb}(R, t|R, 0)dt = \frac{1}{1 + \frac{k_D}{k_{a3D}}}. \tag{2.16}$$

Therefore, the effective rate constant in Equation (2.14) can also be written in terms of the total rebinding probability:

$$k_{eff3D} = k_{a3D}(1 - P_{reb}). \tag{2.17}$$

The above relation was also described previously, but in the context of irreversible and reversible rate constants (Andrews & Bray, 2004).

2.3.2 2D Reaction

In this section, we review the time-dependent rate coefficient for both the irreversible and reversible reaction in 2D space.

2.3.2.1 Irreversible Reaction

The time-dependent rate coefficient for 2D association reaction with radiation boundary condition is given in the Laplace form as (Szabo, 1989):

$$k_{2D}(s) = \frac{k_{a2D}}{s[1 + k_{a2D}\widetilde{g}(s)]}. \tag{2.18}$$

Here, k_{a2D} is the intrinsic rate constant with dimensions of length, L and time, T, given as L^2T^{-1}, and $\widetilde{g}(s)$ is the Green's function for 2D diffusion defined as (Popov & Agmon, 2002)

$$\widetilde{g}(s) = \frac{K_0(\tilde{s})}{2\pi D\tilde{s}K_1(\tilde{s})}. \tag{2.19}$$

K_0 and K_1 are the modified Bessel functions of the second kind, $\tilde{s} = \sqrt{sR^2/D}$, and $D = D_A + D_B$. Equation (2.18) can thus be written as

$$k_{2D}(s) = 2\pi D\tilde{s}\frac{K_1(\tilde{s})}{s[K_0(\tilde{s}) + 2\pi\tilde{s}K_1(\tilde{s})/\kappa]}, \tag{2.20}$$

with $\kappa = k_{a2D}/D$. In the limit of small $\tilde{s}$, we can approximate the modified Bessel functions:

$$\tilde{s}K_1(\tilde{s}) \approx 1 - \frac{2\ln(\tilde{s}e^{\gamma}/2) + 1}{4}\tilde{s}^2 + O(\tilde{s}^4), \tag{2.21}$$

and

$$K_0(\tilde{s}) \approx -\ln(\tilde{s}e^{\gamma}/2) - \frac{2\ln(\tilde{s}e^{\gamma}/2) + 1}{4}\tilde{s}^2 + O(\tilde{s}^4), \tag{2.22}$$

where $\gamma = 0.5772156$ is the Euler constant. Using these approximations, the asymptotic expansion of Equation (2.20) can be expressed as:

$$\begin{aligned}\lim_{s\to 0} sk_{2D}(s) &= \frac{4\pi D}{-2\ln[R\sqrt{s/D}] + \ln 4 + \ln[\exp 2(4\pi/\kappa - \gamma)]} + O(s) \\ &= \frac{4\pi D}{\ln[C_c/s]} + O(s)\end{aligned} \tag{2.23}$$

with $C_c = 4D\exp(4\pi/\kappa - 2\gamma)/R^2$. The corresponding long-time approximation is given as (Ritchie & Sakakura, 1956):

$$k_{2D}(t \gg \frac{R^2}{D}) = 4\pi D(\frac{1}{\ln(C_c t)} - \frac{\gamma}{(\ln(C_c t))^2} - \frac{1.311}{(\ln(C_c t))^3} + \frac{0.25}{(\ln(C_c t))^4} + O(t^{-1}\ln(t)^{-2})), \tag{2.24}$$

where the relative error to the exact form is less than 1% at $t = 100R^2/D$.

The slow logarithmic decay of $k_{2D}(t)$ originates from the modified Bessel function and is fundamentally connected to the recurrence property of the random walk in 2D. A random walker in 3D has an asymptotically finite chance to escape from the origin. However, in the lower dimensions (1D and 2D), the return probability is unity, as described by the Polya's theorem. In the case of 2D, the first passage time of returning to the origin exhibits slow logarithmic decay (Weiss & Rubin, 1982), which in turn causes the slow decay in $k_{2D}(t)$. As a result, unlike in the 3D reaction, the long-time rate coefficient in 2D does not converge to a constant term. This behavior have been observed in experimental studies previously (Melo & Martins, 2006). The lack of an apparent steady-state and the slow decay in the 2D reaction rate have prompted the use of time-dependent rate for the purposes of modeling and analysis of experiments (Naqvi et al., 2000; Melo & Martins, 2006).

(a) ***Steady-state rate constant***

Szabo (Szabo, 1989) provided an alternative way to define a steady-state rate constant by coupling the irreversible bimolecular reaction with the production of A species. Following this strategy, a steady-state expression can be defined in terms of the mean lifetime of A,

τ_m (Szabo, 1989):

$$\begin{aligned} k_{ss2D} &= \frac{1}{[B]\tau_m} \\ &= \left[[B]\int_0^\infty S_{irr,A}^{2D}(t)dt\right]^{-1} \\ &= \left[[B]\int_0^\infty \exp\left(-[B]\int_0^t k_{2D}(\tau)d\tau\right)dt\right]^{-1}. \end{aligned} \tag{2.25}$$

k_{ss2D} can be evaluated using the mean-field self-consistency condition (Szabo, 1989):

$$k_{ss2D} = [s\hat{k}_{2D}(s)]_{s=k_{ss2D}[B]}. \tag{2.26}$$

Substituting the asymptotic form of $k_{2D}(s)$, as defined in Equation (2.23), into Equation (2.26) yields

$$\frac{k_{ss2D}}{2\pi D} \approx \left[\ln 2 - \gamma - \ln[R\sqrt{k_{ss2D}[B]/D}] + 1/\kappa\right]^{-1}. \tag{2.27}$$

Rewriting some variables in terms of the molecule area fraction, $\phi = \pi R^2[B]$ and taking the small concentration limit, $\phi \to 0$ gives the following approximation

$$\begin{aligned} \frac{k_{ss2D}}{2\pi D} &\approx \left[\ln 2 - \gamma - \ln[\sqrt{2\phi y}] + 1/\kappa\right]^{-1} \\ &\approx \left[\frac{1}{2}\ln 2 - \gamma - \frac{1}{2}\ln\phi + 1/\kappa\right]^{-1}. \end{aligned} \tag{2.28}$$

Finally, from Equation (2.28), we obtain the steady-state rate constant for radiation boundary condition as

$$k_{ss2D} \approx \frac{4\pi D}{\ln 2 - 2\gamma - \ln\phi + 4\pi/\kappa}. \tag{2.29}$$

Similar to the 3D effective rate constant in $1/k_{ss3D} = 1/k_{a2D} + 1/(4\pi RD)$, the 2D steady-state rate constant depends on the intrinsic rate k_{a2D} and the relative diffusion coefficient D. Interestingly, the 2D rate constant has the additional dependency on the concentration of B. The steady-state constant given in Equation(2.29) is an extension of the solution based on the absorbing boundary condition provided in Equation (4.10) of (Szabo, 1989).

2.3.2.2 Reversible Reaction

In the SCK model for 2D reversible reaction

$$A + B \underset{k_{d2D}}{\overset{k_{a2D}}{\rightleftharpoons}} AB, \tag{2.30}$$

a bound pair A-B dissociates with the rate constant k_{d2D} (T^{-1}) into A and B, separated at distance R. The survival probability of A, defined as $S_{rev,A}(t)$, can be calculated using the first variant of the multiparticle kernel theory (MPK1) (Sung & Lee, 1999; Popov & Agmon, 2001). Although the closed form solution for $S_{rev,A}(t)$ in 2D is not available, it can be evaluated by numerically solving the normalized deviation defined as

$$\frac{S_{rev,A}(t) - S_{rev,A}(\infty)}{S_{rev,A}(0) - S_{rev,A}(\infty)} = \mathcal{L}^{-1}\left[\frac{1}{s + \lambda\widetilde{F}(s)}\right]. \tag{2.31}$$

Here, the term

$$\widetilde{F}(s) = \frac{k_{d2D}}{\lambda}\hat{F}_{gem}(s) + \frac{[B]k_{a2D}}{\lambda}\hat{F}_{irr}(s;[B]') \tag{2.32}$$

is the diffusion factor function, $S_{rev,A}(\infty) = k_{d2D}/\lambda = 1/(1 + [B]k_{a2D}/k_{d2D})$ is the equilibrium value, $\lambda = k_{d2D} + [B]k_{a2D}$ is the chemical kinetics relaxation rate constant, and $[B]' = \lambda/k_{a2D}$ is the modified concentration. $\hat{F}_{gem}(s) = 1 + k_{a2D}\hat{g}(s)$ contains the 2D

Green's function term $\hat{g}(s)$ as given in Equation (2.19), whereas the function

$$\hat{F}_{irr}(s;c_0) = \frac{c_0 k_{a2D} \hat{S}_{irr,A}(s;c_0)}{1 - s\hat{S}_{irr,A}(s;c_0)}, \tag{2.33}$$

uses the term $\hat{S}_{irr,A}(s;c_0)$, which is the Laplace transform of the irreversible reaction survival probability, $S^{2D}_{irr,A}(t;c_0)$.

2.3.3 1D Volume-surface Adsorption

Before describing the rate for volume-surface adsorption, we first consider the 1D SCK model, where a single immobile species B interacts with multiple mobile species A on a filament according to Equation (2.1). Assuming species A can collide with species B from both sides of B, and there is no self-interaction among A. The time-dependent rate coefficient of this reaction with radiation boundary condition is given as (Szabo, 1989)

$$k_{1D}(t) = k_{a1D} \exp(\kappa^2 Dt/4)\mathrm{erfc}(\kappa\sqrt{Dt/4}), \tag{2.34}$$

with $\kappa = k_{a1D}/D$ denoting the ratio between the intrinsic adsorption rate constant of unit LT^{-1} and the relative diffusion coefficient D. At long time, Equation (2.34) behaves asymptotically as

$$k_{1D}(t \to \infty) \approx 2\sqrt{\frac{D}{\pi t}} + O(t^{-3/2}). \tag{2.35}$$

Next we consider a volume-surface adsorption system that involves an adsorbing plane at $x = 0$ and bulk molecules at $x > 0$. Initially, the molecules of concentration c_0 are distributed uniformly in the bulk and are absent on the surface. For surface adsorption process that follows the radiation boundary condition, the number of molecules adsorbed

to the surface varies as (Equation (3.37) of (Cranket al. , 1979)):

$$N_s(t) = \frac{c_0 S}{\kappa}\left\{\exp(\kappa^2 Dt)\mathrm{erfc}(\kappa\sqrt{Dt}) - 1 + 2\kappa\sqrt{Dt/\pi}\right\}, \tag{2.36}$$

where S is the area of the plane. The corresponding adsorption rate is well-described by the time-dependent adsorption rate coefficient $k_{sa}(t')$:

$$\frac{dN_s(t')}{dt} = k_{sa}(t')c_0 S. \tag{2.37}$$

Note that the adsorption rate coefficient differs from the 1D SCK rate by a factor of two: $k_{sa}(t') = k_{1D}(t)/2$, because in the latter, it occurs on both sides of the plane. At long time, the adsorption rate coefficient becomes

$$k_{sa}(t \to \infty) \approx \sqrt{\frac{D}{\pi t}} + O(t^{-3/2}). \tag{2.38}$$

As the bulk molecules are adsorbed to the surface, a spatial concentration gradient develops in the bulk. The spatial-temporal concentration profile of the bulk molecules $C(x,t)$ follows Equation (3.35) of (Cranket al. , 1979):

$$C(x,t) = c_0\left[\mathrm{erfc}\frac{x}{2\sqrt{Dt}} - \exp\left(\frac{k_{sa}x}{D} + \frac{k_{sa}^2 t}{D}\right)\mathrm{erfc}\left(\frac{x}{2\sqrt{Dt}} + k_{sa}\sqrt{\frac{t}{D}}\right)\right]. \tag{2.39}$$

When adsorbed molecules can dissociate from the surface with a rate k_{sd} (T^{-1}), their number varies according to (Equation (A.12) in (Andrews, 2009))

$$N_s(t) = \frac{c_0 S k_{sa} t\left[c_1\exp(c_2^2)\mathrm{erfc}(c_2) - c_1 + c_2\exp(c_1^2)\mathrm{erfc}(c_1) - c_2\right]}{c_1 c_2(c_2 - c_1)}, \tag{2.40}$$

where

$$c_1 = \frac{k_{sa} - \sqrt{k_{sa}^2 - 4Dk_{sd}}}{2\sqrt{D}}\sqrt{t}, \tag{2.41}$$

and

$$c_2 = \frac{k_{sa} + \sqrt{k_{sa}^2 - 4Dk_{sd}}}{2\sqrt{D}}\sqrt{t}. \tag{2.42}$$

2.4 Random Walk Theory

This section introduces some terminology and definitions from random walk theory. Useful expressions and theirs properties shown here will be used in next chapter.

2.4.1 Terminology

We consider random walk on infinite lattice where the walker hops to nearest neighbor voxels at discrete time intervals. For $n \in \mathbb{N}$, we define $P_n(s_a|s_b)$ as the voxel occupation probability from s_b to s_a, that is, the probability of being at voxel s_a after n steps, given that the walk started at voxel s_b. $F_n(s_a|s_b)$ is the first-passage time distribution from s_b to s_a, that is the probability of arriving at s_b for the first time on the nth step, given that the walk started at site s_a. We denote s_0 as the origin voxel, s_1 as the element of the set of immediate neighboring voxels of s_0, and s_2 as the element of the set of the second nearest neighbor voxels of s_0.

Properties of the voxel occupation probability and the first-passage time distribution are well described through the probability generating functions defined as

$$P(s_a|s_b; z) = \sum_{n=0}^{\infty} P_n(s_a|s_b)z^n, \tag{2.43}$$

and

$$F(s_a|s_b;z) = \sum_{n=0}^{\infty} F_n(s_a|s_b)z^n, \tag{2.44}$$

respectively. The former function, also known as the Lattice Green's function (Hughes, 1995), relates to latter via the relation (Equation (I.18) in (Montroll & Weiss, 1965)):

$$F(s_0|s_0;z) = 1 - \frac{1}{P(s_0|s_0;z)}. \tag{2.45}$$

Another useful quantity is the probability to reach the origin for the jth time at the nth step defined as (I.1.9 in (Montroll & Weiss, 1965)),

$$F_n^j(s_0|s_0) = \sum_{i=1}^{n} F_{n-i}^{j-1}(s_0|s_0)F_i(s_0|s_0), \text{ for } j \in \mathbb{Z}^+, \tag{2.46}$$

where $F_n^1(s_0|s_0) = F_n(s_0|s_0)$. The generating function of $F_n^j(s_0|s_0)$ satisfies the relation (Equation (I.20) in (Montroll & Weiss, 1965)):

$$\sum_{n=0}^{\infty} F_n^j(s_0|s_0)\, z^n = F(s_0|s_0;z)^j, \tag{2.47}$$

2.4.2 Lattice Green's Function

2.4.2.1 FCC Lattice

In our MLM scheme, we use hexagonal close-packed (HCP) lattice for 3D diffusion. Here, we introduce some useful random walk properties on HCP lattice which will be used latter on. Topologically speaking, the random walk properties on HCP lattice is the same as on the face-centered cubic (FCC). The lattice Green's function for the FCC lattice is

given as (Equations(2.6)-(2.9) of (Joyce, 1998))

$$P(s_0|s_0;z) = \left[\frac{2(1+3\xi^2)}{\pi(1-\xi)(1+3\xi)}\right]^2 \mathbf{K}(k_+)\mathbf{K}(k_-), \tag{2.48}$$

$$k_+^2 = \frac{16\xi}{(1-\xi)(1+3\xi)^3}, \tag{2.49}$$

$$k_-^2 = \frac{16\xi^3}{(1-\xi)^3(1+3\xi)}, \tag{2.50}$$

$$\xi = \frac{-1+\sqrt{1+z/3}}{1+\sqrt{1-z}}, \tag{2.51}$$

wherein $\mathbf{K}(k)$ is the complete elliptic integral of the first kind. The corresponding voxel occupation probability is given as (Joyce, 1998)

$$P_n(s_0|s_0) = \frac{1}{12^n}\sum_{j=0}^{n}\binom{n}{j}(-4)^{n-j}b_j, \text{ for } j \in \mathbb{N} \tag{2.52}$$

where

$$b_j = \sum_{k=0}^{j}\binom{j}{k}^2\binom{2k}{k}\binom{2j-2k}{j-k}. \tag{2.53}$$

The first-passage time distribution is related to the voxel occupation probability recursively via:

$$F_n(s_0|s_0) = P_n(s_0|s_0) - \sum_{j=1}^{n-1} P_{n-j}(s_0|s_0)F_j(s_0|s_0), \text{ for } j \in \mathbb{Z}^+. \tag{2.54}$$

2.4.2.2 Triangular Lattice

For 2D MLM, we focus on the triangular lattice where some useful random walk quantities will be introduced. The voxel occupancy probability $P_n(s_0|s_0)$ for the triangular lattice is given as (Guttmann, 2010; Hughes, 1995):

$$P_n(s_0|s_0) = \frac{1}{6^n} \sum_{j=0}^{n} \binom{n}{j} (-3)^{n-j} b_j, \tag{2.55}$$

where

$$b_j = \sum_{k=0}^{j} \binom{j}{k}^2 \binom{2k}{k}. \tag{2.56}$$

The corresponding probability generating function is

$$P(s_0|s_0; z) = \frac{6}{\pi z\sqrt{c}} \mathbf{K}(k') \tag{2.57}$$

where $c = (a+1)(b-1)$,

$$a = \frac{3}{z} + 1 - \sqrt{3 + \frac{6}{z}} \qquad \text{and} \qquad b = \frac{3}{z} + 1 + \sqrt{3 + \frac{6}{z}}, \tag{2.58}$$

and $\mathbf{K}(k')$ is the complete elliptic integral of the first kind with

$$k' = \sqrt{\frac{2(b-a)}{c}}\,. \tag{2.59}$$

The asymptotic expansion of $P(s_0|s_0;z)$ in terms of the asymptotic form for $\mathbf{K}(z)$ is derived as (see Equation (A.198) in (Hughes, 1995)):

$$P(s_0|s_0;z) \approx \frac{\sqrt{3}}{2\pi}\ln[12(1-z)^{-1}]\{1+O(1-z)\}. \tag{2.60}$$

2.4.2.3 One-dimensional Lattice

In 1D lattice, the generating function for the voxel occupancy probability from origin to origin is defined as (Montroll & Weiss, 1965)

$$P(s_0|s_0;z) = \frac{1}{\sqrt{1-z^2}}. \tag{2.61}$$

The corresponding first passage time distribution, obtained through Equation (2.45), is given as

$$F(s_0|s_0;z) = 1-\sqrt{1-z^2}. \tag{2.62}$$

2.5 Anomalous Diffusion

A molecule undergoing normal diffusion is characterized by the linear time-dependency in the mean-square-displacement as described by the Einstein-Smoluchowski equation $\langle r^2(t)\rangle = (2d)D_x t$, where d is the spatial dimension and D_x is the diffusion coefficient. As the intracellular medium is highly crowded with obstacles such as macromolecules and cytoskeletal elements, the diffusion of macromolecules is highly obstructed. This could

leads to anomalous diffusion where the mean-square-displacement scale with time as

$$\langle r^2(t)\rangle \propto t^{\alpha}, \tag{2.63}$$

with $0 < \alpha < 1$ (Saxton, 2002; Vilaseca et al., 2011) . A generalized diffusion coefficient function, $D_{\phi}(t)$, that depend on time and excluded volume ϕ, can describe the influence of crowding on diffusion (Vilaseca et al., 2011). The mean-squared-displacement relationship Equation (2.64) then follows

$$\langle r^2(t)\rangle = (2d)D_{\phi}(t)t^{\alpha}. \tag{2.64}$$

In experimental studies, diffusion behavior inside a cell varies according to the observation time window whereas the exponent α is system specific (Höfling & Franosch, 2013). This behavior can be understood through the theory of obstructed diffusion (Stauffer & Aharony, 2014). For a macromolecule that diffuses in an uniformly dense medium, diffusion is normal when the diffusion distance is smaller than the correlation length. As the diffusion distance becomes comparable to the correlation length, diffusion becomes anomalous. Diffusion becomes normal again when the distance is much larger than the correlation length. As obstacle density approaches the percolation threshold, the transition time from anomalous to normal diffusion diverges: diffusion is anomalous for all time. This is because of the absence of a characteristic length scale in fractal medium (Stauffer & Aharony, 2014). Reaction kinetics in such fractal medium has been shown to exhibit non-classical kinetics (Kopelman, 1988; Berry, 2002; Schnell & Turner, 2004; Haugh, 2009).

2.6 Summary

We have emphasized the necessity of having a spatial stochastic modeling approach to capture the intracellular RD process. Several continuum and lattice-based spatial stochastic simulation methods including the RDME, eGFRD, Smoldyn, and MLM are introduced. We highlighted MLM as being more computational efficient than conventional continuum-based method. In the subsequent chapters, we will compare the accuracy and performance of MLM with the other simulation schemes. We have introduced the SCK model for diffusion-influenced reaction in all spatial dimensions. SCK rate coefficients defined in this chapter will be referred again for comparison with the lattice theory. Important definitions and concepts from random walk theory such as the first passage time are introduced. The SCK model and random walk theory serve as the basis for deriving the lattice rate coefficient in the next chapter. We also described the anomalous transport behavior due to obstacle hindrance, which we will revisit in Chapter 5.

CHAPTER 3: METHODOLOGY

We begin by describing the Spatiocyte simulation scheme, which is the basis for constructing the lattice theory. Significant theoretical contribution of this thesis including the lattice rate coefficient and the rebinding probability are derived in this chapter. Because of the different random walk behavior in 1D, 2D and 3D space, the lattice theory for each spatial dimensions is derived separately. Comparison between the lattice and continuum theory reveals their similarity in the steady-state and long-time asymptotic behavior. By matching the lattice and continuum theory, we obtain the criteria for determining MLM parameters to achieve consistent reaction kinetics behaviors. These criteria include the voxel size and the reaction probability, which will be used in the validation and application in later chapters.

3.1 Spatiocyte Simulation Scheme

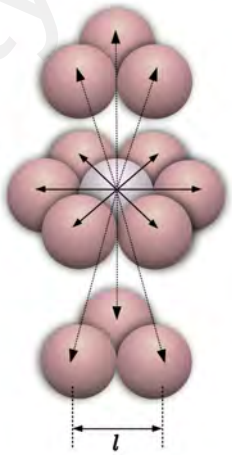


Figure 3.1: A voxel on the HCP lattice has 12 nearest neighbor voxels. The distance between the centers of two adjacent voxels is the voxel size l.

In the Spatiocyte scheme (see Algorithm 1) (Arjunan & Tomita, 2010), space is discretized into HCP lattice because the arrangement allows the highest density of regular sphere voxels in 3D space. The voxel has a diameter l and can be occupied by at most,

a single molecule. At each diffusion time step $t_d = l^2/2dD_x$, where D_x is the diffusion coefficient of molecule x and d is the spatial dimension, a molecule can hop to one of its 12 nearest neighbor voxels in 3D space or 6 nearest neighbor in 2D space (see Figure 3.1) with the step acceptance probability $P_w = 1$. When the destination voxel is occupied by a molecule belonging to a non-reactive species, the move will be rejected because of volume exclusion. If the destination voxel contains a reactant molecule, the corresponding reaction will occur with probability P_a; otherwise, if the reaction fails, the molecule will remain in its original voxel. When a molecule species dissociates into two product molecules, the pair will be placed side-by-side on the lattice. In volume-surface adsorption, a 3D species will absorb to the surface with predefined probability upon collision with the surface voxel.

The acceptance probability is linked to the macroscopic variable via $P_a = \Delta N/Z$. Here ΔN is the microscopic change in the number of product molecules in step interval t_d and Z is the expected number of collisions between A and B in the interval. For diffusion-influenced reactions ($k_a \gg k_D$), the collision rate Z is reduced relative to the production rate ΔN. The acceptance probability $P_a = \Delta N/Z$ would then have the issue of exceeding unity when $\Delta N > Z$. The Spatiocyte scheme overcomes this issue by reducing the simulation interval by a factor of α to $t' = t_d\alpha$. With the reduced interval, the effective number of collisions in t_d, Z is increased. The step and reaction acceptance probabilities are then decreased accordingly to $P_w = \alpha$ and $P'_a = P_a\alpha$, respectively. Algorithm 1 describes how α is set. In summary, the Spatiocyte scheme operates with $\alpha = 1$ when $P_a \leq 1$ (activation-limited case) and with $\alpha < 1$ when $P_a > 1$ (diffusion-influenced case). The MLM simulation scheme has been implemented in the Spatiocyte module of the E-Cell4 software (Kaizu et al., 2018).

Initialization:
$t_{sim} \leftarrow 0$, scheduler $S \leftarrow \{\}$
for each species x **do**
    $\rho_x = \max\{P_{axy}\}$, where xy denotes the pair of reactive species x and y;
    $S \leftarrow t' = t_d\alpha$, where $\alpha = \begin{cases} 1/\rho_x, & \text{for } \rho_x > 1 \\ 1, & \text{for } \rho_x \leq 1 \end{cases}$;
    reaction acceptance probability $P'_{axy} = P_{axy}\alpha$;
    step acceptance probability $P_{wx} = \alpha$;
end
Main loop:
while $S \neq \{\}$ and $t_{sim} < t_{end}$ **do**
    $t_{sim} \leftarrow \tau_x$ = next event in S;
    get species identity x;
    get current voxel location s_0;
    reschedule next event, $\tau_x = \tau_x + t'$
    for each molecule of species x **do**
        choose a random target voxel $s_1 \in$ {nearest neighbor of s_0};
        if s_1 is vacant **then**
            draw $rand$;
            if $rand \leq P_{wx}$ **then** walk succeeded, $s_0 \leftarrow s_1$;
            else
                walk rejected, $s_0 \leftarrow s_0$;
            end
        else if s_1 contains reactant species y **then**
            draw $rand$;
            if $rand \leq P'_{axy}$ **then**
                reaction xy accepted, $s_0 \leftarrow s_1$
            else
                reaction failed and walk rejected, $s_0 \leftarrow s_0$
            end
        else
            walk rejected, $s_0 \leftarrow s_0$;
        end
    end
end

Algorithm 1: Basic outline of the Spatiocyte algorithm for bimolecular reactions. t_{sim} is the current simulation time, t_{end}-t_{sim} is the simulation duration, P_{axy} is the reaction acceptance probability for the reactive pair of species x and y, $t_d = l^2/2dD_x$ is the diffusion (hopping) time step of the current species x, l is the voxel size, D_x is the diffusion coefficient of x, and $rand$ is a random number drawn from the uniform distribution with the interval $[0, 1)$.

3.2 Lattice Theory for Bimolecular Reaction

We derive the lattice theory for bimolecular reaction based on the particle-pair formalism of SCK model as defined in Equation (2.5). We start with the derivation of lattice initial rate constant using the mass action assumption. Consider bimolecular reaction in Equation (2.1) occurring on the lattice space. The number of reactions occurred in a simulation time interval t' is approximately

$$\Delta N = k'_a[A][B]t' = \text{number of reaction occurred on lattice,} \tag{3.1}$$

which give the expression for lattice initial rate constant

$$k'_a = \frac{\Delta N}{[A][B]t'}. \tag{3.2}$$

The discrete-space version of the time-dependent reaction rate coefficient in Equation (2.5) is defined as (Torney & McConnell, 1983; Montroll, 1969)

$$k_m = k'_a \left[1 - \sum_{n=0}^{m} H_n(s_0|s_1) \right], \text{ for } m, n \in \mathbb{N}, \tag{3.3}$$

where m is the simulation step, related to the simulation time via the relation $2dD_x t' = ml^2$ and $H_n(s_0|s_1)$ is the lattice analogue of the rebinding-time probability function $p_{reb}(R, t|R, 0)$ in diffusion step n. At long-time, we have the lattice asymptotic rate coefficient defined as

$$k'(t \to \infty) = \lim_{m\to\infty} k_m = k'_a \left[1 - \sum_{n=0}^{\infty} H_n(s_0|s_1) \right], \tag{3.4}$$

where the summation term corresponds to the total rebinding probability.

To obtain the analytical expression for $H_n(s_0|s_1)$, we consider a reactive pair A and B, which are initially in-contact by occupying adjacent voxels on the lattice. We are interested in the rebinding-time probability distribution as a function of the diffusion step n. Without losing generality, we can fix one of the molecules and diffuse the other with the relative diffusion coefficient D. The rebinding-time probability distribution of A and B is related to the arrival-time probability distribution of a random walker to the origin for the first time, given that the walk started at one of the neighbor voxels of the origin with diffusion coefficient D.

In the following sections, we derive the general expression for the rebinding-time probability distribution. Since the approaches for activation-limited and diffusion-influenced schemes are different, we perform their derivations separately.

3.2.1 Rebinding-time Probability Distribution

3.2.1.1 Activation-limited Scheme

We first consider the rebinding-time probability distribution for the case $P_a = 1$. Let $F_n(s_0|s_0)$ and $F_n(s_0|s_1)$ denote the first-passage time distributions to origin from origin and s_1, respectively. The two probabilities are related via

$$p(s_0 \to s_1)F_n(s_0|s_1) = F_{n+1}(s_0|s_0), \tag{3.5}$$

where $p(s_0 \to s_1) = 1$ is the transition probability from s_0 to s_1 in a single step. This implies that the trajectory we are interested in, which is from an in-contact situation (e.g., A at s_1 and B at s_0) to the rebinding situation (A hops to s_0) in a single step, is equivalent to the 2-step trajectory, $s_0 \to s_1 \to s_0$. Therefore, the rebinding-time probability distribution is fully described by $F_n(s_0|s_1)$ and is related to $F_n(s_0|s_0)$.

For $P_a < 1$, the rebinding-time probability distribution is given by

$$H_n(s_0|s_1) = P_a F^1_{n+1}(s_0|s_0) + P_a(1-P_a)F^2_{n+2}(s_0|s_0) + P_a(1-P_a)^2 F^3_{n+3}(s_0|s_0) + ..., \quad (3.6)$$

. Using Equation (2.46) we can obtain $H_n(s_0|s_1)$ recursively via

$$H_n(s_0|s_1) = P_a \sum_{j=1}^{n} F^j_{n+j}(s_0|s_0)(1-P_a)^{j-1}, \text{ for } j \in \mathbb{Z}^+, \ n \in \mathbb{N}. \quad (3.7)$$

The generating function of $H_n(s_0|s_1)$ is related to the generating function of $F_n(s_0|s_0)$:

$$\begin{aligned} H(s_0|s_1; z) &= \sum_{n=0}^{\infty} H_n(s_0|s_1) z^n \\ &= P_a \sum_{n=0}^{\infty} \sum_{j=1}^{\infty} F^j_{n+j}(s_0|s_0) \, (1-P_a)^{j-1} z^n \\ &= P_a \sum_{j=1}^{\infty} (1-P_a)^{j-1} z^{-j} \sum_{n=0}^{\infty} F^j_{n+j}(s_0|s_0) \, z^{n+j} \\ &= P_a \sum_{j=1}^{\infty} (1-P_a)^{j-1} z^{-j} \sum_{n=0}^{\infty} F^j_{n}(s_0|s_0) \, z^{n} \end{aligned} \quad (3.8)$$

where in the last step we have $\sum_{k=1}^{j-1} F^j_k(s_0|s_0) \, z^n = 0$ since for all k such that $k < j-1$, the return probability is zero. Using property Equation (2.47), we finally obtain

$$\begin{aligned} H(s_0|s_1; z) &= P_a \sum_{j=1}^{\infty} (1-P_a)^{j-1} z^{-j} F(s_0|s_0; z)^j \\ &= \frac{P_a F(s_0|s_0; z)}{F(s_0|s_0; z)(P_a - 1) + z}. \end{aligned} \quad (3.9)$$

3.2.1.2 Diffusion-influenced Scheme

The diffusion-influenced scheme differs from the activation-limited scheme mainly by the presence of non-unity step acceptance probability $P_w = \alpha$. The diffusion step n is therefore no longer the same as the simulation step. Specifically, a successful arrival at a

new target voxel (or a successful reaction attempt with a reactant) after $n = 1$ step could have had multiple k simulation steps in the past with hopping failures (or failed reaction attempts). As a result, the actual simulation time corresponding to n steps is not a single value $nt' = nt_d\alpha$, but follows some distribution.

Eventually we want to obtain the long-time asymptotic of the rate coefficient, which is independent of the transient time-dependent behavior. Hence, we parameterize the rebinding-time according to the eventful step n (which will be incremented after a physical movement or a reaction attempt), rather than the actual simulation step k. We begin with the derivation of two useful return probability function and then proceed with the rebinding-time probability distribution.

(a) ***Return probability*** $F_n(s_1|s_1)$

We denote $P_n(s|s_0)$ as the voxel occupation transition probability from s_0 to s. It is related to $F_n(s|s_0)$ via the convolution relation ((Hughes, 1995), pp. 121)

$$P_n(s|s_0) = \delta_{ss_0}\delta_{n,0} + \sum_{j=1}^{n} F_j(s|s_0)P_{n-j}(s|s), \text{ for } n \in \mathbb{N}. \tag{3.10}$$

If a random walker started at s_0, it must go through s_1 before reaching the destination voxel s, we then have

$$P_n(s|s_0) = \delta_{ss_0}\delta_{n,0} + \delta_{s_0s_1}\delta_{n,1}P_1(s|s_1) + \sum_{j=1}^{n} F_j(s_1|s_0)P_{n-j}(s|s_1). \tag{3.11}$$

Note that $P_n(s|s_1) = P_{n+1}(s|s)$. Thus, with $s_0 = s_1$, we have

$$P_{n+1}(s|s) = \delta_{n1}P_2(s|s) + \sum_{j=1}^{n} F_j(s_1|s_1)P_{n-j+1}(s|s). \tag{3.12}$$

Multiplying both sides with z^{n+1} gives

$$z^{n+1}P_{n+1}(s|s) = \delta_{n1}z^{n+1}P_2(s|s) + \sum_{j=1}^{n} z^j F_j(s_1|s_1)z^{n-j+1}P_{n-j+1}(s|s). \tag{3.13}$$

Taking the sum of both sides from $n = 0$ to infinity gives

$$P(s|s;z) - P_0(s|s) = z^2P_2(s|s) + F(s_1|s_1;z)[P(s|s;z) - P_0(s|s)], \tag{3.14}$$

where

$$P(s|s;z) = \sum_{n=0}^{\infty} z^n P_n(s|s),\ F(s_1|s_1;z) = \sum_{n=0}^{\infty} z^n F_n(s_1|s_1) \text{ and } F_0(s_1|s_1) = 0. \tag{3.15}$$

As such, we have

$$\begin{aligned} F(s_1|s_1;z) &= 1 - \frac{z^2P_2(s|s)}{P(s|s;z) - 1} \\ &= 1 - \frac{z^2P_1(s_0|s_1)}{P(s_0|s_0;z) - 1}. \end{aligned} \tag{3.16}$$

The total return probability to s_1 from s_1 is

$$\begin{aligned} \sum_{n=0}^{\infty} F_n(s_1|s_1) &= \lim_{z\to 1^-} F(s_1|s_1;z) \\ &= 1 - \frac{P_1(s_0|s_1)}{P(s_0|s_0;1^-) - 1}. \end{aligned} \tag{3.17}$$

Using relation Equation (2.45), we have

$$\sum_{n=0}^{\infty} F_n(s_1|s_1) = 1 - (1/F(1) - 1)P_1(s_0|s_1). \tag{3.18}$$

(b) ***Return probability*** $\bar{F}_n(s_1|s_2)$

If we increment the step count n for every successful step to a new voxel, then the first-passage time distribution from s_2 to s_1 at step n is given by

$$\bar{F}_n(s_1|s_2) = \sum_{m=n}^{\infty} \binom{m-1}{n-1} P_w^n (1-P_w)^{m-n} F_n(s_1|s_2), \text{ for } n \in \mathbb{Z}^+, \tag{3.19}$$

where $P_w = \alpha$, is the step acceptance probability. It can be shown that

$$\begin{aligned} \bar{F}_n(s_1|s_2) &= P_w^n F_n(s_1|s_2) \sum_{m=n}^{\infty} \binom{m-1}{n-1} (1-P_w)^{m-n} \\ &= P_w^n F_n(s_1|s_2) \frac{1}{P_w^n} \\ &= F_n(s_1|s_2). \end{aligned} \tag{3.20}$$

(c) ***Derivation of rebinding-time probability distribution***

The rebinding-time probability distribution $G_n(s_0|s_1)$ is defined as

$$G_{n+1}(s_0|s_1) = S_n(s_1|s_1)\, p(s_1 \to s_0), \text{ for } n \in \mathbb{N} \tag{3.21}$$

where $S_n(s_1|s_1)$ is the in-contact probability of a reactive pair after n steps and $p(s_1 \to s_0)$ is the transition probability from s_1 to s_0. The latter can be expressed as

$$\begin{aligned} p(s_1 \to s_0) &= P_a \alpha P_1(s_0|s_1) \sum_{k=0}^{\infty} \{[1 - P_1(s_0|s_1)](1-P_w)\}^k \\ &= \frac{P_a \alpha P_1(s_0|s_1)}{1 - (1-P_w)(1-P_1(s_0|s_1))}, \end{aligned} \tag{3.22}$$

where the nominator term accounts for the probability of hopping to s_0 from s_1 and successfully reacting with the reactant located at s_0 in one diffusion step, while the denominator term comes from the infinite sum representing the total probability of

unsuccessful escape to $s \in \{\text{adjacent voxel of } s_1\} \setminus s_0$ at the previous simulation step[1]. When $P_w = 1$ and $\alpha = 1$ as in the activation-limited case, the reaction probability becomes $p(s_1 \to s_0) = P_a P_1(s_0|s_1)$.

Next, we derive the generating functions of two first-passage time distributions $F_n(s_1|s_1)$ and $F_n(s_1|s_2)$ that correspond to the current scheme. We start from

$$\begin{aligned} F_{n+1}(s_1|s_1) &= \sum_s P_1(s|s_1) F_n(s_1|s), \text{ for } n \in \mathbb{N} \\ &= P_1(s_0|s_1)\delta_{n,1} + P_1(s_1|s_1)\delta_{n,0} + P_1(s_2|s_1) F_n(s_1|s_2), \end{aligned} \tag{3.23}$$

where the first term on the right-hand side relates to the failed reaction attempt $s_1 \to s_0 \to s_1$[2], the second term describes the hop from $s_1 \to s_1$, and the last term accounts for the trajectory $s_1 \to s_2$, which is continued by a series of n steps that have ended up in s_1 again.

From Equation (3.23), we obtain the generating function of $F_n(s_1|s_1)$ as

$$F(s_1|s_1; z) = z^2 P_1(s_0|s_1) + z P_1(s_1|s_1) + z P_1(s_2|s_1) F(s_1|s_2; z). \tag{3.24}$$

Thus we obtain

$$F(s_1|s_2; z) = \frac{F(s_1|s_1; z) - z^2 P_1(s_0|s_1) - z P_1(s_1|s_1)}{z P_1(s_2|s_1)}, \tag{3.25}$$

where $F(s_1|s_1; z)$ is given in Equation (3.16).

Now, we define the probability that a particle is in-contact after n-step as:

$$\begin{aligned} S_n(s_1|s_1) = \gamma_1 S_{n-1}(s_1|s_1) + \sum_{m=0}^{n-1} \gamma_2 S_m(s_1|s_1)\, \bar{F}_{n-m-1}(s_1|s_2) + \delta_{n,0} S_0(s_1|s_1), \\ \text{for } n \in \mathbb{N}, \end{aligned} \tag{3.26}$$

[1] there are k simulation steps in between each diffusion step n

[2] only $s_1 \to s_0$ is considered as a diffusion step, whereas the rejection $s_0 \to s_1$ is not

where the first term accounts for the trajectories $s_1 \rightarrow s_0 \rightarrow s_1$ and $s_1 \rightarrow s_1$, the second term represents the trajectories $s_1 \rightarrow s_2 \rightarrow s_1$ and the last term accounts for the initial condition. In detail, the coefficient

$$\begin{aligned}\gamma_1 &= [(1 - P_a\alpha)P_1(s_0|s_1) + P_w P_1(s_1|s_1)] \sum_{k=0}^{\infty} \{[1 - P_1(s_0|s_1)](1 - P_w)\}^k \\ &= \frac{(1 - P_a\alpha)P_1(s_0|s_1) + P_w P_1(s_1|s_1)}{1 - [1 - P_1(s_0|s_1)](1 - P_w)},\end{aligned} \tag{3.27}$$

accounts for the total probability of arrival at s_1 from a rejected reaction attempt (first sub-term) or from the adjacent neighbor s_1 (second sub-term) given that there was no successful escape to $s \in \{\text{adjacent voxel of } s_1\} \setminus s_0$ at the last simulation step k before the arrival, while the coefficient

$$\begin{aligned}\gamma_2 &= P_w P_1(s_2|s_1) \sum_{k=0}^{\infty} \{[1 - P_1(s_0|s_1)](1 - P_w)\}^k \\ &= \frac{P_w P_1(s_2|s_1)}{1 - [1 - P_1(s_0|s_1)](1 - P_w)},\end{aligned} \tag{3.28}$$

accounts for the total probability of arriving at s_2 from s_1 given that there was no successful escape to $s \in \{\text{adjacent voxel of } s_1\} \setminus s_0$ at the last simulation step k before the arrival, and finally $\bar{F}_n(s_1|s_2) = F_n(s_1|s_2)$ (as shown in Equation (3.20)) denotes the first-passage time distribution of the scheme with step-acceptance probability $P_w = \alpha$.

We multiply Equation (3.26) with z^n:

$$\begin{aligned}S_n(s_1|s_1)\, z^n = \gamma_1 z S_{n-1}(s_1|s_1)\, z^{n-1} + \gamma_2 z \sum_{m=0}^{n-1} S_m(s_1|s_1)\, z^m F_{n-m-1}(s_1|s_2) z^{n-m-1} \\ + \delta_{n,0} S_0(s_1|s_1)\, z^n,\end{aligned} \tag{3.29}$$

and take the sum to infinity to obtain

$$S(s_1|s_1; z) = \gamma_1 z S(s_1|s_1; z) + \gamma_2 z S(s_1|s_1; z) F(s_1|s_2; z) + S_0(s_1|s_1). \tag{3.30}$$

After collecting the terms, we obtain the generating function of $S_n(s_1|s_1)$:

$$S(s_1|s_1;z) = \frac{S_0(s_1|s_1)}{1-\gamma_1 z - \gamma_2 z F(s_1|s_2;z)}. \tag{3.31}$$

Substituting Equations (3.22) and (3.31) into Equation (3.21) gives the rebinding-time probability distribution:

$$G_{n+1}(s_0|s_1) = \frac{P_a \alpha P_1(s_0|s_1)}{1-[1-P_1(s_0|s_1)](1-P_w)} S_n(s_1|s_1), \text{ for } n \in \mathbb{N}, \tag{3.32}$$

with the corresponding probability generating function as

$$G(s_0|s_1;z) = \frac{P_a \alpha P_1(s_0|s_1)}{1-[1-P_1(s_0|s_1)](1-P_w)} S(s_1|s_1;z). \tag{3.33}$$

Now we have the general form of the lattice rebinding-time probability distribution, we can obtain the dimension-specific expression as shown in following subsections.

3.2.2 3D Reaction

We begin with the derivation of lattice initial rate constant which according to Equation (3.2), is defined as

$$k'_{a3D} = \frac{\Delta N V}{N_A N_B t'}. \tag{3.34}$$

The number of reactions in a single step t' is $\Delta N = Z P'_a$, where the average number of encounter in a single step can be approximated as $Z = N_B N_A / N_v$. Here, $N_v = \sqrt{2}V/l^3$ is the total number of voxels in a compartment volume V, and $P'_a = P_a \alpha$ is the actual reaction acceptance probability during the encounter.

In the activation-limited scheme, we have $t' = t_d$ and $P'_a = P_a$. Thus k'_{a3D} is derived as

$$\begin{aligned} k'_{a3D} &= \frac{P'_a l^3}{\sqrt{2}t'} \\ &= 3\sqrt{2}P_a D l, \end{aligned} \tag{3.35}$$

Note that D is the sum of diffusion coefficients of the reacting pair, $D_A + D_B$.

Similary, for the diffusion-influenced scheme, where $t' = t_d\alpha$ and $P'_a = P_a\alpha$, we have

$$\begin{aligned} k'_{a3D} &= \frac{\alpha P_a l^3}{\sqrt{2}\alpha t_d} \\ &= 3\sqrt{2}P_a D l. \end{aligned} \tag{3.36}$$

Also note that the physical dimension of k'_{a3D} satisfies $cm^3 s^{-1}$.

The above derivation for HCP lattice can be generalized to other lattice arrangements:

$$k'_{a3D} = \frac{\pi P_a D l}{\phi_l}, \tag{3.37}$$

where ϕ_l is the packing fraction for the lattice (e.g. $\phi_l = \pi/6$ for the simple cubic lattice).

Next we derive the total rebinding probability and the rate coefficient for the activation-limited and diffusion-influenced schemes.

3.2.2.1 Activation-limited Case ($k_{a3D} \ll k_D$)

(a) ***Total rebinding probability and effective rate constant***

The total rebinding probability of an in-contact pair on lattice is obtained by taking the limit $z \to 1$ in Equation (3.9) as:

$$H_{reb} = \lim_{z\to 1} H(s_0|s_1; z) = \frac{P_a}{P_a + \frac{1}{F(1)} - 1}, \tag{3.38}$$

where $F(1) = F(s_0|s_0; z = 1)$. It was shown previously that the probability generating function of the HCP lattice is topologically equivalent to that of the face-centered cubic (FCC) lattice (Ishioka & Koiwa, 1978). Therefore, we have $F(1) \approx 0.256318$ ((Hughes, 1995), pp. 153) for HCP lattice.

The effective rate constant on lattice is obtained by substituting the initial rate constant defined in Equation (3.35) and the total rebinding probability defined in Equation (3.38) into Equation (3.4) as

$$k'_{eff3D} = 3\sqrt{2}Dl\left(\frac{1}{F(1)} - 1\right)\frac{P_a}{P_a + \frac{1}{F(1)} - 1}\,. \tag{3.39}$$

(b) ***Long-time asymptotic behavior***

The asymptotic behavior of the rebinding-time probability distribution $H_n(s_0|s_1)$ at large n can be estimated directly from the generating function. First we expand the generating function of the return probability $P_n(s|s)$ for the HCP lattice around $z = 1$ up to the $O(1-z)$ term (see Equation (D.8b) in (Montroll & Weiss, 1965) and Equation (A.237) in (Hughes, 1995))

$$P(s|s; z) \approx P(1) - c_1\sqrt{1-z} + O(1-z), \tag{3.40}$$

where $P(1) = P(s|s; z = 1) \approx 1.344661$ and $c_1 = 3^{3/2}/2\pi$. The corresponding expansion of the generating function of $F_n(s|s)$ is

$$\begin{aligned} F(s|s; z) &= 1 - \frac{1}{P(s|s; z)} \\ &\approx 1 - \frac{1}{P(1) - c_1\sqrt{1-z}} \\ &\approx 1 - \frac{1}{P(1)} - \frac{c_1}{P(1)^2}\sqrt{1-z}, \end{aligned} \tag{3.41}$$

where we have ignored the term equal to or higher than $O(1-z)$. Recall that the generating function of the rebinding-time probability distribution for the activation-limited case:

$$\begin{aligned} H(s_0|s_1;z) &= \frac{P_a F(s_0|s_0;z)}{z + F(s_0|s_0;z)(P_a - 1)} \\ &= \frac{P_a F(s_0|s_0;z)}{z[1 - F(s_0|s_0;z)(1-P_a)/z]}. \end{aligned} \tag{3.42}$$

By the expansion of the denominator we have

$$\begin{aligned} H(s_0|s_1;z) &= \frac{P_a F(s_0|s_0;z)}{z}\left\{1 + \frac{(1-P_a)F(s_0|s_0;z)}{z} + \left[\frac{(1-P_a)F(s_0|s_0;z)}{z}\right]^2 + \ldots\right\} \\ &= \frac{P_a}{z}\left\{F(s_0|s_0;z) + \frac{(1-P_a)}{z}F(s_0|s_0;z)^2 + \left[\frac{(1-P_a)}{z}\right]^2 F(s_0|s_0;z)^3 + \ldots\right\}. \end{aligned} \tag{3.43}$$

Substituting Equation (3.41) into $H(s_0|s_1;z)$ and collecting the leading terms gives

$$H(s_0|s_1;z) \approx w\sqrt{1-z} + O(1-z), \tag{3.44}$$

with constant

$$\begin{aligned} w &= -\frac{c_1 P_a}{zP(1)^2}\left\{1 - 2\frac{(P_a-1)(P(1)-1)}{zP(1)} + 3\left[\frac{(P_a-1)(P(1)-1)}{zP(1)}\right]^2 + \ldots\right\} \\ &= -\frac{c_1 P_a}{zP(1)^2}\sum_{n=1}^{\infty} n(-1)^{n+1}\left[\frac{(P_a-1)(P(1)-1)}{zP(1)}\right]^{n-1} \\ &= -\frac{c_1 P_a}{zP(1)^2}\left[1 + \frac{(P_a-1)(P(1)-1)}{zP(1)}\right]^{-2} \\ &= -\frac{c_1 P_a}{z\left[1 + P_a(P(1)-1) + P(1)(z-1)\right]^2} \\ &= -\frac{c_1 P_a}{\left[1 + P_a(P(1)-1)\right]^2}, \end{aligned} \tag{3.45}$$

where we have substitute z with 1 in the last step. By means of singularity analysis of the generating function (see Equation (2.3) of (Flajolet & Odlyzko, 1990)), the corresponding

asymptotic behavior of $H_n(s_0|s_1)$ as $n \to \infty$ is therefore

$$H_n(s_0|s_1) \approx -\frac{w}{2\sqrt{\pi}} n^{-3/2} + O(n^{-5/2}). \tag{3.46}$$

To analyze the long-time asymptotic behavior of rate lattice rate coefficient, we start with expression defined in Equation (3.4)

$$k_m^{3D} = k'_{a3D} \left[1 - \sum_{n=0}^{m} H_n(s_0|s_1) \right], \text{ for } m, n \in \mathbb{N}, \tag{3.47}$$

which can be rewritten as

$$k_m^{3D} = k'_{a3D} \left[1 - \sum_{n=0}^{\infty} H_n(s_0|s_1) + \sum_{n=m}^{\infty} H_n(s_0|s_1) \right]. \tag{3.48}$$

The first summation term is the total rebinding probability while the second term can be evaluated using the Euler-Maclaurin formula:

$$\begin{aligned} \sum_{n=m}^{\infty} H_n(s_0|s_1) &\approx \int_m^{\infty} dn \frac{w}{2\sqrt{\pi}} n^{-3/2} \\ &\approx \frac{w}{\sqrt{\pi m}} \\ &\approx \frac{lw}{\sqrt{6D\pi t}}, \end{aligned} \tag{3.49}$$

where we have used the definition $ml^2 = 6Dt$ in the last step.

Now we have the asymptotic reaction rate as

$$\lim_{t \to \infty} k_{3D}(t)' \approx k'_{a3D} \left[1 - H_{reb} + \frac{lw}{\sqrt{6D\pi t}} \right]. \tag{3.50}$$

After rearrangement we have

$$\begin{aligned}\lim_{t\to\infty} k'_{3D}(t) &\approx k'_{a3D}(1-H_{reb})\left[1+\frac{lw}{(1-H_{reb})\sqrt{6D\pi t}}\right] \\ &\approx k'_{a3D}(1-H_{reb})\left[1+\frac{c_1P_al}{(1+(P(1)-1)P_a)\sqrt{6D\pi t}}\right].\end{aligned} \tag{3.51}$$

Using the definition $k'_{a3D}(1-H_{reb}) = k'_{eff3D}$, and applying the expressions for reaction acceptance probability in Equation (3.35) and voxel size in Equation (3.95), we obtain the long-time approximation as

$$\lim_{t\to\infty} k'_{3D}(t) \approx k'_{eff3D}\left[1+\frac{k_{a3D}R}{(k_{a3D}+k_D)\sqrt{\pi Dt}}\right]. \tag{3.52}$$

Note that the long-time asymptotic rate coefficient on lattice has the exact same form as the continuum theory in Equation (2.13).

3.2.2.2 Diffusion-influenced Case ($k_{a3D} \gg k_D$)

(a) ***Total rebinding probability and effective rate constant***

To derive the total rebinding probability for diffusion-influenced scheme, we use the expression given in Equations (3.33) and (3.31). In the diffusion-influenced scheme of Spatiocyte, we have $P_1(s_0|s_1) = 1/12$, $P_1(s_1|s_1) = 4/12$, $P_1(s_2|s_1) = 7/12$ and $P_w = \alpha$. Using these parameters, we have the following quantities:

$$\gamma_1 = \frac{(1-P_a\alpha)+4\alpha}{12\,[1-11(1-\alpha)/12]}, \tag{3.53}$$

$$\gamma_2 = \frac{7\alpha}{12\,[1-11(1-\alpha)/12]}, \tag{3.54}$$

$$F(s_1|s_2;z) = \frac{F(s_1|s_1;z) - z^2/12 - 4z/12}{7z/12}, \tag{3.55}$$

$$\begin{aligned} F(s_1|s_2;z=1) &= \frac{F(s_1|s_1;z=1) - 1/12 - 4/12}{7/12} \\ &= \frac{8 - 1/F(1)}{7}, \end{aligned} \tag{3.56}$$

where we have used definition Equation (3.18) in Equation (3.56). Using Equation (3.56), we obtain the limit of Equation (3.31) as:

$$\begin{aligned} S(s_1|s_1;z=1) &= \left[1 - \frac{(1-P_a\alpha)+4\alpha}{12\,[1-11(1-\alpha)/12]} - \frac{7\alpha}{12\,[1-11(1-\alpha)/12]} F(s_1|s_2;z=1)\right]^{-1} \\ &= \left[1 - \frac{(1-P_a\alpha)+4\alpha}{12\,[1-11(1-\alpha)/12]} - \frac{7\alpha}{12\,[1-11(1-\alpha)/12]} \frac{8-1/F(1)}{7}\right]^{-1} \\ &= \frac{12\,[1-11(1-\alpha)/12]}{P_a\alpha - \alpha + \alpha/F(1)}. \end{aligned} \tag{3.57}$$

Finally, we substitute Equation (3.57) into Equation (3.33) to obtain

$$\begin{aligned} G(s_0|s_1;1) &= \frac{P_a\alpha/12}{1-11(1-\alpha)/12} \frac{12\,[1-11(1-\alpha)/12]}{P_a\alpha - \alpha + \alpha/F(1)} \\ &= \frac{P_a}{P_a + \frac{1}{F(1)} - 1} \end{aligned} \tag{3.58}$$

Therefore, we have the total rebinding probability as:

$$G_{reb} = G(s_0|s_1;z=1) = \frac{P_a}{P_a + \frac{1}{F(1)} - 1}. \tag{3.59}$$

which is identical to Equation (3.38) in the activation-limited case.

Similarly, by substituting Equation (3.59) into the summation term in Equation (3.4), we get the effective rate constant for the diffusion-influenced case as

$$\begin{aligned} k'_{eff3D} &= k'_{a3D}[1 - G_{reb}], \\ &= 3\sqrt{2}Dl\left(\frac{1}{F(1)} - 1\right)\frac{P_a}{P_a + \frac{1}{F(1)} - 1}\ . \end{aligned} \tag{3.60}$$

which also follows Equation (3.39). Henceforth, we adopt the same notations of the effective reaction rate and total rebinding probability for both the activation-limited and diffusion-influenced cases.

(b) ***Long-time asymptotic behavior***

In the diffusion-influenced scenario, Spatiocyte uses a different approach for hopping and reaction. Simulation progresses with a smaller time step $t' = t_d\alpha$ to resolve fast reaction events. We show that as α becomes smaller, the reaction and hopping events occur in a probabilistic manner that follows exponential time distribution. This property provides us with an approximation to study the time-dependent behavior of the reaction kinetics.

Hopping time distribution Consider a single particle hopping on a completely vacant lattice. Let P_w be the step acceptance probability for a particle heading to a vacant voxel. The probability of successful hopping after m trials is

$$P_h(t = m) = P_w(1 - P_w)^{m-1}, \text{ for } m \in \mathbb{Z}^+, \tag{3.61}$$

The survival probability (no hopping) until mth trial is thus given as

$$\begin{aligned} P_h(t > m) &= \sum_{m}^{\infty} P_w(1 - P_w)^{m-1} \\ &= (1 - P_w)^{m-1}. \end{aligned} \tag{3.62}$$

If we perform the trial every δ sec such that $P_w = \beta_1 \delta$, where $\beta_1 = t_d^{-1}$ is the average hopping rate per second. The survival probability becomes

$$\begin{aligned} P_h(t > m\delta) &= P_h(t > t') \\ &= (1 - \beta_1 \delta)^{m-1}, \end{aligned} \tag{3.63}$$

where $t' = m\delta$. Similarly, we have

$$\begin{aligned} P_h(t > t') &= (1 - \beta_1 \delta)^{\frac{t'}{\delta} - 1} \\ &= \frac{(1 - \beta_1 \delta)^{\frac{t'}{\delta}}}{(1 - \beta_1 \delta)}. \end{aligned} \tag{3.64}$$

Taking the limit of small δ, we then have

$$\begin{aligned} \lim_{\delta \to 0} P_h(t > t') &= \left[\lim_{\delta \to 0} (1 - \beta_1 \delta)^{1/\delta} \right]^{t'} \\ &= \exp(-\beta_1 t'). \end{aligned} \tag{3.65}$$

Since $P_w = \alpha$, when α is small enough, the hopping time distribution of a particle approximates the exponential distribution

$$\psi_h(t) = \exp(-\beta_1 t), \tag{3.66}$$

with $\beta_1 = t_d^{-1}$.

Reaction time distribution Consider a reaction pair at an in-contact situation. The survival probability that they are still at the in-contact situation after n steps is

$$S_n = (1 - P_r - P_e)^n, \text{ for } n \in \mathbb{N}, \tag{3.67}$$

where $P_r = P_a\alpha/12$ is the reaction probability and $P_e = 11P_w/12 = 11\alpha/12$ is the escape probability. Let the simulation trial performed at infinitesimal time δ, such that $t' = n\delta = t_d\alpha$. The survival probability as a function of time is then

$$\begin{aligned} S(t') &= \lim_{\delta\to 0} S_n \\ &= \lim_{\delta\to 0}\left[1 - \frac{\alpha}{12}(P_a + 11)\right]^n \\ &= \lim_{\delta\to 0}\left[1 - \frac{\delta}{t_d}\frac{(P_a + 11)}{12}\right]^{t'/\delta} \\ &= \left[\lim_{\delta\to 0}(1 - \beta\delta)^{1/\delta}\right]^{t'} \\ &= \exp(-\beta t'). \end{aligned} \tag{3.68}$$

where $\beta = (P_a + 11)/12t_d$. Note that the survival probability in this form includes both the probability of reaction and hopping events. Since the two events are independent of each other, the survival probability can be split into two separate terms:

$$S(t') = \exp(-\beta_1 t')\exp\left(-\frac{11\beta_2 t'}{12}\right), \tag{3.69}$$

where $\beta_1 = P_a/12t_d$ is the average reaction rate and $\beta_2 = 1/t_d$ is the average hopping rate. Therefore, the survival probability of the reaction also follows the exponential function

$$\psi_r(t') = \exp(-\beta_1 t'). \tag{3.70}$$

Time dependent survival probability In summary, the survival probability of the reaction and hopping events are (from Equation (3.66) and Equation (3.70))

$$\begin{aligned} &\psi_r(t) = \exp(-\beta_1 t)\ \text{, where } \beta_1 = \frac{P_a}{12t_d}, \\ &\psi_h(t) = \exp(-\beta_2 t)\ \text{, where } \beta_2 = \frac{1}{t_d}. \end{aligned} \tag{3.71}$$

Thus, the survival probability after one step is

$$\psi(t) = \psi_r(t)\psi_h(t) = \exp\{-\beta t\}, \tag{3.72}$$

where $\beta = \beta_1 + \beta_2$. As a consequence, the survival probability of a reactive pair at short time t after step n follows the Poisson distribution:

$$S_n(t) = \frac{(\beta t)^n}{n!}\exp(-\beta t), \text{ for } n \in \mathbb{N}, \tag{3.73}$$

where $S_0(t) = \exp(-\beta t)$.

Long-time rate coefficient Here, we study the time-dependent kinetics of the diffusion-influenced scheme. We start with the definition of continuous rebinding-time probability density, and use it to express the time-dependent rate coefficient. Denoting the continuous rebinding-time probability density after $(n + 1)$ steps as

$$G_{n+1}(t) = \beta_1 S_n(s_1|s_1;t), \text{ for } n \in \mathbb{N}, \tag{3.74}$$

where

$$S_n(s_1|s_1;t) = \delta_{n,0}S_0(s_1|s_1;t) + \int_0^t dt' \sum_{j=0}^{n} S_{n-j}(s_1|s_1;t-t')F_j(s_1|s_1;t), \tag{3.75}$$

is the survival time probability density of a particle that started and ended at s_1 on the nth step. The first term on the right-hand side of Equation (3.75) is the initial probability density $S_0(s_1|s_1;t) = \exp(-\beta t)$, while the last term involves two convolutions: the continuous time convolution and the discrete step convolution nested inside the time convolution.

$F_n(s_1|s_1;t)$ is the first-passage time density at the nth step, defined as

$$F_n(s_1|s_1;t) = F_n(s_1|s_1)\left[\delta_{n,1}\beta_2\exp(-\beta t) + (1-\delta_{n,1})\int_0^t dt'\beta_2\exp(-\beta t')\frac{\beta_2^n(t-t')^{n-1}}{(n-1)!}\exp(-\beta_2(t-t'))\right], \tag{3.76}$$

for $n \in \mathbb{Z}^+$. Intuitively, the first term describes the first-passage time distribution for single step while the second term accounts for the convolution of the probability of time required for the $n-1$ steps after the first step.

The continuous rebinding-time probability density is related to the rate coefficient of the particle-pair formalism through:

$$k'_{3D}(t) = k'_{a3D}\left[1 - \int_0^t dt' G_{n+1}(t')\right]. \tag{3.77}$$

We then take the Laplace transform of $k'_{3D}(t)$ which is easier to work with:

$$s\hat{k}'_{3D}(s) = k'_{a3D}[1 - \hat{G}(s)]. \tag{3.78}$$

Note that $\hat{G}(s)$ is related to the rebinding-time and survival-time probability densities via:

$$\hat{G}(s) = \sum_{n=1}^{\infty}\hat{G}_n(s_1|s_1;s) = \beta_1\sum_{n=1}^{\infty}\hat{S}_n(s_1|s_1;s). \tag{3.79}$$

The corresponding Laplace transform of Equation (3.75) is given as

$$\hat{S}_n(s_1|s_1;s) = \frac{\delta_{n,0}}{s+\beta} + \sum_{j=0}^{n}\hat{S}_{n-j}(s_1|s_1;s)\hat{F}_j(s_1|s_1;s), \tag{3.80}$$

where

$$\hat{F}_n(s_1|s_1;s) = F_n(s_1|s_1)\frac{\beta_2}{s+\beta}\left[\delta_{n,1} + (1-\delta_{n,1})\left(\frac{\beta_2}{s+\beta_2}\right)^n\right]. \tag{3.81}$$

The infinite sum of Equations(3.80) and (3.81) are

$$\begin{aligned}\sum_{n=0}^{\infty}\hat{S}_n(s_1|s_1;s) &= \frac{1}{s+\beta} + \sum_{n=0}^{\infty}\sum_{j=0}^{n}\hat{S}_{n-j}(s_1|s_1;s)\hat{F}_j(s_1|s_1;s),\\ &= \frac{1}{s+\beta} + \sum_{n=0}^{\infty}\hat{S}_n(s_1|s_1;s)\sum_{n=0}^{\infty}\hat{F}_n(s_1|s_1;s),\\ &= \frac{1}{s+\beta}\left[1-\sum_{n=0}^{\infty}\hat{F}_n(s_1|s_1;s)\right]^{-1},\end{aligned} \tag{3.82}$$

$$\begin{aligned}\sum_{n=0}^{\infty}\hat{F}_n(s_1|s_1;s) &= \frac{\beta_2}{s+\beta}\left[F_1(s_1|s_1) + \sum_{n=2}^{\infty}F_n(s_1|s_1)\left(\frac{\beta_2}{s+\beta_2}\right)^n\right]\\ &= \frac{\beta_2}{s+\beta}\left[F_1(s_1|s_1)\frac{s}{s+\beta_2} + \sum_{n=0}^{\infty}F_n(s_1|s_1)\left(\frac{\beta_2}{s+\beta_2}\right)^n\right]\\ &= \frac{\beta_2}{s+\beta}\left[F_1(s_1|s_1)\frac{s}{s+\beta_2} + F\left(s_1|s_1;z=\frac{\beta_2}{s+\beta_2}\right)\right],\end{aligned} \tag{3.83}$$

where $F(s_1|s_1;z)$ is the generating function, $\sum_{n=0}^{\infty}F(s_1|s_1)z^n$ is as defined in (3.16). Hence, we have

$$\sum_{n=0}^{\infty}\hat{S}_n(s_1|s_1;s) = [s+\beta - sF_1(s_1|s_1)z - \beta_2 F(s_1|s_1;z)]^{-1}, \tag{3.84}$$

where $z = \beta_2/(s+\beta_2)$.

Substituting Equation (3.84) into Equation (3.79) and with the final value theorem, we obtain the long-time behavior of $k'_{3D}(t)$ by taking the limit $s \to 0$ in Equation (3.78). Assuming the asymptotic Laplace form of the rate coefficient on lattice ((Agmon & Szabo,

1990) Equation (2.37a)):

$$s\hat{k}'_{3D}(s) \approx k'_{eff3D}(1 + a'_{eff}\sqrt{s/D} + ...). \tag{3.85}$$

We then set $s = 0$ to obtain the effective lattice reaction rate constant:

$$k'_{eff3D} = k'_{a3D}[1 - \hat{G}(0)]. \tag{3.86}$$

Evaluating $\hat{G}(0)$ by referring to Equation (3.79), we then get

$$k'_{eff3D} = \frac{3\sqrt{2}P_a D l}{1 + P_a/(1/F(1) - 1)}, \tag{3.87}$$

which is consistent with the activation-limited result in Equation (3.39).

The second order term of Equation (3.85) is evaluated by expanding $s\hat{k}'_{3D}(s)$ around $s = 0$:

$$\begin{aligned}
\lim_{s\to 0}\frac{d}{d\sqrt{s}}s\hat{k}'_{3D}(s) &= \lim_{s\to 0}\frac{d}{d\sqrt{s}}(-k'_{a3D}\hat{G}(s)) \\
&= \frac{-k'_{a3D}\beta_1}{\sqrt{\beta_2}}\lim_{q\to 0}\frac{d}{dq}\sum_{n=0}^{\infty}\hat{S}_n(s_1|s_1;s) \\
&= k'_{a3D}\beta_1\sqrt{\beta_2}\,[\beta - \beta_2 F(1)]^{-2}\lim_{q\to 0}\frac{d}{dq}F(q) \\
&= \frac{2k'_{a3D}P_a}{\sqrt{\beta_2}}\left[1 + \frac{P_a}{1/F(1) - 1}\right]^{-2}\lim_{q\to 0} q\frac{d}{dz}P(s_0|s_0;z),
\end{aligned} \tag{3.88}$$

where $q = \sqrt{s/\beta_2}$ and $z = 1/(q^2 + 1)$. Thus, by comparing the terms we obtain

$$\begin{aligned}
a'_{eff} &= \frac{\sqrt{D}}{k'_{eff3D}}\frac{2k'_{a3D}P_a b}{\sqrt{\beta_2}}[P_a\{P(s_0|s_0;1) - 1\} + 1]^{-2} \\
&= \frac{\sqrt{2/3}bP_a l}{1 + P_a/(1/F(1) - 1)},
\end{aligned} \tag{3.89}$$

where

$$b = \lim_{q \to 0} q \frac{d}{dz} P(s_0|s_0; z) = \frac{3\sqrt{3}}{4\pi}. \tag{3.90}$$

Applying the definitions of reaction acceptance probability (3.35) and voxel size (3.95), we obtain

$$a'_{eff} = \frac{k_{a3D} R}{k_{a3D} + k_D}. \tag{3.91}$$

Note that the corresponding time domain form of Equation (3.85) is given as

$$k'_{3D}(t) \approx k'_{eff3D} \left[1 + a'_{eff} \sqrt{\pi D t} + \ldots \right]. \tag{3.92}$$

Hence, the long-time behavior of the lattice rate coefficient follows the same form as the continuum theory in Equation (2.13):

$$k'_{3D}(t) \approx k'_{eff3D} \left[1 + \frac{k_{a3D} R}{(k_{a3D} + k_D)\sqrt{\pi D t}} \right]. \tag{3.93}$$

3.2.2.3 Comparison with Continuum-based Theory

Since the effective rate on lattice in Equation (3.60) has the same form as Equation (2.17) of the continuum case, we can match the lattice-based and continuum-based theory by equating the initial rate and total rebinding probability: $k'_{a3D} = k_{a3D}$ and $G_{reb} = P_{reb}$. With the matching in initial rate constants, we obtain the expression for the reaction acceptance probability as

$$P_a = \frac{k_{a3D}}{3\sqrt{2} D l}. \tag{3.94}$$

By matching the total rebinding probability, $G_{reb} = P_{reb}$, the voxel size is found to be about 2% greater than the molecule size:

$$l = \frac{4\pi R}{3\sqrt{2}(\frac{1}{F(1)} - 1)} \approx 1.0209R. \tag{3.95}$$

The Spatiocyte scheme is thus guaranteed to have the same effective rate and total rebinding probability as the continuum framework provided that Equations (3.94) and (3.95) are satisfied. In addition, the expression for lattice effective rate constant follows the same form as in the continuum-based framework:

$$\begin{aligned} k'_{eff3D} &= \frac{k'_{a3D}k'_D}{k'_{a3D} + k'_D} = k'_D G_{reb}, \\ G_{reb} &= \frac{1}{1 + k'_D/k'_{a3D}}, \end{aligned} \tag{3.96}$$

where $k'_D = 3\sqrt{2}lD(1/F(1) - 1)$.

According to Equation (3.95), accurate matching of both the effective rate and the total rebinding probability requires the voxel size to be larger than the molecule size. Nonetheless, we can fix the voxel size to be the same as the molecule size, $l = R$. In this case, it is still possible to match the lattice effective rate coefficient to the continuum-based one by setting the reaction acceptance probability to

$$P_a = (1/F(1) - 1)\left[\frac{3\sqrt{2}(k_{a3D} + k_D)(1/F(1) - 1)}{4\pi k_{a3D}} - 1\right]^{-1}. \tag{3.97}$$

However, this is done at the expense of losing accuracy in the total rebinding probability, since $G_{reb} \neq P_{reb}$.

3.2.2.4 Generalization for Other Lattice Arrangement

A more general expression for the voxel size of any regular lattice arrangement is given as

$$l = \frac{4\phi_l}{\frac{1}{F(1)} - 1} R. \tag{3.98}$$

For example, for the simple cubic lattice we have the voxel length:

$$l = \frac{4\pi/6}{\frac{1}{0.340537} - 1} R = 1.081515R, \tag{3.99}$$

about 8% larger than the molecule size ($F(1)$ for the simple cubic lattice is given in pp. 153 of (Hughes, 1995)).

3.2.2.5 Reversible Reaction

To model the reversible reaction $A + B \underset{k_{d3D}}{\overset{k_{a3D}}{\rightleftharpoons}} C$ accurately, the local detailed balance need to be obeyed. This can be achieved on lattice by choosing a lattice dissociation rate constant k'_{d3D} from the following equilibrium constant relation:

$$K_{eq} = \frac{k'_{a3D}}{k'_{d3D}} = \frac{k_{a3D}}{k_{d3D}}. \tag{3.100}$$

The MLM method can simulate the dissociation reaction as a first-order process with rate k'_{d3D} and place the dissociated pair of molecules at an in-contact condition.

3.2.3 2D Reaction

We begin with the derivation of initial rate constant on 2D lattice. The number of reactions occurred in interval t' according to the continuum-based framework is given as

$$\Delta N = \frac{k'_{a2D} N_A N_B t'}{S}, \tag{3.101}$$

where S is the surface area.

The number of reaction in a step interval t' on lattice can be estimated as

$$\Delta N = \frac{P'_a N_B N_A}{N_{sv}}, \tag{3.102}$$

where $N_{sv} = S\phi_l/(\pi l^2/4)$ is the number of surface voxels and ϕ_l is the packing fraction for the lattice, and $P'_a = P_a\alpha$ is the actual reaction probability during the encounter.

In the activation-limited scheme, we have $t' = t_d$ and $P'_a = P_a$. Thus, we have

$$\begin{aligned} k'_{a2D} &= \frac{P_a S}{t_d N_{sv}}, \\ &= \frac{\pi P_a D}{\phi_l}. \end{aligned} \tag{3.103}$$

Note that the physical unit of k'_{a2D} is $[\mathrm{L}^2\mathrm{T}^{-1}]$, consistent with the continuum rate constant. For triangular lattice we have $\phi_l = \pi\sqrt{3}/6$ and $N_{sv} = 2S/(\sqrt{3}l^2)$. Thus the lattice initial rate is given as

$$k'_{a2D} = 2\sqrt{3} P_a D, \tag{3.104}$$

valid for both the activation-limited and diffusion-influenced schemes.

As for square lattice, we have $\phi_l = \pi/6$, and the initial rate is

$$k'_{a2D} = 6P_a D. \tag{3.105}$$

In homodimerization reaction $A + A \xrightarrow{k'_a} C$, the number of reactions according to continuum framework is

$$\Delta N_C = \frac{k'_{a2D} N_A (N_A - 1) t'}{S}, \tag{3.106}$$

whereas on lattice we have

$$\Delta N_C = \frac{P'_a N_A (N_A - 1)}{2N_{sv}}. \tag{3.107}$$

From these two equations, the lattice rate constant is derived as

$$\begin{aligned} k'_{a2D} &= \frac{P_a S}{2t_d N_{sv}}, \\ &= \frac{\pi P_a D}{2d}, \end{aligned} \tag{3.108}$$

which differs from Equation (3.103) by a factor of 2. As for the triangular lattice, the lattice rate constant is given as

$$k'_{a2D} = \sqrt{3} P_a D, \tag{3.109}$$

in which the relative diffusion coefficient D is defined as the sum of the two diffusion coefficients D_A.

Now we derive the long-time rate coefficient for 2D lattice. The methods used is generalized for any regular lattice arrangement, but we focus on the triangular lattice since it is used to simulate surface-surface reactions in the Spatiocyte scheme. Similar to the 3D case, the derivation for activation-limited ($k_{a2D} \ll D$) and diffusion-influenced ($k_{a2D} \gg D$) reactions is treated separately.

3.2.3.1 Activation-limited Case ($k_{a2D} \ll D$)

In the activation-limited scheme, we start the derivation with the generating function for the rebinding-time probability distribution $H_n(s_0|s_1)$ defined in Equation (3.9). First we obtain an expression for Equation (3.9) in term of $P(s_0|s_0;z)$:

$$\begin{aligned} H(s_0|s_1;z) &= \frac{P_a F(s_0|s_0;z)}{z + F(s_0|s_0;z)(P_a - 1)} \\ &= \frac{P_a F(s_0|s_0;z)}{z[1 - F(s_0|s_0;z)(1-P_a)/z]} \\ &= \frac{P_a}{z}\left\{F(s_0|s_0;z) + \frac{(1-P_a)}{z}F(s_0|s_0;z)^2 + \left[\frac{(1-P_a)}{z}\right]^2 F(s_0|s_0;z)^3 + \ldots\right\}. \end{aligned} \tag{3.110}$$

Write $F(s_0|s_0;z) = 1 - 1/P(s_0|s_0;z)$ as $1-x$ and $q = 1 - P_a$, we have

$$H(s_0|s_1;z) = \frac{P_a}{z}\left\{1 - x + \frac{q}{z}(1-x)^2 + \left[\frac{q}{z}\right]^2 (1-x)^3 + \ldots\right\}, \tag{3.111}$$

in which regular z terms are neglected since $z = 1$.

Finally, by rearranging the generating function in terms of x, we obtain

$$\begin{aligned}
H(s_0|s_1;z) &= P_a\left\{1+q\left[1+q+q^2+\ldots\right]-x\left[1+2q+3q^2+\ldots\right]\right. \\
&\quad \left.+x^2q\left[1+3q+6q^2+\ldots\right]+\ldots\right\} \\
&= P_a\left\{1+\frac{q}{1-q}-x\sum_{n=1}^{\infty}nq^{n-1}+x^2q\sum_{n=1}^{\infty}\frac{n(n+1)}{2!}q^{n-1}+\ldots\right\} \\
&= P_a\left\{\frac{1}{1-q}-\frac{x}{(q-1)^2}-\frac{x^2q}{(q-1)^3}-\frac{x^3q^2}{(q-1)^4}-\ldots\right\} \\
&= P_a\left\{\frac{1}{1-q}-\frac{x}{(q-1)^2}\frac{1}{1-\frac{xq}{q-1}}\right\} \\
&= 1-\frac{x}{P_a\left[1+\frac{x(1-P_a)}{P_a}\right]}.
\end{aligned} \tag{3.112}$$

The generating function $P(s_0|s_0;z)$ for the triangular lattice in asymptotic form is given in Equation (2.60) as

$$P(s_0|s_0;z)\approx\frac{\sqrt{3}}{2\pi}\ln[12(1-z)^{-1}]\{1+O(1-z)\}. \tag{3.113}$$

By substituting Equation (3.113) into Equation (3.112), we obtain the following approximated form:

$$H(s_0|s_1;z)\approx-\frac{b_1}{P_a}\left\{\ln\left(\frac{E}{1-z}\right)\right\}^{-1}, \tag{3.114}$$

where $E=12\exp\{b_1(1/P_a-1)\}$ and $b_1=2\pi/\sqrt{3}$. We then apply singularity analysis (Figure VI.4 of (Flajolet & Sedgewick, 2009)) on Equation (3.114) to obtain the large n behavior:

$$H_n(s_0|s_1)\approx\frac{2\pi}{\sqrt{3}P_a}\frac{1}{n}\left(\frac{1}{(\ln En)^2}-\frac{2\gamma}{(\ln En)^3}+\frac{3\gamma^2-\frac{\pi^2}{2}}{(\ln En)^4}+\ldots\right). \tag{3.115}$$

Using Equation (3.115), we evaluate the discrete sum in Equation (3.3) using the Euler-Mascheroni formula together with the definition of recurrence in 2D random walk: $\sum_{n=0}^{\infty} H_n = 1$. The solution in high order m terms is given by:

$$k_m^{2D} = \frac{2\pi k'_{a2D}}{\sqrt{3}P_a}\left[\frac{1}{\ln Em} - \frac{\gamma}{(\ln Em)^2} + \frac{\gamma^2 - \frac{\pi^2}{6}}{(\ln Em)^3} + \ldots\right]. \tag{3.116}$$

Finally, we apply the definition of initial rate for the triangular lattice and the relation $ml^2 = 4Dt'$ to obtain the time-dependent rate coefficient:

$$k'_{2D}(t) = 4\pi D\left[\frac{1}{\ln C_l t} - \frac{\gamma}{(\ln C_l t)^2} + \frac{\gamma^2 - \frac{\pi^2}{6}}{(\ln C_l t)^3} + \ldots\right], \tag{3.117}$$

where $C_l = 48D\exp\{b_1(1/P_a - 1)\}/l^2$ and l is the voxel size.

3.2.3.2 Diffusion-influenced Case ($k_{a2D} \gg D$)

In the derivation of diffusion-influenced scheme, again we work with the Laplace form for the rate coefficient:

$$\hat{k}'_{2D}(s) = k'_{a2D}[1 - \hat{G}(s)]/s, \tag{3.118}$$

Here $\hat{G}(s)$ is the Laplace form of the rebinding-time probability density on lattice, defined in Equation (3.79) as

$$\hat{G}(s) = \beta_1[s + \beta - sF_1(s_1|s_1)z - \beta_2 F(s_1|s_1; z)]^{-1}, \tag{3.119}$$

where

$$F(s_1|s_1; z) = 1 - \frac{z^2 P_1(s_0|s_1)}{P(s_0|s_0; z) - 1}, \tag{3.120}$$

$P_1(s_0|s_1) = 1/6$, $F_1(s_1|s_1) = 1/3$, $P_1(s_2|s_1) = 1/2$, $z = \beta_2/(s+\beta_2)$, $\beta = \beta_1 + \beta_2$, $\beta_1 = P_a/6t_d$ and $\beta_2 = 1/t_d$. $t_d = l^2/2dD_x$ here refers to the average time interval needed for a molecule with diffusion coefficient D_x to hop across one voxel. By applying the final value theorem, we get the asymptotic form for Equation (3.78) as

$$\begin{aligned}\lim_{s\to 0} s\hat{k}'_{2D}(s) &= k'_{a2D}[1 - \lim_{s\to 0}\hat{G}(s)] \\ &= k'_{a2D}\left[1 - \frac{\beta_1}{\lim_{s\to 0}[s+\beta - sF_1(s_1|s_1)z - \beta_2 F(s_1|s_1;z)]}\right] \\ &= k'_{a2D}\left[1 - \left(1 + \frac{\beta_2}{\beta_1}\lim_{z\to 1}\frac{z^2/6}{P(s_0|s_0;z)-1}\right)^{-1}\right].\end{aligned} \tag{3.121}$$

Finally, by taking the small z expansion together with Equation (3.113), we obtain the asymptotic rate coefficient expression:

$$\begin{aligned}\lim_{s\to 0} s\hat{k}'_{2D}(s) &= \frac{2\pi k'_{a2D}}{P_a\sqrt{3}}\left\{\ln\left[\frac{12\exp\{2\pi(1/P_a - 1)/\sqrt{3}\}}{1-z}\right]\right\}^{-1} \\ &= \frac{4\pi D}{\ln\left[E(1+\frac{4D}{l^2 s})\right]} \\ &\approx \frac{4\pi D}{\ln(C_l/s)}.\end{aligned} \tag{3.122}$$

3.2.3.3 Comparison with Continuum-based Theory

By comparing the lattice and continuum rate coefficient, we found that the asymptotic expression in Equation (3.122) for the diffusion-influenced scheme is the same as its continuum counterpart shown in Equation (2.23), while the time domain expression in Equation (3.117) for the activation-limited scheme is consistent with the continuum counterpart shown in Equation (2.24). To match the lattice and continuum rates, we need to impose the equality $C_l = C_c$. It then implies that the reaction probability should be

chosen as

$$P_a = \left[1 + \frac{\sqrt{3}}{2\pi}\left(\ln(f^2/12) + \frac{4\pi D}{k_{a2D}} - 2\gamma\right)\right]^{-1}, \tag{3.123}$$

where $f = l/R$ denotes the ratio of voxel to the molecule size. Since probability P_a is positive, it sets an additional constraint:

$$\ln f + \frac{2\pi}{\kappa} > -\frac{\pi}{\sqrt{3}} + \frac{\ln 12}{2} + \gamma = C_1 \tag{3.124}$$

To satisfy the last inequality, $f = l/R$ has to be adapted according to the value of κ. Since κ is always positive, we only need to set a lower bound expression for the voxel size:

$$\begin{aligned} \ln f &> C_1 - \frac{2\pi}{\kappa} > C_1, \\ f &> \exp(C_1), \\ l &> \exp(C_1)R \approx 1.005887R. \end{aligned} \tag{3.125}$$

In 3D MLM, accurate reaction kinetics requires the voxel size of HCP lattice to be larger than the molecule by $l \approx 1.02086R$. If an HCP lattice volume compartment is bounded by a triangular lattice surface, the 3D voxel size condition would still satisfy Equation (3.125). Therefore, all surface and volume voxels in the model can adopt the same HCP voxel size.

3.2.3.4 Generalization for Other Lattice Arrangements

The expression of MLM parameter derived for triangular lattice can be generalized to other lattice arrangements that adopt the MLM scheme. In general, the variable C_l in Equation (3.117) takes the form of

$$C_l = 4b_2 D \exp\left\{(1/P_a - 1)/b_1\right\}/l^2, \tag{3.126}$$

where b_1 and b_2 are coefficients present in the highest order term of the generating function $P(s_0|s_0; z)$:

$$P(s_0|s_0; z) \approx b_1 \ln\left(\frac{b_2}{1-z}\right). \tag{3.127}$$

On the other hand, the reaction probability has the following general form:

$$P_a = \left[1 + b_1\left(\ln(1/b_2) + \frac{4\pi D}{k_{a2D}} - 2\gamma\right)\right]^{-1}. \tag{3.128}$$

The expression for the probability has the following constraint on the voxel size:

$$l > \exp\left(\gamma - \frac{1}{2b_1} + \frac{\ln b_2}{2}\right) R. \tag{3.129}$$

Here as an example, we consider the square lattice, a popular lattice choice to simulate surface reactions. The coefficients for square lattice are given as $b_1 = 1/\pi$ and $b_2 = 8$ (Equation (A.187) in (Hughes, 1995)). The corresponding reaction probability is

$$P_a = \left[1 + \frac{1}{\pi}\left(\ln(1/8) + \frac{4\pi D}{k_{a2D}} - 2\gamma\right)\right]^{-1}, \tag{3.130}$$

with the voxel size constrained by

$$l > 1.04722R. \tag{3.131}$$

Therefore, to recapitulate the correct continuum rate, the voxel size in square lattice has to be about 5% larger than the molecule size. This voxel size is substantially larger than the 0.6% required by the triangular lattice. The different voxel size requirements reflect the influence of lattice arrangement on the first-passage time behavior and emphasize the

importance of choosing the right MLM parameters to generate accurate reaction kinetics.

3.2.3.5 Reversible Reaction

Accurate simulation of reversible reaction $A + B \underset{k_{d2D}}{\overset{k_{a2D}}{\rightleftharpoons}} C$ according to the SCK model needs to satisfy the local detailed balance. This is achieved in MLM by adopting a rate constant k'_{d2D} for the dissociation reaction such that the relation

$$\frac{k'_{a2D}}{k'_{d2D}} = \frac{k_{a2D}}{k_{d2D}}, \tag{3.132}$$

is satisfied.

3.2.4 1D Volume-surface Adsorption

We begin by formulating the 1D lattice rate coefficient according to the SCK model and then apply the rate expression to the problem of volume-surface adsorption.

3.2.4.1 1D Reaction

Again, we start with the derivation of initial lattice rate constant on 1D space and continue with the time-dependent lattice rate coefficient. To derive the initial lattice rate constant, we consider the number of reactions in interval t' according to the continuum framework as

$$\Delta N = \frac{k'_{a1D} N_A N_B t'}{L}, \tag{3.133}$$

where L denotes the length of the 1D system.

To be compatible with the continuum framework, we have the following assumptions in the derivation of the lattice rate constant: (i) each voxel can accommodate more than one molecule; (ii) molecules A are static whereas molecules B are mobile with relative

diffusion coefficient D. Then the number of reactions occured in a simulation step on the lattice is approximately

$$\Delta N = \frac{P_a N_A N_B}{N_L}, \tag{3.134}$$

where $N_L = L/l$ denotes the number of lattice voxels in length L. Finally, the initial lattice rate constant is given by

$$\begin{aligned} k'_{a1D} &= \frac{P_a l}{t_d}, \\ &= \frac{2DP_a}{l}, \end{aligned} \tag{3.135}$$

with unit of $[\mathrm{LT}^{-1}]$.

In 1D lattice, the generating function for the voxel occupancy probability from origin to origin is defined as (Montroll & Weiss, 1965)

$$P(s_0|s_0; z) = \frac{1}{\sqrt{1-z^2}}. \tag{3.136}$$

The corresponding first passage time distribution in generating function form is given by

$$F(s_0|s_0; z) = 1 - \sqrt{1-z^2}. \tag{3.137}$$

Substituting Equation (3.137) into Equation (3.9) yields the generating function for the rebinding-time probability distribution

$$H(s_0|s_1; z) \approx \frac{-\sqrt{2(1-z)}}{P_a}, \tag{3.138}$$

where we consider only the highest order term $\sqrt{1-z}$ in the limit of $z \to 1$.

The corresponding large n coefficient is obtained from the generating function according to the rule given in Figure VI.4 of (Flajolet & Sedgewick, 2009) as:

$$H_n(s_0|s_1) \approx \frac{1}{P_a\sqrt{2\pi n^3}}. \tag{3.139}$$

Applying Equation (3.139) to the Noyes' rate formula in Equation (3.3), we obtain the asymptotic form for the 1D rate coefficient:

$$k_m^{1D} \approx k'_{a1D}\frac{\sqrt{2}}{P_a\sqrt{\pi m.}}, \tag{3.140}$$

Using the definitions of initial lattice rate constant and the 1D simulation step size $ml^2 = 2Dt$, we have the rate expression as a function of time:

$$k'_{1D}(t \to \infty) \approx 2\sqrt{\frac{D}{\pi t}}. \tag{3.141}$$

Note that Equation (3.141) shares the same time-dependent form as the continuum-based theory given in Equation (2.35).

3.2.4.2 Adsorption

To derive the initial lattice rate for adsorption on lattice, we consider a cuboid compartment with an adsorbing plane in the middle. Let molecules A diffuse in the bulk with diffusion coefficient D_A and adsorption occurs on both sides of the plane.

According to the continuum theory, the number of adsorbed molecules in time step t' is approximated by

$$\Delta N_s = \frac{2k'_{sa}N_A t' S}{V}, \tag{3.142}$$

where N_s is the number of molecules adsorbed, N_A is the initial number of molecule in the bulk, S is the area of the plane and V is the volume of the cuboid compartment.

In the case of HCP lattice arrangement, the number of adsorption to the plane is

$$\Delta N_s = P_a \frac{2N_{sv}}{N_v} \frac{3}{12} N_A, \tag{3.143}$$

where N_{sv} is the number of surface voxel (triangular lattice), $N_v = \sqrt{2}V/l^3$ is the number of volume voxel (HCP lattice) and P_a is the reaction probability. Note that 3/12 is the probability that a molecule adjacent to the plane hops to the plane in one step and $2N_s/N_v$ is the probability of a randomly distributed molecule A adjacent to the plane.

By equating these two expressions and solve for k_{sa}, we obtain

$$k'_{sa} = \frac{P_a l}{2\sqrt{6} t_d}. \tag{3.144}$$

Finally, with the diffusion time step definition $t_d = l^2/6D_A$, the initial lattice adsorption rate constant is expressed as

$$k'_{sa} = \frac{\sqrt{3} P_a D_A}{\sqrt{2} l}, \tag{3.145}$$

where the unit is $[\mathrm{LT}^{-1}]$.

To obtain the time-dependent rate coefficient for volume-surface adsorption, definitions in Equation (3.145) and the 3D simulation step size relation $nl^2 = 6Dt$ are substituted into Equation (3.140). The resulting adsorption rate coefficient is given as

$$k'_{sa}(t \to \infty) \approx \frac{1}{2\sqrt{2}} \sqrt{\frac{D}{\pi t}}, \tag{3.146}$$

which shares the same long-time scaling behavior with the continuum-based theory in Equation (2.38) up to the same order.

In contrast to the 3D and 2D cases, the long-time expression for the 1D rate coefficient does not depend on the reaction probability and the voxel size. Since the long-time rate coefficient has the same form in both lattice and continuous spaces, we only need to match the initial lattice rate constant k'_{sa} with the adsorption rate constant k_{sa} in continuum.

3.3 Summary

We have derived the theoretical framework for MLM based on the Spatiocyte simulation scheme according to the SCK model. We described the difference in Spatiocyte scheme for simulating activation-limited and diffusion-influenced reaction. Based on these schemes, we have derived the general lattice theory for each case separately. The specific lattice theory for 3D, 2D and 1D reaction are derived based on the general lattice theory. We found that the lattice time-dependent rate coefficient and rebinding probability exhibit consistent analytical form as the continuum-based theory. This has allowed us to compare and match the lattice and continuum theory. From the matching of the two theories, we obtained the expression to determine the MLM parameters for recapitulating the continuum reaction-diffusion behavior. Particularly, we found that the voxel size is required to be larger than the actual molecular size. Lattice theory derived in this chapter will be validated in the next chapter. The MLM expression for simulation derived here will be used in the remaining of this thesis.

CHAPTER 4: RESULTS AND ANALYSIS

In this chapter, we validate the accuracy of lattice theory derived in Chapter 3 with numerical simulation and the continuum-based theory as defined in Chapter 2. We also verify the accuracy of Spatiocyte in simulating irreversible and reversible bimolecular reaction in 1D, 2D and 3D space. The performance of Spatiocyte in simulating diffusion and reaction is benchmarked against several particle-based simulation method introduced in Chapter 2.

4.1 Validation of Lattice Theory

The theoretical expressions for the rebinding probability and rate coefficient derived in Chapter 3 are validated with numerical simulation performed using Spatiocyte. Spatiocyte is included in E-Cell System version 4 (Kaizu et al., 2018), an open-source biochemical simulation environment that supports multiple algorithms, time scales and spatial representations. The Python notebooks used to generate the simulation results reported here are available at https://github.com/wxchew/MLM.

4.1.1 3D Reaction

4.1.1.1 Rebinding Probability

The theoretical rebinding-time probability distribution of a reactive pair initially in-contact, $H_n(s_0|s_1)$ and $G_n(s_0|s_1)$, are validated against numerical results. In the activation-limited case ($k_{a3D}/k_D \ll 1$), the expected first rebinding probability at nth step is obtained using Equation. (3.7), whereas in the diffusion-influenced case ($k_{a3D}/k_D \geq 1$), the probability is calculated from the generating function $G(s_0|s_1; z)$ (defined in Equation (3.33)) as

$$G_n(s_0|s_1) = \left[\frac{1}{n!}\frac{d^n}{dz^n}G(s_0|s_1; z)\right]\Bigg|_{z=0}. \tag{4.1}$$

Table 4.1: Theoretical and simulated rebinding-time probabilities on lattice for activation-limited and diffusion-influenced cases.

	$H_n(s_0\|s_1)$, $P_a = 0.5$			$G_n(s_0\|s_1)$, $P'_a = 2\alpha$, $\alpha = 1/2$		
n	**Theory**	**Simulation**	**Error (%)**	**Theory**	**Simulation**	**Error (%)**
1	0.0416666	0.0416586	0.019	0.1538461	0.1538326	0.009
2	0.0156250	0.0156228	0.014	0.0473373	0.0473431	0.012
3	0.0107784	0.0107779	0.005	0.0313306	0.0313317	0.004
4	0.0074297	0.0074274	0.031	0.0200584	0.0200534	0.026
5	0.0056802	0.0056773	0.049	0.0147588	0.0147496	0.062

Table (4.1) shows the simulated and the expected theoretical values for $n \in [1,5]$ steps. Simulation is performed with $l = 0.01$ μm, volume = $(100\ l)^3$ with periodic boundary, and number of iterations = 1×10^9. The simulation results agree well with the expected values, with discrepancies never exceeding 0.1%. Since the theoretical rebinding-time probability distribution on lattice is validated by simulations, the analytical formulas for the total rebinding probability derived from it, Equations (3.38) and (3.59) are therefore valid.

To illustrate the dependency of total rebinding probability on k_{a3D}/k_D, we obtained the probability at various k_{a3D}/k_D up to $n = 10$. Table 4.2 shows the simulated and the expected theoretical values for various k_{a3D}/k_D ratios. Simulation are performed with $l = 0.01$ μm, $D = 1$ μm^2/s, volume = $(10000\ l)^3$ with periodic boundary, number of iteration = 1×10^9 and $\alpha = 1/P_a$ for the diffusion-influenced cases. Both simulated and theoretical values coincide well, with discrepancies never exceeding 0.03%. Qualitatively, the total rebinding probability increases with larger k_{a3D}/k_D values, consistent with the continuum theory in Equation (2.16).

We then evaluated the rebinding-time probability distribution by recording the time

Table 4.2: Theoretical and simulated rebinding probabilities up to $n = 10$ steps with k_{a3D}/k_D ratios ranging from the highly activation-limited case ($k_{a3D}/k_D = 0.01$) to the strongly diffusion-influenced case ($k_{a3D}/k_D = 100$).

k_{a3D}/k_D	**0.01**	**0.1**	**1**	**10**	**100**
Lattice theory	0.0062657	0.05973879	0.397486	0.874154	0.985988
Simulation	0.0062672	0.05973410	0.397459	0.874126	0.986000
Discrepancy (%)	0.025	0.0078	0.0068	0.0032	0.0012

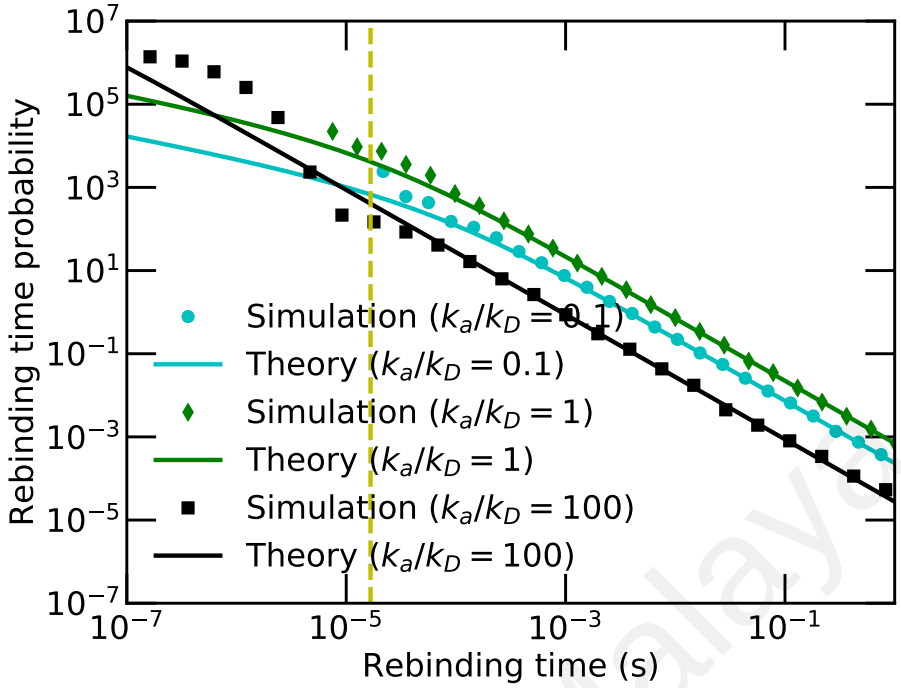


Figure 4.1: The rebinding time of a reactive pair that is initially in-contact. The rebinding time is sampled from simulations with $k_{a3D}/k_D = 0.1$, 1 and 100. Markers show the simulation results of Spatiocyte while solid lines depict the analytical results from the continuum-based theory in Equation (2.11). The vertical dashed line marks the characteristic diffusion time step, t_d.

taken for A and B to associate immediately after a dissociation event. We performed the simulations for a large number of steps and independent runs. The simulation parameters used are $l = 0.01$ μm, volume = $(10000\ l)^3$ with periodic boundary, runs = 10^4, $D_A = 1\mu\text{m}^2\text{s}^{-1}$, $D_B = 0$ and $\alpha = 1/P_a$ for the diffusion-influenced case. Figure 4.1 shows the average number of rebinding events per unit time at $k_{a3D}/k_D = 0.1, 1$ and 100. Lines depicting the rebinding-time probability distribution of the continuum-based theory according to Equation (2.11) are also shown as reference. It is clear that at times larger than t_d, the time-dependent behavior of lattice simulations is consistent with the continuum-based theory. The scaling behavior at long-time, $p_{reb}(R, t|R, 0)(t) \propto t^{-3/2}$ is a well-known characteristic of a 3D random walker returning to the origin (Pfluegl & Silbey, 1998). This result corroborate the asymptotic analysis performed in Subsection 3.2.2.1 paragraph (b).

Note that in the diffusion-influenced case ($k_{a3D}/k_D = 1$ and 100), finer step intervals generate rebindings at times smaller than the diffusion time step t_d, denoted by the vertical dashed line in Figure 4.1. In this temporal regime, MLM behaves differently from the continuum-based framework because the MLM reaction kinetics approximates the Poisson process. This is consistent with the analysis done in Subsection 3.2.2.2 paragraph (b). Despite the difference, the rebinding behavior correctly converges to the continuum-based formalism for times larger than t_d.

4.1.1.2 Irreversible Reaction

In this subsection, we evaluate the accuracy of MLM in simulating bimolecular reaction Equation (2.1) over various k_{a3D}/k_D regimes. We considered an immobile species A and a diffusing species B that are uniformly distributed at initialization with concentrations $[A]$ and $[B]$, respectively. Simulation is performed using Spatiocyte with expression derived in previous chapters. The parameters used in the simulation are the following: volume = $(3.5\ \mu\text{m})^3$ with periodic boundary, $R = 0.01\ \mu$m, $l = 0.01 \times 1.0209\ \mu$m, $D_A = 0$, $D_B = 1\ \mu\text{m}^2\text{s}^{-1}$, $N_a = N_b = 4000$, duration = 0.05 s, runs = 3×10^4 and $\alpha = 1/P_a$ for the diffusion-influenced case. We recorded the surviving fraction of A molecules at each time step.

Figure 4.2 displays the survival probability of A and the expected theoretical curve according to Equation (2.3). From the survival probability, we calculated the time-dependent reaction rate coefficient using (Equation (2.1) in (Szabo, 1989))

$$k(t) = -\frac{1}{[B]S_{irr,A}(t)}\frac{dS_{irr,A}(t)}{dt}. \tag{4.2}$$

We adopted the following discretization scheme for the time derivative to get the discrete

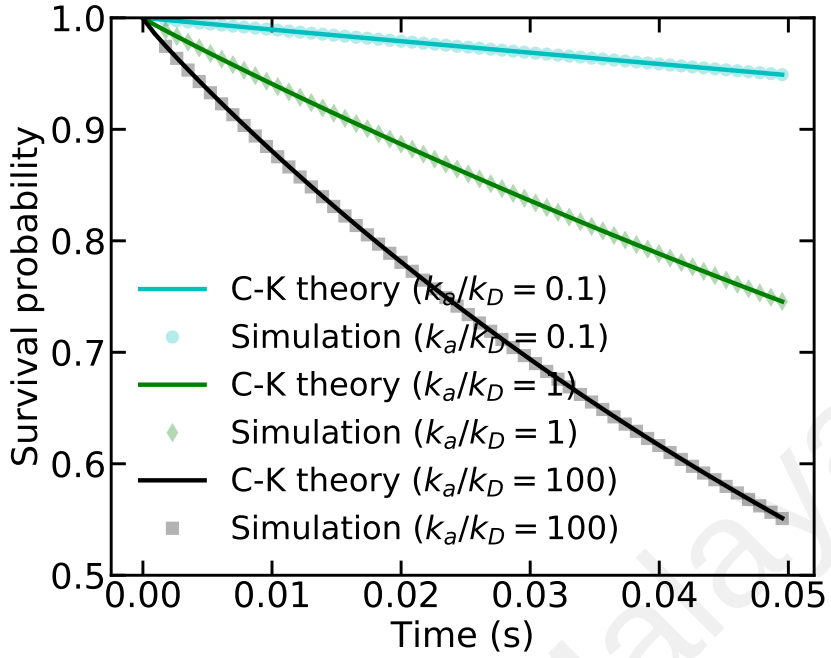


Figure 4.2: Simulated and theoretical Survival probability of A in $A + B \rightarrow B$ with k_{a3D}/k_D = 0.1, 1 and 100.

rate coefficient:

$$k_{n+1} = -\frac{S_{n+2} - S_n}{[B]S_{n+1}\,(t_{n+2} - t_n)}, \text{ for } n \in \mathbb{Z}^+, \tag{4.3}$$

where n is the index of the discretized S_A and t. The boundary cases are computed as

$$k_1 = -\frac{S_2 - S_1}{[B]S_1\,(t_2 - t_1)},\ k_N = -\frac{S_N - S_{N-1}}{[B]S_N\,(t_N - t_{N-1})}, \tag{4.4}$$

where N denotes the final time step. The reaction rate coefficient obtained for various k_{a3D}/k_D ratios are shown in Figure 4.3 along with their corresponding theoretical curves from Equation (2.13).

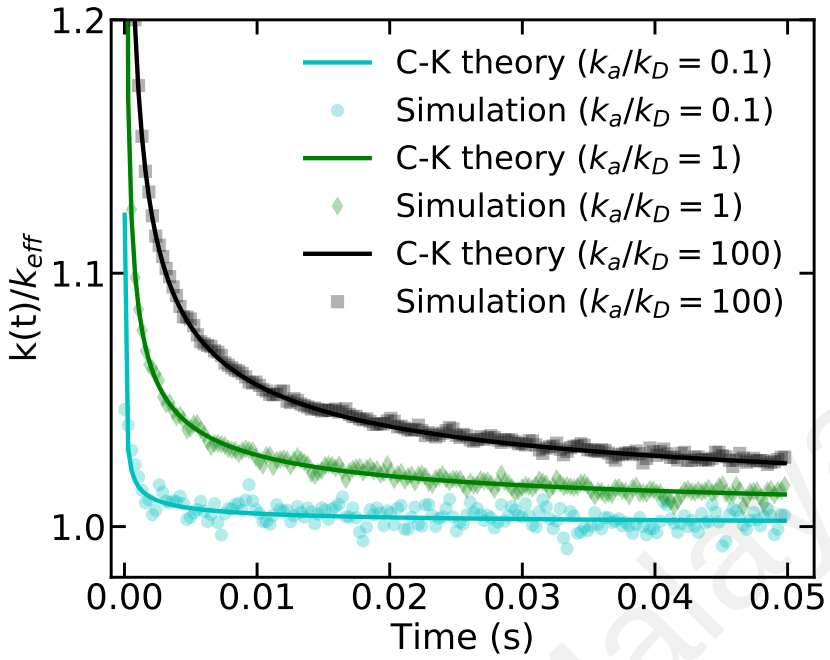


Figure 4.3: Simulated time-dependent rate coefficients of the reaction and the corresponding long-time approximation of SCK theory in Equation (2.13) for 3D reaction $A + B \rightarrow B$.

Recall that the long-time asymptotic variant of the SCK theory as shown in Equation (2.13) has the form

$$k_{3D}(t) \simeq C_1 \left(1 + \frac{C_2}{\sqrt{t}}\right), \tag{4.5}$$

where C_1 and C_2 denote the steady-state rate constant and the time-dependent term, respectively. We fitted Equation (4.5) to the numerical data, omitting early time points to avoid non-steady-state effects. Standard deviation of simulated survival probability data is then used as weight in the fitting procedure. The resulting C_1 and C_2 parameters after fitting are listed in Table 4.3. The theoretical values correspond favorably to the estimated steady-state reaction rate constants and are well within the standard error, thus validating the lattice theory for the effective rate.

Table 4.3: The steady-state rate constant C_1 and the time-dependent term C_2 of reaction (2.1) at various k_{a3D}/k_D were obtained by fitting the simulated reaction rate coefficient with Equation (4.5). Theoretical values from Equation (2.13) are listed for comparison.

k_{a3D}/k_D	**0.1**	**1**	**100**
Theoretical C_1 ($\mu m^3 s^{-1}$)	0.011424	0.062832	0.124420
Simulation	0.011423±0.0012	0.062848±0.0029	0.124459±0.0046
Discrepancy (%)	0.011	0.026	0.032
Theoretical C_2 ($s^{1/2}$)	0.00051	0.00282	0.00559
Simulation	0.00054±0.01	0.00279±0.0052	0.00563±0.004
Discrepancy (%)	5.5	1.04	0.77
m.s.e. of fit	3.4×10^{-7}	2.2×10^{-6}	4.2×10^{-6}

The time-dependent terms are also in good agreement with the theory, especially in the diffusion-limited case, with discrepancy less than 1%. This is consistent with the asymptotic analysis result in Equations (3.52) and (3.93) In the activation-limited case ($k_{a3D}/k_D = 0.1$), the fitted C_2 had the largest deviation from theory because the standard error was also the highest. The low number of data points contributed to the higher standard error.

4.1.2 2D Reaction

4.1.2.1 Irreversible Reaction

In this section, we verify the accuracy of the lattice rate coefficient $k'_{2D}(t)$ with the numerical simulation. We obtained the theoretical rate coefficient from the numerical inverse Laplace transform of Equation (3.121). We simulated the reaction in Equation (2.1) with Spatiocyte at both the activation-limited ($\kappa = 0.01 \cdot 4\pi$) and the diffusion-limited ($\kappa = 100 \cdot 4\pi$) regimes. Simulations were performed with the following parameters: Area = (6.5×6.5) μm^2, $R = 0.01$ μm, $l = 0.01 \times 1.0209$ μm, $D_A = 1$, $D_B = 0$ $\mu m^2 s^{-1}$, $N_a = N_b = 423$, duration = 0.2 s, logging interval=$10t_d$. The number of surviving A is logged and the discrete rate coefficient is calculated.

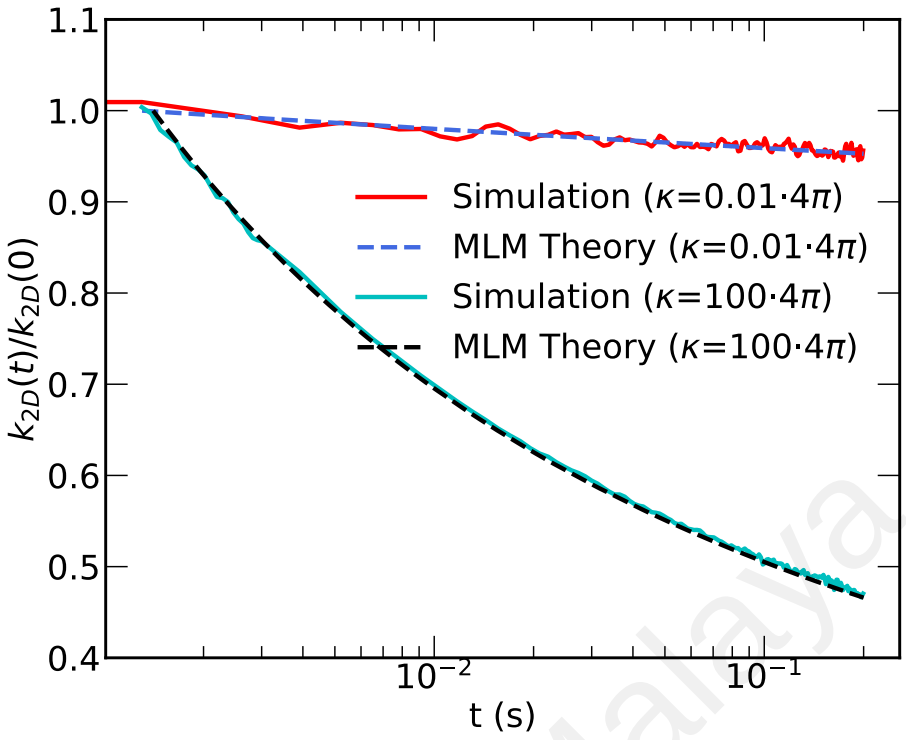


Figure 4.4: Simulated lattice time-dependent rate coefficients (solid lines) compared with MLM theory in Equation (3.121) (dashed lines) for surface-surface reaction $A + B \rightarrow B$.

The simulated rate coefficients shown in Figure 4.4 agree with the theoretical values from Equation (3.117) at the activation-limited ($\kappa = 0.01 \cdot 4\pi$) and diffusion-limited ($\kappa = 100 \cdot 4\pi$) regimes for $t \gg t_d$. For better visualization of the time-dependent behavior of the two extreme cases, the simulated and theoretical lines are normalized by the initial theoretical value.

Next, we compared the simulated survival probability of the same reaction with the continuum-based theory, where the values are numerically evaluated according to

$$S^{2D}_{irr,A}(t,[B]) = \exp\left[-[B]\int_0^t k_{2D}(\tau)d\tau\right]. \tag{4.6}$$

As shown in Figure 4.5, the simulated results overlap almost perfectly with the continuum-based theory, thus, confirming the accuracy of MLM.

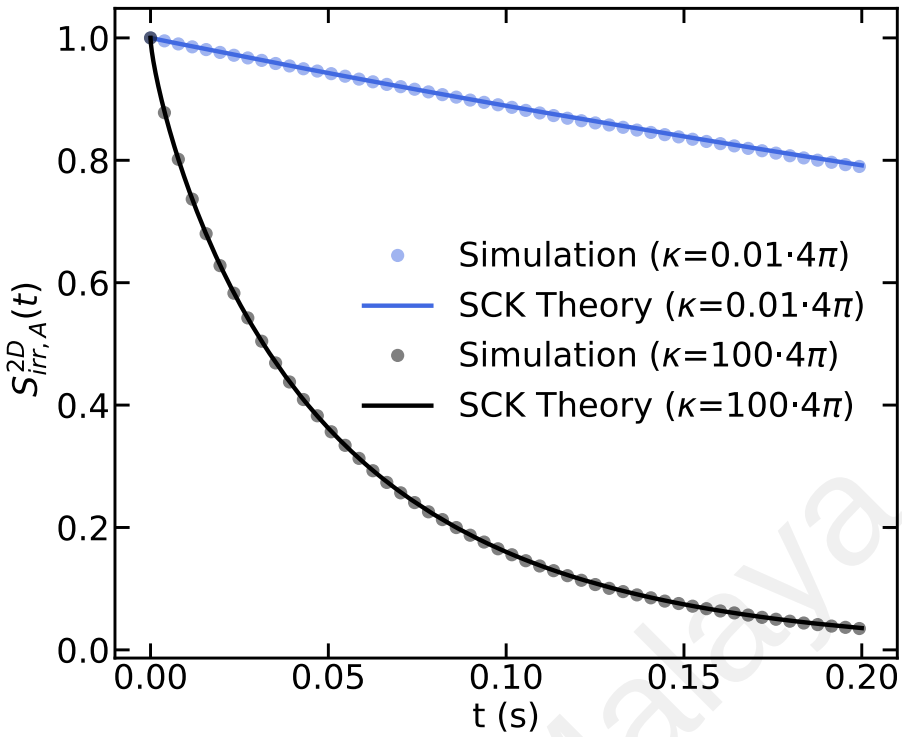


Figure 4.5: Simulated lattice survival probability of A (points) compared with continuum SCK theory in Equation (4.6) (solid lines) in surface-surface reaction $A + B \rightarrow B$. Activation-limited ($\kappa = 0.01 \cdot 4\pi$) and diffusion-limited ($\kappa = 100 \cdot 4\pi$) cases are indicated by the top and bottom lines respectively.

In a previous 2D reaction study on homogeneous lattice (Grima & Schnell, 2006), the diffusion-limited reaction kinetics follows $k_{2D}(t) \sim t^a$, where a is a constant. The kinetics of 2D activation-limited reaction on the other hand is generally described by the rate constant of classical reaction kinetics. As demonstrated by the theoretical and simulated lattice reaction kinetics in this work, $k_{2D}(t)$ of activation-limited reaction decays slowly ($\kappa = 0.04\pi$ in Figure 4.4). Therefore, up to a certain extent, the activation-limited $k_{2D}(t)$ can be approximated by the classical kinetics rate constant. For diffusion-limited reaction, however, the asymptotic logarithmic decay in $k_{2D}(t)$ is significantly more apparent. The behavior at the intermediate time can be well-approximated by $k_{2D}(t) \sim t^{-1/2}$, as shown previously in (Naqvi et al., 2000). In a recent work (Yogurtcu & Johnson, 2015), the logarithmic function of the asymptotic $k_{2D}(t)$ is further simplified to give an effective formula $k_{2D}(t) \sim t^c$, where c is a function of the intrinsic rate and the diffusion coefficient.

The simplified form of the time-dependent rate can be relatively accurate and convenient in describing 2D reaction kinetics. Nonetheless, care should be taken to ensure the assumption, the observation time regime and the physical parameter range are satisfied.

4.1.2.2 Reversible Reaction

We perform numerical simulations to confirm the ability of MLM to correctly reproduce the steady state and time-dependent behaviors in the reversible reaction. Association rates in the activation-limited ($\kappa = 0.01$) and diffusion-limited ($\kappa = 100$) cases were used in the simulation, while the dissociation rate k_{d2D} is set to be 10 times larger than the association rate. Detailed simulation parameters are $k_{d2D} = 10k_{a2D}$, surface area = (6.5×6.5) μm^2 with periodic boundary, $R = 0.01$ μm, $l = 0.01 \times 1.0209$ μm, $D_A = D_C = 0$ μm^{2}s^{-1}, $D_B = 1$ μm^{2}s^{-1}, $N_b = 20, N_b = 401$ and duration =10 s. Simulated result is compared with the MPK1 theory in Equation (2.31), obtained by numerical Laplace transform. The outcome shown in Figure 4.6 indicates good agreement between the simulation and theory for time ranging from t_d until equilibrium.

Since reactions in MLM take place stochastically, we examine the steady-state distribution of molecule number in the reversible homodimerization reaction, $A + A \underset{k_{d2D}}{\overset{k_{a2D}}{\rightleftharpoons}} B$, with association and dissociation rates, $k_{a2D} = 0.001$ $\mu m^2 s^{-1}$ and $k_{d2D} = 1$ s^{-1}, respectively. The simulation schemes of homodimerization and heterodimerization reactions are identical except in their reaction probabilities, which differ by a factor of two. We initialize the simulation with 169 A monomers in a compartment of unit volume with voxel size $l = 0.01$. The steady-state distribution of the number of B dimers obtained from the simulation is shown in Figure 4.7. The result of the simulation is consistent with the analytical solution from the chemical master equation (Equation (43) in (Cianci et al., 2016)). We then investigate the effects of volume exclusion on the steady-state distribution by increasing the voxel size, while retaining the number of molecules and the

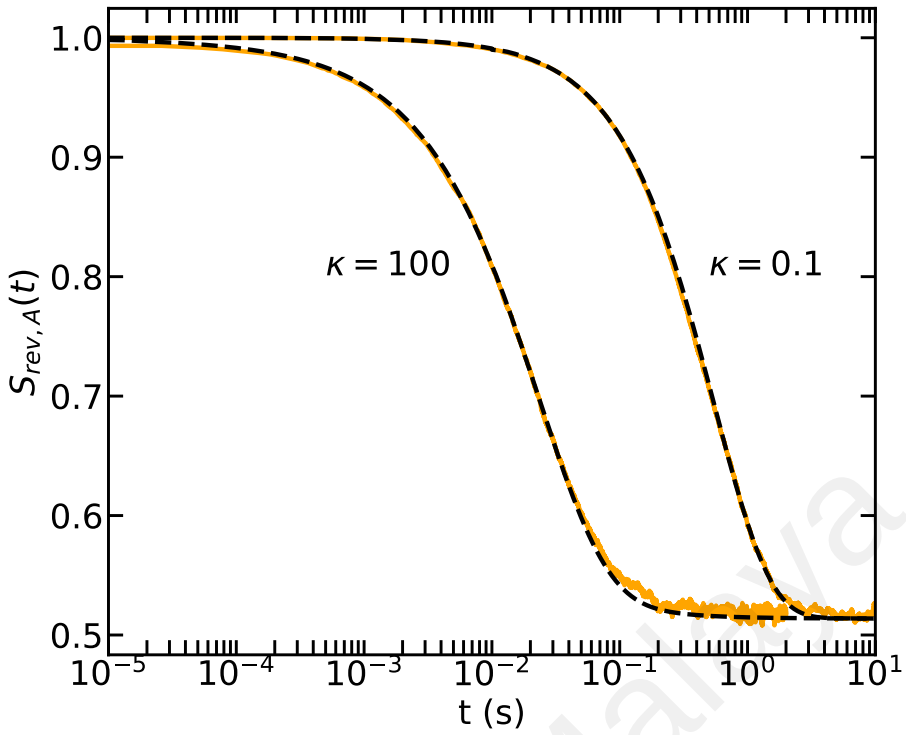


Figure 4.6: Survival probability of A in the surface-surface reaction $A + B \underset{k_{d2D}}{\overset{k_{a2D}}{\rightleftharpoons}} C$. Dashed curves are the values calculated according to the MPK1 theory given in Equation (2.31), solid lines are the simulation results of Spatiocyte. Association rates in the activation-limited ($\kappa = 0.1$) and diffusion-limited ($\kappa = 100$) regimes are chosen.

compartment size. With $l = 0.09$, the steady-state distribution shifts to the right. This result is qualitatively consistent with the crowding-influenced reaction as predicted by the thermodynamic theory (Zhou et al., 2008) and the master equation (Cianci et al., 2016). However, we note that the exact quantitative behavior simulated on a crowded lattice would not necessarily agree with the continuum description in all conditions because of the different degrees of molecule obstruction between continuum and lattice representations, as demonstrated in (Meinecke & Eriksson, 2016). It would require further analysis to exactly match the crowding-influenced on-lattice reaction kinetics with the continuum behavior, which is beyond the scope of the current work.

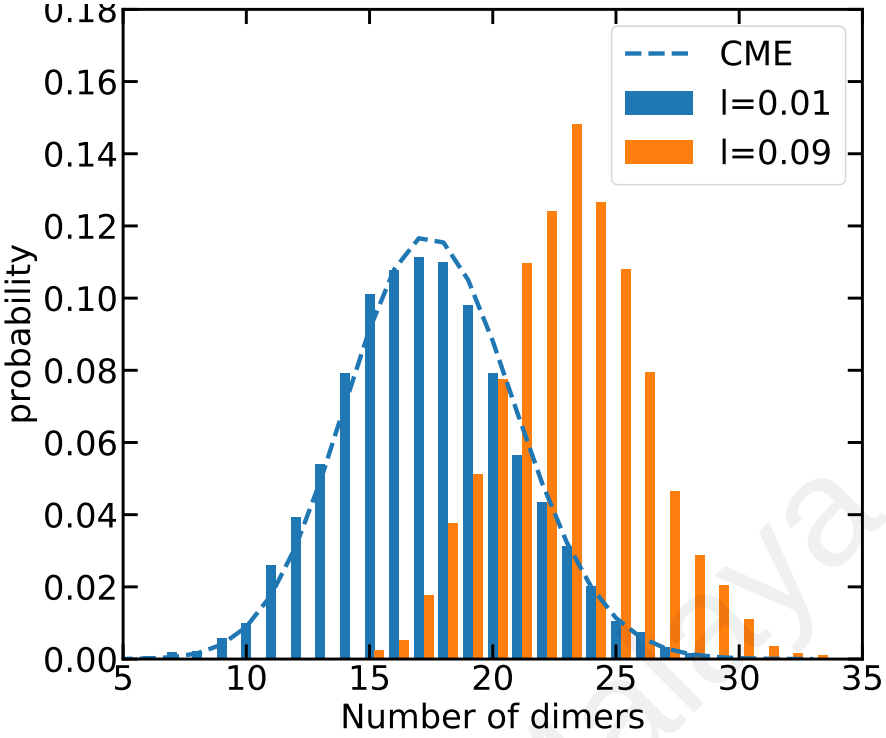


Figure 4.7: Steady-state number distribution of dimers from a reversible homodimerization reaction. The reaction is given by $A + A \underset{k_{d2D}}{\overset{k_{a2D}}{\rightleftharpoons}} B$, with $k_{a2D} = 0.001\ \mu m^2 s^{-1}$ and $k_{d2D} = 1\ s^{-1}$. The histogram on the left is simulated at an uncrowded condition with voxel size $l = 0.01\ \mu m$. Dashed line is the analytical solution of the chemical master equation (CME). Histogram on the right is obtained with the same parameters except with a larger voxel size, $l = 0.09\ \mu m$, resulting in a crowded compartment. The diffusion coefficient of A is 1 $\mu m^2 s^{-1}$, the length of the square compartment is 1 μm and the initial number of A is 169.

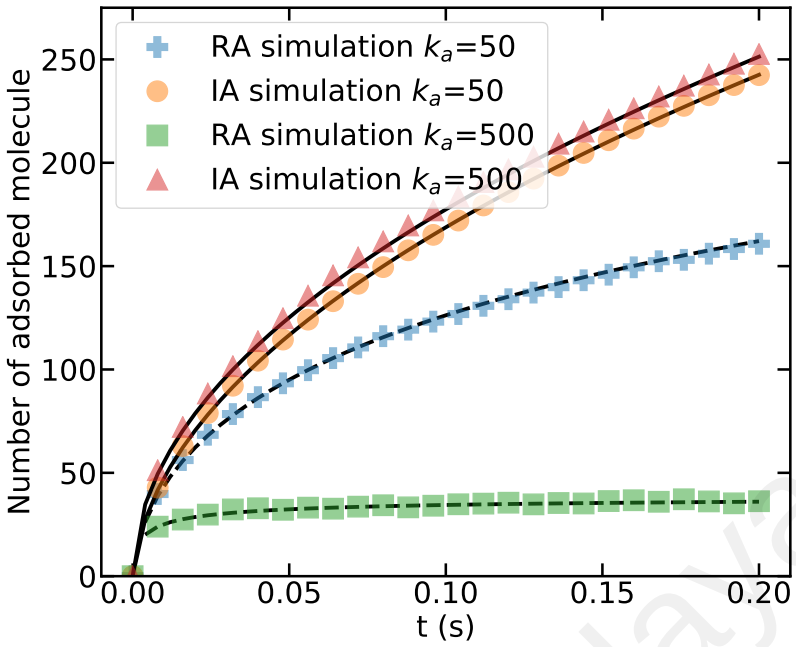
4.1.3 1D Volume-surface Adsorption

4.1.3.1 Irreversible and Reversible Reaction

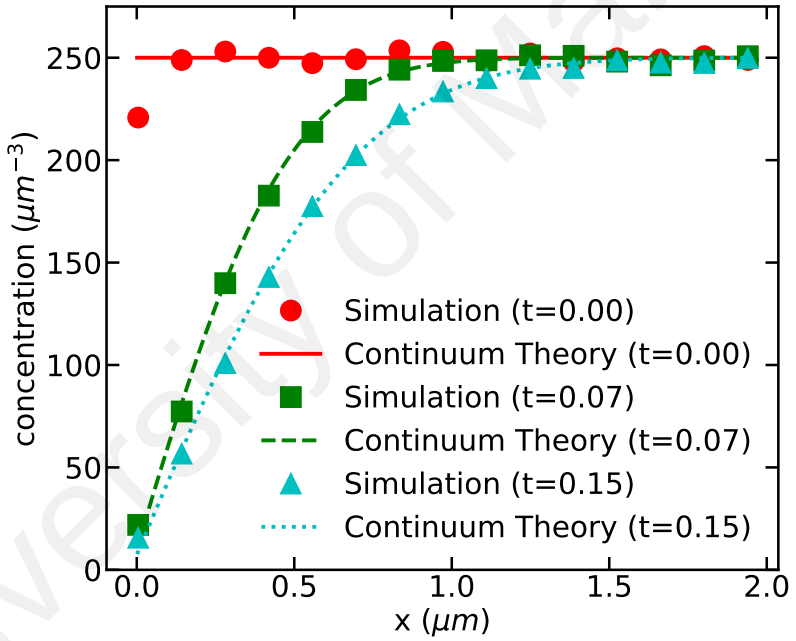
To examine the accuracy of MLM in simulating the adsorption kinetics, we performed Spatiocyte simulations using the derived expression for the reaction probability. We used large number of cytosolic A molecules in a cuboid compartment with a cross sectional area $(1\ \mu m)^2$ and length 4 μm. An adsorbing plane is placed in the middle of the cuboid compartment, allowing adsorption from both sides of the surface. Other simulation used are $l = 0.01\ \mu m$, $D_A = 1\ \mu m^2 s^{-1}$, and initial number of cytosolic molecule $N_a = 1000$. The number of adsorbed molecules at each time step is monitored.

Figure 4.8(a) shows the time series of A on the adsorbing plane for irreversible (adsorption-only) and reversible (adsorption and desorption) reactions. Simulated results agree well with the expected values according to the continuum theories for the irreversible reaction in Equation (2.36) and reversible reaction in Equation (2.40).

The good fit can be seen at both strongly ($k_{sa} = 500\ \mu\text{m}s^{-1}$) and weakly ($k_{sa} = 50\ \mu\text{m}s^{-1}$) adsorbing rates. To examine the spatial-temporal concentration profile, we counted the number of cytosolic molecules near the adsorbing plane in the irreversible adsorption. The resulting concentration profile along the axis perpendicular to the adsorbing plane are shown in Figure 4.8(b). The simulation results coincide very well with the curves of continuum-based theory in Equation (2.39).



(a)



(b)

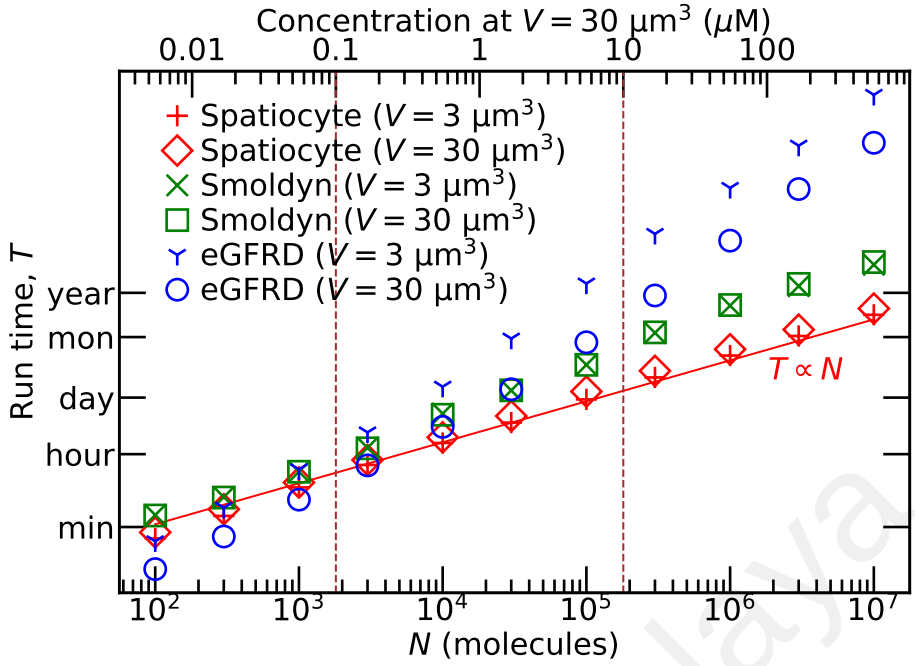
Figure 4.8: (a) Time series of adsorbed molecules simulated with irreversible (IA, triangle and circle markers) and reversible (RA, plus and square markers) adsorptions. In each case, strong (k_{sa} = 500 μms^{-1}) and weak (k_{sa} = 50 μms^{-1}) adsorption rates were tested. In the reversible adsorption, the membrane dissociation rates are k_{sd} = 62.5 and 6250 s^{-1}, corresponding to the association rates k_{sa} = 50 and 500 μms^{-1}, respectively. Solid and dashed lines represent the continuum-based values according to the irreversible and reversible reaction formulas in Equations (2.36) and (2.40), respectively. (b) The concentration profile of cytosolic A along the axis perpendicular to the adsorbing surface at x = 0 for the given time points. The adsorption is irreversible with the rate k_{sa} = 50 μms^{-1}. Theoretical lines shown are according to the continuum-based theory in Equation (2.39).

4.2 Performance of the Numerical Simulation

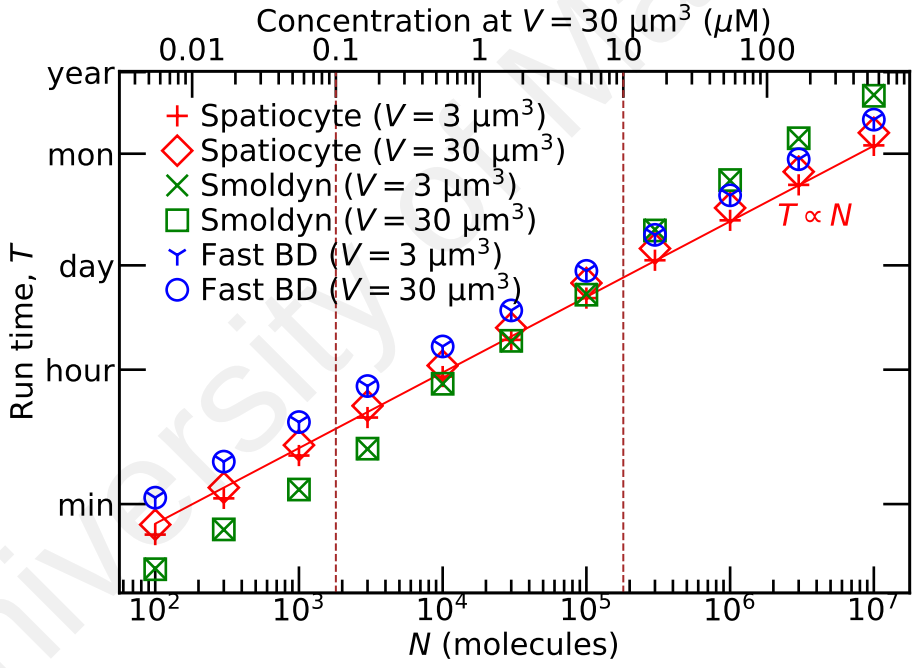
4.2.1 Diffusion

We compared the 3D diffusion performance of MLM using Spatiocyte (git 9757fb3) and three other continuum particle-based simulation methods, Smoldyn (Andrews et al., 2010) (version 2.55), eGFRD (Takahashi et al., 2010) (in E-Cell System version 4.1.4) and fast Brownian dynamics (BD) (Smith & Grima, 2017) (C++ program example in Spatiocyte git 9757fb3). Each model was simulated for a predefined run time, t_r and the resulting simulated time, t_s was recorded. We calculated T, the run time in seconds for 10 s of simulated time with $T = 10t_r/t_s$. t_r was set such that at least hundreds of simulation steps have been completed. The resulting range of t_r was between 1 hour to several days. All simulations were executed on the same server with Intel Xeon Platinum 8180 2.5 GHz (max 3.80 GHz) CPU, 768 GB memory and Ubuntu 18.04 LTS operating system. The performance benchmark models for all tested methods are included in Spatiocyte package (http://spatiocyte.org) as examples.

When the molecules are represented as hard-spheres with volume exclusion, Spatiocyte required shorter run times than Smoldyn in all cases (Figure 4.9(a)). Spatiocyte achieves comparable or better performance than eGFRD in the typical concentration range of cytoplasmic macromolecules (0.1 to 10 μM). For example at 6 μM in volume 30 μm^3, Spatiocyte is about 4.5 and 16 times faster than Smoldyn and eGFRD, respectively. In contrast to eGFRD, Spatiocyte and Smoldyn execution times increase with the number of molecules but not the molecular crowdedness (V = 30 μm^3 vs. 3 μm^3). The simulation times of Spatiocyte scale almost linearly with the number of molecules ($T \propto N$), which is not apparent with Smoldyn and eGFRD. The drastic slowdown of eGFRD at higher concentrations is caused by the shorter time steps required to resolve many molecular interactions that take place in the densely occupied system (Takahashi et al., 2010).



(a)



(b)

Figure 4.9: 3D diffusion performance of particle-based methods. Vertical axis, T shows the run times to diffuse molecules with diffusion coefficient $D_x = 1$ μm^{2}s^{-1} in volume V for 10 s. Bottom horizontal axis, N represents the number of diffusing molecules, while the top axis shows the corresponding concentration at $V = 30$ μm^3. (a) Molecules are represented as hard-sphere particles with radius $r = l/2 = 2.5$ nm. (b) Molecules are dimensionless point particles that can overlap one another. Solid lines depict the ideal scaling for Spatiocyte. Vertical dashed lines indicate the typical concentration range of proteins in the cytoplasm (0.1 to 10 μM).

If molecules are represented as dimensionless point particles, higher diffusion performance is expected since inter-molecular collisions can be ignored. Figure 4.9(b) shows the run times of Spatiocyte, Smoldyn and fast BD when diffusing point particles with the same simulation interval. eGFRD was not considered here since it only supports molecules with physical volume. Spatiocyte and fast BD execution times showed an almost linear scaling with the number of molecules. Although Smoldyn did not scale as well, it had the fastest run times when the number of diffusing molecules was 30,000 or less. Spatiocyte outperformed fast BD in all tests and is on average 2.5 times faster. As expected, the simulation times of all three methods were not affected by the crowdedness in the volume since molecular collisions are disregarded. On average, Spatiocyte takes about 2 times longer to diffuse hard-sphere molecules than point particles.

4.2.2 Reaction

Recently, Andrews (Andrews, 2018) benchmarked the performance of Smoldyn, MCell (Kerr et al., 2008), eGFRD, SpringSaLaD (Michalski & Loew, 2016) and ReaDDy (Schöneberg & Noé, 2013) particle simulators when running the well-known Michaelis-Menten enzymatic reaction. Smoldyn required the least amount of time to complete the benchmark. Running the model on our hardware (see Figure 4.10 for specifications) with the same 1 ms simulation interval, Spatiocyte took 113 s, whereas Smoldyn required 31 s. Since it would take too long for eGFRD to complete the simulation of the original model (Andrews, 2018), we decreased the number of molecules, diffusion coefficients and reaction rates. The execution times of Spatiocyte, Smoldyn and eGFRD when running the model with the new parameters are presented in Figure 4.10. The simulators generated almost identical results. Spatiocyte and Smoldyn had similar run times, whereas eGFRD required about one to two orders of magnitude longer. Although Spatiocyte is about four times slower than Smoldyn when executing the original model, both had very similar times

with the new parameters. Our results indicate that the relative performance of Spatiocyte and Smoldyn depends on the model parameters.

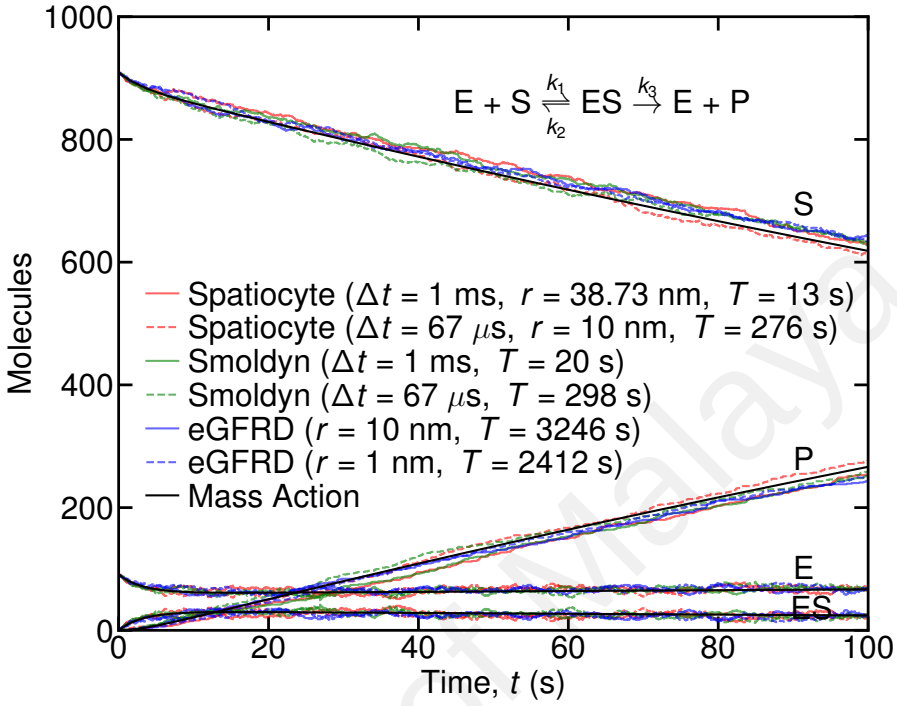


Figure 4.10: Particle simulation performance of the Michaelis-Menten reaction. Original benchmark model from (Andrews et al., 2010; Andrews, 2018) was modified with volume (V) = 90.9 μm^3, diffusion coefficient (D_x) = 1 μm^{2}s^{-1}, k_1 = 0.01 μm^{3}s^{-1}, $k_2 = k_3$ = 0.1 s^{-1}. Molecule or voxel radius (r), simulation or diffusion step interval (Δt) and run time (T) are as indicated. All simulations were executed on the same workstation with Intel Xeon X5680 3.33 GHz CPU, 48 GB memory and Ubuntu 16.04 LTS operating system.

4.3 Summary

We validated the lattice theory including the rebinding probability and reaction rate coefficient as derived in Chapter 3 with numerical simulation. We compared the simulated irreversible and reversible reaction kinetics with the continuum-based theory defined in Chapter 2. We observed consistent behavior in time-dependent kinetics and steady-state distribution between Spatiocyte's simulation result and the continuum-based theory for all spatial dimensions. In the performance comparison with selected particle-based simulation methods, Spatiocyte outperformed under certain parameter ranges. To further examine

the capability of MLM in simulating intracellular RD processes, we apply MLM with the same Spatiocyte configuration used in this chapter to more complicated and realistic scenario in next chapter.

CHAPTER 5: APPLICATIONS

We have validated the accuracy and consistency of MLM in solving diffusion-influenced bimolecular reaction in Chapter 4. In this chapter, we apply MLM to study several biological-inspired models including the production-degradation process, previously studied using lattice-based methods (Erban & Chapman, 2009; Sturrock, 2016; Cianci et al., 2016), and the dual phosphorylation-dephosphorylation cycle of the mitogen-activated protein kinase (MAPK) cascade (Chang & Karin, 2001; Ferrell & Bhatt, 1997; Aoki et al., 2011), a common motif found in signal transduction systems but with a response function that is highly sensitive to the binding kinetics. We study the influence of diffusion on the steady-state and the overall response behavior in these reactions. In order to investigate the influence of obstacle hindrance on diffusion and reaction, we study the mean-squared-displacement and the reaction rate of reactant molecules in the presence of immobile obstacles. We also investigate the contribution of 3D volume-surface reaction, 2D surface-surface reaction and 1D volume-surface adsorption to the overall surface reaction.

5.1 Production-degradation Process

Consider the production and degradation processes of protein A represented by a zero-order production coupled with a second-order degradation:

$$\emptyset \xrightarrow{k_1} A, \; A + B \xrightarrow{k_2} B. \tag{5.1}$$

The concentration of A will go through an initial transient state before settling down at a steady-state equilibrium, $[A] = k_1/(k_2[B])$ that is distributed according to the Poisson distribution (Erban & Chapman, 2009).

According to theory of diffusion-influenced reaction, the equilibrium concentration of A is given as

$$[A] = \frac{k_1}{k_{ss3D}[B]}, \tag{5.2}$$

where k_{ss3D} is the steady-state rate constant given by (Equation (4.5) in (Agmon & Szabo, 1990))

$$k_{ss3D} \approx k_{on3D} \left\{1 + \left[4\pi(R_{eff})^3[B]\right]^{1/2}\right\}, \tag{5.3}$$

with macroscopic rate constant $k_{on3D} = 4\pi D R_{eff}$ and the effective radius $R_{eff} = k_{a3D}R/(k_{a3D} + 4\pi RD)$.

To confirm if MLM can recapitulate the production-degradation process correctly, we simulated the process with Spatiocyte and compared the outcomes with eGFRD and the well-mixed model. Simulation are performed with following parameters: $k_1 = 0.1\ \mu\text{m}^{-3}\text{s}^{-1}$, $k_2 = 0.02\ \mu\text{m}^3\text{s}^{-1}$, $[B] = 1\ \mu\text{m}^{-3}$, runs = 700, duration > 10^4 s to achieve steady-state, and volume = 100 μm^3 with periodic boundary. To generate the results of the well-mixed model, we solved the rate equation using an ordinary differential equation (ODE) solver. The time-series of A is shown in Figure 5.1, while the equilibrium values are provided in Table 5.1. As evident from the figure and table, Spatiocyte results are all in good agreement with both the well-mixed model and eGFRD.

Recently, the Spatiocyte scheme was reported to not only fail to reproduce the expected equilibrium value of A but also generate different values depending on the voxel size (Sturrock, 2016). In the report, the effective bimolecular rate k_2 was used in the calculation of reaction acceptance probability instead of the intrinsic reaction rate constant k_{a3D}, which inevitably caused the deviation from the well-mixed model (see first row of Table 5.1).

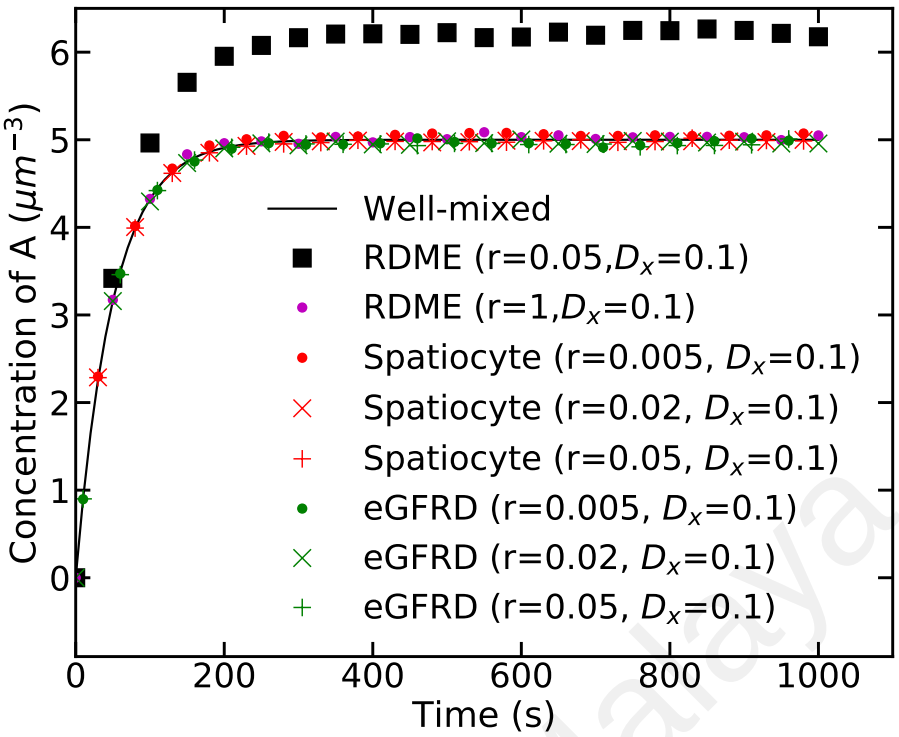


Figure 5.1: Time-series profile of A in production-degradation process showing simulated result with Spatiocyte (using intrinsic rate, k_{a3D}), eGFRD and RDME. $D_A = D_B = 0.1\ \mu\text{m}^2\text{s}^{-1}$, molecule radius $r_A = r_B = r \in \{0.005, 0.02, 0.05\}\ \mu$m. Note that r represents half of the subvolume size in RDME, and the actual molecule radius in Spatiocyte and eGFRD. For comparison, solid line shows the well-mixed model.

Table 5.1: Equilibrium concentration of A in Equation (5.1) simulated with Spatiocyte and eGFRD at different spatial discretizations. k_2 is the effective rate, k_{a3D} is the intrinsic rate, l is the voxel size, $K = 2^{1/6}L/l$ is the compartment length in number of voxels, while L denotes the actual length (Sturrock, 2016). At $l = 0.01$, $K = 521$; at $l = 0.04$, $K = 130$; and at $l = 0.1$, $K = 52$. The well-mixed equilibrium concentration is 5 μm^{-3}. Discrepancy (%) from the well-mixed model is shown in parenthesis.

Simulation scheme	l **= 0.01**	l **= 0.04**	l **= 0.1**
Spatiocyte with k_2	9.014 (80.28)	6.023 (20.46)	5.393 (7.86)
Spatiocyte with k_{a3D}	5.009 (0.18)	4.984 (0.32)	4.990 (0.2)
eGFRD	4.968 (0.64)	4.975 (0.5)	4.950 (1)

As shown in Figure 5.1 and Table 5.1, there was no discrepancy when the intrinsic rate k_{a3D} was used to compute the reaction acceptance probability given in Equation (3.94). The well-known relation between k_{a3D} and k_2 is given by Equation (2.14), wherein k_2 is represented by k_{eff}. Furthermore, just as in the well-mixed and eGFRD models, the resulting equilibrium concentration from Spatiocyte is also independent of the molecule radius or spatial discretization. Conversely, the RDME method deviated substantially from

the well-mixed result when the voxel size is small, which is expected (Fange et al., 2010; Hellander et al., 2012; Isaacson, 2013).

The well-mixed model assumes the time scale of diffusion to be always shorter than that of the reactions. As a result, molecules are expected to be uniformly distributed at all times and reactions can take place independent of spatial localization. The well-mixed assumption is valid when describing activation-limited reactions but when they are diffusion-influenced, the position of molecules should be taken into account. We therefore expected some disparity between the well-mixed model and MLM when the production-degradation process is diffusion-influenced. In Figure 5.2, at smaller diffusion coefficients ($D_x = 0.01, 0.02$), the equilibrium concentrations are indeed lower with Spatiocyte than with well-mixed model. Spatiocyte behavior is consistent with eGFRD, which also accounts for molecule positions. RDME however, has the same outcomes as

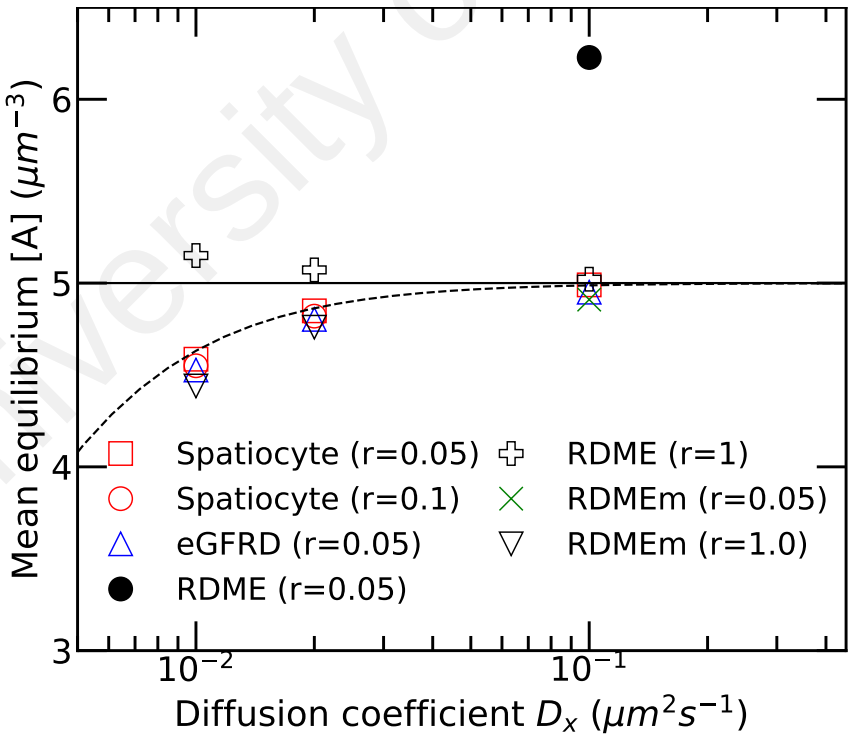


Figure 5.2: Mean equilibrium concentration of A from Spatiocyte, eGFRD, RDME and RDME with modified propensity (RDMEm) with $D_A = D_B = D_x \in \{0.01, 0.02, 0.1\}$ μm^{2}s^{-1}. Solid and dashed lines represent expected results according to the well-mixed model and the microscopic theory respectively.

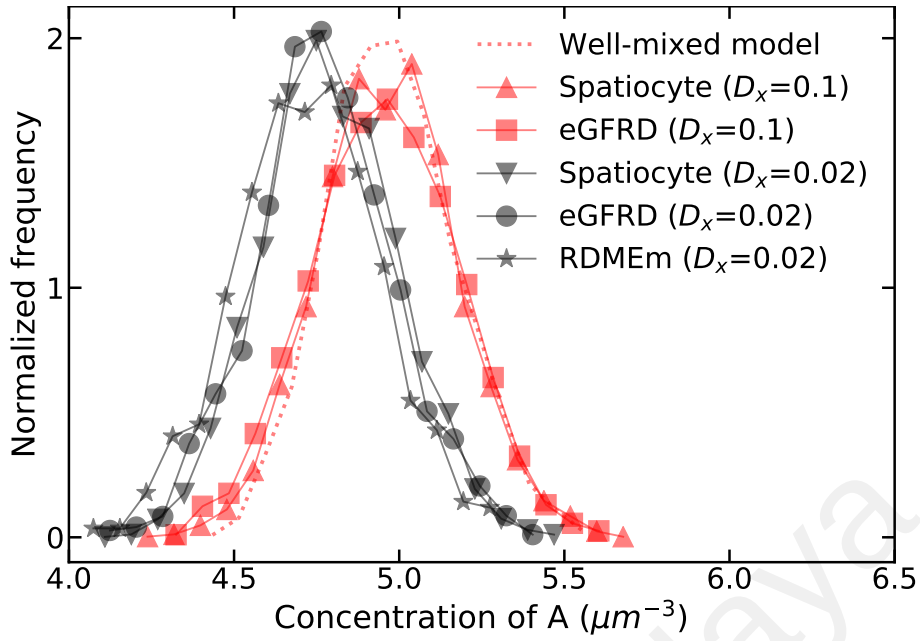


Figure 5.3: Steady-state distribution of A from Spatiocyte and eGFRD with $r = 0.05$ μm and $D_A = D_B = D_x \in \{0.1, 0.02\}$ $\mu m^2 s^{-1}$. RDMEm simulated with $r = 1$ and $D = 0.02$ is also shown for comparison. The frequency is normalized such that the sum over the bin is unity. Dotted line represents the well-mixed model simulated using the Gillespie method.

the well-mixed model.

The reduction in equilibrium value when the diffusion coefficient is decreased was previously described by the microscopic theory of Agmon and Szabo (Agmon & Szabo, 1990). In contrast to the SCK theory, Agmon and Szabo have considered the non-negligible effect of B concentration on the effective reaction rate, especially when the reaction is diffusion-influenced. The slow diffusion of molecules increases the effective contact radius, resulting in higher effective annihilation rate as shown in Equation (5.3). The output of the production-degradation process according to the microscopic theory is shown in Figure 5.2 as a dashed line that coincides with Spatiocyte and eGFRD, further verifying the MLM theory. Given the same diffusion and macroscopic reaction rates, the change in the Spatiocyte voxel size does not affect the equilibrium behavior (at $r = 0.1$ and $r = 0.05$ in Figure 5.2) since the reaction acceptance probability, P_a is adjusted according to the voxel size to obtain the correct macroscopic behavior.

On the other hand, RDME shows large deviation from the expected values at slow diffusion. The inability of conventional rate equation and RDME to correctly capture diffusion-influenced reactions has previously been noted and worked on before (Erban & Chapman, 2009; Fange et al., 2010; Hellander et al., 2012; Isaacson, 2013; Smith et al., 2016). By incorporating the diffusion coefficient into the bimolecular reaction propensity formula (Equation (26) in (Erban & Chapman, 2009)), the equilibrium concentration of RDME shows a better agreement with the expected values (see RDMEm, $r = 1.0$ in Figure 5.2). However, when the reaction is diffusion-limited ($D_x = 0.01, 0.02$), unlike MLM, the subvolume size of RDMEm cannot reach the microsopic resolution, $r = 0.05$. This is because the size is constrained by a critical value (Equation (25) in (Erban & Chapman, 2009)) that preserves the well-mixed condition. At $D_x = 0.01$ for example, the critical subvolume size is about 13 times the molecule diameter, any size smaller is invalid.

We have also examined the fluctuation of A at equilibrium, as depicted in Figure 5.3. At $D_x = 0.1$, the histogram of Spatiocyte matches the distribution curves of eGFRD and the well-mixed model (Gillespie method (Gillespie, 1976)). At much reduced diffusion coefficient ($D_x = 0.02$) however, both Spatiocyte and eGFRD shared similar distributions, with the width becoming narrower and the mean value shifting to the left. With the modified propensity function, RDMEm also exhibited similar distribution. The narrow width and the shifted mean are consistent with the characteristics of the Poisson distribution.

It was reported that MLM would not be able to solve the first-order production-degradation reaction $\emptyset \underset{k_2}{\overset{k_1}{\rightleftharpoons}} A$ accurately because of its spatial discretization scheme (Sturrock, 2016). When the number of total voxels in the compartment, N_v is less than k_1/k_2, the equilibrium concentration deviates from the well-mixed model. This deviation however, is a direct consequence of the volume exclusion property of MLM. Since each voxel can only occupy a single molecule, there would be an insufficient number of vacant

voxels to accommodate new molecules when the degradation rate is not sufficiently fast. The maximum occupancy on HCP lattice simply reflects the maximum physical occupancy of voxel-sized molecules in the compartment because the HCP arrangement packs the highest density of sphere voxels (Szpiro, 2003). Just as in the cellular compartment, no more molecules can be added into the system when the number of generated molecules exceeds available free space. Moreover, since only about 34% of the cell volume is occupied by macromolecules (Zimmerman & Trach, 1991), it is also an unlikely scenario to fully occupy the voxels of HCP lattice with macromolecules. With the multi-algorithm implementation of Spatiocyte (Arjunan & Takahashi, 2017), we can use the Gillespie's Next-Reaction method (Gibson & Bruck, 2000) to simulate small molecules that are in large abundance and are homogeneously distributed. In this case, the equilibrium result is independent of spatial discretization since the method assumes the well-mixed condition.

5.2 Dual Phosphorylation-dephosphorylation

In mean-field models, the spatio-temporal correlation of microscopic rebinding events is not resolved explicitly because the correlation usually does not cause a significant impact on the dynamics at the macroscopic scale. One case where the correlation does influence the macroscopic response is the dual phosphorylation-dephosphorylation cycle of the MAPK cascade (Chang & Karin, 2001; Ferrell & Bhatt, 1997; Aoki et al., 2011), shown in Figure 5.4. The substrate MAPK (K in Figure 5.4) is phosphorylated in a two-step process by the MAPK kinase (KK) and dephosphorylated by a phosphatase P. The phosphorylation and dephosphorylation processes proceed according to the Michaelis-Menten kinetics and exhibit distributive property (Ferrell & Bhatt, 1997), wherein the enzymes must unbind from the substrate before they can rebind and modify the second site. Upon phosphorylation or dephosphorylation, the respective enzymes are inactivated (denoted as KK* and P*), and reactivated (KK or P) after some time τ_{rel}. When the reactivation time is short and the

enzyme-substrate reaction is diffusion-limited, the newly dissociated enzyme and substrate are close enough to rebind instead of escaping into the bulk. Processive behavior caused by rebindings of the same enzyme results in higher overall phosphorylation rate than the distributive case where the dissociated molecules can escape rebinding (Ferrell & Bhatt, 1997; Aoki et al., 2011). Such microscopic spatio-temporal correlation has been shown to change the response sensitivity of the phosphorylation state, which can cause the subsequent removal of ultra-sensitivity or bi-stability in the system (Elf & Ehrenberg, 2004; Takahashi et al., 2010).

Rebinding events taking place within very short time scales are difficult to be captured by RDME because of the fine spatial resolution required. To test whether MLM can resolve such events faithfully, we use Spatiocyte to simulate the dual phosphorylation cycle with the same parameters from (Takahashi et al., 2010). Simulation are performed with parameters as following: molecule size $l = 0.0025 \times 1.0209$ μm, diffusion coefficient D_x, [KK] + [P] = 60, duration = 200 s, volume = 1 μm^3 with periodic boundary. Distributive and processive models are represented by Equations (1-5) of (Takahashi et al., 2010), and were solved using ODE solver. Figure 5.5 displays the steady-state response curves of Spatiocyte and reference theoretical models. Note that since the reactivation time τ_{rel} is equal to or less than the diffusion time step t_d (given in Figure 5.5), the molecules can rebind soon after dissociation. The simulation result coincides very well with the switch-like response curve of the distributive model at fast diffusion ($D_x = 4$ μm^{2}s^{-1}), whereas at much slower diffusion ($D_x = 0.06$ μm^{2}s^{-1}), it converges to the graded response curve of the processive model. The influence of diffusion on the response curve can be understood through the rebinding events. When diffusion is slow, reactions become more diffusion-limited and rebinding occurs at higher frequency. The ensuing processive-like mechanism then leads to the loss of the switch-like response curve. Conversely, in the

limit of fast diffusion as assumed in the mean-field model, a sharper switch-like response curve is recovered because of fewer rebindings.

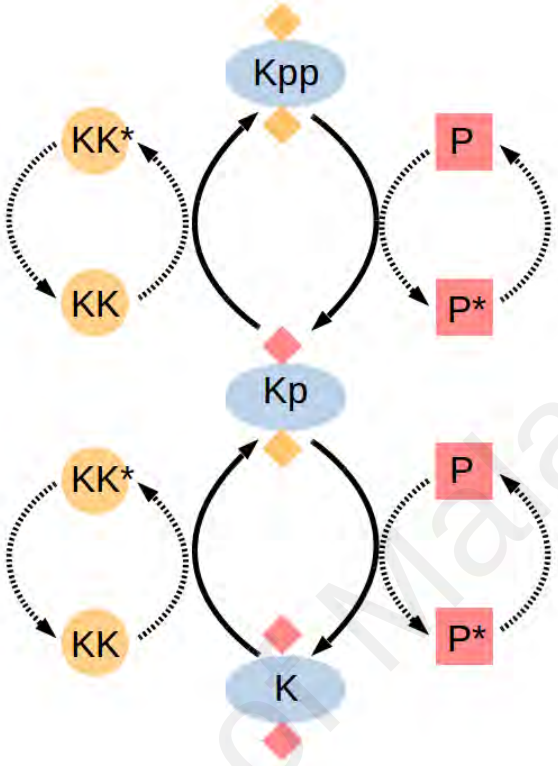


Figure 5.4: Reaction model showing MAPK (K) is first activated into Kp and then Kpp by MAPKK (KK) in two phosphorylation steps. Kpp is also deactivated by phosphatase (P) in two dephosphorylation steps to become K again. Enzymes KK and P become inactive immediately after reacting with their respective substrates and then relax back to the active state after some delay τ_{rel}.

The parameter ranges examined so far have a stable steady-state as demonstrated by the response curves in Figure 5.5. When the total concentration of the substrate is increased five-fold, the mean-field theory generates hysteresis, shown by the dotted and dash-dotted lines. The dotted line represents the response when initialized with $[K_{pp}]/[K]_{total} = 1$, whereas the dash-dotted line has the initial condition $[K_{pp}]/[K]_{total} = 0$. MLM produced similar responses when the diffusion is fast ($D_x = 4$) (diamond markers in Figure 5.5). However, as diffusion slowed down to $D_x = 0.06$, the bistability is lost (triangle markers). Bistable states appear when the diffusion is fast and the substrate concentration relative to enzyme is high. For example, at the initial state when almost all substrates are in the

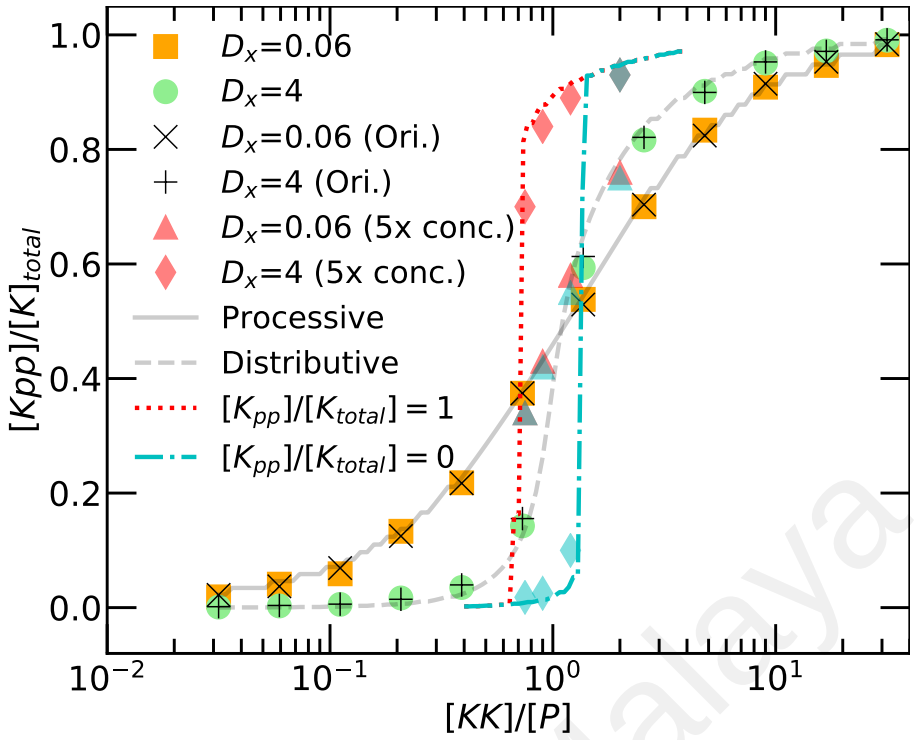


Figure 5.5: Fraction of Kpp in response to MAPKK/phosphatase ratio at steady-state. Circle and square markers denote simulation result using Spatiocyte with $D_x = 4\ \mu\text{m}^2\text{s}^{-1}$ and $D_x = 0.06\ \mu\text{m}^2\text{s}^{-1}$, respectively. Dashed and solid lines represent distributive and processive mechanism models, respectively. Cross and plus markers show the results from the original Spatiocyte scheme, wherein the voxel and molecule sizes are exactly the same. We used a short reactivation time, $\tau_{rel} = 1\ \mu\text{s}$, relative to t_d (for comparison $t_d \approx 1\ \mu\text{s}$ when $D_x = 4\ \mu\text{m}^2\text{s}^{-1}$, $t_d \approx 70\ \mu\text{s}$ when $D_x = 0.06\ \mu\text{m}^2\text{s}^{-1}$) with the total number of substrates, $\text{K}_{\text{total}} = 120$. Hysteresis responses from mean-field distributive model with five-fold substrate concentration ($\text{K}_{\text{total}} = 600$) are indicated by dotted and dash-dotted lines with initial conditions $[\text{K}_{\text{pp}}]/[\text{K}]_{\text{total}} = 1$ and $[\text{K}_{\text{pp}}]/[\text{K}]_{\text{total}} = 0$, respectively. Diamond and triangle markers represent Spatiocyte responses with five-fold substrate concentration at the indicated diffusion coefficient, D_x.

unphosphorylated form, most kinase will be bound to the substrates rapidly. Hence, a substrate that has been phosphorylated once is more likely to be dephosphorylated by free phosphatase than to be phosphorylated the second time by scarce and fast diffusing kinase. The inverse situation where all substrates are in the phosphorylated form would also respond similarly to phosphatase. On the other hand, when diffusion is slow, the kinase activity becomes processive because of the high rebinding probability. As a result, molecules are more likely to be phosphorylated or dephosphorylated consecutively before they could be disrupted by antagonistic enzymes from the bulk. This example highlights how local spatio-temporal correlation can change the binding behavior and results in a

different global response than the one predicted by the mean-field model.

As a side remark, in the original Spatiocyte scheme (Arjunan & Tomita, 2010), the voxel adopts the size of the diffusing molecules. However, as we found in Section 3.2.2.3, the voxel needs to be about 2% larger than the molecule size for the total rebinding probability and the effective rate constant to be exactly the same as in the continuum-based theory. Despite the 2% difference in voxel sizes, both new and original schemes displayed very good fit with the expected dual phosphorylation cycle response curves in Figure 5.5. To be fully consistent with the continuum-based theory however, the size should be set according to Equation (3.95). The voxel size is not hard-coded to be the same as the molecule size and can be easily specified in the Spatiocyte model file (Arjunan & Takahashi, 2017).

5.3 Bimolecular Reaction with Volume Exclusion

Anomalous diffusion due to excluded volume has been shown to generate non-classical reaction kinetics on 2D (Berry, 2002; Schnell & Turner, 2004) and 3D lattices (Pitulice et al., 2014). Here, we use MLM on HCP lattice to examine the effects of volume exclusion on the bimolecular reaction $E + S \rightarrow \emptyset$ in the presence of uniformly distributed immobile obstacles. E and S have the radius 5 nm and diffusion coefficient, $D_0 = 1\ \mu\text{m}^2\text{s}^{-1}$, where D_0 is the diffusion coefficient in non-crowded dilute condition. Bimolecular intrinsic reaction rate constant $k_{a3D} = 10k_D$ is chosen such that the reaction is diffusion-limited. Excluded volume is quantified by the lattice occupancy of the obstacles, $\phi = N_o/N_v$, where N_o and N_v are the numbers of obstacles and total voxels, respectively. Simulation is carried out in a periodic cubic compartment with length $L = 1\ \mu\text{m}$ for a duration of $1000t_d$. Reactants have dilute concentrations, $[S] = 5[E] = 0.001N_v$ and are placed randomly at the beginning of simulation.

We first consider the effects of immobile obstacles on diffusing molecules. We calculate the time-dependent diffusion coefficient from the mean-squared displacement of simulated particle trajectories. The time-dependent diffusion coefficient in Figure 5.6 indicates that the diffusion is anomalous at short times and normal at long times. The crossover time from anomalous to normal diffusion depends on the volume occupancy. The reduced long-time diffusion coefficient is well-described by (Saxton, 1989; Vilaseca et al., 2011),

$$D' = D_0(1 - \phi/\phi_p), \tag{5.4}$$

where $\phi_p \approx 0.77$ is the percolation threshold for HCP lattice. We confirmed that the long-time diffusion coefficients obtained for ϕ in Figure 5.6 (dashed lines) are consistent with D' in Equation (5.4).

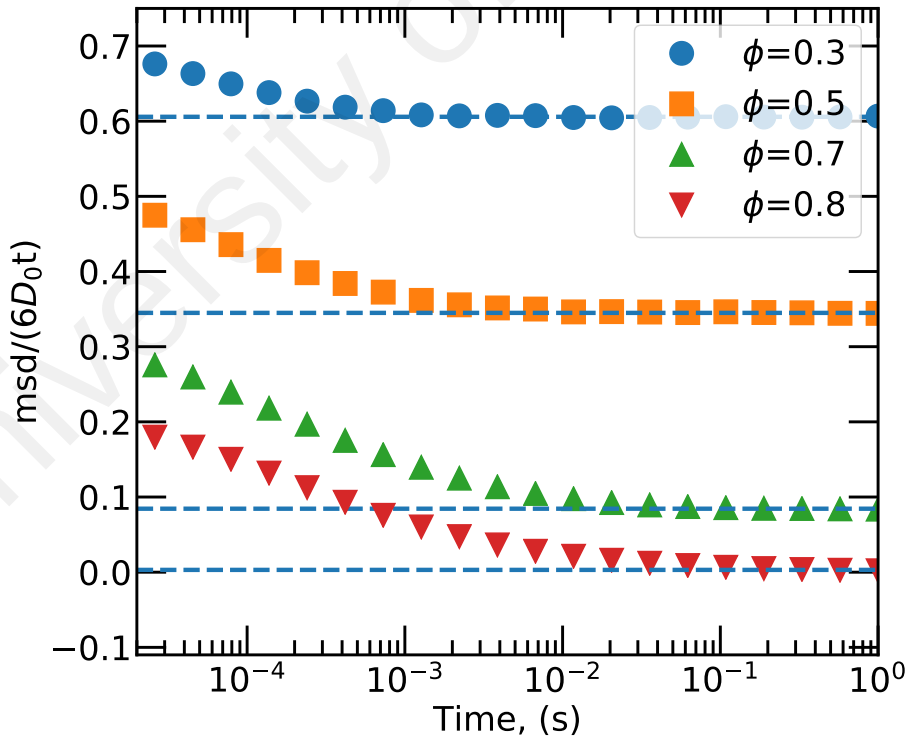


Figure 5.6: Time-dependent diffusion coefficient of tracer molecules in the presence of immobile obstacles at volume occupancy, ϕ. The diffusion coefficient at a time point is determined from the mean-squared displacement of simulated particle trajectories. Dashed lines denote the diffusion coefficient at long-time as predicted by $D' = D_0(1 - \phi/\phi_p)$.

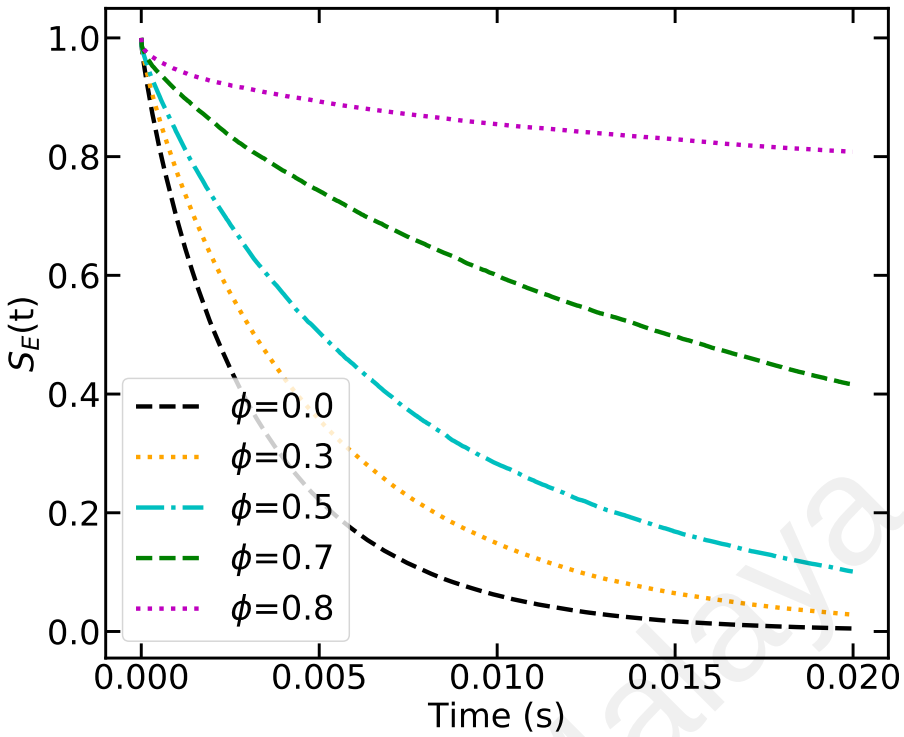


Figure 5.7: Survival probability of E in reaction $E + S \rightarrow \emptyset$ at various volume occupancy ϕ.

Figure 5.7 shows that the survival probability of E decays slower when the volume occupancy, ϕ is increased. From the survival probability, we can calculate the rate coefficient according to Equation (4.3) to obtain the kinetics. We replaced the constant concentration term $[B]$ in Equation (4.3) with the time varying term $[E](t)$ in the equation. For the dilute case ($\phi = 0$) in Figure 5.8, there is a good agreement for the simulated $k(t)$ with the SCK rate coefficient in Equation (2.12). As ϕ increases to 0.3 and 0.5, the overall reaction rate decreases, and thus progressively diverges from the SCK rate. Despite the discrepancy, the rates can still conform to the SCK theory when the long-time diffusion coefficient in Equation (5.4) is used.

As the volume occupancy approaches the percolation threshold (Figure 5.8, $\phi = 0.7$), the kinetics begins to deviate from the SCK theory. The deviation is strongest at $\phi = 0.8$, which is beyond the percolation threshold. Note that at lower volume occupany ($\phi = 0.3, 0.5$), the anomalous to normal diffusion crossover time in Figure 5.6 is faster than the observation time in Figure 5.8. Here, the kinetics is well described by the long-time effective diffusion

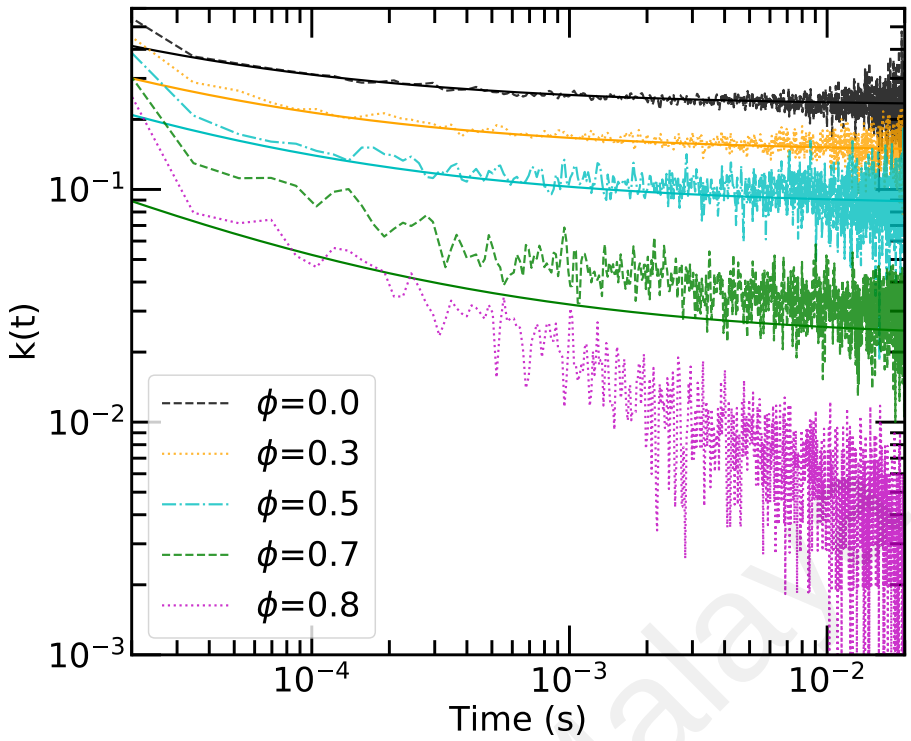


Figure 5.8: Time-dependent reaction rates (dashed lines) at various volume occupancy ϕ. Solid lines represent SCK theory with the long-time diffusion coefficient calculated in Figure 5.6.

coefficient. However, when the crossover time is comparable to the observation time because of the increased volume occupancy (Figure 5.6, $\phi = 0.7$), the effects of anomalous diffusion is visible in the kinetics (Figure 5.8, $\phi = 0.7$). At above the percolation threshold ($\phi = 0.8$), anomalous diffusion does not crossover to normal diffusion. As a result, the long-time diffusion coefficient eventually decays to zero. In these highly crowded cases, the SCK theory fails to describe the kinetics.

Grima and Schnell (Grima & Schnell, 2006) have shown that reaction kinetics, either classical or non-classical, is not determined by the heterogeneity of the accessible space but rather by the reaction probability and the initial condition. In the SCK model, reaction follows classical kinetics when it is activation-limited ($k_{a3D}/k_D \ll 1$) but non-classical kinetics is observed when it is diffusion-influenced ($k_{a3D}/k_D \gg 1$). The non-classical behavior in the latter is well-described by Equation (2.12) using microscopic parameters. The corresponding long-time behavior up to the second order term scales according to

Equation (2.13), which has the same general form of the Zip-Mandelbrot equation proposed by Schnell and Turner (Schnell & Turner, 2004; Pitulice et al., 2014). The Zip-Mandelbrot equation is valid for long-time kinetics whereas the SCK rate in Equation (2.12) describes the kinetics for all time ranges.

Here, we have studied the kinetics of bimolecular reaction in the presence of immobile obstacles with MLM. When the total volume occupied by obstacles is much smaller than the percolation threshold and the observation time scale is longer than the anomalous to normal diffusion crossover time, the kinetics is still reproducible with the SCK theory and Equation (5.4). However, it deviates from the theory when the volume occupancy nears or crosses the percolation threshold, wherein anomalous diffusion dominates and the diffusion coefficient approaches zero at the long-time limit. Therefore, to better describe the non-classical kinetics analytically, we should incorporate the anomalous diffusion induced by fractal medium into the theory either phenomenologically (Kopelman, 1988; Schnell & Turner, 2004; Pitulice et al., 2014) or by extending the SCK model using a generalized diffusion equation (Barzykin & Tachiya, 1993; Sung et al., 2001).

5.4 Surface Reaction Pathways

A cytosolic molecule can react with a membrane-bound reactant via two possible pathways: it can either perform 3D diffusion in the cytoplasm and then directly react with the membrane-bound reactant exposed to the cytosol or it can bind first to the membrane and then perform 2D diffusion before reacting with the reactant. Both of these pathways are often adopted simultaneously in the cell. Previous works have investigated how each pathway contributes to the overall process (Adam & Delbrück, 1968; Berg & Purcell, 1977; Axelrod & Wang, 1994; Kholodenko et al., 2000). Here we apply the Spatiocyte scheme with the derived MLM expressions to simulate surface reactions comprising all dimensions. We study the contribution of each pathway to the overall reaction rate under

the influence of different diffusivity and reactivity.

We consider a cuboid compartment of dimension $H \times L \times L$, depicting the cytoplasmic volume. The top surface of the cuboid is reflective, whereas the bottom surface represents an absorbing lipid membrane. Each of these surfaces has the area $L \times L$. Within the system, there are two elementary species, A and B, with radius $r = 0.005\ \mu m$. A_c denotes the cytosolic state of A that diffuses freely in the bulk at a rate of D_c. A_c can reversibly associate with the membrane to become A_m:

$$A_c \underset{k_{sd}}{\overset{k_{sa}}{\rightleftharpoons}} A_m. \tag{5.5}$$

The ratio of the membrane association constant over the dissociation constant is the equilibrium constant, $k_{sa}/k_{sd} = K_{eq}$. Upon the adsorption onto the membrane, A_m performs 2D diffusion at a rate of D_m. On the membrane, B molecules are initialized to be immobile and randomly distributed with concentration $[\mathrm{B}]_0$.

A can react with B via the 3D pathway:

$$A_c + B \xrightarrow{k_{a3D}} AB, \tag{5.6}$$

or the 2D pathway:

$$A_m + B \underset{k_r}{\overset{k_{a2D}}{\rightleftharpoons}} AB. \tag{5.7}$$

$k_{a\{2D,3D\}}$ denotes the intrinsic association rate constants for 2D and 3D reactions, whereas k_r represents the dissociation rate constant.

To quantify the dominance of the 2D pathway, we measured the fraction of the 2D equilibrium rate in the total reaction rate, as in (Axelrod & Wang, 1994):

$$f_{2D} = \frac{k_{on2D}}{k_{on2D} + k_{on3D}} = \frac{1}{1 + k_{on3D}/k_{on2D}}. \tag{5.8}$$

$k_{on\{2D,3D\}}$ represents the macroscopic effective rates for the 2D and 3D association reactions. The k_{on3D}/k_{on2D} ratio is calculated using the simulated equilibrium concentrations according to the formula

$$\frac{k_{on3D}}{k_{on2D}} = \frac{1}{[A_c]_{eq}} \left(\frac{k_r [AB]_{eq}}{k_{a2D} [B]_{eq}} - [A_m]_{eq} \right), \tag{5.9}$$

which is derived by solving the rate equations for Equations (5.7) and (5.6) at equilibrium.

We examined the dominance of the 2D pathway with changes in D_c/D_m, $[\mathrm{B}]_0$, and the association reaction probability, $P_{a\{2D,3D\}}$ for the 2D and 3D pathways. We fixed other variables such as the sizes of the system and molecule, k_r, K_{eq}, k_{a2D}/k_r, and the initial concentration $[A_c]$. We used the typical cytosolic rate for D_c (10 $\mu m^2 s^{-1}$) with D_c/D_m ratio ranging from 1 to 1000. $K_{eq} = 0.15\ \mu m$, $[A_c] = 5\ \mu M$, $k_{a2D}/k_r = 0.001\ \mu m^2$ and $k_{sd} = 10\ s^{-1}$ are within the biologically realistic values (Fulbright & Axelrod, 1993; Tolentino et al., 2008). The simulation compartment size is $L = 1\ \mu m$ with $H = 2L$ and voxel size $l = 0.01 \times 1.0209\ \mu$m.

From the simulation results in Figure 5.9, we can observe the overall decreasing trend of f_{2D} as the ratio D_c/D_m increases. The actual value of f_{2D} depends on the reaction probability and the concentration of reactant, $[\mathrm{B}]_0$,. When the association reaction is diffusion-limited ($P_{a2D} = P_{a3D} = 1$) and the reactant concentration is low ($[\mathrm{B}]_0$=100 μm^{-2}), f_{2D} becomes more than 50% for D_c/D_m values between 1 and 30. When D_c/D_m is larger than 30, the 3D pathway becomes dominant. At very high $[\mathrm{B}]_0$=500 μm^{-2},

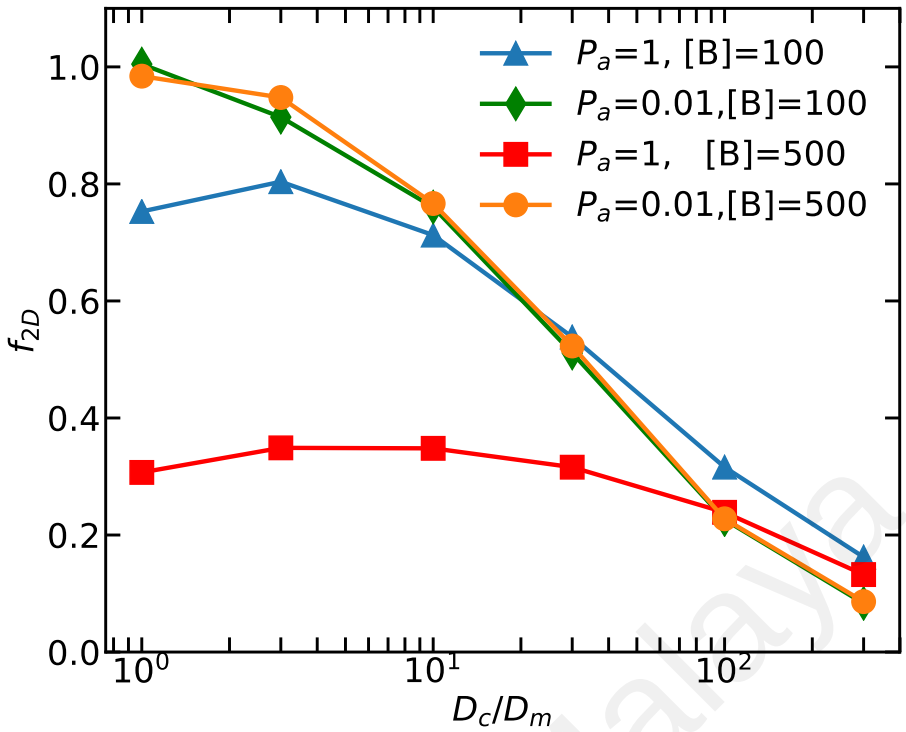


Figure 5.9: Contribution of 2D reaction pathway as indicated by the fraction f_{2D} plotted against D_c/D_m ratio for different values of reaction probability $P_a = P_{a2D} = P_{a3D}$ and concentration of membrane-associated reactant [B] (unit μm^{-2}).

the 3D pathway is dominant for all ratios of D_c/D_m. When the association reaction is activation-limited ($P_{a2D} = P_{a3D} = 0.01$), f_{2D} is larger than 50% for D_c/D_m ratio in between 1-30, and is less than 50% for higher ratio of D_c/D_m. Unlike the diffusion-limited case, f_{2D} for activation-limited reaction is less sensitive to changes in reactant concentration $[B]_0$.

In typical intracellular environment, membrane-associated molecule diffuses 10-100× slower than the cytosolic molecule. In such regime, our simulation result implies the following: the 2D reaction pathway dominates the overall reaction provided that the concentration of membrane-associated reactant is sparse, the species diffusion on the membrane is fast and the reaction is activation-limited. On the other hand, 3D reaction pathway becomes dominant when the diffusion of membrane species is extremely slow or when the membrane-associated reactant is highly abundant and the reaction probability upon collision is high.

5.5 Summary

We have applied MLM to study selected reaction-diffusion process typically encountered in intracellular system. In the protein production-degradation process, we found that diffusion has a non-negligible influence on the mean concentration and distribution at equilibrium. In the case of dual phosphorylation-dephosphorylation cycle, diffusion affects the rebinding behavior between the substrate and enzyme, which can lead to changes in the overall steady-state response. When diffusion is highly obstructed by immobile obstacles, we found that diffusion becomes anomalous, for time regime that depends on the obstacle concentration. The obstructed diffusion alters the reaction kinetics of bimolecular reaction, deviated from the SCK description. For surface reaction in typical intracellular condition, we found that the dominance of 2D and 3D reaction pathways in the overall reaction is highly influenced by the concentration of membrane-associated reactant, the diffusivity on the membrane and also the reactivity. Overall, we show that MLM is able to reflect the microscopic nature of intracellular RD processes.

CHAPTER 6: DISCUSSION AND CONCLUSIONS

6.1 Summary and Overall Discussion

Realistic modeling of intracellular RD process requires the consideration of the stochastic, multi-compartmental and finite molecular size characteristics of the system. The main advantage of MLM as a reaction-diffusion simulation method is its ability to capture these characteristic without incurring high computational cost. In this work, we have successfully established the accuracy and consistency of MLM in simulating RD process. We have derived the on-lattice reaction rate coefficient for diffusion-influenced reaction in 1D, 2D, and 3D spatial dimensions. The accuracy and performance of Spatiocyte, a MLM simulation scheme, has been benchmarked with continuum-based theory and particle-based simulation method. We have also applied MLM to study several biological reaction-diffusion process and observed interesting behavior that arose from the microscopic nature of the underlying process.

In the analysis of lattice theory, we have shown that MLM exhibit the same long-time behavior as the continuum-based theory for 3D volume-volume and volume-surface reaction, 2D surface-surface reaction and 1D volume-surface adsorption. By matching the lattice and continuum theoretical expression for the lattice effective rate constant, total rebinding probability, and long-time rate coefficient, we obtained the reaction probability formula in terms of the physical and lattice parameters. We also found that the 3D voxel size should be larger than the molecule size by about 2% in the case of HCP lattice to be quantitatively consistent with continuum reaction kinetics. On the other hand, 2D voxel size should be larger than the molecule at least by about 0.6% for the triangular lattice and by 5% for the square lattice. The constraint in 2D meets the minimum voxel size requirement of the corresponding lattice arrangement in 3D. If the voxel size is exactly the

same as the molecule, the simulated time-dependent reaction kinetics will deviate from the expected behavior in continuum. Such deviations should be carefully considered especially when simulating reactions containing nonlinear terms.

Aside from theoretical analysis, we have validated the accuracy of MLM using numerical simulation as well. Spatiocyte exhibit asymptotic time-dependent behavior that agrees with the SCK theory for both activation- and diffusion-limited bimolecular reaction in all spatial dimensions. In the case of 1D volume-surface adsorption, the spatiotemporal behavior of MLM agrees with the continuum description when the correct expression for reaction probability was used. Performance study of Spatiocyte show that it only takes minutes for the Spatiocyte algorithm to simulate thousands of molecules with a time step of μs for a duration of seconds on a single CPU core. Apart from simple bimolecular reaction, we have applied MLM in several application studies in which MLM displayed good consistencies with eGFRD, a particle-based simulator.

MLM can be easily implemented in practice, for example, Spatiocyte scheme only takes physical parameters comprising the molecule size, diffusion coefficient and intrinsic reaction rate as input, and generates time-series outputs such as molecule copy number and trajectory. These outputs can be compared directly with experimental measurement such as those obtained from single molecule microscopy. At present, Spatiocyte supports surface reaction with various geometries at the cellular scale. It has been used to study the influence of microscopic effects on cellular behavior. This include the formation of a high density ring over the entire cell membrane as a result of transient membrane association and rebinding of proteins in bacteria (Arjunan & Tomita, 2010), the clustering of proteins on the red blood cell membrane from oxidative stress (Shimo et al., 2015) and the oligomerization of receptors and its influence on ligand binding kinetics (Watabe et al., 2018). As the spatiotemporal resolution of imaging technique advances (Hell et al., 2015),

time-dependent reaction kinetics and molecular trajectories will become more accessible. The high resolution experimental data coupled with efficient microscopic simulation technique such as MLM, provide a complementary way to postulate and investigate the mechanism underlying biological RD processes.

6.2 Limitations of MLM

MLM captures the effects of excluded volume naturally but comparing on-lattice behavior with the continuum one is not straightforward since the influence of volume exclusion and the resulting reaction kinetics vary according to the lattice arrangement (Grima & Schnell, 2006; Meinecke & Eriksson, 2016). Moreover, since all diffusing species in this work have the same molecule size, it is not possible to replicate the effects of relative size of interacting molecules. To minimize such lattice artifacts and to better approximate the volume exclusion in the continuum, we can improve the size representation of each molecule on lattice by occupying multiple voxels as in the SVTA approach or by employing a hybridized on- and off-lattice approach. Higher spatial resolution for representing molecule would generate a more realistic diffusion behavior in a crowded environment. Alternatively, we can introduce a density-dependent hopping rate as adopted by two previous RDME methods (Grima & Schnell, 2007; Cianci et al., 2017).

6.3 Suggestions for Future Work

Realistic simulation of intracellular reaction-diffusion processes should also incorporate the influence of inter-molecular potentials such as van der Waals and hydrodynamic forces. By employing contact interactions on lattice as proposed by Fernando et al. (Fernando et al., 2010) or the SVTA approach with interaction potentials (Gillespie et al., 2014), it may be possible to incorporate the above forces in MLM. In conclusion, the theoretical framework presented in this work serves as a building block for further development and

integration of MLM-based algorithms. For future research, MLM could act as a bridge for connecting the microscopic simulation method such as Brownian and molecular dynamics with the macroscopic method such as the reaction-diffusion or rate equation description. With that, a multi-scale simulation framework can be established.

REFERENCES

Adam, G. & Delbrück, M. (1968). Reduction of dimensionality in biological diffusion processes. *Structural Chemistry and Molecular Biology*, *198*, 198–215.

Agbanusi, I. C. & Isaacson, S. A. (2014). A comparison of bimolecular reaction models for stochastic reaction–diffusion systems. *Bulletin of Mathematical Biology*, *76*(4), 922–946.

Agmon, N. & Szabo, A. (1990). Theory of reversible diffusion-influenced reactions. *The Journal of Chemical Physics*, *92*(9), 5270–5284.

Almeida, P. F. & Vaz, W. L. (1995). Lateral diffusion in membranes. *Handbook of Biological Physics*, *1*(C), 305–357.

Andrews, S. S. (2009). Accurate particle-based simulation of adsorption, desorption and partial transmission. *Physical Biology*, *6*(4), Article#046015.

Andrews, S. S. (2016). Smoldyn: particle-based simulation with rule-based modeling, improved molecular interaction and a library interface. *Bioinformatics*, *33*(5), 710–717.

Andrews, S. S. (2018). Particle-based stochastic simulators. In D. Jaeger & R. Jung (Eds.), *Encyclopedia of Computational Neuroscience* (pp. 1–5). New York: Springer.

Andrews, S. S., Addy, N. J., Brent, R., & Arkin, A. P. (2010). Detailed simulations of cell biology with Smoldyn 2.1. *PLoS Computational Biology*, *6*(3), Article#e1000705.

Andrews, S. S. & Bray, D. (2004). Stochastic simulation of chemical reactions with spatial resolution and single molecule detail. *Physical Biology*, *1*(3), Article#137.

Aoki, K., Yamada, M., Kunida, K., Yasuda, S., & Matsuda, M. (2011). Processive phosphorylation of ERK MAP kinase in mammalian cells. *Proceedings of the National Academy of Sciences*, *108*(31), 12675–12680.

Arjunan, S. N. V. & Takahashi, K. (2017). Multi-algorithm particle simulations with Spatiocyte. In D. Kihara (Ed.), *Protein Function Prediction. Methods in Molecular Biology* (pp. 219–236). New York: Humana Press.

Arjunan, S. N. V. & Tomita, M. (2010). A new multicompartmental reaction-diffusion modeling method links transient membrane attachment of *E. coli* MinE to E-ring formation. *Systems and Synthetic Biology*, *4*(1), 35–53.

Axelrod, D. & Wang, M. (1994). Reduction-of-dimensionality kinetics at reaction-limited cell surface receptors. *Biophysical Journal*, *66*(3), 588–600.

Bar-Even, A., Paulsson, J., Maheshri, N., Carmi, M., O'Shea, E., Pilpel, Y., & Barkai, N. (2006). Noise in protein expression scales with natural protein abundance. *Nature Genetics*, *38*(6), 636–643.

Baras, F. & Mansour, M. M. (1996). Reaction-diffusion master equation: A comparison with microscopic simulations. *Physical Review E*, *54*(6), Article#6139.

Barzykin, A. & Tachiya, M. (1993). Diffusion-influenced reaction kinetics on fractal structures. *The Journal of Chemical Physics*, *99*(12), 9591–9597.

Bashardanesh, Z. & Lötstedt, P. (2017). Efficient Green's function reaction dynamics (GFRD) simulations for diffusion-limited, reversible reactions. *Journal of Computational Physics*, *357*, 78–99.

Berg, H. C. & Purcell, E. M. (1977). Physics of chemoreception. *Biophysical Journal*, *20*(2), 193–219.

Berg, O. G. (1978). On diffusion-controlled dissociation. *Chemical Physics*, *31*(1), 47–57.

Berry, H. (2002). Monte Carlo simulations of enzyme reactions in two dimensions: Fractal kinetics and spatial segregation. *Biophysical Journal*, *83*(4), 1891–1901.

Bittig, A. T. & Uhrmacher, A. M. (2017). ML-Space: Hybrid spatial Gillespie and particle simulation of multi-level rule-based models in cell biology. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *14*(6), 1339–1349.

Boulianne, L., Al Assaad, S., Dumontier, M., & Gross, W. (2008). GridCell: a stochastic particle-based biological system simulator. *BMC Systems Biology*, *2*(1), Article#66.

Byrne, M. J., Waxham, M. N., & Kubota, Y. (2010). Cellular dynamic simulator: An event driven molecular simulation environment for cellular physiology. *Neuroinformatics*,

8(2), 63–82.

Carslaw, H. S. & Jaeger, J. C. (1959). *Conduction of Heat in Solids, 2nd edition*. Oxford: Clarendon Press.

Chang, L. & Karin, M. (2001). Mammalian MAP kinase signalling cascades. *Nature*, *410*(6824), 37–40.

Chew, W.-X., Kaizu, K., Watabe, M., Muniandy, S. V., Takahashi, K., & Arjunan, S. N. V. (2018). Reaction-diffusion kinetics on lattice at the microscopic scale. *Physical Review E*, *98*, Article#032418.

Chew, W.-X., Kaizu, K., Watabe, M., Muniandy, S. V., Takahashi, K., & Arjunan, S. N. V. (2019). Surface reaction-diffusion kinetics on lattice at the microscopic scale. *Physical Review E*, *99*, Article#042411.

Cianci, C., Smith, S., & Grima, R. (2016). Molecular finite-size effects in stochastic models of equilibrium chemical systems. *The Journal of Chemical Physics*, *144*(8), Article#084101.

Cianci, C., Smith, S., & Grima, R. (2017). Capturing Brownian dynamics with an on-lattice model of hard-sphere diffusion. *Physical Review E*, *95*(5), Article#052118.

Collins, F. C. & Kimball, G. E. (1949). Diffusion-controlled reaction rates. *Journal of Colloid Science*, *4*(4), 425–437.

Cowan, A. E., Moraru, I. I., Schaff, J. C., Slepchenko, B. M., & Loew, L. M. (2012). Spatial modeling of cell signaling networks. *Methods in Cell Biology*, *110*, 195–221.

Crank, J. et al. (1979). *The Mathematics of Diffusion*. London: Oxford university press.

Devreotes, P. N., Bhattacharya, S., Edwards, M., Iglesias, P. A., Lampert, T., & Miao, Y. (2017). Excitable signal transduction networks in directed cell migration. *Annual Review of Cell and Developmental Biology*, *33*(1), 103–125.

Dix, J. A. & Verkman, A. (2008). Crowding effects on diffusion in solutions and cells. *Annual Review of Biophysics*, *37*, 247–263.

Doi, M. (1976). Stochastic theory of diffusion-controlled reaction. *Journal of Physics A: Mathematical and General*, *9*(9), 1479–1495.

Donev, A., Yang, C.-Y., & Kim, C. (2018). Efficient reactive Brownian dynamics. *The Journal of Chemical Physics*, *148*(3), Article#034103.

Drawert, B., Engblom, S., & Hellander, A. (2012). URDME: a modular framework for stochastic simulation of reaction-transport processes in complex geometries. *BMC Systems Biology*, *6*(1), Article#76.

Drawert, B., Hellander, A., Bales, B., Banerjee, D., Bellesia, G., Daigle Jr, B. J., Douglas, G., Gu, M., Gupta, A., Hellander, S., et al. (2016). Stochastic simulation service: Bridging the gap between the computational expert and the biologist. *PLoS Computational Biology*, *12*(12), Article#e1005220.

Earnest, T., Cole, J. A., & Luthey-Schulten, Z. (2018). Simulating biological processes: Stochastic physics from whole cells to colonies. *Reports on Progress in Physics*, *81*(5), Article#052601.

Edelstein-Keshet, L. (1988). *Mathematical Models in Biology*. Philadelphia: SIAM.

Elf, J. & Ehrenberg, M. (2004). Spontaneous separation of bi-stable biochemical systems into spatial domains of opposite phases. *Systems Biology*, *1*(2), 230–236.

Ellis, R. J. (2001). Macromolecular crowding: Obvious but underappreciated. *Trends in Biochemical Sciences*, *26*(10), 597–604.

Ellis, R. J. & Minton, A. P. (2003). Cell biology: Join the crowd. *Nature*, *425*(6953), 27–28.

Engblom, S., Hellander, A., & Lötstedt, P. (2017). Multiscale simulation of stochastic reaction-diffusion networks. In D. Holcman (Ed.), *Stochastic Processes, Multiscale Modeling, and Numerical Methods for Computational Cellular Biology* (pp. 55–79). New York: Springer.

Erban, R. & Chapman, S. J. (2009). Stochastic modelling of reaction–diffusion processes: algorithms for bimolecular reactions. *Physical Biology*, *6*(4), Article#046001.

Fange, D., Berg, O. G., Sjöberg, P., & Elf, J. (2010). Stochastic reaction-diffusion kinetics in the microscopic limit. *Proceedings of the National Academy of Sciences*, *107*(46), 19820–19825.

Fernando, A. E., Landman, K. A., & Simpson, M. J. (2010). Nonlinear diffusion and exclusion processes with contact interactions. *Physical Review E*, *81*(1), Article#011903.

Ferrell, J. E. & Bhatt, R. R. (1997). Mechanistic studies of the dual phosphorylation of mitogen-activated protein kinase. *Journal of Biological Chemistry*, *272*(30), 19008–19016.

Flajolet, P. & Odlyzko, A. (1990). Singularity analysis of generating functions. *SIAM Journal on Discrete Mathematics*, *3*(2), 216–240.

Flajolet, P. & Sedgewick, R. (2009). *Analytic Combinatorics*. UK: Cambridge University press.

Fulbright, R. M. & Axelrod, D. (1993). Dynamics of nonspecific adsorption of insulin to erythrocyte membranes. *Journal of Fluorescence*, *3*(1), 1–16.

Gibson, M. A. & Bruck, J. (2000). Efficient exact stochastic simulation of chemical systems with many species and many channels. *The Journal of Physical Chemistry A*, *104*(9), 1876–1889.

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

Gillespie, D. T., Hellander, A., & Petzold, L. R. (2013). Perspective: Stochastic algorithms for chemical kinetics. *The Journal of Chemical Physics*, *138*, Article#170901.

Gillespie, D. T., Seitaridou, E., & Gillespie, C. A. (2014). The small-voxel tracking algorithm for simulating chemical reactions among diffusing molecules. *Journal of Chemical Physics*, *141*(23), Article#234115.

Gomez, D. & Klumpp, S. (2015). Biochemical reactions in crowded environments: Revisiting the effects of volume exclusion with simulations. *Frontiers in Physics*, *3*, Article#45.

Grima, R. & Schnell, S. (2006). A systematic investigation of the rate laws valid in intracellular environments. *Biophysical Chemistry*, *124*(1), 1–10.

Grima, R. & Schnell, S. (2007). A mesoscopic simulation approach for modeling intracellular reactions. *Journal of Statistical Physics*, *128*(1-2), 139–164.

Gruenert, G., Ibrahim, B., Lenser, T., Lohel, M., Hinze, T., & Dittrich, P. (2010). Rule-based spatial modeling with diffusing, geometrically constrained molecules. *BMC Bioinformatics*, *11*(1), Article#307.

Guigas, G. & Weiss, M. (2016). Effects of protein crowding on membrane systems. *Biochimica et Biophysica Acta - Biomembranes*, *1858*(10), 2441–2450.

Guttmann, A. J. (2010). Lattice Green's functions in all dimensions. *Journal of Physics A: Mathematical and Theoretical*, *43*(30), Article#305205.

Halatek, J., Brauns, F., & Frey, E. (2018). Self-organization principles of intracellular pattern formation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1747), Article#20170107.

Hattne, J., Fange, D., & Elf, J. (2005). Stochastic reaction-diffusion simulation with MesoRD. *Bioinformatics*, *21*(12), 2923–2924.

Haugh, J. M. (2009). Analysis of reaction-diffusion systems with anomalous subdiffusion. *Biophysical Journal*, *97*(2), 435–442.

Hell, S. W., Sahl, S. J., Bates, M., Zhuang, X., Heintzmann, R., Booth, M. J., Bewersdorf, J., Shtengel, G., Hess, H., Tinnefeld, P., et al. (2015). The 2015 super-resolution microscopy roadmap. *Journal of Physics D: Applied Physics*, *48*(44), Article#443001.

Hellander, S., Hellander, A., & Petzold, L. (2012). Reaction-diffusion master equation in the microscopic limit. *Physical Review E*, *85*(4), 1–5.

Hellander, S., Hellander, A., & Petzold, L. (2015). Reaction rates for mesoscopic reaction-diffusion kinetics. *Physical Review E*, *91*(2), Article#023312.

Hepburn, I., Chen, W., Wils, S., & De Schutter, E. (2012). STEPS: efficient simulation of stochastic reaction–diffusion models in realistic morphologies. *BMC Systems*

Biology, *6*(1), Article#36.

Hoffmann, M., Fröhner, C., & Noé, F. (2019). ReaDDy 2: Fast and flexible software framework for interacting-particle reaction dynamics. *PLOS Computational Biology*, *15*(2), 1–26.

Höfling, F. & Franosch, T. (2013). Anomalous transport in the crowded world of biological cells. *Reports on Progress in Physics*, *76*(4), Article#046602.

Howard, M. & Rutenberg, A. D. (2003). Pattern formation inside bacteria: Fluctuations due to the low copy number of proteins. *Physical Review Letters*, *90*(12), Article#128102.

Hughes, B. (1995). *Random Walks and Random Environments: Random walks*. Oxford: Clarendon Press.

Imbihl, R. & Ertl, G. (1995). Oscillatory kinetics in heterogeneous catalysis. *Chemical Reviews*, *95*(3), 697–733.

Isaacson, S. A. (2013). A convergent reaction-diffusion master equation. *Journal of Chemical Physics*, *139*, Article#054101.

Ishioka, S. & Koiwa, M. (1978). Random walks on diamond and hexagonal close packed lattices. *Philosophical Magazine A*, *37*(4), 517–533.

Joyce, G. S. (1998). On the cubic modular transformation and the cubic lattice Green functions. *Journal of Physics A: Mathematical and General*, *31*(22), Article#5105.

Kaizu, K., De Ronde, W., Paijmans, J., Takahashi, K., Tostevin, F., & Wolde, P. R. (2014). The Berg-Purcell limit revisited. *Biophysical Journal*, *106*(4), 976–985.

Kaizu, K., Nishida, K., Sakamoto, Y., Kato, S., Niina, T., Nishida, N., Koizumi, M., Aota, N., & Takahashi, K. (2018). ecell/ecell4: An integrated software environment version 4.1.3. retrieved from https://github.com/ecell/ecell4.

Kalay, Z. (2012). Reaction kinetics in the plasma membrane. *Biotechnology Journal*, *7*(6), 745–752.

Kalay, Z., Fujiwara, T. K., & Kusumi, A. (2012). Confining domains lead to reaction bursts: Reaction kinetics in the plasma membrane. *PLoS One*, *7*(3), Article#e32948.

Kerketta, R., Halász, Á. M., Steinkamp, M. P., Wilson, B. S., & Edwards, J. S. (2016). Effect of spatial inhomogeneities on the membrane surface on receptor dimerization and signal initiation. *Frontiers in Cell and Developmental Biology*, *4*, Article#81.

Kerr, R. A., Bartol, T. M., Kaminsky, B., Dittrich, M., Chang, J.-C. J., Baden, S. B., Sejnowski, T. J., & Stiles, J. R. (2008). Fast Monte Carlo simulation methods for biological reaction-diffusion systems in solution and on surfaces. *SIAM Journal on Scientific Computing*, *30*(6), 3126–3149.

Kholodenko, B. N., Hoek, J. B., & Westerhoff, H. V. (2000). Why cytoplasmic signalling proteins should be recruited to cell membranes. *Trends in Cell Biology*, *10*(5), 173–178.

Klann, M. & Koeppl, H. (2012). Spatial simulations in systems biology: from molecules to cells. *International Journal of Molecular Sciences*, *13*(6), 7798–7827.

Klann, M., Koeppl, H., & Reuss, M. (2012). Spatial modeling of vesicle transport and the cytoskeleton: the challenge of hitting the right road. *PLoS One*, *7*(1), Article#e29645.

Kopelman, R. (1988). Fractal reaction kinetics. *Science*, *241*(4873), 1620–1626.

Lagerholm, B. C. & Thompson, N. L. (1998). Theory for ligand rebinding at cell membrane surfaces. *Biophysical Journal*, *74*(3), 1215–1228.

Lehnert, T. & Figge, M. T. (2017). Dimensionality of motion and binding valency govern receptor-ligand kinetics as revealed by agent-based modeling. *Frontiers in Immunology*, *8*, Article#1692.

Li, X. & Holmes, W. R. (2018). Biophysical attributes that affect CaMKII activation deduced with a novel spatial stochastic simulation approach. *PLoS Computational Biology*, *14*(2), 1–32.

Lombardo, S. J. & Bell, A. T. (1991). A review of theoretical models of adsorption, diffusion, desorption, and reaction of gases on metal surfaces. *Surface Science Reports*, *13*(1-2), 3–72.

Loose, M., Fischer-Friedrich, E., Herold, C., Kruse, K., & Schwille, P. (2011). Min protein patterns emerge from rapid rebinding and membrane interaction of MinE. *Nature Structural & Molecular Biology*, *18*(5), 577–583.

Lukkien, J., Segers, J., Hilbers, P., Gelten, R., & Jansen, A. (1998). Efficient Monte Carlo methods for the simulation of catalytic surface reactions. *Physical Review E*, *58*(2), Article#2598.

Meinecke, L. & Eriksson, M. (2016). Excluded volume effects in on-and off-lattice reaction-diffusion models. *IET Systems Biology*, *11*(2), 55–64.

Melo, E. & Martins, J. (2006). Kinetics of bimolecular reactions in model bilayers and biological membranes. a critical review. *Biophysical Chemistry*, *123*(2-3), 77–94.

Michalski, P. J. & Loew, L. M. (2016). SpringSaLaD: A spatial, particle-based biochemical simulation platform with excluded volume. *Biophysical Journal*, *110*(3), 523–529.

Milo, R. & Phillips, R. (2015). *Cell Biology by the Numbers*. London and New York: Garland Science.

Montalenti, F. & Ferrando, R. (2000). Probability of dimer reassociation in two dimensions. *Physical Review E*, *61*(4), Article#3411.

Montroll, E. W. (1969). Random walks on lattices. iii. calculation of first passage times with application to exciton trapping on photosynthetic units. *Journal of Mathematical Physics*, *10*(4), 753–765.

Montroll, E. W. & Weiss, G. H. (1965). Random walks on lattices. ii. *Journal of Mathematical Physics*, *6*(2), 167–181.

Mugler, A., Bailey, A. G., Takahashi, K., & Wolde, P. R. (2012). Membrane clustering and the role of rebinding in biochemical signaling. *Biophysical Journal*, *102*(5), 1069–1078.

Murray, J. D. (2001). *Mathematical Biology. II Spatial Models and Biomedical Applications Interdisciplinary Applied Mathematics Vol. 18*. New York: Springer.

Naqvi, K. R., Martins, J., & Melo, E. (2000). Recipes for analyzing diffusion-controlled

reactions in two dimensions: Time-resolved and steady-state measurements. *The Journal of Physical Chemistry B*, *104*(50), 12035–12038.

Naqvi, K. R., Mork, K. J., & Waldenstrom, S. (1980). Diffusion-controlled reaction kinetics. Equivalence of the particle pair approach of Noyes and the concentration gradient approach of Collins and Kimball. *The Journal of Physical Chemistry*, *84*(11), 1315–1319.

Nicolau Jr, D. V., Hancock, J. F., & Burrage, K. (2007). Sources of anomalous diffusion on cell membranes: A Monte Carlo study. *Biophysical Journal*, *92*(6), 1975–1987.

Noyes, R. M. (1954). A treatment of chemical kinetics with special applicability to diffusion controlled reactions. *The Journal of Chemical Physics*, *22*(8), 1349–1359.

Noyes, R. M. (1961). Effects of diffusion rates on chemical kinetics. *Progress in Reaction Kinetics and Mechanism*, *1*, 129–160.

Opplestrup, T., Bulatov, V. V., Gilmer, G. H., Kalos, M. H., & Sadigh, B. (2006). First-passage Monte Carlo algorithm: Diffusion without all the hops. *Physical Review Letters*, *97*(23), Article#230602.

Pablo, M., Ramirez, S. A., & Elston, T. C. (2018). Particle-based simulations of polarity establishment reveal stochastic promotion of Turing pattern formation. *PLoS Computational Biology*, *14*(3), 1–25.

Pfluegl, W. & Silbey, R. J. (1998). Long-time properties of random walks with a single trap. *Physical Review E*, *58*(4), Article#4128.

Phillips, R., Theriot, J., Kondev, J., & Garcia, H. (2012). *Physical Biology of the Cell*. London and New York: Garland Science.

Pitulice, L., Vilaseca, E., Pastor, I., Madurga, S., Garcés, J. L., Isvoran, A., & Mas, F. (2014). Monte Carlo simulations of enzymatic reactions in crowded media. Effect of the enzyme-obstacle relative size. *Mathematical Biosciences*, *251*, 72–82.

Plimpton, S. J. & Slepoy, A. (2005). Microbial cell modeling via reacting diffusive particles. *Journal of Physics: Conference Series*, *16*, 305–309.

Popov, A. V. & Agmon, N. (2001). Three-dimensional simulations of reversible bimolecular reactions: The simple target problem. *The Journal of Chemical Physics*, *115*(19), 8921–8932.

Popov, A. V. & Agmon, N. (2002). Exact solution for the geminate ABCD reaction. *The Journal of Chemical Physics*, *117*(12), 5770–5779.

Rao, C. V., Wolf, D. M., & Arkin, A. P. (2002). Control, exploitation and tolerance of intracellular noise. *Nature*, *420*(6912), 231–237.

Ridgway, D., Broderick, G., Lopez-Campistrous, A., Ru'aini, M., Winter, P., Hamilton, M., Boulanger, P., Kovalenko, A., & Ellison, M. J. (2008). Coarse-grained molecular simulation of diffusion and reaction kinetics in a crowded virtual cytoplasm. *Biophysical Journal*, *94*(10), 3748–3759.

Ritchie, R. & Sakakura, A. (1956). Asymptotic expansions of solutions of the heat conduction equation in internally bounded cylindrical geometry. *Journal of Applied Physics*, *27*(12), 1453–1459.

Roberts, E., Stone, J. E., & Luthey-Schulten, Z. (2013). Lattice microbes: High-performance stochastic simulation method for the reaction-diffusion master equation. *Journal of Computational Chemistry*, *34*(3), 245–255.

Ryan, K. R. & Shapiro, L. (2003). Temporal and spatial regulation in prokaryotic cell cycle progression and development. *Annual Review of Biochemistry*, *72*(1), 367–394.

Sanford, C., Yip, M. L. K., White, C., & Parkinson, J. (2006). Cell++ - simulating biochemical pathways. *Bioinformatics*, *22*(23), 2918–2925.

Saxton, M. J. (1987). Lateral diffusion in an archipelago. The effect of mobile obstacles. *Biophysical Journal*, *52*(6), 989–997.

Saxton, M. J. (1989). Lateral diffusion in an archipelago. Distance dependence of the diffusion coefficient. *Biophysical Journal*, *56*(3), 615–622.

Saxton, M. J. (2002). Chemically limited reactions on a percolation cluster. *The Journal of Chemical Physics*, *116(1)*, 203–208.

Saxton, M. J. (2007). Modeling 2D and 3D diffusion. In A. M. Dopico (Ed.), *Methods in Membrane Lipids* (pp. 295–321). New York: Springer.

Saxton, M. J. (2008). A biological interpretation of transient anomalous subdiffusion. ii. Reaction kinetics. *Biophysical Journal*, *94*(3), 760–771.

Saxton, M. J. & Jacobson, K. (1997). Single-particle tracking: Applications to membrane dynamics. *Annual Review of Biophysics and Biomolecular Structure*, *26*(1), 373–399.

Sayyidmousavi, A. & Rohlf, K. (2018). Reactive multi-particle collision dynamics with reactive boundary conditions. *Physical Biology*, *15*(4), Article#046007.

Schmit, J. D., Kamber, E., & Kondev, J. (2009). Lattice model of diffusion-limited bimolecular chemical reactions in confined environments. *Physical Review Letters*, *102*(21), Article#218302.

Schnell, S. & Turner, T. E. (2004). Reaction kinetics in intracellular environments with macromolecular crowding: Simulations and rate laws. *Progress in Biophysics and Molecular Biology*, *85*(2-3), 235–260.

Schöneberg, J. & Noé, F. (2013). ReaDDy-a software for particle-based reaction-diffusion dynamics in crowded cellular environments. *PLoS One*, *8*(9), Article#e74261.

Shimo, H., Arjunan, S. N. V., Machiyama, H., Nishino, T., Suematsu, M., Fujita, H., Tomita, M., & Takahashi, K. (2015). Particle simulation of oxidation induced band 3 clustering in human erythrocytes. *PLoS Computational Biology*, *11*(6), Article#e1004210.

Smith, S., Cianci, C., & Grima, R. (2016). Analytical approximations for spatial stochastic gene expression in single cells and tissues. *Journal of The Royal Society Interface*, *13*(118), Article#20151051.

Smith, S. & Grima, R. (2017). Fast simulation of Brownian dynamics in a crowded environment. *Journal of Chemical Physics*, *146*(2), Article#024105.

Smith, S. & Grima, R. (2018). Spatial stochastic intracellular kinetics: A review of modelling approaches. *Bulletin of Mathematical Biology*, *(special issue)*, 1–50.

Smoluchowski, M. (1917). Mathematical theory of the kinetics of the coagulation of colloidal solutions. *Zeitschrift Fur Physikalische Chemie*, *92*, 129–168.

Sokolowski, T. R., Paijmans, J., Bossen, L., Miedema, T., Wehrens, M., Becker, N. B., Kaizu, K., Takahashi, K., Dogterom, M., & Wolde, P. R. (2019). eGFRD in all dimensions. *Journal of Chemical Physics*, *150*(5), Article#054108.

Stauffer, D. & Aharony, A. (2014). *Introduction to Percolation Theory: 2nd edition*. Florida: CRC press.

Stiles, J. R., Bartol Jr, T. M., Salpeter, E. E., & Salpeter, M. M. (1998). Monte Carlo simulation of neuro-transmitter release using MCell, a general simulator of cellular physiological processes. In J. M. Bower (Ed.), *Computational Neuroscience* (pp. 279–284). New York: Springer.

Sturrock, M. (2016). Stochastic reaction-diffusion algorithms for macromolecular crowding. *Physical Biology*, *13*(3), Article#036010.

Sung, J. & Lee, S. (1999). Nonequilibrium distribution function formalism for diffusion-influenced bimolecular reactions: Beyond the superposition approximation. *The Journal of Chemical Physics*, *111*(3), 796–803.

Sung, J., Park, H., & Lee, S. (2001). Green's function approach to nonclassical reaction kinetics in fractal media. *Physical Review Letters*, *86*(11), 2463–2466.

Szabo, A. (1989). Theory of diffusion-influenced fluorescence quenching. *The Journal of Physical Chemistry*, *93*(19), 6929–6939.

Szpiro, G. G. (2003). *Kepler's Conjecture*. New Jersey: Wiley.

Tachiya, M. (1983). Theory of diffusion-controlled reactions: Formulation of the bulk reaction rate in terms of the pair probability. *Radiation Physics and Chemistry*, *21*(1-2), 167–175.

Takahashi, K., Tănase-Nicola, S., & Wolde, P. R. (2010). Spatio-temporal correlations can drastically change the response of a MAPK pathway. *Proceedings of the National Academy of Sciences*, *107*(6), 2473–2478.

Tolentino, T. P., Wu, J., Zarnitsyna, V. I., Fang, Y., Dustin, M. L., & Zhu, C. (2008). Measuring diffusion and binding kinetics by contact area FRAP. *Biophysical Journal*, *95*(2), 920–930.

Torney, D. C. & McConnell, H. M. (1983). Diffusion-limited reaction rate theory for two-dimensional systems. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, *387*(1792), 147–170.

Tsourkas, P. K. & Raychaudhuri, S. (2011). Monte Carlo investigation of diffusion of receptors and ligands that bind across opposing surfaces. *Annals of Biomedical Engineering*, *39*(1), 427–442.

van Zon, J. S. & Wolde, P. R. (2005). Simulating biochemical networks at the particle level and in time and space: Green's function reaction dynamics. *Physical Review Letters*, *94*(12), Article#128103.

Vilaseca, E., Isvoran, A., Madurga, S., Pastor, I., Garcés, J. L., & Mas, F. (2011). New insights into diffusion in 3d crowded media by Monte Carlo simulations: Effect of size, mobility and spatial distribution of obstacles. *Physical Chemistry Chemical Physics*, *13*(16), 7396–7407.

Watabe, M., Arjunan, S. N., Chew, W. X., Hiroshima, M., Kaizu, K., Kozuka, J., Ueda, M., & Takahashi, K. (2018). An application of bioimage simulation: Cooperative binding measurement. Retrieved on feb 2019 from https://arxiv.org/abs/1802.10080.

Weiss, G. H. & Rubin, R. J. (1982). Random walks: Theory and selected applications. *Advances in Chemical Physics*, 363–505.

Wettmann, L. & Kruse, K. (2018). The Min-protein oscillations in *Escherichia coli*: An example of self-organized cellular protein waves. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1747), Article#20170111.

Yogurtcu, O. N. & Johnson, M. E. (2015). Theory of bi-molecular association dynamics in 2D for accurate model and experimental parameterization of binding rates. *Journal of Chemical Physics*, *143*(8), 1–17.

Zhdanov, V. P. & Höök, F. (2015). Kinetics of enzymatic reactions in lipid membranes containing domains. *Physical Biology*, *12*(2), Article#026003.

Zhou, H. X., Rivas, G., & Minton, A. P. (2008). Macromolecular crowding and confinement: Biochemical, biophysical, and potential physiological consequences. *Annual Review of Biophysics*, *37*, 375–397.

Zhou, Y. & Hancock, J. F. (2018). Deciphering lipid codes: K-Ras as a paradigm. *Traffic*, *19*(3), 157–165.

Zimmerman, S. B. & Trach, S. O. (1991). Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of *Escherichia coli*. *Journal of Molecular Biology*, *222*(3), 599–620.

LIST OF PUBLICATIONS AND PAPERS PRESENTED

1. **Chew, W.-X**., Kaizu, K., Watabe, M., Muniandy, S. V., Takahashi, K., & Arjunan, S. N. V. (2019). Surface reaction-diffusion kinetics on lattice at the microscopic scale. *Physical Review E, 99*, Article#042411.

2. **Chew, W.-X**., Kaizu, K., Watabe, M., Muniandy, S. V., Takahashi, K., & Arjunan, S. N. V. (2018). Reaction-diffusion kinetics on lattice at the microscopic scale. *Physical Review E,* 98, Article#032418.