UNIVERSITY OF CALIFORNIA, SAN DIEGO

Stochastic Modeling of Advection-Diffusion-Reaction Processes in Biological Systems

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Engineering Sciences (Mechanical Engineering)

by

TaiJung Choi

Committee in charge:

Daniel M. Tartakovsky, Chair
Shankar Subramaniam, Co-Chair
Marcos Intaglietta
Ratneshwar Lal
Mano R. Maurya
Sutanu Sarkar

2013

The dissertation of TaiJung Choi is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____
Co-Chair

_____
Chair

University of California, San Diego

2013

DEDICATION

To "my family, my advisers, my friends".

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to acknowledge Professor Daniel Tartakovsky and Shankar Subramaniam for their support as advisors during my doctoral period. I would also like to appreciate Doctor Mano Maurya for his help and advices. Through many research topics and multiple drafts for journal papers and dissertation, their guidance has proved to be very priceless and invaluable to me.

The text of this dissertation includes the reprints of the following papers, either accepted or submitted for consideration at the time of publication. The dissertation author was the primary investigator and author of these publications.

Chapter 2

TaiJung Choi, Mano R. Maurya, Daniel M. Tartakovsky, Shankar Subramaniam (2010), 'Stochastic Hybrid Modeling of Intracellular Calcium Dynamics'. *J. Chem. Phys., 133, 165101.*

Chapter 3

TaiJung Choi, Mano R. Maurya, Daniel M. Tartakovsky, Shankar Subramaniam (2012), 'Stochastic Operator-Splitting Method for Reaction-Diffusion Systems'. *J. Chem. Phys., 137, 184102.*

Chapter 4

TaiJung Choi, Mano R. Maurya, Daniel M. Tartakovsky, Shankar Subramaniam (2013), 'Stochastic Modeling of Chemotaxis-Diffusion-Reaction Processes'. *Under preparation.*

VITA

2012	Ph.D. in Mechanical and Aerospace Engineering, University of California, San Diego.
2004	M.S. in Mechanical and Aerospace Engineering, Seoul National University, Seoul, Korea
2002	B.S. in Mechanical Engineering, Korea University, Seoul, Korea

JOURNAL PUBLICATIONS

TaiJung Choi, Mano R. Maurya, Daniel M. Tartakovsky, Shankar Subramaniam (2010), 'Stochastic Hybrid Modeling of Intracellular Calcium Dynamics'. *J. Chem. Phys., 133, 165101.*

TaiJung Choi, Mano R. Maurya, Daniel M. Tartakovsky, Shankar Subramaniam (2012), 'Stochastic Operator-Splitting Method for Reaction-Diffusion Systems'. *J. Chem. Phys., 137, 184102.*

TaiJung Choi, Mano R. Maurya, Daniel M. Tartakovsky, Shankar Subramaniam (2013), 'Stochastic Modeling of Chemotaxis-Diffusion-Reaction Processes'. *Under preparation.*

SELECT PRESENTATIONS

TaiJung Choi, Mano R. Maurya, Daniel M. Tartakovsky, Shankar Subramaniam, 'Stochastic Simulation of Cytosolic Calcium Dynamics', (2009) *AIChE Annual Meeting, Nashville, TN, November 8-13.*

Mano R. Maurya, TaiJung Choi, Daniel M. Tartakovsky, Shankar Subramaniam, 'Timescale Analysis of Cytosolic Calcium Dynamics', (2010) *AIChE Annual Meeting, Salt Lake City, UT, November 7-12.*

TaiJung Choi, Mano R. Maurya, Daniel M. Tartakovsky, Shankar Subramaniam, 'Stochastic Operator Splitting Method for Biological Systems', (2011) *AIChE Annual Meeting, Minneapolis, MN, October 16-21.*

POSTER PRESENTATIONS

TaiJung Choi, Mano R. Maurya, Daniel M. Tartakovsky, Shankar Subramaniam, 'Stochastic Modeling of Calcium Dynamics in RAW 264.7 Cells', (2010) *Research Expo, University of California, San Diego, April 15.*

TaiJung Choi, Mano R. Maurya, Daniel M. Tartakovsky, Shankar Subramaniam,

'Stochastic Operator-Splitting Method for Reaction-Diffusion Process', (2011) *Re-*

search Expo, University of California, San Diego, April 14.

ABSTRACT OF THE DISSERTATION

Stochastic Modeling of Advection-Diffusion-Reaction Processes in Biological Systems

by

TaiJung Choi

Doctor of Philosophy in Engineering Sciences (Mechanical Engineering)

University of California, San Diego, 2013

Daniel M. Tartakovsky, Chair
Shankar Subramaniam, Co-Chair

This dissertation deals with complex and multi-scale biological processes. In general, these phenomena can be described by ordinary or partial differential equations and treated with deterministic methods such as Runge-Kutta and alternating direction implicit algorithms. However, these approaches cannot predict the random effects caused by the low number of molecules involved and can result in severe stability and accuracy problem due to wide range of time or length scales depending upon the system being studied. In the first part of the dissertation, therefore, we develop the stochastic hybrid algorithm for complex reaction networks. Deterministic models of biochemical processes at the subcellular level might become inadequate when a cascade of chemical reactions is induced by a few

molecules. Inherent randomness of such phenomena calls for the use of stochastic simulations. However, being computationally intensive, such simulations become infeasible for large and complex reaction networks. To improve their computational efficiency in handling these networks, we present a hybrid approach, in which slow reactions and fluxes are handled through exact stochastic simulation and their fast counterparts are treated partially deterministically through chemical Langevin equation. The classification of reactions as fast or slow is accompanied by the assumption that in the time-scale of fast reactions, slow reactions do not occur and hence do not affect the probability of the state. In the second and third part of the dissertation, we employ stochastic operator splitting algorithm for (chemotaxis-)diffusion-reaction processes. The reaction and diffusion steps employ stochastic simulation algorithm and Brownian dynamics, respectively. Through theoretical analysis, we develop an algorithm to identify if the system is reaction-controlled, diffusion-controlled or is in an intermediate regime. The time-step size is chosen accordingly at each step of the simulation. We apply our algorithm to several examples in order to demonstrate the accuracy, efficiency and robustness of the proposed algorithm comparing with the solutions obtained from deterministic partial differential equations and Gillespie multi-particle method. The third part deals with application of the stochastic-operator splitting approach to model the chemotaxis of leukocytes as part of the inflammation process during wound healing. We analyze both chemotaxis as well as the diffusion process as a drift phenomenon. We use two dimensionless numbers, Damköhler and Pèclet num-

ber, in order to analyze the system. Damköhler number determines if the system is reaction-controlled or drift controlled and Pèclet number identifies which phenomenon is dominant between diffusion and chemotaxis.

Chapter 1

Introduction

1.1 Multi-scale modeling in biology

Biological systems involve various processes taking place at a wide range of spatial and temporal scales. Biological systems have spatial scales which range from kilometers, e.g. the habitat of animals, to micrometers such as phenomena at the cellular level. The spatial scales can range from meters to microns in dealing with a single organism. Within the intercellular space, depending upon the context, the spatial domain of interest can vary from nanometers to micrometers. Similar phenomena is observed for the temporal scale. For example, population fluctuation of some animal group can be detected at the time scale of years whereas events such as cell division occur on a scale of hours and molecular chemical reactions take place within milliseconds to minutes. Specially at subcellular volumes when the number of molecules can be low so that the continuum approximation becomes

invalid, stochastic effects become important. The interplay of stochasticity with the multiplicity in the spatial and temporal domains is complex. Not accounting the different time and spatial scales in the modeling and simulation of stochastic systems results in errors and/or large simulation times. The development of robust mathematical techniques for the modeling and simulation of stochastic biological systems with multiscale temporal and spatial scales is the main focus of this dissertation. The stochasticity is quantitatively modeled through the use of random variables. In addition, this dissertation also deals with approaches to account for multiple spatial and temporal scales. In this dissertation, several biological systems are used to demonstrate the effectiveness of methodologies developed. The biological systems include (1) regulation of the dynamics of intracellular calcium ion levels, (2) molecular diffusion and reactions in *E coli* and (3) leukocyte chemotaxis through the tissue during the inflammation phase of wound healing process. These case studies are linked with each other in various ways and at various biological scales (from intracellular to tissue level) and serve as excellent test-beds in multi-scale mathematical modeling and quantitative systems biology.

1.2 Stochasticity in biology

Over the last few decades, in the field of molecular biology, the importance of stochasticity has been increasingly recognized and outstanding developments have led to a better understanding of biological systems at the subcellular level.

At the level of micro-scale systems such as the interactions between molecules, e.g. DNA, mRNA, protein, small molecules, it follow an important law of physics, i.e., randomness or fluctuations in a system are inversely proportional to the square root of the number of particles [1]. Therefore, a lower number of molecules (or low concentration) results in high fluctuation which is largely due to thermal oscillations. For example, in processes such as gene transcription/regulation and signal transduction [2], number of molecules involved in the chemical reactions is usually low, e.g., a single DNA template, tens of mRNA molecules and around hundred molecules of transcription factors. Such stochastic effects arising due to the inherent nature of biochemical interactions are often termed as intrinsic noise. In addition, there exists an extrinsic noise as well caused by random fluctuations in other factors such as the number of ribosomes, the stage of the cell cycle, mRNA degradation, and the cellular environment [3]. Yarchuk *et al.* showed that protein production occurs in short bursts at random time intervals rather than in a continuous manner [4]. In addition, spatial randomness plays an important role during processes such as *E. coli* movement [5], tumor growth [6] and leukocyte chemotaxis [7].

1.3   Hybrid algorithms for multi-scale systems

1.3.1   Motivation

Intracellular signaling is an important event in cellular life that mediates most of cell functions, such as adaptation in response to environmental changes, metabolism, cellular growth and proliferation. Mathematical modeling, traditionally based on ordinary differential equations, helped to explain and illustrate many of these complex phenomena, including the bistability and graded versus switch-like response in intracellular signaling [8] and sub-population variability [9]. ODE-based formulations offer accurate predictions of biochemical dynamics with large numbers of molecules, but are expected to fail if the numbers of reacting molecules become exceedingly small. When this occurs, randomness associated with the dynamics of individual molecules becomes important and calls for a probabilistic (stochastic) description. Chapter 2 provides an example of such modeling in the context of intracellular calcium dynamics.

All ODE-based models, and most of stochastic models of the type discussed above, are based on the assumption of a perfectly mixed (homogeneous) system, in which every point (or volume) in space has the same concentration (or number of molecules) of reacting species. This assumption becomes invalid when the number of reacting molecules becomes small and transport also take place in heterogeneous crowded environments. In Chapters 3 and 4, we develop computational methods to deal with such spatial heterogeneity in the context of

molecular diffusion and reactions in *E. coli* (Chapter 3) and leukocyte chemotaxis (Chapter 4). These two biological phenomena illustrate the complexity of most cellular processes by exhibiting multiple time and length scales, randomness and spatial inhomogeneity. These chapters present a new stochastic hybrid algorithm for multi-scale systems and a new stochastic operator splitting algorithm for (chemotaxis-)diffusion-reaction systems, respectively.

1.3.2   Temporal multi-scale processes

In Chapter 2, we present a novel algorithm for the stochastic simulation of multi-scale (time-domain) biochemical processes. The methodology is applied to study intracellular calcium dynamics in mouse RAW 264.7 macrophage cells. Intracellular signaling plays an important role in cellular life that regulates most of its functions, such as adaptation in response to environmental changes and regular functions including metabolism, cellular growth and proliferation. Mass balance for chemical reactions which is described by ordinary differential equations (ODE) is generally applied to analyze these chemical reactions. These ODE-based formulations can predict accurately the dynamics of biochemical pathways with large numbers of molecules of all reacting species. However, it might fail in the case that the concentrations of involved chemical species become exceedingly small [10]. In this case, it is necessary to apply stochastic analysis which treat chemical reactions as random events. A chemical master equation (CME) yields an exact probabilistic description of multi-species reactions, but its high dimensionality renders it com-

putationally prohibitive. Gillespie algorithm [11], a good approximation of CMEs, deals with all possible reactions using uniformly distributed random variables in [0,1]. A tau-leap algorithm [12] or chemical Langevin equation (CLE) [13] can further approximate CME using Poisson random variables and Gaussian random variables, respectively. Implicit in these and other approximations of the SSA is a trade-off between computational speed-up and accuracy, which undermines their use in complex multi-scale biochemical phenomena involving fast and slow reactions. Therefore, we present a hybrid algorithm in which slow and fast reactions are identified, they can be reclassified during simulation in response to changes in concentrations, and we can deal with complex fluxes that cannot be modeled explicitly through reactions.

1.3.3 Temporal and spatial multi-scale processes

In Chapters 3 and 4, we investigate diffusion-reaction systems in various biological systems such as CheY diffusion in *E. coli* (Chapter 3) and inflammation process during wound healing (Chapter 4). In addition to randomness from small number of molecules, we have stochasticity due to inhomogeneity of molecules across the space. In order to simulate the spatial variation, mesh/grid-based approaches are used. The number of molecules within each voxel can be low resulting in stochasticity. Therefore, we have to deal with two types of randomness, i.e., in the temporal domain and in the spatial domain. Partial differential equations (PDEs) can predict accurately the dynamics of spatially heterogeneous systems

composed of chemical species with high concentration. However, similar to ODE-based models, they fail to account for the randomness inherent in a system comprised of small number of molecules. A number of simulation methods have been developed for the simulation of reaction-diffusion systems. The Green's function reaction dynamics [14] and Smoldyn's algorithm [15], employs Brownian dynamics to track the diffusion of molecules and assume that bimolecular reactions can take place when two molecules exist within a certain distance. These requirements necessitate the tracking of individual particles and/or distances between them, which makes such algorithms computationally expensive. MesoRD [16] and the Gillespie multi-particle (GMP) methods [17, 18] trade representational accuracy for computational efficiency. They are based on a reaction-diffusion master equation [19], which generalizes a chemical master equation developed for well-mixed chemical reactions by discretizing the space into a collection of cells and treating each cell as a well mixed system. MesoRD [16] treats diffusion as a unimolecular reaction whose reaction rate is related to the corresponding diffusion coefficient. The GMP method [17, 18] employs an operator-splitting scheme in which the Gillespie algorithm and cellular automata [20] handle reaction and diffusion processes, respectively.

We have developed a stochastic numerical algorithm to simulate reaction-diffusion processes with a small number of non-uniformly distributed molecules. It employs an operator-splitting, in which the Gillespie algorithm [11] and Brownian dynamics are used to simulate reaction and diffusion processes, respectively. Our

algorithm is conceptually similar to the GMP method in that it relies on operator-splitting. However, it offers a number of computational advantages in terms of both accuracy and efficiency. First, the cellular automata used in the GMP method restrict a particle's movement during one fixed time-step to the adjacent cells only, while Brownian motion places no restrictions on the distance particles can travel during one time-step, thus gaining in computational efficiency. Second, Brownian dynamics provides a more accurate representation of diffusion than cellular automata. Third, our algorithm offers the flexibility of adaptive selection of the time-step sizes for operator-splitting, depending on whether the system is reaction- or diffusion-controlled.

We have also studied the inflammation process during wound healing in which leukocyte cells sense the gradient of chemoattractants from the wound site and chemotax in the direction of higher concentration while also diffusing randomly. It involves three processes, diffusion, reaction and chemotaxis. In addition to the diffusion-reaction processes explained above, we also have to deal with the chemotaxis process. In order to identify the drift (chemotaxis + diffusion) time scale, Péclet number is introduced. Damköhler number decides if system is diffusion or reaction controlled.

1.4   Conclusions and future directions

In biological systems, randomness is caused by small number of molecules and spatial inhomogeneity. Therefore, we have developed stochastic hybrid algorithms for multi-scale reaction systems, in which chemical reactions are classified as fast or slow according to propensity functions and chemical species are classified as low or high based on the number of molecules. We applied our hybrid algorithm to intracellular calcium dynamics in mouse macrophage cells by applying Gillespie algorithm or Chemical Langevin equation appropriately according to system's state. Next, we have developed stochastic operator-splitting method for (chemotaxis)-diffusion-reaction systems. Proper selection of time step is very important because drift (chemotaxis, diffusion) time constant and reaction time constant may be significantly different. Furthermore, one needs to identify the dominant process between diffusion and chemotaxis. Hence, we use Damköhler and Péclet numbers. In this dissertation, these novel methodologies have been developed and applied to interesting biological systems in order to verify accuracy, efficiency and robustness of the proposed algorithms. We have applied these approaches to several biological systems of low to moderate complexity.

In the future, these approaches can be tested on more complex and realistic systems. For example, stem cells exist during all phases of development, e.g., embryonic stem (ES) cells during the embryonic stage and adult stem cells after all the organs are formed. Stem cells are characterized by two important abilities,

viz., renew themselves and differentiate into a variety of distinct lineages. ES cells are omnipotent or pluripotent i.e., they have the ability to generate all embryonic tissues. Due to their potential to regenerate tissue damaged due to disease or injury, stem cell-based therapies for various degenerative diseases are being developed [**?**]. Stem cell properties are governed by a complex set of interactions between signaling from the extracellular and intercellular environment and the dynamics of core transcriptional machinery. Spatial variability plays an important role in this process. Therefore, we can apply our operator-splitting approach to stochastic stem-cell fate decision modeling in order to quantitatively study the molecular mechanism of ES cells [**?**]. In another application, we will consider all processes from rolling to chemotaxis of leukocytes based on our stochastic operator-splitting method because spatial variation is important for the leukocyte movement inside the capillary, across endothelial cell wall and through the tissue. We will employ different boundary and initial conditions for these three connected spatial zones. We will perform a comprehensive quantitative analysis of leukocyte motion during wound healing process by accounting for blood flow, wall shear stress and contact force between endothelial cells.

Chapter 2

Stochastic Hybrid Modeling of Intracellular Calcium Dynamics

2.1   Introduction

Intracellular signaling is an important event in cellular life that mediates most of its functions, such as adaptation in response to environmental changes and regular functions including metabolism, cellular growth and proliferation. Mathematical modeling has helped to explain and illustrate many of these complex phenomena, including the bistability and graded versus switch-like response in intracellular signaling [8], auto-catalysis as a mechanism of positive feedback in the cell cycle [21], and sub-population variability [9]. Much of this modeling is done in a deterministic setting, and involves systems of coupled ordinary differential equations (ODEs) describing the rate of change of components (reactants and products)

11

of the biochemical reactions and other processes involved in the pathway.

ODE-based formulations provide accurate predictions of the dynamics of biochemical pathways with large numbers of molecules of all reacting species, but might fail when the concentrations of reactants and/or products become exceedingly small so than only a few molecules (less than 10 in some cases) are involved [10]. Indeed, for small volumes and small concentrations that often characterize sub-cellular processes, the very concept of concentration breaks down. When this occurs, randomness associated with the dynamics of individual molecules becomes pronounced, necessitating the use of probabilistic (stochastic) models. A chemical master equation (CME) yields an exact probabilistic description of multi-species reactions, but its high dimensionality renders it computationally prohibitive.

Gillespie's stochastic simulation algorithm (SSA) [11] provides an exact sampling of the solution of the CME, thus providing highly accurate results with sufficient sampling. The computational efficiency of the SSA can be increased by adopting, for example, a tau-leap algorithm [12] or its continuous-limit approximation in the form of a chemical Langevin equation (CLE) [13]. Implicit in these and other approximations of the SSA is a trade-off between computational speed-up and accuracy, which undermines their use in complex multi-scale biochemical phenomena involving fast and slow reactions. A quasi-steady-state approximation [22], which neglects the fast reactions by assuming that a subset of chemical species is at steady state at the timescale of interest, is efficient but clearly inexact.

Some of the more recent contributions in this area include: (1) speed-up of

computation through a binomial tau-leaping approach [23] and k-skip method [24], (2) time-scale/reaction partitioning based on the propensity values [25], a hybrid approach [26] and quasi-steady-state approximation [27], (3) partial-propensity-based approach [28] and (4) alternative formulations of CLE [29]. Besides, [30] has developed an approach to perform stochastic simulation of reaction systems with time-delays. [31] have developed a software called Biomolecular Network Simulator to study various aspects of stochastic simulation of complex biomolecular reaction networks. [32] have presented a detailed analysis of issues in simplification of Michaelis-Menten formulation into a single-step reaction in stochastic simulation. [33] have developed a methodology for parametric sensitivity analysis in stochastic simulation of reaction networks. By no means this is an exhaustive list.

Hybrid methods, e.g., by [26], which we pursue here, address the multi-scale nature of reactive systems by identifying fast and slow reactions, and simulating the former with a CLE and the latter with Gillespie's SSA. This approach significantly reduces simulation time without compromising the accuracy of the outputs. We present a hybrid algorithm in which slow and fast reactions are identified *a priori*, they can be reclassified during simulation in response to changes in concentrations, and we can deal with complex fluxes that cannot be modeled explicitly through reactions. An example of such as flux, in the model of cytosolic calcium dynamics, is the flux of $[Ca^{2+}]$ from the endoplasmic reticulum to the cytosol through inositol 1,4,5- trisphosphate receptor channels (please see the expression for J_{ch} in Section 2.3.3).

We have used the dynamics of cytosolic calcium as a case study to test our approach. The cytosolic calcium dynamics and its mathematical descriptions are briefly discussed in Section 3.2 to motivate the development of a multi-scale stochastic hybrid algorithm (SHA) in section 2.3, which consists of the following steps. Section 2.3.1 contains a formulation of the calcium dynamics model used in our analysis. In Section 2.3.2, we compare the performance of existing stochastic approaches, i.e., the Gillespie's SSA, a tau-leap algorithm, and a chemical Langevin equation. In Section 2.3.3, we present the SHA, which consists of deterministic and stochastic components, explicitly accounts for the presence of slow and fast reactions, and incorporates complex fluxes that cannot be modeled through reactions explicitly. An approach to handle reactions with complex rate expressions is also presented in this section explaining why the existing approaches to deal with complex rates laws such as Michaelis-Menten mechanism [22, 27, 32] may not be directly applicable. The practical implementation of the SHA to the cytosolic calcium dynamics model [9] is presented in Section 2.3.4. Section 2.4 contains the results of stochastic simulations of cytosolic calcium dynamics, whose biological implications are further discussed in Section 2.5.

2.2   Dynamics of Cytosolic Calcium

Cytosolic calcium is a second messenger that plays an important role in intracellular signaling. Dynamic changes in intracellular calcium serve both as an

important indicator of cellular events and as a quantitative measure of cellular response to stimuli. In addition to affecting gene regulation, calcium regulates the activity of many proteins such as calmodulin [34], calreticulin [35, 36, 37] and calcineurin [38]. Through such regulation, cytosolic calcium affects many functions including muscle contraction, fertilization, learning and memory, among many others.

2.2.1   Biological mechanisms and pathways

Following [9], we consider a signaling network for calcium dynamics (Fig. 2.1), which represents the ligand-induced release of calcium from the ER into cytosol, binding of calcium (Ca_i) to proteins (Pr) in the cytosol (shown) and in the ER (not shown) and other calcium exchange fluxes to/from the ER, the extra-cellular space and mitochondria. In the basal state, the channel flux from the ER is very small and, along with the leakage flux from the ER, is balanced by the Ca^{2+} uptake back into the ER by the sarco(endo)plasmic reticulum calcium ATPase (SERCA) pump; the net flux across the mitochondria and the PM is zero; and the Ca^{2+} outflux from the cytosol to the extracellular matrix (ECM) is mediated by the plasma membrane calcium ATPase (PMCA) pump and the Na^+/Ca^{2+} exchanger (NCX). The influx across the plasma membrane consists of a non-specific leakage flux and an [IP$_3$]-dependent specific flux, which combines many fluxes including the entry through store-operated channels in response to the ER depletion and other effects [39]. Ca^{2+} binds to buffer proteins in all three compartments, the

cytosol, the ER and the mitochondria, for which rapid buffering kinetics suggested earlier [40, 41] is used. For a more detailed analysis of the perturbation of the calcium network, we refer the reader to [42]. Maurya et al. [9] developed a kinetic model for calcium signaling in mouse macrophage-like RAW 264.7 cell and simulated the calcium dynamics for the ligand Complement 5a (C5a). In non-excitable cells, such as macrophages, ligand-induced release of calcium from the endoplasmic reticulum (ER) is the main initiator of calcium dynamics. Upon stimulation with C5a, the C5a receptor (C5aR) becomes activated leading to activation of G-protein, $G_{\alpha,i}$ followed by activation of phospholipase C (PLC) β (PLCβ). The net result is increased hydrolysis of phosphatidylinositol 4,5-bisphosphate (PIP$_2$) into inositol 1,4,5-trisphosphate (IP$_3$) and increase in the levels of cytosolic calcium ([Ca^{2+}]$_i$) due to the opening of IP3 receptor (IP$_3$R) channels on the endoplasmic (or sarcoplasmic) reticulum (ER/SR) membrane [43]. The concentration of calcium in the cytosol is in sub-micromolar range whereas it can be 10's to 100's micromolar (μM) in the ER [43]. Hence, upon opening of the IP$_3$R channels, the large gradient of calcium between the ER and the cytosol results in a burst (large peak) of [Ca^{2+}]$_i$ response [43]. Through a positive feedback mechanism, also known as calcium-induced calcium release (CICR) [44, 45], more Ca^{2+} is released from the ER into the cytosol. Most of the calcium released binds to various proteins, such as calmodulin (CaM). Calcium is also pumped back to the ER by the SERCA pump. Some calcium is also expelled to the extracellular space through the Na^{2+}/Ca^{2+} exchanger (NCX) and the PMCA pump. The resulting calcium

current facilitates the cross-talk between calcium dynamics and action potential in cardiac pacemaker cells [46]. Calcium exchange between cytosol and mitochondria also has been observed at elevated level of $[Ca^{2+}]_i$.

2.2.2 Mathematical representations of calcium dynamics

Mathematical models of cytosolic calcium dynamics were developed for both excitable [47, 48, 49, 50, 51] cells and non-excitable [41, 47, 52] cells. Many of these models deal with spatial distribution of calcium by employing two- or three-dimensional partial-differential equations [53]. Most of such models rely on non-specific (independent of cell-type) parameter values and provide qualitative (rather than quantitative) predictions of the behavior of various cell types. Moreover, they fail to capture the calcium dynamics in RAW 264.7 cells without parameter-tuning [9].

The Maurya et al. [9] model overcomes these limitations by using experimental measurements in RAW cells to constrain parameter values. The model neglects molecular diffusion, the presence of IP_3R clusters, and local-concentration effects in the mechanism for calcium release from the ER [54], all of which are accounted for in the work by [55, 56, 49]. On the other hand, it includes detailed mechanisms of G-protein coupled receptor and G-protein activation and inactivation, which are absent in the Refs. [39, 41, 52, 53]. The model enables the analysis of the effects of single and multiple knockdowns of proteins and sub-populational variability, i.e., to account for the fact that different cell-populations, when trig-

gered by the same strength of a stimulus, result in quantitatively and qualitatively different responses (different peak heights, rise-time, etc.) [57]. Hence, we adopt the signaling network identified by [9] as the basis for the present analysis. The focus of the modeling studies is on the sensitivity analysis of the peak-height of cytosolic Ca^{2+} to stochastic versus deterministic simulation.

2.3   Materials and Methods

A mathematical representation of the signaling network identified by [9] is presented in Section 2.3.1. The performance of standard stochastic simulation algorithms is compared in Section 2.3.2. A new hybrid algorithm that significantly improves the computational efficiency of the standard stochastic algorithms is presented in Section 2.3.3. The application of the hybrid algorithm to the cytosolic calcium dynamics model [9] is presented in Section 2.3.4.

2.3.1   The mathematical model of cytosolic calcium dynamics

A system of ordinary differential equations (ODEs) that describe the cytosolic calcium dynamics [9] accounts for the chemical reactions grouped into the four modules in Fig. 2.1B. The receptor module (box 1) consists of the reactions 1-11 responsible for receptor activation, desensitization of the ligand-bound active receptor due to its phosphorylation, internalization of the ligand-bound phospho-

rylated receptor and receptor recycle. The GTPase cycle module (box 2) consists of reactions 12-16 corresponding to activation and deactivation of G-protein (G-protein is active when $G_{\beta\gamma}$ and $G_{\alpha,i}T$ are separated). The IP_3 module (box 3) includes activation of $PLC\beta$ upon binding of $G_{\beta\gamma}$ and cytosolic Ca^{2+} and subsequently catalyzed hydrolysis of PIP_2 into IP_3 and DAG. Reactions 19 and 20 capture IP_3 metabolism, i.e. its degradation/conversion to/from other inositolphosphates and back to PIP_2, with only one intermediate pseudo-species, namely $IP_{3,p}$ or IP_3 product (Fig. 2.1A) [41]. Positive feedback effects from calmodulin constitute the fourth module (box 4).

The cytosol and other compartments are assumed to be well-mixed. The state variables are described by a set of ODEs [58] involving the Ca^{2+} fluxes between different cellular compartments and other fluxes due to reactions. The 15 state variables (concentrations) used to model the details of ligand-induced generation of IP_3 are [L], [R], [LR], $[G_{\beta\gamma}]$, [GRK], $[LR_p]$, $[R_p]$, $[LR_i]$, $[R_{p,i}]$, $[R_{\text{pool}}]$, $[G_{\alpha,i}T]$, $[G_{\alpha,i}D]$, $[PIP_2]$, $[IP_3]$ and [CaM]. [X] represents concentration of species X. These differential equations involve fluxes only related to reactions modeled explicitly. Calcium dynamics introduces four additional state variables: $[Ca^{2+}]_i$, $[Ca^{2+}]_{ER}$, h and $[Ca^{2+}]_{mit}$, where $[Ca^{2+}]_{ER}$ and $[Ca^{2+}]_{mit}$ denote the concentrations of free Ca^{2+} in the ER and mitochondria, respectively; and h is the fraction of IP_3R to which calcium is not bound at the inhibitory site (IP_3 and calcium may or may not be bound at the other two sites, respectively) [59]. These differential equations deal with flux expressions due to complex lumped mechanisms which

cannot be modeled through reactions explicitly. Thus, the model by [9, 60] has 19 state variables. The quantities of all chemical species are in terms of their concentrations, normalized with respect to a unit volume of the cytosol. The model involves 65 reaction-rate parameters, including both simple and complex reaction fluxes and other flux exchanges between different compartments.

In this analysis, we focus on the calcium dynamics in the regimes with exceedingly small concentrations of relevant chemical compounds. To give an example, for dose response, corresponding to the lowest dose of the ligand C5a, the number of the molecules is 180 (0.1% of 30nM concentration). In another case, in sensitivity analysis of $G_{\beta\gamma}$, the number of molecules of $G_{\beta\gamma}$(total pool) considered is 2500 at 5% level of nominal value. Corresponding to this, the number of molecules of free $G_{\beta\gamma}$ is 10. In such regimes, the fidelity of continuum (ODE-based) descriptions might be compromised, and stochastic effects become important.

2.3.2 Comparison of computational efficiency of stochastic simulation algorithms

For the sake of completeness, in Appendix A, we present a brief overview of existing stochastic algorithms, namely Gillespie algorithm, tau-leap method and chemical Langevin equation. To compare their performance, we have applied these three algorithms to an enzymatic reaction satisfying the Michaelis-Menten rate law

(example taken from [61]),

$$S + E \xrightarrow{c_1} C, \qquad C \xrightarrow{c_2} S + E, \qquad C \xrightarrow{c_3} P + E, \qquad (2.1)$$

where S, E, C, and P denote the substrate, enzyme, enzyme-substrate complex, and product, respectively, or the number of their molecules. Fig. 2.2 shows the temporal evolution of $S(t)$ and $P(t)$ from their initial levels $S(0) = 312, E(0) = 125$ and $P(0) = 0$, computed with the three approaches for stochastic simulation described above. The three algorithms yield similar predictions, with the tau-leap and CLE algorithms giving nearly indistinguishable solutions.

Fig. 2.2A shows time-course of one realization from each method. Although the single time-courses show good agreement, time-course of mean and standard deviation of 1024 realizations are computed as well in order to ensure that they have similar statistical characteristics. Fig. 2.2B-C show excellent agreement among three algorithms in terms of mean and standard deviation. Next three histograms show probability distribution of the number of molecules of S sampled at t=10 second (Fig. 2.2D-F). The three histograms have almost same values of the mean ([Gillespie, Tau-leap, CLE] = [140.40, 139.25, 139.89]) and standard deviation ([Gillespie, Tau-leap, CLE] = [5.84, 6.09, 6.06]).

Table 2.1 demonstrates the scalability of the three stochastic algorithms with the number of molecules involved in the simulation of Eq. (2.1). As the initial number of molecules, $S(0)$ and $P(0)$, increases 100-fold, the computational time of the Gillespie algorithm increases almost 100-fold, while the run times of

the tau-leap and CLE algorithms remain practically unchanged. The run times reported represent Matlab simulations carried out on a Windows based PC with 2.1GHz Intel dual core processor and 2GB RAM.

2.3.3 A multi-scale hybrid approach

While the use of the CLE is appealing due to its computational efficiency, its accuracy suffers as the number of molecules involved in the chemical reactions becomes small. Likewise, the Gillespie algorithm is attractive due to its accuracy but it becomes inefficient when the number of chemical reactions and/or molecules becomes large. This dichotomy calls for the use of a hybrid approach (described in Section 2.3.3 below) in which fast reactions are tackled with the CLE, and the Gillespie algorithm is employed to simulate slow reactions.

An additional complication in modeling the cytosolic calcium dynamics arises from the presence of fluxes in which reactions are either absent or modeled implicitly and, hence, are not readily amenable to the stochastic formulations described above. These fluxes are modeled deterministically via ODEs as described in Section 2.3.3, giving rise to a stochastic-deterministic hybrid approach. Besides, the rate expressions for some reactions are complex. These rate expressions are a combination (function) of one or more law of mass action kinetics, Michaelis-Menten kinetics or Hill-dynamics-based terms. A stochastic treatment of such reactions in terms of propensity functions is described in Section 2.3.3. Our new multi-scale hybrid approach accounts for all these three scenarios.

Multi-scale approach

In many complex biochemical systems, including the cytosolic calcium dynamics, some reactions occur very frequently over short time-intervals, while others seldom occur. In deterministic ODE-based models, the Jacobian matrix, which is a function of both the reaction rate constants and the species concentrations, can be used to classify species as fast or slow. In particle-based stochastic simulations, the system proceeds through firing of reactions and hence the speed of both the reactions and species is important. To call a reaction "slow" or "fast", the knowledge of reaction rate constants alone is not sufficient. Indeed, a reaction with a large reaction rate constant cannot be classified as "fast" if they involve small numbers of reactant species. The approach presented below is, essentially, based on the previous work of [26] and [62] (see also the contribution of [63]).

Following [26], we classify a j-th reaction as fast if the following two constraints on the propensity function (Eq. (A.2)) and the number of molecules of each species involved in the reaction are simultaneously satisfied,

$$a_j[\mathbf{X}(t)]\mathrm{d}t \gg \alpha, \qquad 1 \leq j \leq M \tag{2.2a}$$

and

$$X_i(t) > \beta|\nu_{j_i}|, \qquad 1 \leq i \leq N, \tag{2.2b}$$

where ν_{j_i} are the components of the vector $\boldsymbol{\nu}_j$. The coefficients $\alpha > 1$ and β serve to specify how many reactions occur and how many molecules exist within

dt, respectively. Both α and β can vary with a system's size. For the simulations reported in Section 2.4, the values of α and β are based on trial and error. We tried the following combinations: $(\alpha, \beta) = \{(3,000, 16,000), (3,000, 15,000), (2,000, 16,000), (4,000, 16,000)\}$. Values of beta less than 16,000 result in negative number of molecules of at least one component. Thus, values of β have a significant effect on classification of reactions as slow or fast. However, values of α have weaker effect as revealed by little change in computation time. This is because the range of α is wide so that these values are not critical in deciding fast or slow reactions. As a result, we found that $\alpha = 3,000$ and $\beta = 16,000$ provide good computational efficiency and maintain the positivity of the number of molecules.

Suppose that at a time t the system state is denoted as $\mathbf{X}(t)$, and the system consists of M_s slow and M_f fast reactions $(M_s + M_f = M)$: $\mathcal{M} = \mathcal{M}_s \cup \mathcal{M}_f$, $\overline{\overline{\mathcal{M}_s}} = M_s$ and $\overline{\overline{\mathcal{M}_f}} = M_f$. Let the probability of the system state be denoted by $P[\mathbf{X}; t]$. Then, $P[\mathbf{X}; t]$ can be rewritten as the joint probability $P_{s,f}[\mathbf{X}; t]$, which is in turn expressed in terms of the conditional probability as $P_{s,f}[\mathbf{X}; t] = P_{s|f}[\mathbf{X}; t] P_f[\mathbf{X}; t]$. This allows one to approximate the rate of change of $P[\mathbf{X}; t]$ [62],

$$\frac{\mathrm{d}P[\mathbf{X}; t]}{\mathrm{d}t} = \frac{\mathrm{d}P_{s|f}[\mathbf{X}, t]}{\mathrm{d}t} P_f[\mathbf{X}; t] + \frac{\mathrm{d}P_f[\mathbf{X}; t]}{\mathrm{d}t} P_{s|f}[\mathbf{X}; t], \qquad (2.3)$$

with

$$\frac{\mathrm{d}P[\mathbf{X}; t]}{\mathrm{d}t} \approx \frac{\mathrm{d}P_f[\mathbf{X}; t]}{\mathrm{d}t} P_{s|f}[\mathbf{X}; t]. \qquad (2.4)$$

This approximation is justified by the fact that, at the time-scale of interest, the probability of the occurrence of slow reactions (conditioned on the occurrence of

the fast reactions) does not change with time, so that its derivative is approximately zero.

The approximation in Eq. (2.4) provides a theoretical foundation [62] for the following hybrid strategy:

- Use the CLE to model fast reactions for which $a_j\tau$ $(j \in \mathcal{M}_f)$ are large and the number of molecules of all the reactants is large to warrant the continuum approximation (see Appendix A.3).

- Employ the Gillespie algorithm to describe the remaining slow reactions. For the slow reactions, instead of Gillespie algorithm one can use the Gibson and Bruck's next reaction method [64] as the latter is about 5 times faster for the same level of accuracy.

These criteria for classifying reactions as slow or fast, and the corresponding numerical methods to be used to model each reaction, are summarized in Table 2.2.

Deterministic modeling of non-reaction fluxes

Previously, [65] have used an adaptive approach to select between deterministic and stochastic approaches depending upon the number of molecules involved [66]. However, in the present work, the need to use deterministic equations arises when one or more of the fluxes involved in the corresponding ODEs cannot be modeled as reactions. This does not depend on the number of molecules. Examples of such fluxes include complex inter-organelle transport of molecules such as, in

our model, movement of Ca^{2+} from endoplasmic reticulum to the cytosol through IP_3R channels (J_{ch} in Eq. (2.5)). One can argue that this particular flux could be modeled using the 12 reversible reactions proposed by [67] and later simplified by [59]. However, in some cases the detailed mechanisms are not known and flux approximation is the only option.

The calcium dynamics model [9] includes four coupled ODEs for the state variables $[Ca^{2+}]_{ER}$, $[Ca^{2+}]_i$, h and $[Ca^{2+}]_{mit}$, which contain fluxes whose underlying mechanisms involve many reactions that are not modeled explicitly. These processes are treated deterministically in our algorithms. Consider, for example, the rate of change of $[Ca^{2+}]_{ER}$ (the other three ODEs can be found here [9]),

$$\frac{d[Ca^{2+}]_{ER}}{dt} = \frac{\beta_{ER}}{\rho_{ER}}(J_{SERCA} - J_{ch} - J_{ER,leak}). \qquad (2.5)$$

In Eq. (2.5), the rapid binding of calcium to buffer proteins is modeled implicitly through β_{ER}, the ratio of free calcium to total (free and bound) calcium in the ER; and the use of ρ_{ER}, the volumetric ratio of the ER and the cytosol, obviates the need to specify the ER volume explicitly. The calcium fluxes through the SERCA pump back to the ER, J_{SERCA}, through the IP_3R channel from the ER to the cytosol, J_{ch}, and due to the calcium leakage from the ER, $J_{ER,leak}$, are prescribed as nonlinear functions of the state variables $[Ca^{2+}]_{ER}$, $[Ca^{2+}]_i$, h and $[Ca^{2+}]_{mit}$.

The complexity of the fluxes of the state variables $[Ca^{2+}]_{ER}$, $[Ca^{2+}]_i$, h and $[Ca^{2+}]_{mit}$ complicate their modeling with the stochastic simulation algorithms described above. For example, the expression for J_{ch} is given by:

$$J_{\text{ch}} = v_{max,ch} \times \left(\left[\frac{[\text{IP}_3]}{[\text{IP}_3] + K_{IP3}} \right] \times \left[\frac{[\text{Ca}^{2+}]_\text{i}}{[\text{Ca}^{2+}]_\text{i} + K_{act}} \right] \times h \right)^3 \times \left([\text{Ca}^{2+}]_{\text{ER}} - [\text{Ca}^{2+}]_\text{i} \right)$$

(2.6)

So, in our hybrid approach, the corresponding four ODEs are integrated via a first-order Euler scheme after all other quantities are updated using the multi-scale stochastic method described in Section 2.3.3. The coupling of continuum (ODE-based) and stochastic (particle-based) descriptions requires relating the concentrations to numbers of molecules. For the cytosolic calcium dynamics in RAW 264.7 cells considered in this study, we use a cytosolic volume $V = 10$pL or a cell diameter of 27μm. Then the concentrations, e.g., the concentration of ligand, $[L] = 30$nM, can be related to the numbers of molecules, as follows

$$30\text{nM} = 30 \times 10^{-9} \times \frac{6.022 \times 10^{23}}{\text{L}} \times 10^{-11}\text{L} = 180,660 \, \text{molecules}.$$

(2.7)

Reactions with complex rate expressions

Some explicitly modeled reactions have complex rate laws which are actually functions of Michaelis-Menten (M-M) or Hill dynamics-based complex rate expressions.

We studied four methods for stochastic simulation presented in the literature to perform course-graining and handle complex rate laws such as Michaelis-Menten rate law for a single reaction and coupled reactions with Michaelis-Menten rate law. The first such contribution is the quasi-steady-state approximation

(QSSA) approach of [22]. [68] have carried out in-depth analysis of using QSSA under different conditions through the use of singular perturbation analysis. More recently, [27] have extended the QSSA by analyzing the conditions under which the standard QSSA might fail. They have utilized the total QSSA (TQSSA) and have shown that under certain conditions the method of [22] fails. They have applied the TQSSA approach to a single Michaelis-Menten mechanism, the Goldbeter-Koshland (GK) ultrasensitive switch system involving two coupled Michaelis-Menten mechanisms and a bistable system composed of two GK switches. The approach requires solving quadratic equations to solve for the propensity for slow reactions for use with the standard Gillespie algorithm. For these cases, the results are outstanding in that the mean temporal responses obtained from the TQSSA and the standard Gillespie algorithm are indistinguishable. The work of [32] deals with a detailed analysis of the issues in simplification of Michaelis-Menten formulation into a single-step reaction in stochastic simulation.

All these are successful approaches in handling systems with one or a few reactions. However, these approaches have not been applied on more complex systems involving many reactions (say, about 20 or more) with both simple and complex rate laws. Some of the rate laws in our model are much more complex than even the most complex examples presented in these contributions because in our case, the corresponding mechanisms are highly lumped representations of the underlying detailed mechanisms. If one were to consider the detailed mechanisms, the parameters would be unknown.

To handle such rate laws, here we provide two examples of the calculation of the propensity functions (Eq. (A.2)) for such reactions. In lieu of the first example, we consider the forward component of the lumped-enzymatic reaction 3 in box 1 of Fig. 2.1B,

$$[\text{GRK}; \text{Ca}_i] \, \text{LR} \xrightarrow{k_{f,3}} \text{LRp}, \tag{2.8}$$

which is facilitated by the presence of enzymes GRK and Ca_i. The corresponding flux can be written as [60]

$$v_5 = k_{f,3}[\text{LR}][\text{GRK}]MM_f(K_{m,\text{Ca}_i,3}, [\text{Ca}^{2+}]_i), \tag{2.9}$$

where where $k_{f,3}$ is 2nd-order rate constant, $K_{m,\text{Ca}_i,3}$ is the Michaelis parameter, and $MM_f(K_m, x) = x/(K_m + x)$ is the Michaelis-Menten (M-M) rate expression. Recognizing that LR acts as a reactant and recalling Eq. (A.2), we compute the propensity function a_5 for reaction of Eq. (2.8) by first determining the number of possible combinations of reactant molecules in the reaction of Eq. (2.8) as $h_5 = [LR]N_A V$, and the corresponding specific probability rate constant as $c_5 = k_{f,3}[\text{GRK}]MM_f(K_{m,\text{Ca}_i,3}, [\text{Ca}^{2+}]_i)$. Hence, the propensity function $a_5 = c_5 h_5$ is given by

$$a_5 = v_5 N_A V. \tag{2.10}$$

2.3.4 Application to cytosolic calcium dynamics in RAW cells

This multi-scale hybrid approach was applied to the cystolic calcium dynamics with parameter values and initial conditions taken from [9]. The system consists of 28 irreversible reactions and 26 species, which are represented by the state vector

$$\mathbf{X} = [\text{L}, \text{R}, \text{LR}, \text{G}_{\beta\gamma}, \text{GRK}, \text{GRK.G}_{\beta\gamma}, \text{Ca}_i^{2+}, \text{LR}_\text{p}, \text{R}_\text{p}, \text{LR}_\text{i}, \text{ARR}, \text{R}_\text{p,i}, \text{R}_\text{pool},$$

$$\text{G}_\text{i}\text{D}, \text{T}, \text{G}_{\alpha,\text{i}}\text{T}, \text{G}_{\alpha,\text{i}}\text{D}, \text{A}, \text{PIP}_2, \text{IP}_3, \text{PLC}_\beta, \text{IP3}_\text{p}, \text{X}_{\text{PIP2,gen}}, \text{CaM},$$

$$\text{Ca2.CaM}, \text{Ca2.CaM.GRK}]^T. \tag{2.11}$$

The multi-scale hybrid algorithm is needed because the numbers of molecules of some of these species are close to 0 while others have above 10^6 molecules, and because the propensity functions $a_j(\mathbf{X})$ $(j = 1, \ldots, 28)$ vary from 0 to over 10^4.

Before the ligand is added, the system is simulated for 1000 sec so that the system reaches a steady state. At time $t = 1000$ sec, ligand C5a is applied to cells and binds to its receptor (C5aR), which leads to the increase in IP_3 levels. The simulation consists of two phases: before adding ligand and after adding ligand. At $t = 0$, the species R, $\text{G}_{\beta\gamma}$, GRK, Ca_i^{2+}, R_pool, T, $\text{G}_{\alpha,\text{i}}\text{D}$, A, PIP_2, PLC_β, $\text{X}_{\text{PIP2,gen}}$ and CaM are present. Other species have zero concentration.

At the first time step, $\tau = 8.0361 \times 10^{-7}$ sec. Reactions 14, 17, 18 and 21 in Fig. 2.1B are considered to be fast, while the remaining reactions are taken to be slow (see approximation 2.2b). The second time step is calculated based on the

reaction rates and number of molecules obtained from first time step, etc.

All simulations reported in Section 2.4 were carried out on the linux-based Triton Cluster at San Diego Supercomputer Center (SDSC), with parallelization accomplished by using Microsoft's Star-P program. The number of processors used varied between 8 and 256 depending upon the number of realizations generated. On an average, the simulation time for each realization was 15 hrs. The total single-processor equivalent of simulation time for all the results is about 50,000 hrs.

2.4 Results

Comparison of response of Ca_i^{2+} from stochastic and deterministic simulation is presented in Section 2.1. Briefly, in the limit of large number of molecules of reacting species, stochastic and deterministic simulations yield nearly identical results. Below, we compare other features of the response as predicted by stochastic versus deterministic simulation.

2.4.1 Dose response

Dose response, which is a measure of efficacy of a ligand [9], is presented in Fig. 2.3. Rather than relying on commonly used saturating dose levels to generate dose-response curves, we choose only sub-basal (very low) doses. This enables us to identify differences between the dose responses of $[Ca^{2+}]_i$ predicted by deterministic

and stochastic simulations, respectively. Fig. 2.3A demonstrates the temporal evolution of the dose responses of $[Ca^{2+}]_i$ to the basal dose of $[C5a] = 30$ nM and its 0.1%, 1%, 10%, and 50% fractions. The peak height of cytosolic Ca^{2+} increases with the dose of ligand, a finding that is made explicit in Fig. 2.3C.

The stochasticity effects and differences in $[Ca^{2+}]_i$ responses obtained from the deterministic and stochastic simulations are explored in Figs. 2.3B and 2.3C. Note that in Fig. 2.3A the dose responses computed with the two approaches are nearly identical, with the deterministic predictions shifted to the right by 100 sec to improve visibility. Fig. 2.3B demonstrates the importance of stochasticity (randomness) for small numbers of ligand molecules (e.g., 0.1% C5a), when the peak height varies substantially between realizations. Although the ensemble mean of the peak-height of $[Ca^{2+}]_i$ response from these realizations visually overlaps with that from deterministic prediction, quantitatively, they are different as expressed through "normalized response difference (NRD)" in Fig. 2.3C.

As the number of molecules becomes very small, the concept of "concentration" loses its rigor and deterministic simulations can be expected to introduce modeling errors. This effect is elucidated in Fig. 2.3D, where the relative error or "normalized response difference (NRD)" ($\mathcal{E}$) between the deterministic and stochastic solutions of $[Ca^{2+}]_i$ response is shown. $\mathcal{E}$ is computed as,

$$\mathcal{E} \equiv \frac{|\text{deterministic} - \text{ensemble avg}|}{\max(\text{deterministic,ensemble avg})} \times 100\%. \tag{2.12}$$

Fig. 2.3D shows that $\mathcal{E}$ decreases as the dose of C5a increases, indicating the di-

minishing effects of randomness (stochasticity). The NRD varies from $\mathcal{E} = 7\%$ at the 0.1% dose to almost zero at the full dose of 30 nM. These results demonstrate that at lower doses, stochastic simulations are needed and that the ensemble average of multiple realizations provides a more accurate prediction of the system behavior then does the deterministic output. Further analysis of this phenomenon is presented below.

2.4.2 Convergence of stochastic simulations at low doses

Figs. 2.4A-D show the histograms of the peak-value of calcium response, $[Ca^{2+}]_i$, due to the 0.1% dose of C5a. The histograms in Figs. 2.4A-D represent respectively 16, 64, 256, and 512 realizations of the stochastic hybrid algorithm, using 20 bins in each case. The vertical dotted line in each panel corresponds to the mean computed from the corresponding number of realizations, and the solid curves are the Gaussian distributions whose mean and variance are computed from the same realizations. Although the central limit theorem applies to the distribution of the mean of a random variable instead of the distribution of the random variable itself, it is interesting to note that the shape of the computed distributions approaches the Gaussian distribution as the number of realizations increases from 16 in Figs. 2.4A to 512 in Figs. 2.4D.

To find out if the central limit theorem is applicable to the peak-value of $[Ca^{2+}]_i$ response, the mean of 4, 8, 16 or 32 realizations was computed. This was repeated in each case to generate 1024 such mean values. The histogram of the

mean values is shown in Figs. 2.4E-H. All the four histograms are similar to a Gaussian distribution and the standard deviation from these distributions indeed decreased proportional to $1/\sqrt{N_r}$, N_r being the number of realizations used to compute the mean.

2.4.3 Random variability of the $[\text{Ca}^{2+}]_i$ response at low doses

The number of molecules of C5a at 0.1% dose is about 180. The number of molecules of cytosolic Ca^{2+} is of the order of 300,000. The number of molecules of free $G_{\beta\gamma}$ is about 10,000 and that of the phosphorylated receptor still bound to the ligand (LR_p) is about 60. Fig. 2.4I shows how standard deviation (σ) of the $[\text{Ca}^{2+}]_i$ response varies across 16 realizations. Fig. 2.4J shows the variation of the normalized standard deviation $\tilde{\sigma}$, defined as: $\tilde{\sigma} = \sigma/H$, where $H = h - b$ is the difference between the basal level of calcium response b and the peak level h. It is clear from Fig. 2.4J that the normalized standard deviation $\tilde{\sigma}$ increases as the C5a dose decreases, indicating the increasing importance of randomness (stochasticity). This is because as the C5a dose (the number of C5a molecules) decreases, fewer C5a molecules participate in chemical collisions and hence the enhanced relative importance of stochasticity. One implication of this is that more stochastic realizations are needed to accurately estimate the mean response or the variability in response. From experimental view point, a larger population of cells is needed to get a stable reading for mean calcium response.

2.4.4 Sensitivity analysis

In this study we have focused on the perturbations in the initial pool of certain species. Quantification of parametric uncertainty in the reaction rate constants used in the Gillespie and other algorithms described above can be carried out following the procedure described in [69]. A similar analysis could be performed with respect to perturbations in the rate parameters while keeping the C5a dose and the initial pool of all species at their nominal levels. Since the number of molecules is sufficiently large under these conditions, the results of sensitivity analysis using stochastic simulation are similar to those obtained using deterministic simulation. As an example, results of sensitivity analysis of $[Ca^{2+}]_i$ response for changes in k_1.

The sensitivity of $[Ca^{2+}]_i$ response to variations in $[G_{\beta\gamma}]$ is shown in Fig. 2.5. In this discussion, IC refers to initial condition, which is generally also the total pool of protein/species being considered. These concentrations were changed, one at a time, by factors of 10^{-3}, 10^{-2}, 0.05, 0.1, 0,2, 0.5, and 0.75 of their respective base values. For each concentration change, a new basal level (steady state) was computed by allowing the system to evolve for 1000 sec before ligand addition, at which time 30 nM of C5a ligand was applied. Note that 10% of a base value means a 90% knockdown of the species/gene in question. Shift of basal level before ligand addition and the peak-height from basal level are the main focus of this sensitivity analysis.

Figs. 2.5A-C provide an analysis of the $[Ca^{2+}]_i$ response to changing doses of IC:$[G_{\beta\gamma}]$, which varies from its base value to the 1/20, 1/5, 1/2, and 3/4 fractions thereof. The number of molecules involved at 1/20 level of IC:$[G_{\beta\gamma}]$ is: $G_{\alpha,i}D$: 46,000, $G_{\alpha,i}T$: 5,100, free $G_{\beta\gamma}$: 16, GRK.$G_{\beta\gamma}$: 10, LR_p, 1,400, R_p, 15, IP_3: 260,000 and free cytosolic Ca^{2+}: 290,000. Figs. 2.5B-C reveal that the $[Ca^{2+}]_i$ response is very sensitive to the changes in IC:$[G_{\beta\gamma}]$. Its peak height decreases by 90% as IC:$[G_{\beta\gamma}]$ is reduced by 50%, and becomes negligible when $[G_{\beta\gamma}]$ drops below 20% of its base value (Fig. 2.5B). The relative error between the $[Ca^{2+}]_i$ responses predicted by deterministic and stochastic simulations, $\mathcal{E}$ (Fig. 2.5C) becomes very large when the concentration $[G_{\beta\gamma}]$ drops below 20% of its base value, indicating the importance of randomness, which is caused by small numbers of molecules of $G_{\beta\gamma}$.

We have also studied how the mean peak-height and NRD change when different numbers of realizations are used. Fig. 2.5B show the mean peak-height obtained from 8, 16, 32 realizations and deterministic simulation. The curves are almost indistinguishable. Difference for [5%, 20%, 50%, 75%, 100%] of IC:$[G_{\beta\gamma}]$ is [1.1775 0.15677 0.16032 0.10425 0.10491]%; the large difference being less than 1.2%. Essentially, 16 realizations are sufficient to compute the mean with good accuracy that is what is used in other simulations as well.

2.4.5 Calcium response to protein knockdown

Since the stochastic hybrid algorithm enables us to predict cytosolic calcium dynamics when only a few molecules of reacting species are present, we are in a position to explore the effects of proteins' knockdown on calcium response. Figs. 2.6 and 2.7 show the $[\text{Ca}^{2+}]_i$ response to knockdown of proteins PLCβ and GRK, respectively. Fig. 2.8 shows the $[\text{Ca}^{2+}]_i$ response to knockdown of protein GRK and perturbation of (knockdown of the protein related to) $V_{\text{max,PM,IP3dep}}$. To model a protein's knockdown, we first reduced its basal level, and then computed a new basal level (steady state) by evolving the system for 1000 sec, at which time 30 nM of C5a ligand was applied.

Figs. 2.6A and 2.6B show the $[\text{Ca}^{2+}]_i$ response to the 50%, 80%, 90%, and 99% knockdown of PLCβ for 0.1% and 10% doses of IC:[R], respectively. The number of molecules involved at 0.1% dose of IC:[R] and 90% knockdown of PLCβ is: total PLCβ: 3,400, $G_{\alpha,i}D$: 17,000, $G_{\alpha,i}T$: 350, free $G_{\beta\gamma}$: 14,000, GRK.$G_{\beta\gamma}$: 3,700, LR$_p$, 225, R$_p$, 2, IP$_3$: 270,000 and free cytosolic Ca^{2+}: 297,000. Fig. 2.6C provides a temporal snapshot of the $[\text{Ca}^{2+}]_i$ peak heights corresponding to different combinations of the PLCβ and IC:[R] levels. Both the peak height and basal levels of $[\text{Ca}^{2+}]_i$ decrease as the knockdown level of PLCβ increases. The deterministic and stochastic simulations yield similar results with NRD less than 4% (Fig. 2.6D). This clearly suggests that it may not be necessary to carry out stochastic simulation to model knockdown of PLCβ. For experiments, the

implication is that a relatively smaller population of cells may be sufficient to get a stable readout if other experimental factors can be controlled.

Figs. 2.7A and 2.7B present the $[Ca^{2+}]_i$ response to the 50%, 80%, 90%, and 99% knockdown of GRK for 0.1% and 10% doses of IC:[R], respectively. The number of molecules involved at 0.1% dose of IC:[R] and 90% knockdown of GRK is: free GRK: 1,500, $G_{\alpha,i}D$: 9,200, $G_{\alpha,i}T$: 400, free $G_{\beta\gamma}$: 10,000, GRK.$G_{\beta\gamma}$: 400, LR_p, 44, R_p, 1, IP_3: 400,000 and free cytosolic Ca^{2+}: 301,000. The largest peak height occurs at lowest [GRK] and highest [R] (Fig. 2.7C), which is qualitatively opposite to the response due to the PLCβ. Fig. 2.7D demonstrates that either deterministic or stochastic simulations can be used to investigate this behavior, with the maximum NRD $\mathcal{E}$ of about 1.5%, which occurs at low [R] and is practically independent of the level of GRK.

Fig. 2.8 demonstrates the $[Ca^{2+}]_i$ response to various degrees of simultaneous knockdown of protein GRK and the protein related to $V_{\max,\mathrm{PM,IP3dep}}$. Knockdown of GRK has a more pronounced effect on $[Ca^{2+}]_i$ response than does $V_{\max,\mathrm{PM,IP3dep}}$. The relative importance of the two knockdowns does not change at different levels of KD. This suggests the robustness of the system response over a large range of perturbations.

2.5 Summary and Discussion

In summary, we have integrated the existing techniques for multiscale stochastic simulation with deterministic simulation to deal with complex reactions systems and have applied it to studying calcium dynamics in macrophage cells. When the concentration of reactants is sufficiently large, the stochastic method yields time-course profiles identical to those obtained from a deterministic model (ensemble average of 16 or more realizations). However, at lower number of molecules of one or more species, measurable relative difference in $[Ca^{2+}]_i$ responses predicted by the two approaches is obtained, especially for the case of $G_{\beta\gamma}$, thus suggesting the necessity of using stochastic simulation as opposed to deterministic simulation for studying system dynamics at sub-cellular levels. Dose response analysis revealed that while the normalized response difference (NRD) between $[Ca^{2+}]_i$ responses predicted by deterministic and stochastic simulations is negligible at the full dose of 30nM (shown) or higher doses including saturating doses (not shown), it increases with decreasing doses. At 0.1% dose, it is as high as 7%. These results are emphasized again in the sensitivity analysis of the parameters used in the simulation and in the knockdown analysis of reacting protein components.

2.5.1 Methodological novelty

We have developed a hybrid approach to stochastic simulation, in which slow reactions and fluxes are handled through exact stochastic simulation and

their fast counterparts are treated partially deterministically through the chemical Langevin equation. The classification of reactions as fast or slow is accompanied by a partial equilibrium assumption, according to which a population of slow species is not altered by fast reactions. Our new approach also handles reactions with complex rate expressions such as functions of Michaelis-Menten kinetics and power-law kinetics by developing mathematical expressions for their propensity functions and microscopic fluxes. Fluxes which cannot be modeled explicitly through reactions are handled deterministically.

2.5.2 Sensitivity analysis

With decreasing IC:[R], lesser $[G_{\beta\gamma}]$ is available [42] (Figs. 2.5A-C), which results in reduced activation of PLCβ and as a consequence reduced hydrolysis of PIP$_2$ into IP$_3$. Hence, the increase in cytosolic $[Ca^{2+}]$ is smaller. The sensitivity curve for IC:[R] in Fig. 2.5B is nonlinear. This is because the ligand and the receptor bind in 1:1 stoichiometry, and the nominal value of IC:[R] ($\sim$40nM) is larger than the nominal (100%) level of C5a (30nM). Thus, for a small decrease (say, 10%) in IC:[R], about 36nM [R] is present. Since 36nM is still larger than 30nM, the dynamics of [LR] remains almost the same and so does the peak height of the temporal response of $[Ca^{2+}]_i$. Basal level does not change in our model since the receptor comes into play only after adding the ligand. In reality, there is a little decrease of $[Ca^{2+}]$ in cytosol due to the little basal activity, but it is compensated by the basal hydrolysis rate of PIP$_2$ and hence is unobservable.

The peak height of the $[Ca^{2+}]_i$ response decreases with decreasing value of IC:$[G_{\beta\gamma}]$, and no baseline shift is observed (Fig. 2.5D). In the absence of perturbation, at early times, the concentrations $[G_{\beta\gamma}] = 8.28e^{-3}$ μM and $[G_{\alpha,i}D] = 8.12e^{-3}$ μM are almost equal. However, if IC:$[G_{\beta\gamma}]$ is decreased, there is little free $G_{\beta\gamma}$ left. Since this directly affects the rate of PIP_2 hydrolysis, no IP_3 can be generated. Due to this effect, with decreasing IC:$[G_{\beta\gamma}]$, the peak-height of $[Ca^{2+}]_i$ decreases much more sharply. Although not shown in Figs. 2.5D-E, if IC:$[G_{\beta\gamma}]$ increases beyond 100% of base case, then the excess $G_{\beta\gamma}$ is present in the free form, hence both the basal level and peak-height increase till saturation. This is similar to the decrease in IC:$[G_{\alpha,i}D]$ shown in Figs. 2.5G-I and briefly discussed below.

Sensitivity analysis of IC:$[G_{\alpha,i}D]$ shows biphasic response of $[Ca^{2+}]_i$: large baseline shift and low peak height at substantially low IC:$[G_{\alpha,i}D]$ (Fig. 2.5G, upper panel) and a small baseline shift (increase) and the corresponding nominal increase of peak height at relatively smaller perturbations ([90% 85% 80%] of IC:$[G_{\alpha,i}D]$, Fig. 2.5G, lower panel). At substantially low $[G_{\alpha,i}D]$, large amount of free $[G_{\beta\gamma}]$ results in a large basal level shift and with the basal level at this plateau, little additional increase in $[Ca^{2+}]_i$ is observed, i.e. this results in a low peak-height of $[Ca^{2+}]_i$ upon ligand addition.

The NRD increases with decreasing IC:[R]. The behavior of NRD for decrease in IC:$[G_{\beta\gamma}]$ is similar to that for decrease in IC:[R] except that it is drastically larger at very low values (more than 80% NRD at 5% IC:$[G_{\beta\gamma}]$). While the NRD in the sensitivity analysis of IC:[R] is under 2% for all changes, it is up to 90% in

the perturbation of IC:$[G_{\beta\gamma}]$. There are three reasons for this drastic difference: (1) stochastic effects are prominent at low concentrations, (2) the system is very sensitive to large decreases in $[G_{\beta\gamma}]$ as compared to in $[R]$ or $[G_{\alpha,i}D]$, and (3) the NRD is normalized by the peak-height (Eq. (2.12)). Since peak-height is very low at low $[G_{\beta\gamma}]$, the NRD gets amplified.

2.5.3 Knockdown (KD) analysis

Our results show reduced G-protein activity and $[Ca^{2+}]_i$ response upon KD of the receptor. KD of $G_{\beta\gamma}$ results in a sharp decrease in calcium levels and KD of $G_{\alpha,i}D$ results in considerably large increases in basal level of $[Ca^{2+}]_i$ (inferred from sensitivity analysis). KD of GRK results in increased and prolonged mobilization of calcium since the receptor remains active for a longer time. Thus, GRK regulates G-protein activity strongly. Similar to $G_{\beta\gamma}$, knockdown of PLCβ shows a sharp decrease in $[Ca^{2+}]_i$. This is because IP$_3$ generation is catalyzed by the active complex of Ca^{2+}, PLCβ and $G_{\beta\gamma}$. As the knockdown level of PLCβ increases, both the peak height and basal levels of $[Ca^{2+}]_i$ decrease since less IP$_3$ is generated. Qualitatively, the knockdown response of PLCβ is similar to that of the knockdown response of $G_{\beta\gamma}$ since both play a similar role in IP$_3$ generation.

In contrast to the KD response of PLCβ, as KD level of GRK increases, peak height of $[Ca^{2+}]_i$ increases strongly (Figs. 2.7A-B). This is because the phosphorylation induced through reactions 3 and 4 decreases as KD level of GRK increases. Moreover, the time to return to steady state also increases considerably

since the receptor remains active for longer time and relatively more $G_{\beta\gamma}$ is present in the free active state. The basal level increases slightly relative to peak-height only at low IC:[R] (0.1%, Fig. 2.7A). At moderate IC:[R] (10%, Fig. 2.7B), the increase in basal level is negligible as compared to the peak-height.

$V_{\text{max,PM,IP3dep}}$ affects $J_{\text{PM,IP3dep}}$ (IP$_3$-dependent in-flux to cytosol across the plasma membrane) in a proportional manner. Double perturbation of GRK and $V_{\text{max,PM,IP3dep}}$ has revealed that for increase in their KD levels, GRK and $V_{\text{max,PM,IP3dep}}$ have opposite effects on $[\text{Ca}^{2+}]_i$. Reduction of $V_{\text{max,PM,IP3dep}}$ results in decrease of $[\text{Ca}^{2+}]_i$ because $J_{\text{PM,IP3dep}}$ is reduced (the lower three time-courses shown with light colored lines in Fig. 2.8). On the contrary, KD of GRK increases $[\text{Ca}^{2+}]_i$ response because phosphorylation of the active receptor is reduced (Fig. 2.8, time-course shown with light continuous line (100% $V_{\text{max,PM,IP3dep}}$ and 100% GRK) and time-course shown with dark continuous line (100% $V_{\text{max,PM,IP3dep}}$ and 50% GRK)). The qualitative nature of the response does not change at different KD levels of the protein GRK and the protein related to $V_{\text{max,PM,IP3dep}}$ suggesting that the system is robust to such perturbations.

The main features of the KD response are summarized in Table 2.3.

2.5.4 Stochastic effects at low molecular numbers

In the base case (30nM C5a), there is good agreement between $[\text{Ca}^{2+}]_i$ responses predicted by deterministic and stochastic simulation. However, at low doses of the ligand or proteins such as the receptor and GRK, stochastic effects

become prominent resulting in up to 2-4% NRD for low concentrations of the receptor, GRK and $G_{\alpha,i}D$, up to 7% NRD for dose response and up to 90% NRD for low concentration of $G_{\beta\gamma}$. Although the absolute value of fluctuations is larger in the case of higher doses resulting in a higher peak $[Ca^{2+}]_i$ value, the normalized standard deviation of the response increases with decreasing dose.

2.5.5   Deriving statistics from stochastic simulation

We also found that with more realizations, the computed distribution of the ensemble mean of the peak-height approaches a normal distribution when the number of realizations used to compute the mean increases, as would be mandated by the central limit theorem. Our results suggest that when 20 bins are used, about 250 realizations are sufficient to derive an approximate distribution; results from 512 realizations are good in terms of reaching a normal distribution. Statistics related to low order moments of the distribution, such as mean and standard deviation could be computed accurately even with lesser number of realizations (about 16 realizations to compute the mean and about 128 realizations for the standard deviation) at least for the cytosolic calcium response. For other systems some trial may be involved. These results can be potentially used for deciding the number of realizations needed to compute meaningful statistics in stochastic simulations, at least for similar systems with a similar number of components.

TaiJung Choi, Mano R. Maurya, Daniel M. Tartakovsky, Shankar Subramaniam (2010), 'Stochastic Hybrid Modeling of Intracellular Calcium Dynamics'. J. Chem. Phys., 133, 165101

Table 2.1: The run-time scalability of the Gillespie, tau-leap, and chemical Langevin equation algorithms as a function of the number of molecules.

	Initial number of molecules	
	S: 312, E: 125	S: 31200, E: 12500
Method	**Computation time** (s)	**Computation time** (s)
Gillespie algorithm	0.892	100.3
Tau-leap algorithm	0.235	0.354
CLE	0.003	0.003

Table 2.2: Criteria used to identify slow and fast reactions and corresponding numerical method. Column 2 and 3 list the scale and simulation method in the "scale (method)" format.

# of molecules of species involved	Reaction propensity	
	High	Low
Large	Fast (CLE)	Slow (Gillespie)
Small	Slow (Gillespie)	Slow (Gillespie)

Table 2.3: Summary of results of KD response. The change in the features of calcium response listed is for increase in KD-level (decrease in IC:[.]) of the protein. Qualitative nature of the features is mostly independent of the level of [R].

Protein/variable name	Basal level	Peak height
PLCβ	decreases	decreases, convex
GRK	very small increase	increases, linear
$V_{\mathrm{max,PM,IP3dep}}$	no change	small decrease

Figure 2.1: A simplified model for calcium signaling including calcium influx, ER, and mitochondrial exchange and storage. (A) Ligand C5a binds to its receptor on plasma membrane (PM) and activates G protein G_i. The free $G_{\beta\gamma}$ binds to PLCβ and increases its activity which accelerates the phosphorylation of PIP_2 into IP_3 and DAG. IP_3 binds to its receptor IP_3R on the ER membrane. Thus, calcium from the ER is released into the cytosol. Other fluxes between cytosol and mitochondria or ECM are also shown. (B) Receptor module (box 1), GTPase cycle module (box 2), IP_3 generation module (box 3) and feedback module (box 4).

Figure 2.2: Temporal evolution of the concentrations of substrate, $S(t)$, and product, $P(t)$, computed using the Gillespie, tau-leap, and chemical Langevin equation (CLE) approaches. (A) shows time-course of 1 realization from each method. (B) and (C) show the time-course of mean and standard deviation from 1024 realizations, which show excellent agreement among the 3 different methods. (D)-(F) show histograms and probability distribution of the number of molecules of S sampled at t = 10 second. The shapes of the three histograms are very similar.
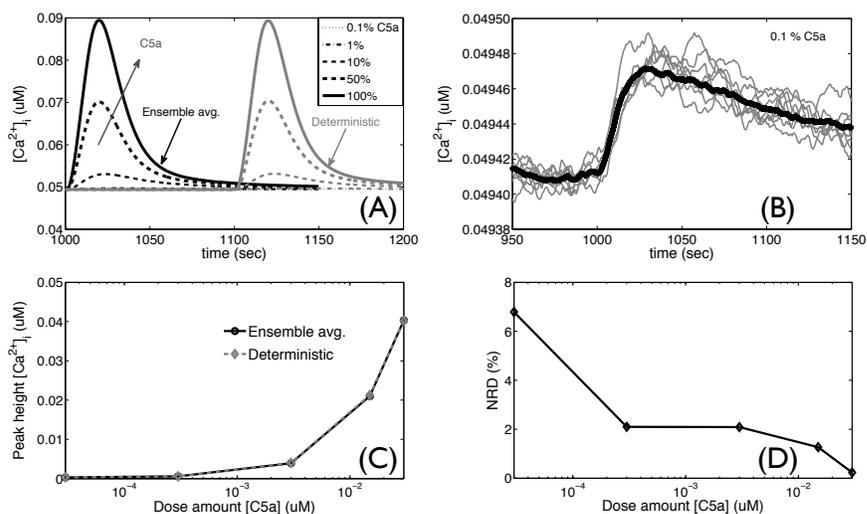
Figure 2.3: Dose response. (A) Comparison between ensemble average of 16 realizations and deterministic simulation. For better contrast, the time-course from deterministic simulation is shifted by 100s. (B) Comparison between ensemble average and individual realizations in stochastic simulation for 0.1% (of 30nM) strength of the ligand C5a. (C) Comparison of the dose response (peak heights): The difference is quite small as compared to the scale of peak-height. (D) At lower doses, the normalized response difference (NRD) is larger indicating the stochastic effects. The NRD decreases with increasing dose as the number of the molecules of C5a becomes several hundreds or more.
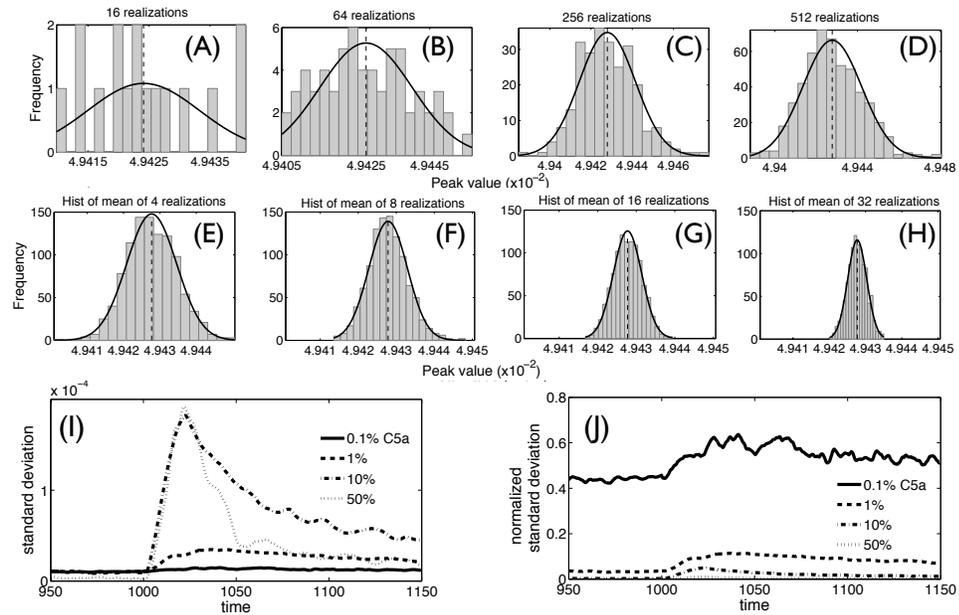
Figure 2.4: Revelation of stochastic effects at low doses. (A)-(D) Distributions of peak-height for the 0.1% dose of C5a computed from 16, 64, 256, and 512 realizations, respectively. The dotted vertical line represents the mean value and the solid curves denote theoretical Gaussian distributions. As the number of realizations increases, the shape of the histogram approaches a Gaussian distribution. (E)-(H) The mean of 4, 8, 16 or 32 realizations was computed. All the four histograms are similar to a Gaussian distribution and the standard deviation from these distributions indeed decreased proportional to $1/\sqrt{N_r}$, N_r being the number of realizations used to compute the mean. (I) The standard deviation computed from 16 realizations for several doses. Contrary to the expectation, higher doses result in larger absolute standard deviations. (J) The normalized standard deviation decreases as the dose is increased, signifying the effect of randomness at lower doses.
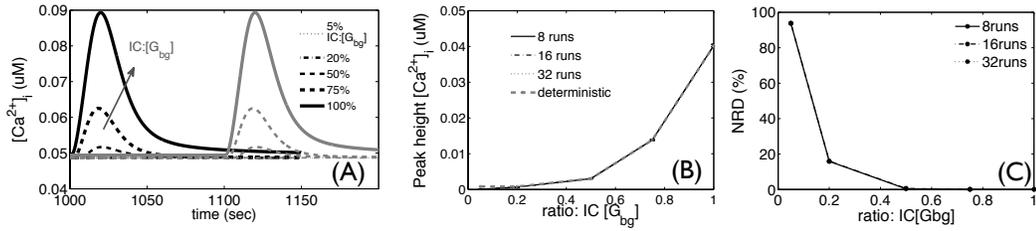
Figure 2.5: Sensitivity analysis. (A)-(C) *Response of* $\left[\text{Ca}^{2+}\right]_i$ *to changes in* IC:[G$_{\beta\gamma}$]. The decrease in the peak height due to decrease in IC:[G$_{\beta\gamma}$] is much more pronounced than that caused by the same decrease of IC:[R]. (C) NRD $\mathcal{E}$ is extremely high at very low IC:[G$_{\beta\gamma}$], suggesting significant stochastic effects at low numbers of molecules of [G$_{\beta\gamma}$]. (B) and (C) also show the effect of using different number of realizations for computing the mean. Such differences are small (see text) indicating that 16 realizations are sufficient for computing the mean.
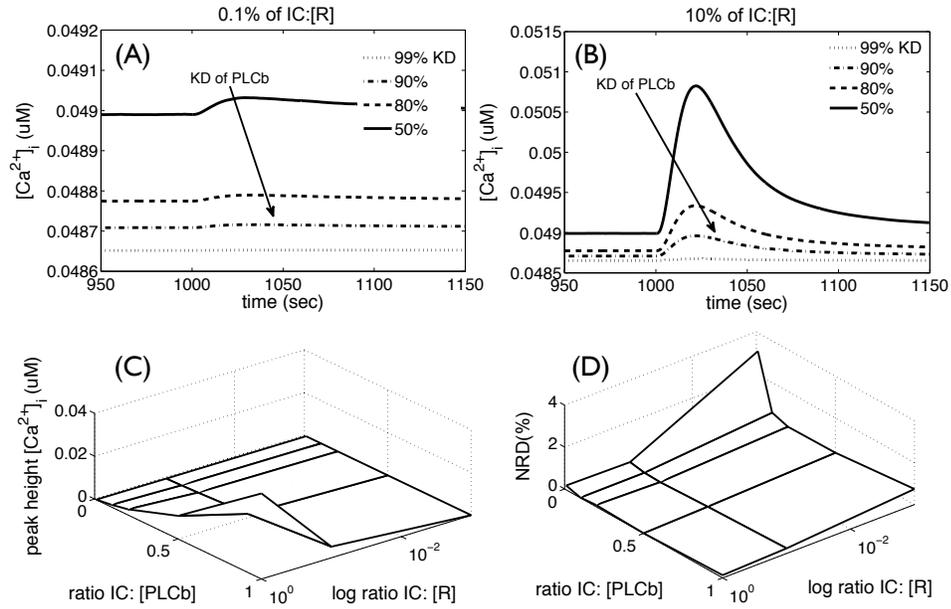
Figure 2.6: Knockdown response of PLCβ. (A)-(B) The $[\text{Ca}^{2+}]_i$ response to the 50%, 80%, 90%, and 99% knockdown of PLCβ for 0.1% and 10% levels of IC:[R], respectively. As the knockdown rate of PLCβ increases, both the basal level and peak height of $[\text{Ca}^{2+}]_i$ decrease, because the IP$_3$ production decreases due to decrease in PLCβ. (C) Peak-height of $[\text{Ca}^{2+}]_i$ response corresponding to different combinations of the PLCβ and IC:[R] levels. Peak height increases with high amount of IC:[R] and [PLCβ]. (D) NRD $\mathcal{E}$ is insignificant and decreases as doses of R and PLCβ increase.
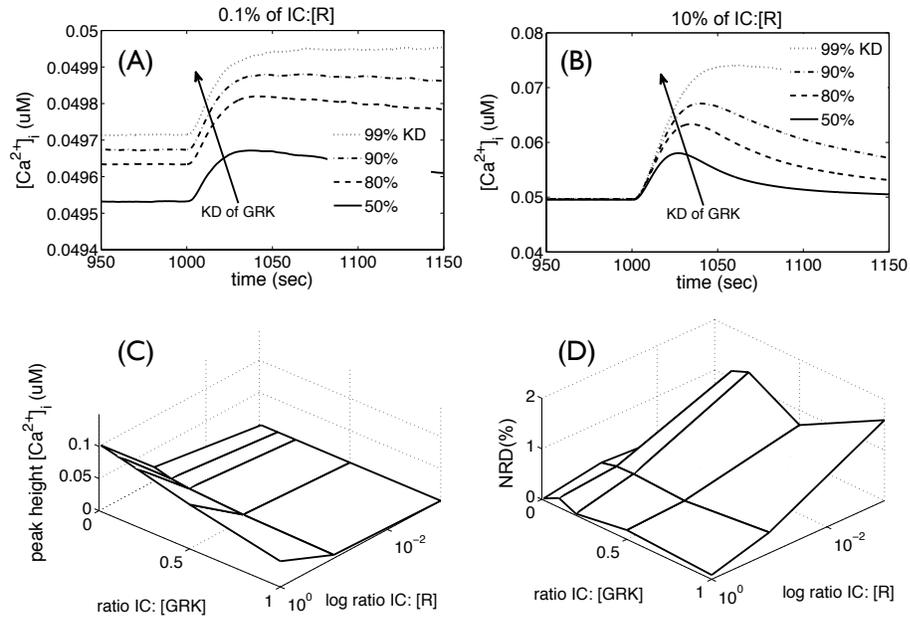
Figure 2.7: Knockdown response of GRK. (A)-(B) The $[Ca^{2+}]_i$ response to the 50%, 80%, 90%, and 99% knockdown of GRK for 0.1% and 10% levels of IC:[R], respectively. (C) Peak-height of $[Ca^{2+}]_i$ response corresponding to different combinations of [GRK] and IC:[R] levels. (D) NRD $\mathcal{E}$ is insignificant, reaching its maximum of about 1.5% at low IC:[R].
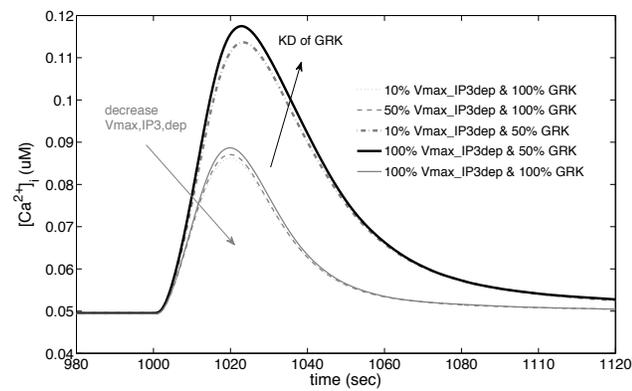
Figure 2.8: The $[Ca^{2+}]_i$ response to the simultaneous knockdown of GRK and gene/protein related to $V_{max,IP3dep}$. Knockdown of GRK and reduction of $V_{max,PM,IP3dep}$ have opposite effects on the $[Ca^{2+}]_i$ response. The response is much more sensitive to knockdown of GRK than to decrease in $V_{max,PM,IP3dep}$.

Chapter 3

Stochastic Operator Splitting Approach for Reaction-Diffusion Processes

3.1 Introduction

Randomness plays an important role in the behavior of many biological phenomena, such as cellular signaling and gene regulatory networks [8, 70, 71]. While deterministic ordinary differential equations (ODEs) often provide accurate predictions of the dynamics of biochemical pathways with large numbers of reacting molecules, they fail when the concentrations of reactants and/or their products

become small and the law of mass action becomes invalid. When this occurs, the randomness associated with the dynamics of individual molecules becomes pronounced, necessitating the use of stochastic simulations. Standard stochastic techniques, e.g., Gillespie's stochastic simulation algorithm [11] and its computationally efficient modifications [12, 72], are routinely used to model biochemical reactions in such systems. Such algorithms assume that reactants and their products are well-mixed, i.e., distributed uniformly in space.

The latter assumption is problematic when the number of molecules is small. This is especially so in crowded environments with complex internal geometry, wherein stochasticity and spatial variability are inseparable. Partial differential equations (PDEs) provide accurate macroscopic predictions of the dynamics of spatially heterogeneous systems with large numbers of molecules. Yet, similar to ODE-based models, they fail to account for the randomness inherent in a system comprised of small numbers of molecules. It is essential that computational methods for reaction-diffusion systems with small numbers of molecules are capable of handling both stochasticity and heterogeneity.

A number of micro- and meso-scale methods have been developed for the simulation of reaction-diffusion systems. The micro-scale approaches, e.g., the Green's function reaction dynamics [14] and Smoldyn's algorithm [15], are based on Brownian dynamics and require the reacting molecules to diffuse within a certain distance from each other in order for bimolecular reactions to take place. The latter requirement necessitates the use of a numerical mesh and the tracking of in-

dividual particles and/or distances between them, which renders such algorithms computationally expensive. Mesoscopic approaches, e.g., MesoRD [16] and the Gillespie multi-particle (GMP) method [17, 18], trade representational accuracy for computational efficiency. They are based on a reaction-diffusion master equation [19], which generalizes a chemical master equation developed for well-mixed chemical reactions by discretizing the space into a collection of cells and treating each cell as a well mixed system. MesoRD [16] treats diffusion as a unimolecular reaction whose reaction rate is related to the corresponding diffusion coefficient. The GMP method [17, 18] employs an operator-splitting scheme in which the Gillespie algorithm and cellular automata [20] handle reactions and diffusion, respectively.

We present a numerical algorithm to simulate stochastic reaction-diffusion processes with a small number of non-uniformly distributed molecules. It employs an operator-splitting, in which the Gillespie algorithm (or its accelerated versions) and Brownian dynamics (or the Smoluchowski equation) are used to simulate reactions and diffusion, respectively. Our algorithm is conceptually similar to the GMP method in that it relies on operator-splitting. However it offers a number of computational advantages in terms of both accuracy and efficiency. First, the cellular automata used in the GMP method restrict a particle's movement during one fixed time-step to the adjacent cells only; while Brownian motion places no restrictions on the distance particles can travel during one time-step, thus gaining in computational efficiency. Second, Brownian dynamics provides a more accurate representation of diffusion than cellular automata. Third, our algorithm offers the

flexibility of an "on-the-fly" adaptive selection of the time-step size for operator-splitting, depending on whether the system is reaction- or diffusion-controlled. The outline of this manuscript is as follows.

Our stochastic operator-splitting approach is described in Section 3.2. This section contains a brief description of the stochastic simulation algorithm for modeling reactions and a comparative analysis of the two approaches—Brownian motion and cellular automata—to deal with diffusion. It also contrasts our operator-splitting algorithm with that used in the GMP method (Section 3.2.2). Section 3.3 presents three computational examples, which demonstrate the accuracy and robustness of the proposed algorithm. The first example (Section 3.3.1) considers diffusion of two chemical species undergoing an irreversible bimolecular reaction in order to validate our algorithm and to analyze its performance and accuracy in terms of the time-step and the cell size. This is done by comparing the stochastic simulation results with solutions of the corresponding deterministic PDEs. The detailed comparison elucidates the effects of the finite (small) number of molecules and space-time discretization on the simulation accuracy and efficiency. The second example (Section 3.3.2) models an idealized gene expression system [14]. It serves to investigate the performance of our algorithm in reaction- and diffusion-controlled regimes and the effects of local inhomogeneity. The third example (Section 3.3.3) considers reaction and diffusion of CheY molecules through the cytoplasm of *Escherichia coli* during chemotaxis [73]. In addition to its biochemical significance, this example poses additional computational challenges by introducing a specific

local structure. In all the three cases, we demonstrate that our algorithm outperforms the GMP method in terms of computational time. In Section 3.4, we summarize the simulation results and provide conclusions.

3.2   Methods: Numerical approach

3.2.1   Operator-splitting method

We consider M species that undergo diffusion and N (bio)chemical reactions. Spatio-temporal evolution of their concentrations $\{c_i(\mathbf{x}, t)\}_{i=1}^{M}$ can be described by a system of reaction-diffusion equations,

$$\frac{\partial c_i}{\partial t} = D_i \nabla^2 c_i + f_i(c_1, \ldots, c_M), \qquad i = 1, \ldots, M \tag{3.1}$$

where D_i is the molecular diffusion coefficient of the i-th species and f_i is the corresponding net production rate through reactions. Our focus is on reaction-diffusion systems with small numbers of molecules, in which continuum representations such as Eq. 3.1 are inadequate. Such phenomena are typically handled with stochastic simulations. While stochastic, particle-based methods for modeling both reactions in well-mixed environments (e.g., the Gillespie algorithm [11]) and diffusion of chemically inert molecules (e.g., Brownian dynamics) are relatively mature, the same cannot be said about chemical reactions in spatially heterogeneous (reaction-diffusion) systems.

We propose an operator-splitting method that enables one to take advan-

tage of the considerable advances in modeling chemical reactions and molecular diffusion by treating these two phenomena separately. We use the (modified) Gillespie algorithm and Brownian dynamics to represent the reaction and diffusion steps, respectively, in lieu of their continuum representations in Eq. 3.1. The relative order of these steps is determined dynamically depending on whether the system is in diffusion- or reaction-controlled state.

This *raison d'être* for employing an operator splitting is different from the use of operator-splitting algorithms to model deterministic reaction-diffusion systems. In the latter case, the goal is to handle the stiffness of the reaction-diffusion equations in which diffusion and reaction processes have different time scales. A typical operator-splitting method for solving deterministic reaction-diffusion equations employs an implicit method to handle the (stiff) reaction simulations and an explicit method to handle diffusion. Examples of deterministic operator-splitting approaches include the Douglas-Gunn alternating direction implicit (ADI) method [74] and the method of lines (MOL) [75]. The former applies an explicit Euler scheme to diffusion and an implicit Crank-Nicholson method to reactions. The latter converts partial-differential equations (PDEs) into ordinary differential equations (ODEs) by discretizing the spatial derivatives and leaving the time variable continuous.

We employ an operator-splitting algorithm [76] to approximate Eq. 3.1 with

$$\frac{\partial c_i'}{\partial t} = D_i \nabla^2 c_i' \tag{3.2a}$$

$$\frac{\partial c_i''}{\partial t} = f_i(c_1'', \ldots, c_M''). \tag{3.2b}$$

during the time interval $[t, t + \Delta t]$. Here $c_i'(t) = c_i(t)$ and $c_i''(t) = c_i'(t + \Delta t)$, so that the concentration of the i-th species at the end of the time-step Δt is $c_i''(t) = c_i(t + \Delta t)$. Fig. 3.1A provides a graphical representation of this operator-splitting algorithm. The resulting stochastic operator-splitting algorithm will enable us to analyze the effects of intrinsic noise in spatially heterogeneous biological systems (Section 3.3). Our implementation of the reaction process using Gillespie algorithm, the diffusion process using either Brownian dynamics or cellular automata, and the GMP algorithm is described in Appendix B. Briefly, in Gillespie algorithm [11], to advance the system from state $\mathbf{X}(t)$, two random numbers r_1 and r_2 distributed uniformly on the unit interval $[0, 1]$ are generated. Then, a discrete random value j and continuous random value τ are selected probabilistically as

$$\tau = \frac{1}{a_{\text{sum}}} \ln \left(\frac{1}{r_1} \right), \qquad \sum_{j'=1}^{j-1} a_{j'} \leq r_2 a_{\text{sum}} \leq \sum_{j'=1}^{j} a_{j'}. \tag{3.3}$$

where a_{sum} is the sum of all propensity functions. The system state at $t + \tau$ is updated according to $\mathbf{X}(t + \tau) = \mathbf{X}(t) + \boldsymbol{\nu}_j$ where the entries of the vector ν_j are the change in the number of molecules of various species due to the j-th reaction [11].

In Brownian dynamics, a species diffuses from its current location $\mathbf{X}(t) \in \mathbb{R}^3$ at time t to its new location at time $(t + \triangle t)$ according to [77]: $\mathbf{X}(t + \triangle t) = \mathbf{X}(t) + \sqrt{2D_i \triangle t}\,\boldsymbol{\xi}$ where $\boldsymbol{\xi} = (\xi_1, \xi_2, \xi_3)^T$ is a normal random displacement vector.

In cellular automata, the ith species can diffuse to one of its neighboring cells (Figs. 3.1 B-C) during the time interval equal to its diffusion-time constant τ_{D_i} given by $\tau_{D_i} = (\Delta x)^2/(2D_i d)$.

3.2.2 Algorithms for the stochastic operator-splitting method

To deal with reaction-diffusion systems composed of a small number of molecules, we propose the following stochastic operator-splitting algorithm.

1. Lattice: The space is discretized into a lattice of cells. Within each cell (lattice element), each species is assumed to be distributed homogeneously.

2. System state: Determine whether the system is at diffusion- or reaction-controlled state to decide the time-step size Δt_j at the j-th time-step.

3. Diffusion process: Diffusion of species between cells is modeled via Brownian dynamics with a fixed time-step by treating the space as a continuum.

4. Reaction process: Reactions within each cell are simulated via the Gillespie algorithm or its accelerated versions.

5. Time is increased by the time-step size and the above steps are repeated till the final desired time.

Dynamic identification of system's state

A key feature of our algorithm is its ability to determine at each time-step the system's state (reaction- or diffusion-controlled) and to set the time-step size accordingly. For an i-th cell ($i = 1, \ldots, C$ where C is the number of cells in a numerical grid) at the j-th time-step Δt_j, we define a macroscopic time constant

$$T_{R_{ij}} = \frac{1}{a_{\text{sum}}^{ij}}, \qquad a_{\text{sum}}^{ij} \equiv \sum_{k=1}^{N} a_k(\mathbf{X}_{ij}), \tag{3.4}$$

where $\mathbf{X}_{ij}$ is the state $\mathbf{X}$ of the i-th cell at the j-th time-step and $a_k(\mathbf{X}_{ij})$ is the propensity function for the k-th reaction. At each time-step, we find the minimum value of the macroscopic time constants over all the cells,

$$T_{R_j}^{\text{min}} \equiv \min_i T_{R_{ij}}, \tag{3.5}$$

and define

$$\tau_{R_j} = T_{R_j}^{\text{min}} \ln\left(\frac{1}{r}\right). \tag{3.6}$$

Figs. 3.2A and 3.2B show a frequency chart of $\ln(1/r)$ and the corresponding cumulative probability distribution. They reveal that the cumulative probability of $\ln(1/r) \leq 1$ is 0.63 (also see Table 3.1), i.e., the probability of $\tau_{R_j} \leq T_{R_j}^{\text{min}}$ is 63%. Then, a time fraction

$$F \equiv \frac{T_{R_j}^{\text{min}}}{\tau_D} \tag{3.7}$$

can be used to classify the system as reaction- or diffusion-controlled as explained below.

It follows from Eq. 3.6 and Eq. 3.7 that

$$\frac{\tau_{R_j}}{\tau_D} = F \ln\left(\frac{1}{r}\right), \tag{3.8}$$

which allows one to compute the cumulative probability of $\tau_{R_j}/\tau_D \leq 1$ as

$$P\left[\frac{\tau_{R_j}}{\tau_D} \leq 1\right] = P\left[F \ln\left(\frac{1}{r}\right) \leq 1\right] = P\left[\ln\left(\frac{1}{r}\right) \leq \frac{1}{F}\right] = 1 - e^{-1/F}. \tag{3.9}$$

This is the same as the waiting time probability in the Gillespie algorithm [11]. It becomes clear that the magnitude of F determines the state of the system. For example, $F = 1$ corresponds to $P[\ln(1/r) \leq 1] = 0.63$ (Fig. 3.2B), so that $P[\tau_{R_j}/\tau_D \leq 1] = 0.63$ as well. In other words, $F = 1$ implies that $\tau_{R_j} \leq \tau_D$ in about 63% cases (Table 3.1), i.e., the system is diffusion-controlled. Similarly, $F = 0.5$ (even faster reactions) translates into $P[\ln(1/r) \leq (1/F) = 2] = 0.86$ (Fig. 3.2B) and $P[\tau_{R_j}/\tau_D \leq 1] = 0.86$. We classify a system as diffusion-controlled, if $P[\tau_{R_j}/\tau_D \leq 1] \geq 0.5$. According to Table 3.1, this corresponds to $F \leq 1/\ln(2) = 1.44$. We introduce a parameter $0 < k_1 \leq 1/\ln(2)$ and say that the system is diffusion-controlled if $F < k_1$. The smaller the value of k_1, the more stringent the criterion becomes. Essentially, as the probability of $\tau_R < \tau_D$ increases, i.e., k_1 increases, the system becomes more diffusion-controlled. Similarly, we define a related parameter k_1' so that if $F > k_1'$ then the system is reaction-controlled.

In diffusion-controlled systems, many reactions may be fired during Δt_j. We set the time-step $\Delta t_j = k_2 \tau_D$, where k_2 is a tunable parameter representative of the cut-off (or critical) value of $\ln(1/r)$ for a desired cumulative probability

(Fig. 3.2B and Table 3.1). For example, $k_2 = 2$ corresponds to 0.86 probability of a reaction taking place during Δt_j.

For the reaction-controlled system, τ_{R_j} (or $T_{R_j}^{\min}$) is much larger than τ_D. For example, $k_1' = 3$ corresponds to $P[\ln(1/r) \leq (1/F) = 1/3] = 0.28$ (Fig. 3.2B), i.e., $P[\tau_{R_j}/\tau_D \leq 1] = 0.28$. To ensure the firing of some reactions, larger Δt_j should be chosen. Based on several simulations, we found that $\Delta t_j = 10\tau_D$ provides good results.

We also define an intermediate regime that is characterized by values of F that prevent one from classifying a system as being diffusion- or reaction-controlled. In this regime, the time-step Δt_j should be chosen between $k_2\tau_D$ and $10\tau_D$. Our numerical experimentation suggests that setting $k_2' = 3$ provides a good balance between accuracy and computational efficiency.

Algorithm

A detailed algorithm for the numerical implementation of the above steps of our stochastic operator-splitting method is provided below.

---

1. For a given space dimension d and cell size Δx, calculate the diffusion time $\tau_{D_i} = (\Delta x)^2/(2D_i d)$ of diffusing species $i = 1, \ldots, M$ and set $\tau_D = \min\{\tau_{D_i}\}$.

2. Initialize $t = 0$.

3. While $t \leq t_{\text{final}}$

   (a) Define whether system is diffusion- or reaction-controlled at every time-

step.

- Calculate $T_{R_j}^{\min}$ through Eq. 3.5.

- Calculate F through Eq. 3.7.

(b) Compute the time-step according to the classification of the system. The multiplicative factors k_1, k_1', k_2 and k_2' are selected based on Fig. 3.2 and Table 3.1.

    i. If $F < k_1$ (diffusion-controlled),

        - Set $\Delta t_j = k_2 \tau_D$.

    ii. Elseif $k_1 < F < k_1'$ (mixed zone),

        - Set $\Delta t_j = k_2' \tau_D$.

    iii. Else $F > k_1'$ (reaction-controlled)

        - Set $\Delta t_j = 10 \tau_D$.

(c) Reset $t_{\text{old}} = t$.

(d) Perform the diffusion step first followed by the reaction step.

    i. Diffusion step: Use Brownian dynamics to advance the species with time-step Δt_j.

    ii. For each cell: Reaction step:

        A. While $(t - t_{\text{old}}) \le \Delta t_j$

        Calculate τ_R using Eq. 3.3.

          - If $\Delta t_j \ge \tau_R$, find which reaction takes place within τ_R using Eq. 3.3. Update the number of molecules of different species and

time as

$$\mathbf{x} \leftarrow \mathbf{x} + \boldsymbol{\nu}_j, \qquad t \leftarrow t + \tau_R. \tag{3.10}$$

- Else, do not update the state vector since no reaction was fired.

end *while*

    B. Reset $t = t_{\text{old}}$ for the next cell.

    end *for*

(e) Set: $t = t_{\text{old}} + \Delta t_j$ (synchronize t across all cells).

end *while*

3.2.3   Comparison of our method with GMP method

The GMP method [17] provides an alternative implementation of the operator-splitting approach shown in Fig. 3.1A. While our approach relies on Brownian dynamics, the GMP method models diffusion with cellular automata. This difference is significant and has far-reaching implications. First, the time step in a cellular automaton is fixed and determined by Eq. (B.2) in Appendix B in terms of the diffusion coefficient and cell size. This is because during one time step molecules in cellular automata can move from a cell only to its immediate neighbors. By relying on Brownian dynamics, our approach allows the time step to vary between the diffusion (τ_D) and reaction ($\bar{T}_R$ or $\bar{\tau}_R$) time scales. This significantly speeds up the simulations, especially when the diffusion coefficient is large and/or the cell

size is small. Second, the GMP method uses the diffusion times for each species to determine when their respective molecules move from one cell of the lattice to the adjacent cells. In our algorithm, diffusing molecules of all species move during the same time step.

The following synthetic example demonstrates the salient features of both Brownian dynamics and cellular automata. We place 18 molecules of a substance P in the bottom-left cell of a lattice and allow them to diffuse towards its center cell/element (Fig. 3.3A). Fig. 3.3 shows the average number of molecules at the center cell as a function of time for several degrees of mesh refinement. Mesh refinement (the increased number of cells in each direction) does not significantly affect the accuracy of the simulation results (Fig. 3.3B-D) but increases the computational time (Table 3.2). Figs. 3.3E and F reveal that the Brownian dynamics reproduces a solution of the corresponding diffusion equation more accurately than the cellular automata does. This is because a particle in Brownian dynamics can move any distance in any directions while the cellular automata limits its displacement to 9 adjacent cells.

Finally, for a given degree of accuracy the Brownian dynamics simulations provide a significant computational speed-up relative to their cellular automata counterparts (Table 3.2). This is because the diffusion time τ_D in the cellular automata is fixed by the lattice size, whereas Brownian dynamics allow for larger time steps Δt. The run time of the Brownian dynamics simulations reported in Table 3.2 correspond to the same time step $\Delta t = 10$ s regardless of the lattice

size, while τ_D in the cellular automata simulations varied with the mesh size and diffusion coefficient in accordance with Eq. B.2 in Appendix B. As a result, the cellular-automata simulation time increases significantly with the number of cells in each direction (L).

3.3 Results: Case studies

We start with a synthetic example (Section 3.3.1) to validate our stochastic operator-splitting method by comparing its results with both the GMP approach and a deterministic solution of the underlying reaction-diffusion equation computed with COMSOL. Next, we use our algorithm to model a gene expression system (Section 3.3.2) and CheY diffusion in *E. coli* (Section 3.3.3). The last two examples were carried out on a Linux-based Triton cluster of the San Diego Supercomputer Center at University of California, San Diego.

3.3.1 Synthetic reaction-diffusion case study

Suppose that at time $t = 0$, A_0 molecules of species A and B_0 molecules of species B are distributed uniformly over the left-half of the computational domain in Fig. 3.4A. At $t > 0$, they diffuse into the rest of the domain and undergo a (bio)chemical reaction whose reaction product is species C,

$$A + B \xrightarrow{k} C. \tag{3.11}$$

The three species are assumed to have the same molecular diffusion coefficient $D = 10^{-13}$ m^2/s. In a more biologically realistic case study of CheY diffusion in *E. coli* (Section 3.3.3), the diffusion coefficients are different for different species. Here, we use a forward reaction rate constant of $k = 3 \times 10^9$ M^{-1}s^{-1}. The computational volume is $V = 10^{-15}$ L ($V = 10^{-18}$ m^3).

Performance analysis

First, we simulate diffusion (no reactions) with the cellular automaton and Brownian dynamics. Fig. 3.4B demonstrates that the numbers of molecules of A and B predicted with cellular automata approach their equilibrium values faster than those computed with the Brownian dynamics. The results from Brownian dynamics approach are more accurate and are in good agreement with the PDE solution. This is because Brownian dynamics provides a better approximation of the diffusion process. Furthermore, Brownian dynamics simulations are computationally more efficient than their cellular automaton counterparts (Table 3.2).

Second, we analyze the impact of a time-step and lattice size on simulations of the full reaction-diffusion system. A numerical solution (obtained using COMSOL software) of the corresponding PDEs is treated as a yardstick. First we analyze the effect of lattice size and diffusion constant on computational time (Table 3.3) and then we study the accuracy. For a fixed diffusion constant $D = 10^{-12}$ m^2/s, as L increases, τ_D (or Δt) decreases and computational time increases for both algorithms. However, for any L, our method is faster than the GMP method.

Similarly, for increasing diffusion constant for a given cell size ($L = 8$), τ_D (or Δt) decreases and computational time increases for both algorithms (Table 3.3 and Fig. 3.5). Our algorithm is faster than the GMP method because our algorithm can apply larger time-steps according to the state of the system. For example, for $D = 10^{-12}$ m^2/s, $\tau_D(= 2.6 \times 10^{-3}s)$ is 10 times larger than that for $D = 10^{-11}$ m^2/s. Hence, the computational time for $D = 10^{-12}$ m^2/s is about 10 times smaller than that for $D = 10^{-11}$ m^2/s. In Fig. 3.5, as D increases, τ_D (or Δt) decreases (Fig. 3.5A) and computational time increases (Fig. 3.5B) for both algorithms. For $D = 10^{-14}$ m^2/s, the system transitions from diffusion-controlled ($\Delta t = k_2\tau_D$; $k_2 = 2$) to reaction-controlled regime during the time-course. For $D \geq 10^{-13}$ m^2/s, the system becomes reaction-controlled ($\Delta t = 10\tau_D$). This explains the increase in the absolute value of the slope of Δt or computational time vs. D plots at $D = 10^{-13}$ m^2/s for our method.

As should be expected, the accuracy of our stochastic operator-splitting algorithm increases as the time-step and/or the cell size become smaller for both the diffusion-controlled ($D = 10^{-13}$ m^2/s) and reaction-controlled ($D = 10^{-12}$ m^2/s) scenarios in Figs. 3.6A, B. The results are based on average of 8 realizations. The relative error-rate, defined as the ratio of the integrated absolute difference between a method and the PDE solution to the integrated absolute value of the PDE solution over the time-course (ratio of the areas),

$$\text{Relative error-rate} = \frac{\sum_t |(\text{method - deterministic PDE})|}{\sum_t |(\text{deterministic PDE})|}, \qquad (3.12)$$

is shown in Table 3.3. For a given D, as L increases, the time-step and the relative error-rate decrease for both methods. The smaller the time-step, the smaller the errors introduced by the operator-splitting procedure. However, for any L, our algorithm is more accurate than the GMP method.

Third, we investigate the impact of ordering the diffusion and reaction steps on the simulation accuracy (Figs. 3.6C, D). Both diffusion-controlled and reaction-controlled systems are considered. If the reaction step is selected to be the first part of the operator-splitting algorithm then a diffusion process does not contribute to the system evolution during the first time-step. Hence, the reaction-first approach introduces larger errors if there is excessive inhomogeneity at the beginning. Thus, the diffusion-first (followed by the reaction step) approach is suited for both diffusion-controlled as well as reaction-controlled processes.

We further compare the accuracy of the results from our algorithm and the GMP method. Simulation results in Fig. 3.6B demonstrate an excellent agreement between our solution and the PDE solution, while the GMP method significantly underestimates both the peak number of molecules and the time it takes for the system to equilibrate (Fig. 3.6E). This finding is consistent with the results shown in Fig. 3.4B, which reveal that the number of molecules estimated with cellular automata reach their equilibrium levels faster than those computed with Brownian dynamics.

Effect of number of molecules

Having established the agreement between our stochastic (discrete) operator-splitting algorithm and its continuum (PDE-based) counterpart for a large number of molecules, we proceed to analyze their ability to handle reaction-diffusion systems composed of small numbers of molecules. The premise here is that the smaller the number of molecules, the more inadequate the deterministic (continuum) models become and the more pronounced are the stochastic effects.

We rely on an absolute difference rate (DR),

$$\text{DR} = \frac{|\text{our method - deterministic PDE}|}{\text{deterministic PDE}}, \tag{3.13}$$

to quantify the difference between the concentrations (relative numbers of molecules) computed with the two approaches. As expected, the DR decreases as the initial number of reacting molecules increases (Fig. 3.6F). It drops from $\text{DR} \approx 0.1$ for $A_0 = 60$ to $\text{DR} \approx 0.01$ for $A_0 = 600$ or 6000. Hence, stochastic and deterministic simulations yield similar results, when the number of molecules becomes large. This expected result is consistent with many other studies of randomness in reacting perfectly-mixed systems [70].

3.3.2 Gene expression case study

The van Zon and ten Wolde model [14] of gene regulation serves as an ideal model system for studying the stochasticity effects due to both the low number of molecules and the spatial inhomogeneity. Similar to Fig. 3.3A, RNAp molecules

initially occupy the left-bottom cell of a numerical mesh, and at $t > 0$ diffuse towards a DNA molecule that is fixed in the center cell ("operator site"). Upon reaching the operator site, the RNAp molecules bind with DNA with a forward reaction rate constant k_a, forming the DNA-RNAp complex and this complex can dissociate with a backward rate constant k_d. In addition, it can produce a mRNA at a production rate constant k_{prod} and mRNA degrade with a decay rate constant k_{dec}. In the following, we use A, B, C and P to denote DNA, RNAp, DNA-RNAp and the produced mRNA, respectively.

Assuming that RNAp is the only diffusing species (i.e., DNA-RNAp and the produced mRNA do not leave the operator site), and that the molecular diffusion coefficient and reaction rates are constant (i.e, neglect anomalous diffusion due to the crowding effect and hydrodynamic effect), a continuum representation of the process is provided by a system of three ordinary differential equations and one partial differential equation,

$$\frac{d[A]}{dt} = -k_a[A][B] + k_d[C] + k_{prod}[C] \tag{3.14}$$

$$\frac{\partial[B]}{\partial t} = D\nabla^2[B] - k_a[A][B] + k_d[C] + k_{prod}[C] \tag{3.15}$$

$$\frac{d[C]}{dt} = k_a[A][B] - k_d[C] - k_{prod}[C] \tag{3.16}$$

$$\frac{d[P]}{dt} = k_{prod}[C] - k_{dec}[P], \tag{3.17}$$

where the square brackets denote concentrations of the respective species.

Reaction- vs. diffusion-limited processes

We set the molecular diffusion coefficient to $D = 10^{-12}$ m^2/s. Then the average time for RNAp molecules to arrive at the operator site is 0.04 s, i.e., RNAp molecules diffuse quickly throughout the system that becomes "well-mixed". Diffusion does not have a significant impact on the system's dynamics since the system is reaction-controlled. In other words, $D = 10^{-12}$ m^2/s can result in a reaction-limited (reaction-controlled) system.

Let us define a dimensionless Damköhler number Da as the ratio of typical diffusion (τ_D) and reaction ($\bar{\tau}_R$) time scales,

$$Da = \frac{\tau_D}{\bar{\tau}_R}. \tag{3.18}$$

A system is diffusion-limited if $Da \gg 1$ and reaction-limited otherwise. For $D = 10^{-12}$ m^2/s, the average diffusion time $\tau_D \in [10^{-2}$ s, 10^{-1} s]. Since $\bar{\tau}_R$ is of the same order of magnitude, the system is reaction-limited. On the other hand, the diffusion coefficient $D = 10^{-15}$ m^2/s corresponds to $Da \sim 10^3$, resulting in the diffusion-limited behavior.

Fig. 3.7A demonstrates the salient features of these two transport regimes with $L = 5$. For $D = 10^{-12}$ m^2/s, the number of protein molecules computed with the Gillespie algorithm (a perfectly mixed system with no diffusion) and with our operator-splitting algorithm are in close agreement. For $D = 10^{-15}$ m^2/s, diffusion becomes important with the protein beginning to burst around 20 s after the RNAp molecules encounter DNA at the central operator site. Our results differ from their

counterparts obtained by the Gillespie algorithm mainly in terms of fluctuations.

Time-step selection

The magnitude of the molecular diffusion coefficient D affects the choice of the time-step Δt in the stochastic operator-splitting algorithm. Fig. 3.7A and 3.7B show the number of mRNA molecules computed for a wide range of the diffusion coefficients, $10^{-15} \leq D \leq 10^{-12}$ with $L = 5$ and $L = 20$, respectively. In the case study with $L = 5$ and $D = 10^{-12}$ m^2/s, the time scales are $\bar{\tau}_R = 0.012$ s and $\tau_D = 6 \times 10^{-3}$ s; fast diffusion quickly homogenizes the system so that its behavior is reaction-controlled. Our numerical experiments suggest that setting $\Delta t = 10\tau_D$ decreases the simulation time and guarantees that a reasonable number of reactions take place during the simulation time-step.

For small diffusion coefficients ($D = 10^{-15}$ m^2/s), $\tau_D = 6.67$ s and $\bar{\tau}_R = 0.012$ s, which means that almost all $\tau_R < \tau_D$. To ensure that a sufficient number of reactions take place during the time interval Δt, we selected $\Delta t = 2\tau_D$. Similar rules are applied for $L = 20$ as well.

Comparison with GMP method and stochastic effect

The mRNA production, predicted with the GMP algorithm and our approach on the meshes with several degrees of refinement ($L = 5, 10, 20$), are shown in Fig. 3.8A and B, respectively. In a display of the lack of self-consistency, the finest mesh ($L = 20$) results in predictions that are quantitatively wrong in that

the system fails to reach its equilibrium state of about 1,000 proteins. It is worthwhile recalling that in the GMP algorithm, Δt is defined as the minimal diffusion time that cannot be adjusted. In the system under consideration, $\bar{\tau}_R > \Delta t$ so that all the reactions cannot take place during the time interval Δt ($\bar{\tau}_R = 0.65$ s and $\tau_D = 0.417$ s). These results demonstrate one of the advantages of our algorithm: unlike the GMP algorithm, our approach is capable of handling different mesh sizes by adapting appropriate time-steps.

The effects of stochasticity (noise) become apparent in predictions averaged over a smaller number of realizations (Fig. 3.8C). As should be expected from the central limit theorem, the standard deviation from the mean prediction decreases as $1/\sqrt{N_r}$. By ignoring the spatial variability, the Gillespie algorithm dampens considerably the noise present in the system. The protein production continues to fluctuate in time even after it reaches its equilibrium (steady-state) value because it depends on the frequency of the encounter of RANp and DNA in the central cell. The statistics of the equilibrium protein production, i.e., its mean μ, standard deviation σ, and noise level (coefficient of variation) $\nu = \sigma/\mu$, are presented in Table 3.6.

3.3.3 CheY diffusion case study

System description

As a final example, we consider a chemotaxis pathway in *E. coli*. A mathematical model of this process has been developed earlier [73]. The species included in the model [73] and their simplified spatial arrangement within a cell are presented in Table 3.8 and Fig. 3.9A, respectively. Table 3.7 lists a set of reactions considered in this model. The diffusing species are CheY, CheYp and CheZ. The species CheA* (active CheA) and CheAp do not diffuse into the cytoplasm, being confined in the inner receptor cluster. The molecules of CheY and CheA* are phosphorylated in the receptor cluster located on the anterior cell wall. Once diffused into the cytoplasm, the CheYp molecules bind with four flagellar motors FliM1,...,FliM4 and the FliM·CheYp complex is produced. The four motors are located on the side walls, ordered FliM1 to FliM4 from the anterior wall (Fig. 3.9A). The reactions in the FliMs induce *E. coli*'s forward or backward motion and/or rotation.

The diffusion step in our stochastic operator-splitting algorithm is implemented in a way that the molecules reaching the cell's surface are reflected back into the cell without loss of momentum. The diffusion step is followed by the reaction step, which employs the Gillespie algorithm to simulate reactions between the molecules within each cell of a numerical mesh. We investigate the effects of varying the length of cell, L_{sv} (*sv* denotes subvolume) and time-step Δt on the

performance of our algorithm, and compare it to that of the GMP algorithm.

Both GMP and our approach are conceptually different from the Smoldyn method [73]. The latter approach simulates diffusion with Brownian dynamics and keeps track of individual molecules. Unlike our algorithm, it allows for multi-molecular reactions between two or three molecules only within a certain radius from each other. This reduces the computational speed and increases storage requirements, because positions of all molecules have to be stored and distances between all molecules must be calculated at each step in order to check if reactions can take place.

Simulation results

The time-course of FliM·CheYp complexes simulated with both the GMP algorithm and our stochastic operator-splitting approach is shown in Figs. 3.9B, C. The GMP algorithm overestimates the equilibrium levels of the FliM·CheYp complexes and underestimates the transition-to-equilibrium times in both Fig. 3.9B (M1 and M2) and Fig. 3.9C (M3 and M4). As established in the two previous computational examples, this discrepancy is due to the errors associated with the cellular automaton treatment of diffusion in the GMP algorithm. In addition to being more accurate, our approach is also computationally more efficient than the GMP algorithm. In both algorithms, the reaction step consumes close to 99% of the total computational time. Therefore, in order to reduce the simulation time, larger Δt should be selected because the execution of the Gillespie algorithm accounts

for most of the computational time. For a given mesh size L_{sv}, the time-step in the GMP algorithm is fixed by the molecular diffusion coefficient D, while in our algorithm it is more flexible according to whether the system is reaction-controlled or diffusion-controlled. The simulation time for our algorithm is 12hrs, whereas it is 26 hrs for the GMP method.

3.4   Summary and Discussion

Complex multi-scale biological systems can be analyzed with microscopic approaches, such as Greens-function reaction dynamics and the Smoldyn algorithm. These methods are accurate albeit computationally expensive and often prohibitive. On the other hand, macroscopic kinetic modeling approaches that use PDEs are amenable to numerical computation, but fail to model the physics of systems with small number of molecules accurately. Mesoscopic approaches, e.g., reaction-diffusion master equation and MesoRD, discretize space into a collection of lattice elements and extend the chemical master equation normally used in well-mixed chemical reactions into the stochastic regime for inhomogeneous systems. To facilitate faster and more accurate solutions within the mesoscopic scale framework, we have developed a stochastic simulation method which is based on operator-splitting for modeling the reaction-diffusion system. In our methodology, the time-step size is chosen automatically at each step depending upon whether the system is reaction- or diffusion-controlled. We use the Gillespie stochastic sim-

ulation algorithm for modeling the reactions and Brownian dynamics approach for modeling the diffusion process. We thus account for both spatial heterogeneity and the fluctuation in concentrations arising from the small number of molecules. Our method yields highly accurate results and has the merit of modeling both the reaction and diffusion processes in the system.

In order to validate accuracy and efficiency of our algorithm, a simple reaction-diffusion system, $A+B \rightarrow C$, is studied first. We concluded that Brownian dynamics provides much more accurate results while being faster than a Cellular automata approach. For example, Table 3.3 reveals that the error-rates for our approach and the GMP algorithm are about 1% and 5%, respectively. The average speed-up by using our method is about 5 times as compared to the GMP method for a wide range of the values of the diffusion constant. Moreover, we compared the stochastic ensemble average with the deterministic result and found out that our results converge to the deterministic result when smaller Δt and larger L are used in the simulation. We also concluded that the fluctuations become larger in case of smaller number of molecules and spatial inhomogeneity.

Towards modeling biologically realistic systems, a simplified gene expression system and CheY diffusion in *E. coli* bacteria are studied. In gene expression case study, the system is classified based on the Damköhler number, Da. If it is larger than 1, it is regarded as a diffusion-limited system and reaction-limited otherwise. In order to simulate the system accurately, the time-step, Δt, should be selected according to the dominant process. In addition, noise levels concerning

the number of molecules and number of realizations are studied. It is shown that as the number of molecules or number of realizations become smaller, the noise level increases. We then simulated a more complicated system, viz., CheY diffusion in *E. Coli*, through both the GMP method and our operator-splitting algorithm. We have shown that the operator-splitting approach provides more accurate results and is faster as compared to the GMP algorithm. For a more accurate analysis of movement of *E. coli* bacteria, the chemotaxis process in which molecules move toward higher or lower concentration according to the concentration gradient should also be analyzed [78].

In conclusion, we present a hybrid numerical method, also known as, operator-splitting method, for stochastic reaction-diffusion process with a small number of heterogeneously distributed molecules. Our approach is conceptually similar to the GMP algorithm that applies Gillespie algorithm for reaction process and Cellular automata for diffusion process. However, our method provides computational advantages in terms of accuracy and efficiency. First, molecules in Brownian dynamics can move freely without the restriction of lattice or time-step whereas molecules in Cellular automata move only to the adjacent lattices during the fixed time-step. Second, Brownian dynamics offers a more accurate simulation result than the cellular automata approach. Third, our algorithm has the flexibility of changing time-steps, depending on whether the system is reaction- or diffusion-controlled.

TaiJung Choi, Mano R. Maurya, Daniel M. Tartakovsky, Shankar Subramaniam (2012), 'Stochastic Operator-Splitting Method for Reaction-Diffusion Systems'. J. Chem. Phys., 137, 184102

Table 3.1: The tunable parameter k_1 is used as a criterion to decide if the system is diffusion- or reaction-controlled. As the probability of τ_R being less than τ_D increases, the system becomes more diffusion-controlled. The other parameter, k_2, is related to the probability of a reaction taking place during Δt. As k_2 increases, the probability of a reaction occurrence during Δt increases. In our algorithm, $k_1 = 0.5$, $k_1' = 3$, $k_2 = 2$, and $k_2' = 3$ are used. Please refer to Fig. 3.2.

F, k_1 or k_1'	Relation	Meaning
0.5	$\bar{T}_R = 0.5\tau_D$	86% of τ_R is less than τ_D
1	$\bar{T}_R = \tau_D$	63% of τ_R is less than τ_D
1.44	$\bar{T}_R = 1.44\tau_D$	50% of τ_R is less than τ_D
3	$\bar{T}_R = 3\tau_D$	28% of τ_R is less than τ_D
k_2	Probability for the reaction to occur during Δt	
1	63%	
2	86%	
3	95%	

Table 3.2: Comparison of computational time for cellular automata and Brownian dynamics. The total of 2048 realizations are considered in order to emphasize the difference in computational time. L is the number of cells along each axis, d is the spatial dimension, and τ_D is the diffusion time constant. Brownian dynamics uses the same time step $\Delta t = 10$ s for all cell sizes, whereas cellular automata has different time steps depending on the cell size and the value of diffusion coefficient D. The simulation time increases with L. Brownian dynamics is more efficient than cellular automata.

		Cellular automata	Brownian dynamics
	$D = 10^{-15}$ (m^2/s)	$\Delta t = \tau_D$ (s)	$\Delta t = 10$ (s)
L	$\tau_D = (\Delta x)^2/(2Dd)$	Computational time (s)	
5	6.67	2.98	6.72
10	1.667	13.39	6.73
20	0.417	55.09	6.68

Table 3.3: A synthetic reaction-diffusion system $A+B \rightarrow C$ with diffusion constant $D = 10^{-12}$ m^2/s. The system is reaction-controlled for $L = 4$ and 8. In these cases, we set the average value of all Δt to $\overline{\Delta t} = 10\tau_D$. Computational time increases with smaller Δt or larger L (smaller cell size). For $L = 16$, the system undergoes transitions between the mixed and diffusion-controlled regimes. In this case, $\overline{\Delta t} \in [2\tau_D, 10\tau_D]$. As L increases, τ_D (or Δt) decreases and computational time increases for both algorithms. However, for any L, our method is faster than the GMP method. The relative error-rate is also shown (see Eq. 3.12 and related text). As L increases, the relative error-rate decreases for both methods. However, our algorithm is more accurate than the GMP method. In a similar way, for a given L, as D increases, τ_D (or Δt) decreases and computational time increases for both algorithms.

$D = 10^{-12}$ m^2/s		Comp. time (s) ($\overline{\Delta t}$ (s))		Error-rate (%)	
L	τ_D (s)	GMP	Our method	GMP	Our method
4	1×10^{-2}	16	1.4 (1×10^{-1})	5.2	1.17
8	2.6×10^{-3}	167	32 (2.6×10^{-2})	4.8	0.95
16	6.5×10^{-4}	4602	4022 (1.4×10^{-3})	4.7	0.91
$L = 8$		Computation time (s) ($\overline{\Delta t}$ (s))			
D (m^2/s)	τ_D (s)	GMP		Our method	
10^{-11}	2.6×10^{-4}	1660		304 (2.6×10^{-3})	
10^{-12}	2.6×10^{-3}	167		33 (2.6×10^{-2})	
10^{-13}	2.6×10^{-2}	17		3.6 (2.6×10^{-1})	
10^{-14}	2.6×10^{-1}	1.7		0.7 (1.56)	

Table 3.4: Gene expression case study: DNA has 1 molecule and RNAp has 18 molecules. C is the DNA·RNAp complex. System volume is 1×10^{-15} l and diffusion coefficient of RNAp is $D = 10^{-12}$ m^2/s (reaction-limited system) or $D = 10^{-15}$ m^2/s (diffusion-limited system). Abbreviation: I.C.: initial condition.

Reaction			Rate	I.C. [nM]
DNA + RNAp	$\xrightarrow{k_a}$	C	3×10^9 M^{-1}s^{-1}	1.67
C	$\xrightarrow{k_d}$	DNA +RNAp	21.5 s^{-1}	30
C	$\xrightarrow{k_{prod}}$	P + DNA + RNAp	89.55 s^{-1}	0
P	$\xrightarrow{k_{dec}}$	ϕ	0.04 s^{-1}	0

Table 3.5: Gene expression case study: Reaction time is averaged over 256 realizations of a simplified gene expression process. As cell sizes become smaller reaction times increase, since fewer molecules in each cell imply lower probability for reactions to take place within a cell. The time-step Δt in the GMP method equals τ_D, whereas Δt in our method can vary according to the system classification as reaction- or diffusion-controlled. Since the cases of $L = 5$ and 10 are diffusion-controlled, we set $\overline{\Delta t} = 2\tau_D$. For $L = 20$, the system changes from diffusion- to reaction-controlled as time progresses.

$D = 10^{-15}$ m^2/s			GMP	Our method
L	Δx (μm)	$\bar{\tau}_R$	$\tau_D = \frac{\Delta x^2}{2Dd}$ (s)	$\overline{\Delta t}$ (s)
5	0.2	0.012	6.67	13.33
10	0.1	0.095	1.667	3.33
20	0.05	0.65	0.417	2.34

Table 3.6: Gene expression case study. According to the central limit theorem, noise level or standard deviation decreases as $1/\sqrt{N_r}$. The mean values remain around 1000. The standard deviation predicted with our algorithm is much higher than that computed with the Gillespie algorithm, because our algorithm accounts for randomness due to both a small number of molecules and spatial inhomogeneity.

Number of realizations N_r	Mean value μ	Standard deviation σ	Noise level ν
4	1006.3	324.3	0.322
16	1015.8	149.6	0.147
64	1010.4	81.9	0.088
256	1004.4	41.2	0.041
256 (Gillespie)	1001.4	1.76	0.0018

Table 3.7: CheY diffusion case study: k_f and k_b denote respectively forward and backward reaction rate constants for the *E. coli* system. Unimolecular and bimolecular reaction rates have dimensions $[\text{s}^{-1}]$ and $[\text{M}^{-1}\text{s}^{-1}]$, respectively. i denotes the index of flagellar motors.

Compartment	Reaction	Reaction constant
Receptor cluster	CheA* $\rightarrow$ CheAp	$k_f = 3.4 \times 10^1 \text{ s}^{-1}$
	CheAp+CheY $\rightarrow$ CheA*+CheYp	$k_f = 10^8 \text{ M}^{-1}\text{s}^{-1}$
Cytoplasm	CheY $\rightleftharpoons$ CheYp	$k_f = 5.0 \times 10^{-5} \text{ s}^{-1}$ $k_b = 8.5 \times 10^{-2} \text{ s}^{-1}$
	CheZ+CheYp $\rightarrow$ CheZ+CheY	$k_f = 1.6 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$
FliM$_i$ $(i = 1, \ldots, 4)$	FliM$_i$+CheYp $\rightleftharpoons$ FliM$_i$·CheYp	$k_f = 5.0 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$ $k_b = 2.0 \times 10^1 \text{ s}^{-1}$

Table 3.8: CheY diffusion case study: 13 species and 13 reactions in the *E. coli* system. Only CheY, CheYp and CheZ molecules can diffuse; others are fixed within their original cells. Initial values are expressed in terms of number of molecules, and i denotes the index of flagellar motors.

Species	Initial # of molecules	Diffusion constant
CheA*	1260	position fixed
CheAp	0	position fixed
CheY	8200	$D = 10^{-11}$ m^2/s
CheYp	0	$D = 10^{-11}$ m^2/s
CheZ	1600	$D = 6 \times 10^{-12}$ m^2/s
FliM$_i$ $(i = 1, \ldots, 4)$	34	position fixed
FliM$_i$·CheYp $(i = 1, \ldots, 4)$	0	position fixed

(A)

$$c(x,t) \xrightarrow[\Delta t]{\text{Diffusion}} c'(x,t)$$

$$c'(x,t) \xleftarrow[\Delta t]{\text{Reaction}} \begin{array}{l} c''(x,t) \\ \equiv c(x,t+\Delta t) \end{array}$$

(B)

		N		
	W	C	E	
		S		

(C)

|				
	NW	N	NE	
	W	C	E	
	SW	S	SE	
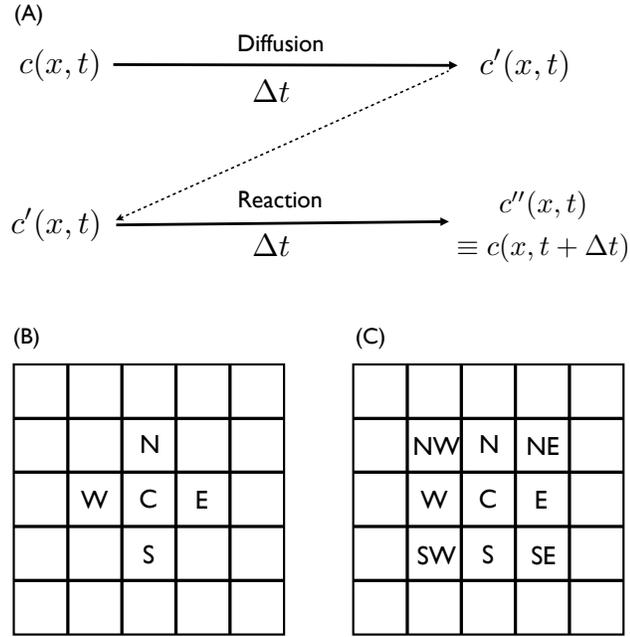
Figure 3.1: (A) Schematic representation of the diffusion-reaction operator-splitting. The final value after diffusion process at time $t + \Delta t$ is used as the initial value for the reaction process. Final value of reaction process is the final value at the end of diffusion-reaction process. (B)-(C) Cellular automata neighborhoods in $d = 2$ dimension: in the von Neumann automata the probability of staying in a cell or diffusing to its neighbors is $1/5$ (B), in the Moore automata this probability is $1/9$ (C).
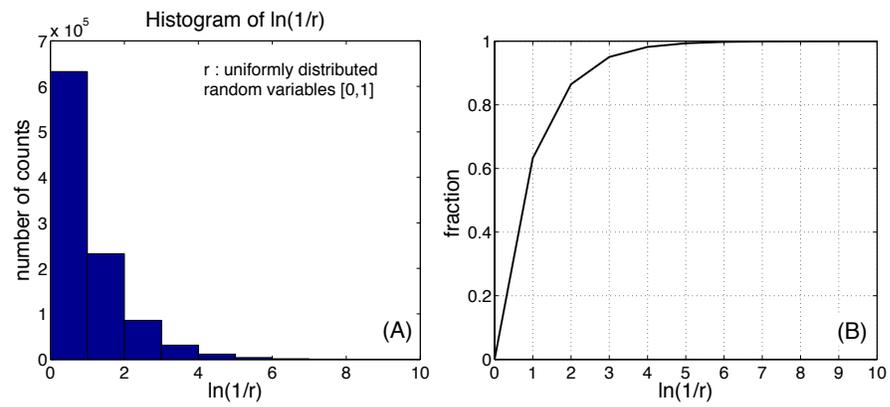
Figure 3.2: (A) Histogram of $\ln(1/r)$, where r is a uniformly distributed random variable in [0,1]. (B) Cumulative fraction of counts out of total counts (1 Million). About 63% of the numbers have values less than 1 and 86% of the numbers are less than 2.
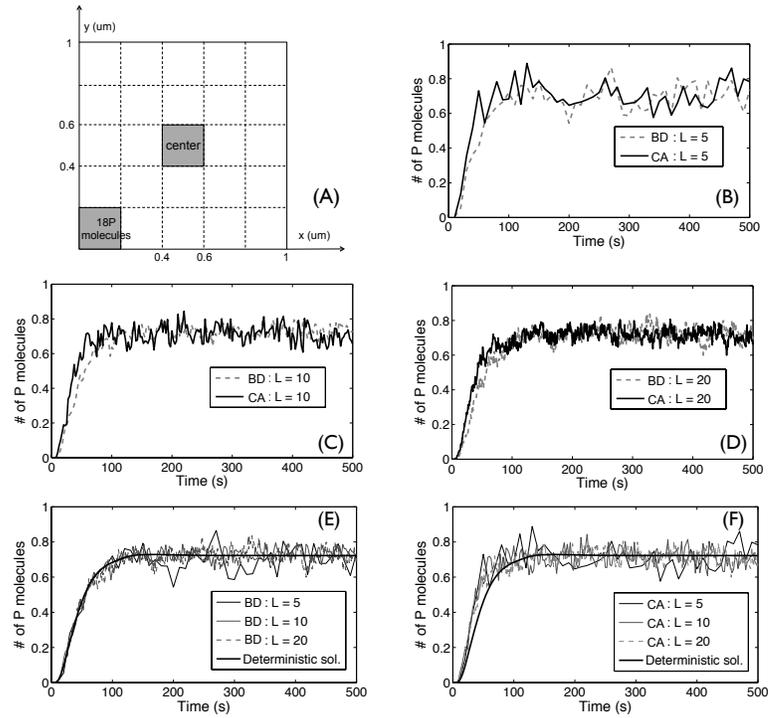
Figure 3.3: Temporal evolution of the count of molecules at the center cell averaged over 512 realizations of cellular automata (CA) and Brownian dynamics (BD) for several degrees of mesh refinement (L denotes the number of cells in each direction). $D = 10^{-15}$ m^2/s, $L_x = L_y = 1$ μm. (A) Initially, 18 P molecules are placed into the bottom-left cell. As time increases, they diffuse and number of P's in the center cell is counted. (B)-(D) For various values of L, the cellular automata simulation results have faster rising times than those of Brownian dynamics. (E)-(F) The simulation results are independent of the cell size (L). The Brownian dynamics results are in better agreement with the deterministic PDE solution than those of cellular automata.
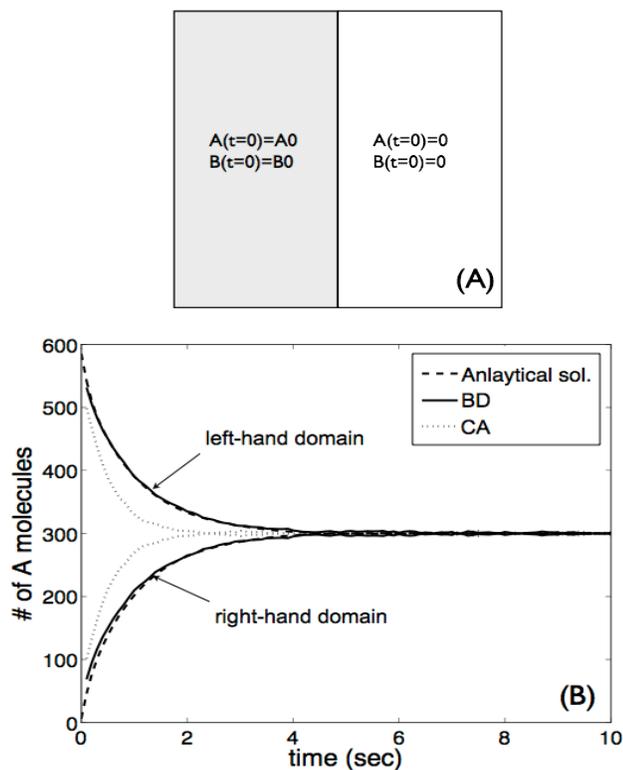
Figure 3.4: $A + B \rightarrow C$ case study: (A) Initially, species A and B exist only in left-hand side. All A and B molecules and their product P diffuse with the same diffusion constant. (B) Comparison of results from analytical solution, cellular automata (CA) and Brownian dynamics (BD). The Brownian dynamics results agree with the analytical solution, while the cellular automata results do not. The increasing curves represent the number of molecules in the right half of the domain and the decreasing ones in the left half.
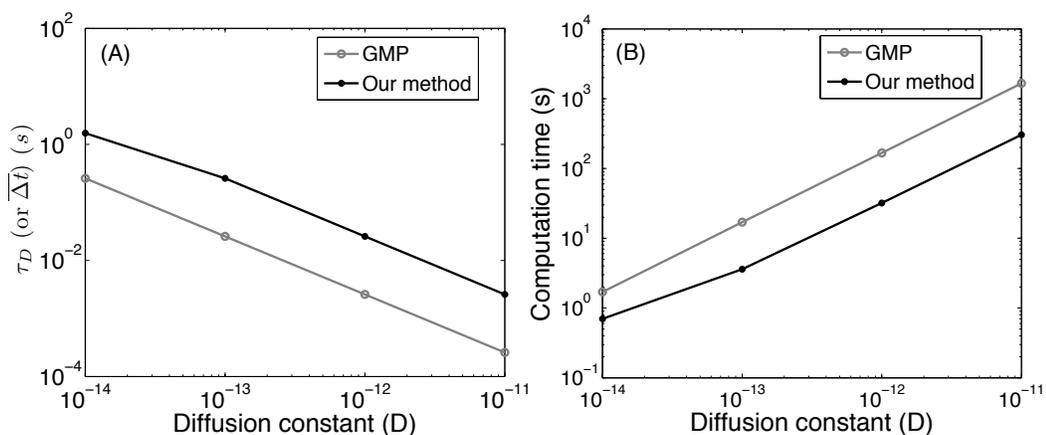
Figure 3.5: $A + B \rightarrow C$ case study: Effect of diffusion constant, D (m^2/s), on (A) τ_D (or Δt) and (B) computational time for our method and the GMP method. As D increases, τ_D (or Δt) decreases and computational time increases for both algorithms. For $D = 10^{-14}$ m^2/s, the system transitions from diffusion-controlled ($\Delta t = k_2 \tau_D$; $k_2 = 2$) to reaction-controlled regime during the time-course. For $D \geq 10^{-13}$ m^2/s, the system becomes reaction-controlled ($\Delta t = 10 \tau_D$), explaining the increase in the absolute value of the slope of Δt or computational time vs. D plots at $D = 10^{-13}$ m^2/s for our method.
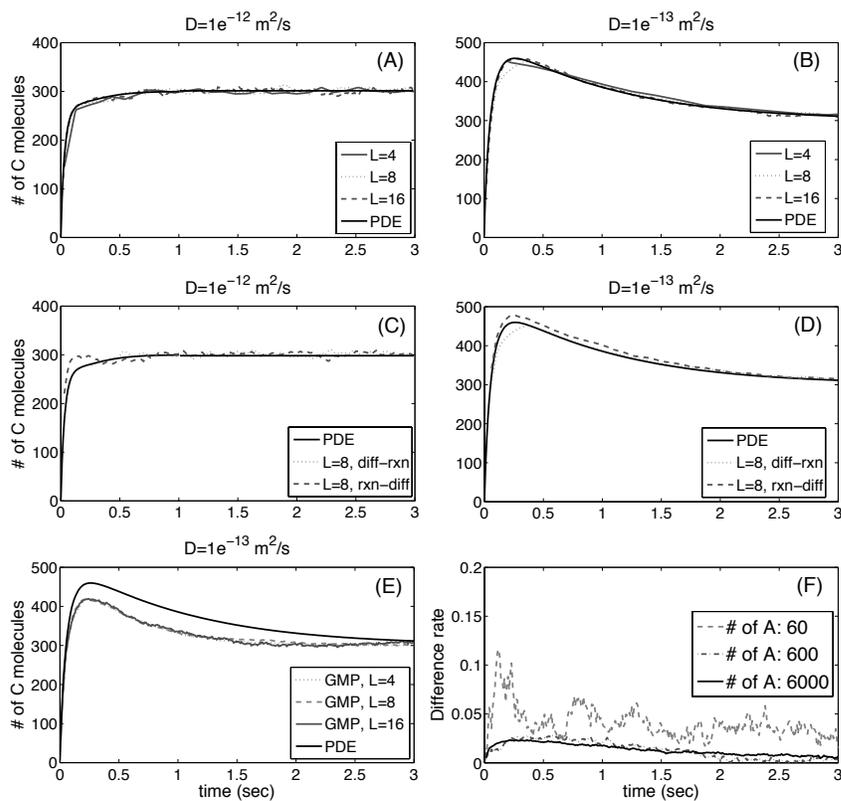
Figure 3.6: $A + B \to C$ case study: (A)-(B) As cell size becomes finer, the results from our approach converge to the numerical solution. (A) shows the results of reaction-controlled system, whereas (B) is for a diffusion-controlled system. (C)-(D) Dashed line is the result of reaction-first and diffusion-later order and dotted line is the reverse order. Diffusion-reaction ordering has better agreement with PDE solution than the reaction-diffusion order. (E) Black line denotes the result of PDE solution. Three gray lines are for the GMP method. (F) As the initial number of molecules gets lower, the difference rate increases and the fluctuations also increase.
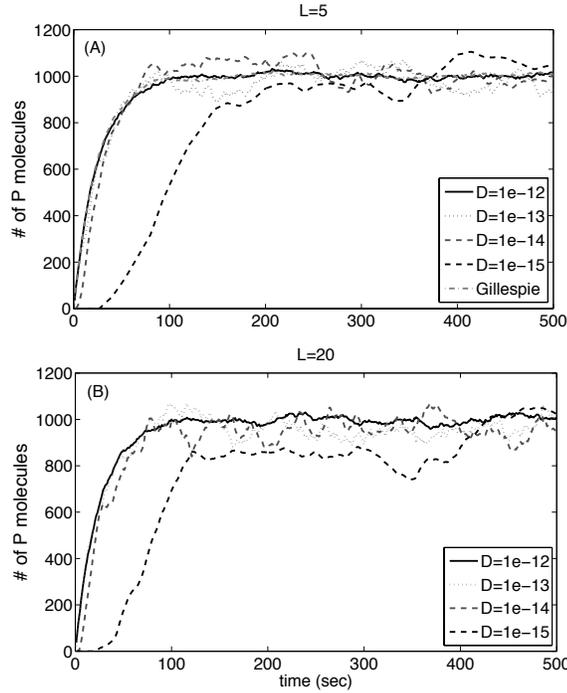
Figure 3.7: Gene expression case study: (A) Dash-dotted line shows the result of Gillespie algorithm which deals with only reaction process. For $D = 10^{-12}$ m^2/s with $L = 5$, our results are similar to those obtained with the Gillespie algorithm, because it is a reaction-limited process so that diffusion does not have serious impact on the system. On the contrary, the case of $D = 10^{-15}$ m^2/s with $L = 5$ exhibits long time lag to reach the steady-state value due to diffusion effect and has much larger fluctuations. (B) In case of $L = 20$, the mesh is much finer than the above cases. The results are similar to those in (A) as our algorithm is able to adjust time-steps according to the system characteristics even though L is increased from 5 to 20.
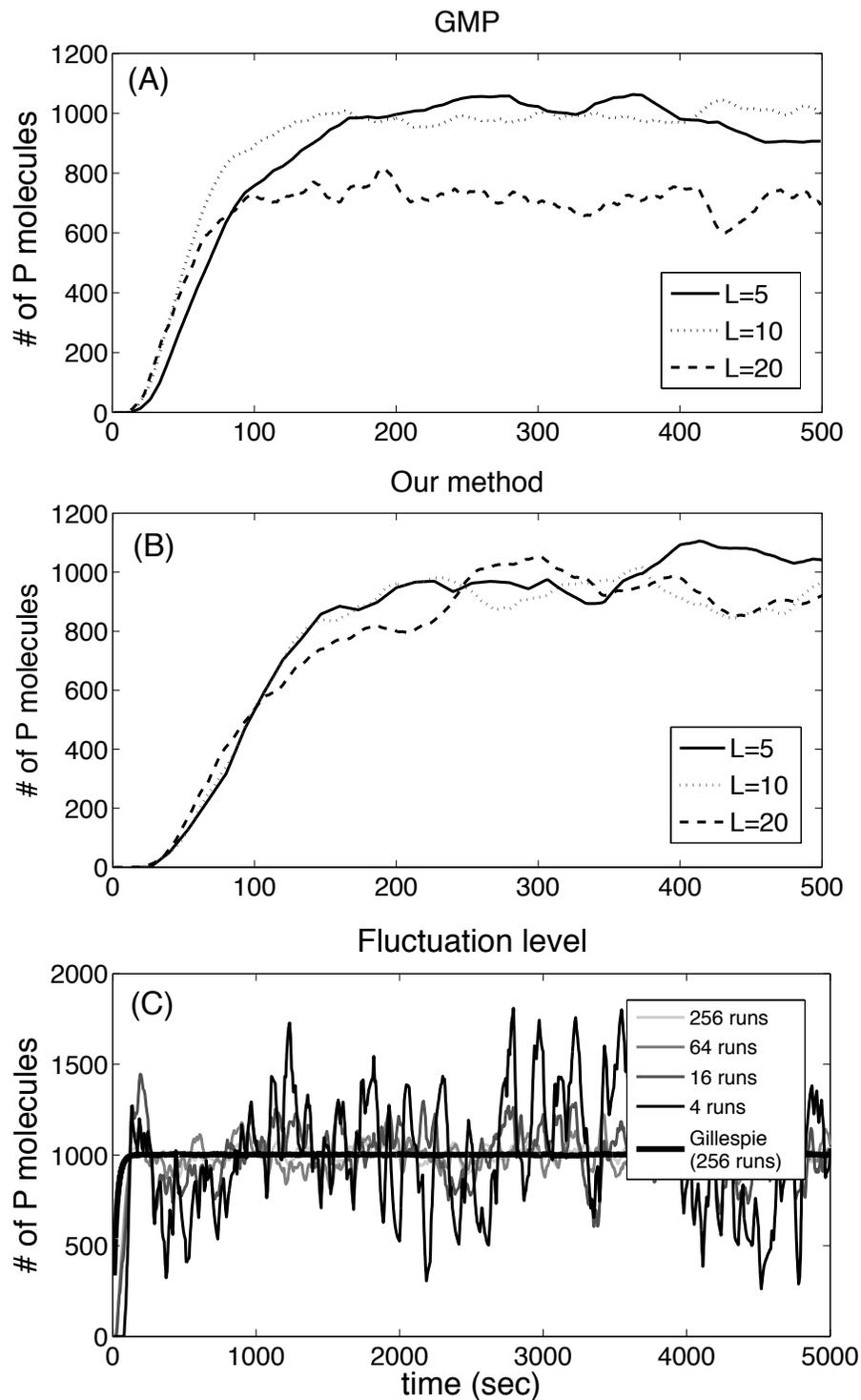
Figure 3.8: Gene expression case study: (A) The result of GMP algorithm for various L and the corresponding Δt $(= \tau_D)$ values. Diffusion constant has a fixed value, $D = 10^{-15}$ m^2/s. For $L = 20$ ($\bar{\tau}_R = 0.65$ s and $\tau_D = 0.417$ s), the number of P molecules does not reach its steady-state value of around 1000
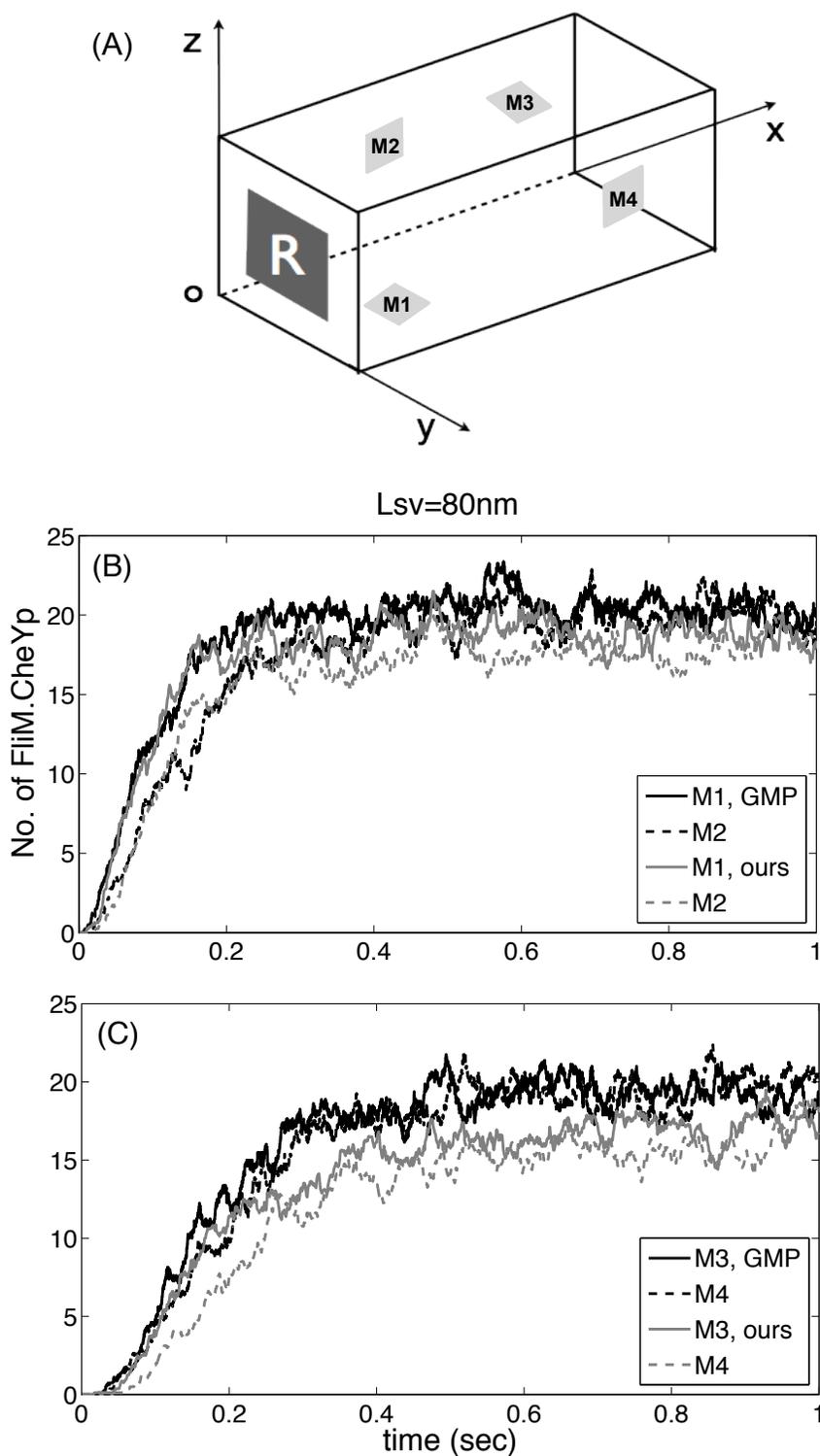
Figure 3.9: CheY diffusion case study: (A) *E. coli* has length [2.48 0.88 0.88] μm along x, y and z direction. R denotes receptor cluster located on the anterior wall and is represented by [x_min x_max; y_min y_max; z_min z_max] = [0 0.08; 0.16 0.64; 0.16 0.64] μm. CheY molecules are phosphorylated in the receptor cluster and

Chapter 4

Stochastic Models of Chemotaxis-Diffusion-Reaction Processes in Wound Healing

4.1 Introduction

Complexity of biochemical systems is manifested through the presence of multiple species, reactions and processes, and of multiple spatial and temporal scales. Cells respond to stimulus by activating myriad biochemical networks, altering the cellular state and ending up in an altered phenotype. In the context of inflammation-mediated immune response, the early cellular signaling events occur

in seconds to minutes, leading to a complex cascade of spatially- and temporally-correlated events that culminate in the final immune response after a few days. Traditional approaches model immune responses or other biological processes at each temporal or spatial scale. For instance, molecular simulations deal with the movement of atoms in proteins on the milliseconds time scale, kinetic modeling of reaction networks models processes on the time scales of seconds to hours, while heuristic models of biological systems capture processes that occur on the time scale of days. We rely on some of the recent advances in computational methods and power, and in acquisition and assimilation of data across multiple scales in biological systems, to develop and apply novel multi-scale simulation algorithms to model leukocyte movement during inflammation and wound healing.

Leukocytes respond to skin or tissue injury through an inflammation process that involves their adhesion and rolling along endothelium cells, transmigration and chemotaxis. The movement and physiology of leukocytes after transmigrations are governed by three processes: chemotaxis, diffusion and chemical reactions. Chemotaxis is represented by an advection term in advection-diffusion-reaction equations. Its strength depends on the concentration gradient of chemoattractants that are released from an injury site.

Deterministic ordinary differential equations (ODEs) often provide accurate predictions of the dynamics of biochemical pathways (complex reaction chains with large numbers of reacting molecules). However, they fail when concentrations of reactants and/or their products become small and the law of mass action be-

comes invalid. Under these conditions, randomness in the dynamics of individual molecules becomes pronounced and requires stochastic simulations. Stochasticity plays a critical role in the behavior of many micro-scale biological phenomena, such as cellular signaling and gene regulatory networks [8, 70, 71]. Standard stochastic techniques, e.g., the Gillespie stochastic simulation algorithm [11] and its computationally efficient modifications [12, 70, 72], can be used to model biochemical reactions in such systems. These algorithms are applicable to systems in which reactants and their products are distributed uniformly in space.

Spatial inhomogeneity has to be considered in crowded environments with complex internal geometry. If the number of molecules is large, partial differential equations (PDEs) provide accurate macroscopic predictions of the dynamics of spatially distributed systems. However, similar to ODE-based models, deterministic PDEs fail to predict random events in systems comprised of small numbers of molecules. Computational methods for reaction-diffusion systems with small numbers of molecules must be able to handle both stochasticity and heterogeneity. Our stochastic numerical algorithm [79] described in Chapter 3 employs an operator splitting, in which the Gillespie algorithm (or its accelerated versions) and Brownian dynamics (or the Smoluchowski equation) are used to model reactions and diffusion, respectively. It employs adaptive time steps to handle both reaction-controlled and diffusion-controlled systems. To model inflammation processes, we enhance this algorithm by incorporating another spatial effect due to chemotaxis.

Four phases of wound healing are briefly discussed in Section 4.2. Our

mathematical model of inflammation is described in Section 4.3. It is followed by introduction of a stochastic operator-splitting approach in Section 4.4. This section also contains a brief description of our stochastic simulation algorithm for modeling reactions and chemotaxis-diffusion. Sections 4.5 and 4.6 contain simulation results and conclusions, respectively.

4.2   Wound healing

Wound healing is a complicated process of tissue repair after injury or damage. It is divided into four sequential but overlapping phases: hemostasis, inflammation, proliferation, and remodeling [80].

4.2.1   Hemostasis

Upon injury in the tissue, a cascade of biochemical events takes place to repair the damage. Within few minutes after injury, in order to control bleeding, platelets aggregate at the injury site to form a fibrin clot. Next, vasoconstriction is followed by vasodilation that causes blood vessels to become porous by releasing histamine. Increased porosity of blood vessels facilitates transmigration of leukocytes from bloodstream to the wound site.

4.2.2 Inflammation

Within an hour of wounding, polymorphonuclear neutrophils (PMNs) arrive at the wound site to phagocytize debris and kill bacteria [80]. As fewer inflammatory factors are secreted, numbers of neutrophils and macrophages are reduced at the wound site and inflammation process ends. However, if the inflammatory phase lasts for too long, e.g., due to the presence of dirt or other foreign objects, it could lead to a chronic wound [81]. This inflammation process includes the leukocytes movement through rolling, adhesion, transmigration and chemotaxis ([82] and Fig. 4.1).

From rolling to chemotaxis

A leukocyte adhesion cascade during the inflammatory phase consists of rolling, firm adhesion and transmigration. Each of these steps plays an important role in effective leukocyte recruitment to the damaged tissues. Rolling, which is mediated by selectin family (L-, P-, E-selectins), is characterized by leukocyte movement with a typical velocity of approximately $5 \sim 10$ μm/s and rolling distance on the order of 100 μm. Adhesion, which is facilitated by integrins and their binding molecules, is a much slower process characterized by a velocity on the order of $v = 10$ μm/min and crawled distance on the order of 10 μm [83]. After firm adhesion, leukocytes transmigrate (with a certain probability) into the tissue through the endothelium cells whose thickness is about $1 \sim 2$ μm. Once in the tissue, leukocytes move toward a wound area driven by diffusion and chemotaxis.

All these processes have a wide range of time- and length-scales, which makes it necessary to develop a multi-scale modeling framework.

Capture and rolling

The first contact of a leukocyte with the activated endothelium is referred to as "capture" or "tethering". Once leukocytes are captured, they may adhere to the endothelium and begin to roll. It is common to refer to a leukocyte as rolling when its rolling velocity is less than half of the blood stream velocity. This process is mediated by the selectin family (L-, E-, P-selectins), and transitory bonds between selectins and their ligands are formed and broken during the rolling motion [82].

Firm adhesion

E-selectin participates in the conversion of rolling to firm adhesion. This process is facilitated by integrins (LFA-1, Mac-1, VLA-4), which are activated by chemokines and soluble factors released by endothelial cells. In the activated state, integrins bind tightly to corresponding receptors expressed on endothelial cells. As a result, despite the shear force of blood flow, the leukocytes are nearly immobilized [82].

Transmigration

After firm adhesion, leukocytes migrate to a wound area in the tissue by extending pseudopodia and passing either through gaps between endothelial cells

(paracellular) or directly through the cells (transcellular) [84]. Paracellular transmigration is not well understood. It is thought that they migrate out of blood vessels in a homophilic way, assisted by bindings of VE-cadherin, PECAM-1, CD99 and JAM family molecules [85] found on the leukocyte and endothelial cell surfaces.

Chemotaxis

After transmigrating from blood vessels, leukocyte cells (neutrophils arrive first) move along the concentration gradient of chemicals (chemokines or cytokines, e.g., fMLP or N-formyl-methionine-leucine-phenylalanine) that are produced at and diffused from the wound site. Chemotaxis causes leukocytes to migrate preferentially in the concentration gradient direction with some probability called chemotactic index (CI). (In addition to chemotaxis, leukocytes also diffuse down their concentration's gradient.)

4.2.3   Proliferation

The proliferative phase is represented by angiogenesis, collagen deposition, tissue formation and epithelialization [81]. During proliferation, the wound is rebuilt with new granulation tissue that is comprised of collagen and extracellular matrix. The process known as angiogenesis leads to development of a new network of blood vessels and healthy granulation tissue. Its efficacy depends on the fibroblast receiving sufficient levels of oxygen and nutrients supplied by blood vessels. Epithelial cells resurface the wound, a process known as epithelialization.

4.2.4 Remodeling

The remodeling phase can last several months or even years. During this phase, collagen is remodeled and complete wound contraction takes place. Wound strength increases from about 20% to 80% within 3 weeks to two years after injury [86].

4.3 A mathematical model of inflammation

Let $b(\mathbf{x}, t)$, $a(\mathbf{x}, t)$ and $c(\mathbf{x}, t)$ denote the concentrations of bacteria, attractant and leukocyte, respectively. Under certain assumptions, these concentrations change due to the mechanisms that are described by the following equations [87].

- Bacteria diffuse, reproduce, and are destroyed when they come in contact with leukocytes:

$$\frac{\partial b}{\partial t} = \mu_b \nabla^2 b + (k_g - k_d c)b, \tag{4.1}$$

where μ_b and k_g are respectively the diffusion coefficient and growth rate of bacteria. This formulation assumes linear kinetics of the bacteria consumption by leukocytes, with the consumption rate k_d.

- The chemoattractant is produced by bacterial metabolism and diffuses:

$$\frac{\partial a}{\partial t} = D \nabla^2 a + k_p b, \tag{4.2}$$

where D and k_p are the diffusion coefficient and production rate of the chemoattractant, respectively.

- The leukocytes are chemotactically attracted to the attractant and die when they digest the bacteria:

$$\frac{\partial c}{\partial t} = -\nabla \cdot \mathbf{J}_c - (g_0 + g_1 b)c, \qquad \mathbf{J}_c = -\mu \nabla c + \chi c \nabla a. \qquad (4.3)$$

Here the reaction term represents linear kinetics of leukocytes destruction due to both natural death in the tissue (with rate constant g_0) and interactions with bacteria (with rate constant g_1, which we set to $g_1 = k_d$ in the subsequent analysis). The equilibrium flux of leukocytes, J_c, is parameterized by the diffusion coefficient of leukocytes μ and the chemotaxis coefficient χ.

4.3.1 Diffusion

Chemoattractant and bacteria diffusion

We treat diffusion of chemoattractant and bacteria as continuous processes that are governed by diffusion equations. These are solved by finite differences to yield the concentrations of chemoattractant and bacteria, $b(\mathbf{x}, t)$ and $a(\mathbf{x}, t)$. The concentration of chemoattractant affects chemotaxis of leukocytes, such that leukocytes move towards the wound site where the concentration of chemoattractant is highest.

Leukocyte diffusion

Motility of a cell population, i.e., diffusion of leukocytes, is described by the motility coefficient (diffusion coefficient) μ. Random motility of a single cell

is characterized by its speed $\tilde{\mathbf{v}}$ and persistence time T_p. The latter is defined as the time interval during which a cell moves without changing direction [88]. The population and single-particle motility coefficients are related by [7]

$$\mu = \frac{1}{2}T_p v^2, \qquad v = |\tilde{\mathbf{v}}|. \tag{4.4}$$

4.3.2 Chemotaxis

At the continuum level, chemotactic velocity $\mathbf{V}_c$ is the product of "oriental bias" ϕ and macroscopic leukocyte velocity $\mathbf{v} = \sum_i v_i \mathbf{e_i}$ [88],

$$\mathbf{V}_c = \sum_i \phi_i v_i \mathbf{e_i} = \sum_i \chi_i \frac{\partial a}{\partial x_i}. \tag{4.5}$$

Since the oriental bias is defined by [88]

$$\phi_i = \frac{\chi_0 \dfrac{dN_b}{da} \dfrac{\partial a}{\partial x_i}}{1 + \chi_0 \dfrac{dN_b}{da} \dfrac{\partial a}{\partial x_i}}, \tag{4.6}$$

the chemotaxis coefficient has the form

$$\chi_i = v_i \frac{\chi_0 \dfrac{dN_b}{da}}{1 + \chi_0 \dfrac{dN_b}{da} \dfrac{\partial a}{\partial x_i}} \mathbf{e_i}, \tag{4.7}$$

where $i(= 1, 2$ and $3)$ is the number of dimensions.

The constant χ_0 is called chemotactic sensitivity and $0 \leq \phi \leq 1$, such that $\phi = 0$ represents purely random process and $\phi = 1$ corresponds to perfectly directed motion. A Michaelis-Menten relationship

$$N_b = \frac{N_T a}{K_d + a} \tag{4.8}$$

is used to relate the number of bound receptors on the cell membrane, N_b, to the concentration of attractant, a [87]. Here K_d is the receptor dissociation constant, and N_T is the total number of cell receptors on the cell membrane.

4.3.3 Reactions

We assume that bacteria and leukocyte have negligible volume. Bacteria (B) reproduce with the rate k_g, leukocytes (C) die with the rate g_0, and bacteria and leukocytes react and die with reaction rate k_d or g_1.

$$B \xrightarrow{k_g} 2B, \qquad C \xrightarrow{g_0} \phi \qquad B + C \xrightarrow{k_d(=g_1)} \phi. \tag{4.9}$$

The reaction constant k_d is defined as [?]

$$k_d = 4\pi(\mu_B + \mu)(r_B + r_C), \tag{4.10}$$

where μ_b and μ are the diffusion constant of bacteria and leukocytes and r is the radius. We set $k_d = g_1$, since it is assumed that both bacteria and leukocytes die when they enter the same neighborhood of radius $r_B + r_C$. In the simulations reported below we set $r_B = 1$ μm and $r_C = 10$ μm. Values of the other model parameters (reaction rates, diffusion coefficients, and chemotactic constant) are listed in Table 4.1.

4.4   Numerical approach

We extend our stochastic operator splitting approach [79] (Chapter 3), in which reactions are handled through a stochastic simulation algorithm (e.g., Gillespie's [11]) and diffusion is treated through Brownian dynamics. Leukocytes moves through chemotaxis and diffusion. Bacteria and leukocytes undergo biochemical reactions. Spatio-temporal evolution of leukocytes concentrations can be described by a system of chemotaxis-diffusion-reaction equations,

$$\frac{\partial c}{\partial t} + \nabla \cdot (\mathbf{V_c} c) = \mu \nabla^2 c - (g_0 + g_1 b)c \tag{4.11}$$

where $\mathbf{V_c}$ is the chemotactic velocity vector and μ is diffusion coefficient and reactions terms are explained in Eq. 4.3.

We employ an operator-splitting algorithm to approximate Eq. 4.11 with

$$\frac{\partial c'}{\partial t} = \mu \nabla^2 c' - \nabla \cdot (\mathbf{V_c} c') \tag{4.12a}$$

$$\frac{\partial c''}{\partial t} = -(g_0 + g_1 b)c''. \tag{4.12b}$$

during the time interval $[t, t + \Delta t]$. Here $c'(t) = c(t)$ and $c''(t) = c'(t + \Delta t)$, so that the concentration of leukocytes at the end of the time-step Δt is $c''(t) = c(t + \Delta t)$.

4.5   Simulation results

4.5.1   Chemotaxis and diffusion

Brian and his colleagues conducted many experiments and showed how chemotaxis, chemokinesis and diffusion affect the movement of leukocytes [88] . Based on the same concept, we start simulations by modeling leukocyte movement without considering the reactions involved. To focus on the interplay between the chemoattractant concentration and the leukocyte movement, we solve only (4.2) and (4.3).

Since the diffusion coefficient of the chemoattractant is much higher than that of leukocytes $(D \gg \mu)$, the chemoattractant diffuses through the tissue much faster than leukocytes move. Moreover, the number of chemoattractant molecules is large enough to ignore stochastic fluctuations in their movement and to warrant the reliance on the deterministic diffusion equation (4.2). Its deterministic solution is to be coupled with the stochastic simulations of the leukocytes movement.

One-dimensional migration

To be concrete, we take the simulation domain to be $\Omega = [x_s, x_e] = [0, 2 \times 10^{-3}]$ m, and impose the following initial and boundary conditions

$$\text{I.C.s}: a(x,0) = 0, \quad c(\frac{x_e}{2}, 0) = c_0, \tag{4.13a}$$

$$\text{B.C.s}: a(x_e, t) = a_0, \quad \frac{\partial a}{\partial x}|_{x_s, x_e} = 0, \quad \frac{\partial c}{\partial x}|_{x_s, x_e} = 0, \tag{4.13b}$$

where a_0 is 1×10^{-9} M and c_0 is 10000, the initial number of leukocytes. Here, reaction term $k_p b$ is considered as a source at $x = x_e$ and has a value of a_0. Stochastic simulations of leukocyte migration replace the deterministic (one-dimensional) chemotaxis-diffusion equation (4.3) with its stochastic counterpart

$$X(t + \Delta t) = X(t) + V_c \Delta t + \sqrt{2\mu \Delta t} \xi \tag{4.14}$$

where Δt is the time step, V_c is computed from the one-dimensional versions of (4.5) and (4.7) and ξ is a standard Gaussian random variable.

In Fig. 4.2, we study the effect as strength of chemotaxis increases and all plots are captured at $t = 10000$ sec. Initially, ten thousands of leukocytes exist in the middle of the x axis and source of chemoattractant exists at the end of x axis. In case that there exists only diffusion process, blue-colored line looks to be normally distributed. When weak chemotaxis process (χ_0) is added to the system, plot moves to the right, which means more leukocytes moves to the right where concentration of chemoattractant is higher. As chemotactic sensitivity increases, chemotactic index, ϕ, increases. Accordingly, larger number of leukocytes moves toward chemoattractant source point as chemotactic sensitivity increases.

Two-dimensional migration

Consider leukocyte migration in a two-dimensional domain $\Omega = [x_s, x_e] \times [y_s, y_e]$. The two-dimensional versions of (4.2) and (4.3) are subject to initial and

boundary conditions

$$\text{I.C.s}: a(x, y, 0) = 0, \quad c(0.4x_e, 0.1y_e, 0) = c_0, \tag{4.15a}$$

$$\text{B.C.s}: a(\frac{x_e}{2}, y_s, t) = a_0, \quad \frac{\partial a}{\partial x}\Big|_{x_s, x_e} = \frac{\partial a}{\partial y}\Big|_{y_s, y_e} = 0, \tag{4.15b}$$

$$\frac{\partial c}{\partial x}\Big|_{x_s, x_e} = \frac{\partial c}{\partial y}\Big|_{y_s, y_e} = 0, \tag{4.15c}$$

where $[x_s, x_e] = [y_s, y_e] = [0, 4 \times 10^{-3}]$ m and a_0 is 1×10^{-8} M and c_0 is 3, the initial number of leukocytes. We use the alternating direction implicit (ADI) method [74] to solve the deterministic two-dimensional equation (4.2). The ADI method solves two one-dimensional problems at each time step:

$$\frac{a_{ij}^{n+1/2} - a_{ij}^n}{D\Delta t/2} = \frac{a_{i+1j}^{n+1/2} - 2a_{ij}^{n+1/2} + a_{i-1j}^{n+1/2}}{\Delta x^2} + \frac{a_{ij+1}^n - 2a_{ij}^n + a_{ij-1}^n}{\Delta y^2} \tag{4.16a}$$

$$\frac{a_{ij}^{n+1} - a_{ij}^{n+1/2}}{D\Delta t/2} = \frac{a_{i+1j}^{n+1/2} - 2a_{ij}^{n+1/2} + a_{i-1j}^{n+1/2}}{\Delta x^2} + \frac{a_{ij+1}^{n+1} - 2a_{ij}^{n+1} + a_{ij-1}^{n+1}}{\Delta y^2}. \tag{4.16b}$$

The reaction term, $k_p b$, can be thought as a source term in the wound site. Leukocyte displacements in the x and y directions are computed stochastically with (4.14).

Figure 4.3 shows the geometry and the simulation results for leukocyte movement. Leukocyte movement due to chemotaxis increases as chemotaxis constant increases in the same manner as 1D simulation. At $t = 5000$ sec, leukocytes with stronger chemotaxis $(10\chi_0)$ arrive at the wound site, meanwhile those with weaker chemotaxis (χ_0) are still moving forward to wound site.

4.5.2 Chemotaxis, diffusion and reactions

In this section, reactions are considered as well as chemotaxis and diffusion. We employ operator splitting algorithm explained in Sec. 3.2.1. Leukocyte migration can be characterized by (dimensionless) Pèclet number

$$Pe = \frac{V_c h}{D} = \frac{\tau_D}{\tau_C} \tag{4.17}$$

where h is the mesh (or lattice) size, and τ_D and τ_C are diffusion and convection time scales, respectively. If $Pe > 1$, chemotaxis is slower than diffusion. Since this regime is dominated by diffusion, we designate τ_D as the drift time constant. Inversely, if $Pe < 1$, τ_C is chosen as the drift time constant. Finally, we define Damköhler number as

$$Da = \frac{\min(\tau_C, \tau_D)}{\tau_R} \tag{4.18}$$

The value of the Damköhler number determines whether the system is drift controlled or reaction controlled in the operator splitting algorithm [79].

Figure 4.4 shows the three-dimensional geometry of the simulation domain, $\Omega = [x_s, x_e] \times [y_s, y_e] \times [z_s, z_e]$.

$$\text{I.C.s}: a(x,y,z,0) = 0, \quad c(x,y,z,0) = c_0, \tag{4.19a}$$

$$\text{B.C.s}: a(\frac{x_e}{2}, y_s, z_s, t) = a_0, \quad \frac{\partial a}{\partial x}|_{x_s,x_e} = \frac{\partial a}{\partial y}|_{y_s,y_e} = \frac{\partial a}{\partial z}|_{z_s,z_e} = 0, \tag{4.19b}$$

$$\frac{\partial c}{\partial x}|_{x_s,x_e} = \frac{\partial c}{\partial y}|_{y_s,y_e} = \frac{\partial c}{\partial z}|_{z_s,z_e} = 0, \tag{4.19c}$$

where $[x_s, x_e] = [y_s, y_e] = [z_s, z_e] = [0, 1 \times 10^{-3}]$ m and a_0 is 1×10^{-8} M and c_0 is 500, the initial number of leukocytes which are positioned arbitrarily inside the domain,

Ω. Chemoattractant diffusion is solved by the three-dimensional ADI method in the same method as 3D ADI. Leukocytes diffuse randomly and chemotax in the system. Bacteria are fixed in the small rectangle and can react with leukocytes when they meet in the same mesh.

In Table 4.2, we can see that chemotactic index increases as strength of chemotaxis increases. As a result, more bacteria are digested through more frequent reactions. However, the change of random motility have no significant effect upon the number of bacteria. It means that chemotaxis sensitivity has a stronger effect to inflammation than diffusion (random motility).

4.6   Summary and Discussion

We apply stochastic operator splitting method to inflammation process during wound healing. Mathematical modeling is described by partial differential equations comprised of chemotaxis, diffusion and reaction processes. It is very difficult and improper to employ deterministic approach due to a wide variety of temporal and spatial scales. Therefore, we analyze chemoattractant diffusion equation using deterministic ADI method because molecular diffusion is much faster than leukocytes and bacteria and number of chemoattractant molecules is reasonably high. However, leukocytes movement and reactions with bacteria are analyzed stochastically by operator splitting algorithm. We consider Péclet and Damköler number in order to decide drift time scale and if system is diffusion or

drift controlled.

In order to verify if simulation works properly, diffusion and chemotaxis are studied first without reaction in 1D and 2D. We figure out that more leukocytes move forward to wound site as chemotactic constant increase. Next, reactions are analyzed with drift in 3D simulation. Similarly, more leukocytes move to wound site as chemotactic constant increases. As a result, more reactions take place and more bacteria are digested by leukocytes.

Table 4.1: Constants for reaction, diffusion and chemotaxis.

Notation	Description	Value [Unit]
χ_0	chemotactic sensitivity	$4e^{-8}$ [m/receptor]
μ_b	random motility coeff. of bacteria	0
μ	random motility coeff. of leukocytes	calculated in Eq. 4.4 [m^2/s]
D	diffusion constant of chemoattractants	$1e^{-9}$ [m^2/s]
k_g	generation rate of bacteria	$1.4e^{-4}$ [s^{-1}]
$k_d(= g_1)$	decay rate of bacteria/leukocytes	calculated in Eq. 4.10 [M^{-1}s^{-1}]
k_p	production rate of bacteria	- [s^{-1}]
g_0	generation rate of leukocytes	$2e^{-6}$ [s^{-1}]

Table 4.2: Initial number of bacteria and leukocytes are 50 and 500. 32 simulations are conducted for 10000 sec. As chemotactic constant increases, chemotactic index increases, which means more leukocytes move forward to wound site. As a result, more reactions take place and more bacteria are killed. However, the change of random motility have no significant effect upon the number of bacteria. It means that chemotaxis has a stronger effect to inflammation than diffusion (random motility).

| Strength of chemotaxis	Chemotactic index	Mean of number of bacteria
$100\chi_0$	$0.4 \sim 0.5$	3.5
$10\chi_0$	$0.3 \sim 0.4$	26.5
χ_0	~ 0.05	48.8
Random motility coeff.	Chemotactic index	Mean of number of bacteria
10μ	~ 0.05	48.1
1μ	~ 0.05	48.8
0.1μ	~ 0.05	50
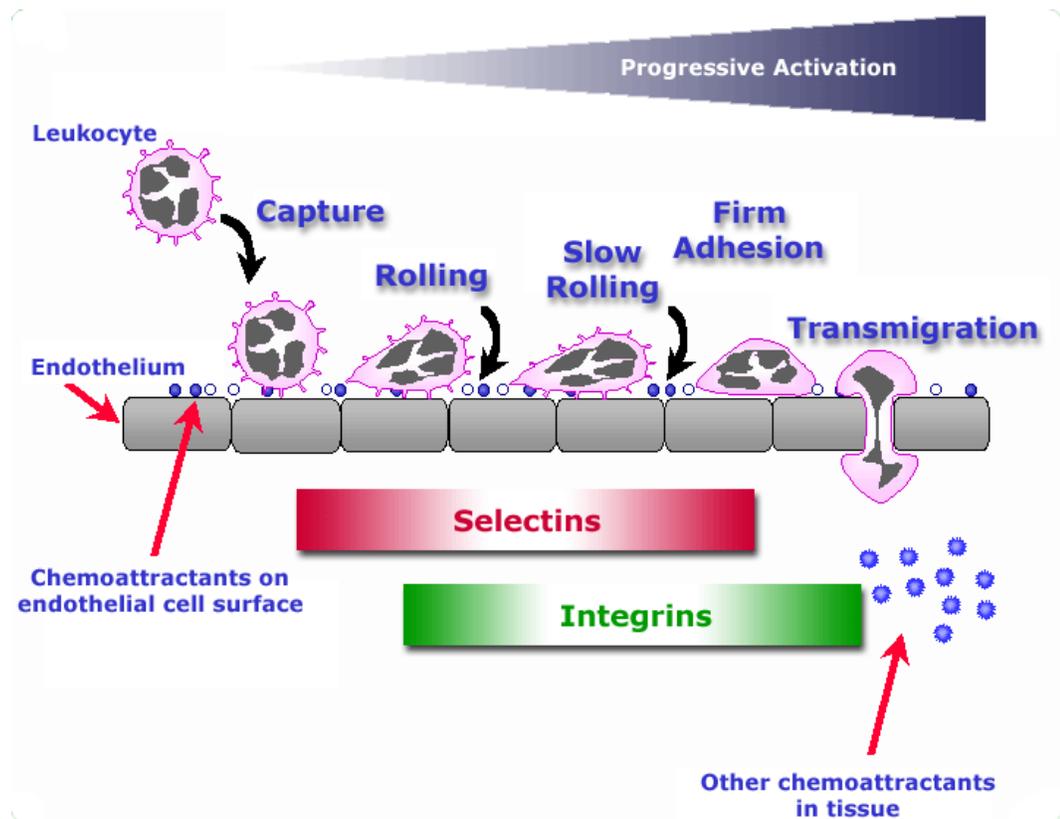
Figure 4.1: Leukocytes flow along the blood stream. When injury occurs in the tissue, they begin to roll and adhere on endothelium cells. Next, they transmigrate through endothelial cells by the effect of histamine. Transmigrated leukocytes move toward lesions by chemotaxis and diffusion processes.
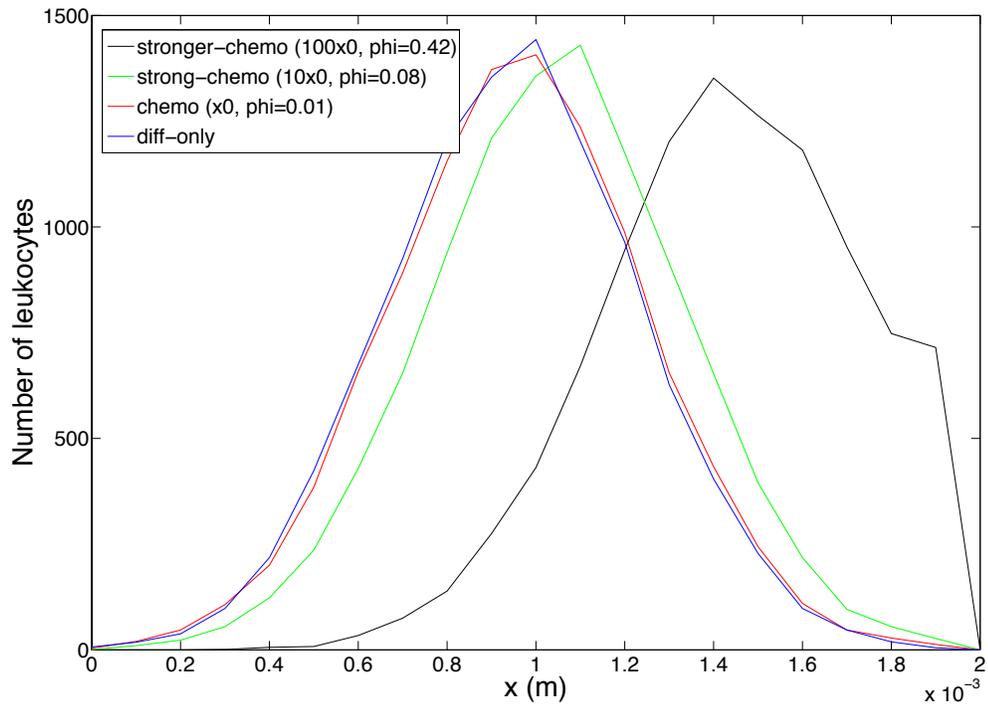
Figure 4.2: 1D diffusion and chemotaxis; Initially, 10000 leukocytes sit on the center of x-axis and chemoattractant source exist in the right end of x-axis. Chemotactic index, ϕ, increases as chemotactic coefficient, χ_0, increases. Therefore, larger number of leukocytes moves toward chemoattractant source point. All plots are captured at $t = 10000$ sec.
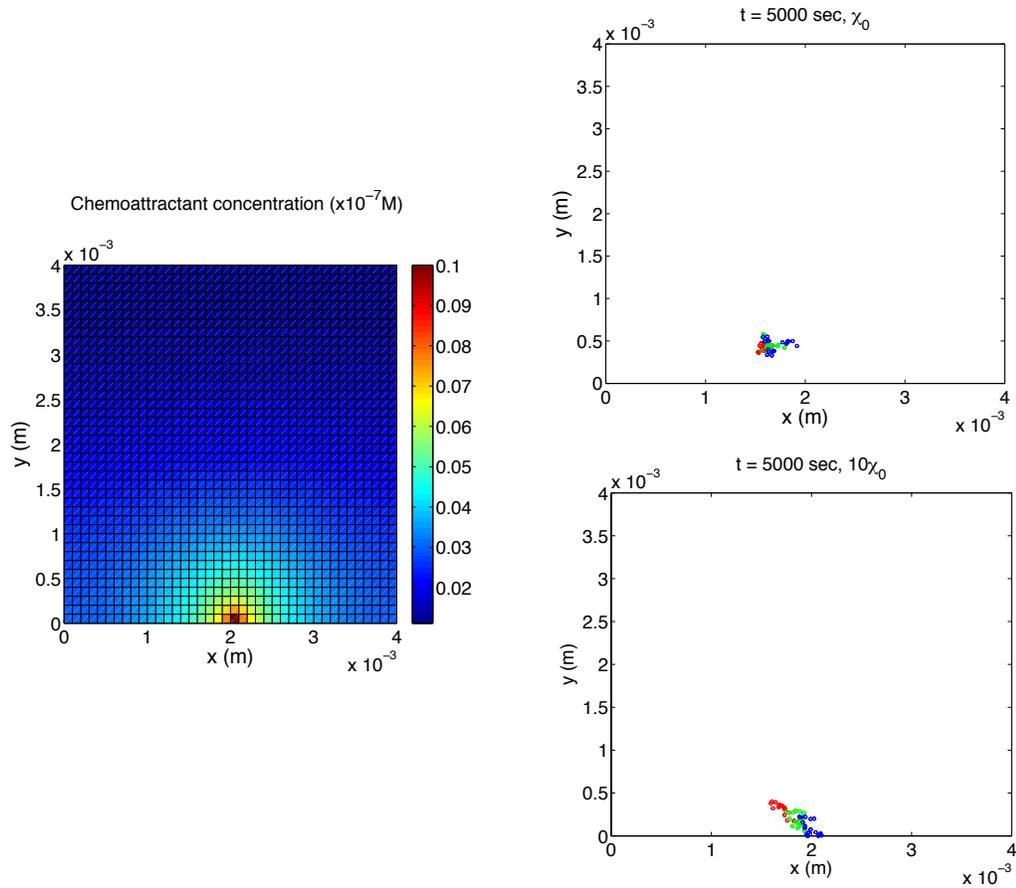
Figure 4.3: 2D diffusion and chemotaxis; (A) It shows the concentration distribution of chemoattractant. The center of x axis is the source point. (B)-(C) Leukocytes moves toward chemoattractant source point as chemotactic coefficient increases. The positions of leukocytes are traced from $t = 0$ to 5000 sec. The color is changed from red ($t = 0 \to 1667$ sec) to green ($t = 1667 \to 3333$ sec) and to blue ($t = 3333 \to 5000$ sec) as simulation time increases.
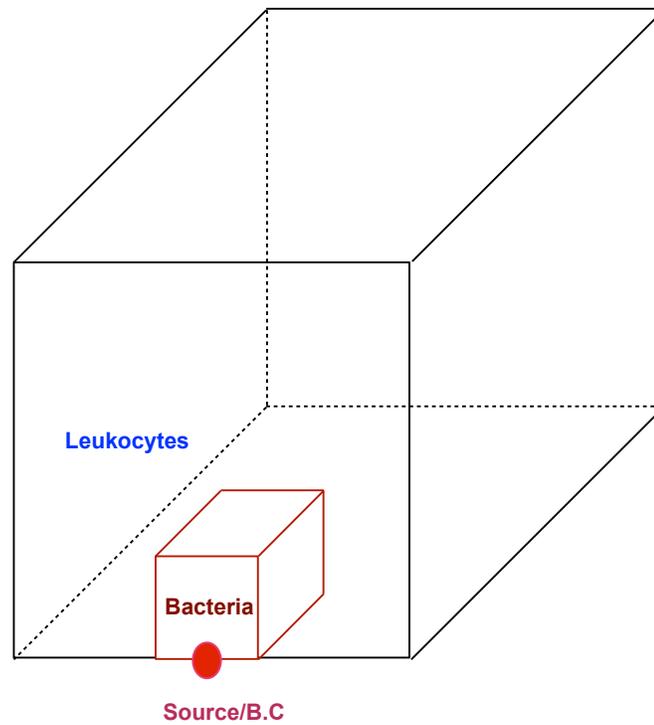
Figure 4.4: 3D diffusion, chemotaxis and reactions; Leukocytes diffuse randomly in the whole domain. Bacteria are fixed in small cube and considered as a source of chemoattractant in the center of x axis.

Chapter 5

Conclusions

This dissertation leads to the following major conclusions:

1. We have developed a hybrid approach to stochastic simulation, in which slow reactions and fluxes are handled through exact stochastic simulation and their fast counterparts are treated partially deterministically through the chemical Langevin equation. The classification of reactions as fast or slow is accompanied by a partial equilibrium assumption, according to which a population of slow species is not altered by fast reactions. Our new approach also handles reactions with complex rate expressions such as functions of Michaelis-Menten kinetics and power-law kinetics by developing mathematical expressions for their propensity functions and microscopic fluxes. Fluxes which cannot be modeled explicitly through reactions are handled deterministically.

2. We have integrated the existing techniques for multiscale stochastic simula-

tion with deterministic simulation to deal with complex reactions systems and have applied it to studying calcium dynamics in macrophage cells. When the concentration of reactants is sufficiently large, the stochastic method yields time-course profiles identical to those obtained from a deterministic model (ensemble average of 16 or more realizations). However, at lower number of molecules of one or more species, measurable relative difference in $[Ca^{2+}]_i$ responses predicted by the two approaches is obtained, especially for the case of $G_{\beta\gamma}$, thus suggesting the necessity of using stochastic simulation as opposed to deterministic simulation for studying system dynamics at sub-cellular levels. Dose response analysis revealed that while the normalized response difference (NRD) between $[Ca^{2+}]_i$ responses predicted by deterministic and stochastic simulations is negligible at the full dose of 30nM (shown) or higher doses including saturating doses (not shown), it increases with decreasing doses. At 0.1% dose, it is as high as 7%. These results are emphasized again in the sensitivity analysis of the parameters used in the simulation and in the knockdown analysis of reacting protein components.

3. Complex multi-scale biological systems can be analyzed with microscopic approaches, such as Green's-function reaction dynamics and the Smoldyn algorithm. These methods are accurate albeit computationally expensive and often prohibitive. On the other hand, macroscopic kinetic modeling approaches that use PDEs are amenable to numerical computation, but fail to model the physics of systems with small number of molecules accurately. Mesoscopic

approaches, e.g., reaction-diffusion master equation and MesoRD, discretize space into a collection of lattice elements and extend the chemical master equation normally used in well-mixed chemical reactions into the stochastic regime for inhomogeneous systems. To facilitate faster and more accurate solutions within the mesoscopic scale framework, we have developed a stochastic simulation method which is based on operator-splitting for modeling the reaction-diffusion system. In our methodology, the time-step size is chosen automatically at each step depending upon whether the system is reaction- or diffusion-controlled. We use the Gillespie stochastic simulation algorithm for modeling the reactions and Brownian dynamics approach for modeling the diffusion process. We thus account for both spatial heterogeneity and the fluctuation in concentrations arising from the small number of molecules. Our method yields highly accurate results and has the merit of modeling both the reaction and diffusion processes in the system. We also have a better result in terms of simulation speed over GMP method. The average speed-up by using our method is about 5 times as compared to the GMP method for a wide range of the values of the diffusion constant in case of simple reaction-diffusion system, $A + B \rightarrow C$. In more complicated system, CheY diffusion in *E. coli*, the simulation time for our algorithm is 12hrs, whereas it is 26 hrs for the GMP method.

4. We applied stochastic operator splitting method to inflammation process during wound healing. The mathematical modeling is described by partial differ-

ential equations comprised of chemotaxis, diffusion and reaction processes. It is difficult and improper to employ deterministic approach due to a wide variety of temporal and spatial scales. Therefore, we analyze chemoattractant diffusion equation using deterministic ADI method because molecular diffusion is much faster than leukocytes and bacteria and number of chemoattractant molecules is reasonably high. However, leukocyte movement and reactions with bacteria are analyzed stochastically by operator splitting algorithm. We consider Péclet and Damköler number in order to decide drift time scale and if system is diffusion- or drift- controlled. In the studies on diffusion and chemotaxis in 1D and 2D, we figure out that more leukocytes move forward to wound site as chemotactic constant increase. In 3D simulation with drift and reactions, more leukocytes move to wound site as chemotactic constant increases and thus more reactions take place and more bacteria are digested by leukocytes.

Appendix A

Existing algorithms for stochastic simulation

Three algorithms, namely Gillespie algorithm, tau-leap method and chemical Langevin equation are reviewed. These algorithms are applied to a well-stirred biochemical system (molecules of each species are spread uniformly throughout a fixed control volume) at thermal equilibrium comprising M different chemical reactions and N different types of chemical species.

At any time t, the population of molecules within a fixed cellular volume V is uniquely described by a state vector $\mathbf{X}(t)$,

$$\mathbf{X}(t) = \{X_1(t), X_2(t), \ldots, X_N(t)\}^T, \tag{A.1}$$

where $X_i(t)$ is the number of molecules of the i-th species $(i = 1, \ldots, N)$. By definition, X_i are non-negative integers. The state vector $\mathbf{X}(t)$ changes whenever

one of the M types of reactions occur.

Let $P[\mathbf{X}; t]$ denote the probability of the system being in the state represented in Eq. (A.1) at time t. Furthermore, let $a_j(\mathbf{X})$ denote the *propensity function* of the j-th reaction channel, which is defined through $a_j(\mathbf{X})\mathrm{d}t$, the probability that the j-th reaction will occur during a (sufficiently small) time interval $[t, t + \mathrm{d}t]$ given the system state $\mathbf{X}(t)$ at time t. The propensity function $a_j(\mathbf{X})$ can be expressed as [61]

$$a_j(\mathbf{X}) = c_j h_j(\mathbf{X}), \qquad j \in \mathcal{M}, \qquad \mathcal{M} = \{1, 2, \ldots, M\}, \tag{A.2a}$$

where $c_j > 0$ is the specific probability rate constant of the j-th reaction, and $h_j(\mathbf{X})$ is the number of possible combinations of reactants in the j-th reaction. The former is given by

$$c_j = \begin{cases} k_j & \text{for monomolecular reactions} \\[2ex] \frac{2k_j}{N_A V} & \text{for bimolecular reactions with identical reactants} \\[2ex] \frac{k_j}{N_A V} & \text{for bimolecular reactions with different reactants} \end{cases} \tag{A.2b}$$

where $N_A = 6.022 \times 10^{23}$ mol^{-1} is Avogadro's constant, and k_j is the macroscopic

reaction rate constant [89]. The latter has the form

$$
h_j(\mathbf{X}) = \begin{cases} 1 & \text{for reactions of type: } \phi \to \text{product(s) [no reactants]} \\[1.5em] X_i & \text{for monomolecular reactions } [A \to B] \\[1.5em] X_i(X_i - 1)/2 & \text{for reactions with identical reactants } [A + A \to C] \\[1.5em] X_i X_k & \text{for reactions with different reactants } [A + B \to C] \end{cases}
$$

(A.2c)

for some $1 \leq j \leq M$, and $1 \leq i, k \leq N$ with $i \neq k$.

Let the time-interval $\mathrm{d}t$ in the definition of $a_j(\mathbf{X})$ (Eq. (A.2)) be small enough that at most one reaction occurs during $[t, t + \mathrm{d}t)$, then $P[\mathbf{X}; t]$ satisfies an ODE that is commonly called a chemical master equation (CME) [61],

$$
\frac{\mathrm{d}P[\mathbf{X}; t]}{\mathrm{d}t} = \sum_{j=1}^{M} \{a_j(\mathbf{X} - \boldsymbol{\nu}_j)P[\mathbf{X} - \boldsymbol{\nu}_j; t] - a_j(\mathbf{X})P[\mathbf{X}; t]\}, \tag{A.3}
$$

where $\boldsymbol{\nu}_j = (\nu_{j_1}, \ldots, \nu_{j_N})^T$ is a vector whose entries are the numbers of molecules of each species added to or removed from the volume V due to the j-th reaction. For complex cellular processes, such as calcium dynamics in the signaling network described in Section 2.3.1, high-dimensionality of the CME (Eq. (A.3)) renders its solutions computationally prohibitive. The standard simulation algorithms described below serve to overcome the computational burden associated with solving Eq. (A.3).

A.1 Gillespie algorithm

Let $P_0[\tau|\mathbf{X}, t]$ denote the conditional probability that no reactions occur during the time interval $[t, t + \tau)$ provided that the system is at state $\mathbf{X}$ at time t. Furthermore, let us assume that the reacting system is Markovian, i.e., the probability that no reactions occur during $[t, t + \tau + \mathrm{d}\tau)$ equals the product of probability that no reactions occur during $[t, t+\tau)$ and probability that no reactions occur during $[t+\tau, t+\tau+\mathrm{d}\tau)$. Then the definition of the propensity function implies that [61]

$$P_0[\tau + \mathrm{d}\tau|\mathbf{X}, t] = P_0[\tau|\mathbf{X}, t]\left[1 - a_{\mathrm{sum}}(\mathbf{X})\mathrm{d}\tau\right], \qquad a_{\mathrm{sum}}(\mathbf{X}) \equiv \sum_{j=1}^{M} a_j(\mathbf{X}). \quad (\text{A.4})$$

Taking the limit as $\mathrm{d}\tau \to 0$ and solving the resulting ODE, we obtain

$$P_0(\tau|\mathbf{X}, t) = \mathrm{e}^{-a_{\mathrm{sum}}(\mathbf{X})\tau}. \qquad (\text{A.5})$$

Using the definition of P_0 and a_j, it can be shown [61] that the joint probability density function $p(\tau, j|\mathbf{x}, t)$, which describes the probability that the next reaction will be the j-th reaction and will occur during $[t + \tau, t + \tau + \mathrm{d}\tau)$ given the present state of the system $\mathbf{X}(t)$, is given by $p(\tau, j|\mathbf{X}, t) = P_0[\tau|\mathbf{X}, t]a_j(\mathbf{X})$. Accounting for Eq. (A.5), we obtain

$$p(\tau, j|\mathbf{X}, t) = \frac{a_j(\mathbf{X})}{a_{sum}(\mathbf{X})}a_{sum}(\mathbf{X})\mathrm{e}^{-a_{\mathrm{sum}}(\mathbf{X})\tau}. \qquad (\text{A.6})$$

The ratio $a_j(\mathbf{X})/a_{sum}(\mathbf{X})$ represents the density of a discrete random variable, and serves to determine the next reaction. The remainder of the right-hand-side of Eq.

(A.6), $a_{sum}(\mathbf{X})\exp[-a_{\text{sum}}(\mathbf{X})\tau]$ is the exponential density function of a continuous random variable, which corresponds to the time at which the next reaction will occur.

To advance the system from state $\mathbf{X}(t)$, the Gillespie algorithm generates two random variables r_1 and r_2 distributed uniformly on the unit interval $[0, 1]$. According to Eq. (A.6), a discrete random value j and continuous random value τ are selected as

$$\tau = \frac{1}{a_{\text{sum}}}\ln\left(\frac{1}{r_1}\right), \qquad \sum_{j'=1}^{j-1} a_{j'} \leq r_2 a_{\text{sum}} \leq \sum_{j'=1}^{j} a_{j'}. \tag{A.7}$$

The system is then updated according to $\mathbf{X}(t+\tau) = \mathbf{X}(t) + \boldsymbol{\nu}_j$.

A faster algorithm for exact stochastic simulation has been presented by [64], called "next reaction method". This approach is about an order of magnitude faster than the Gillespie algorithm discussed above.However, this approach does not scale as well as the tau-leap algorithm discussed below as the number of molecules increases.

A.2 Tau-leap algorithm

The tau-leap algorithm [12] can be used to increase the computational efficiency of the Gillespie algorithm when it is used to simulate large reactive systems consisting of many reactions and molecules. This algorithm allows many reactions to take place simultaneously during a time interval $[t, t+\tau)$. Let $K_j(\tau|\mathbf{X}, t)$ denote the number of times j-th reaction $(1 \leq j \leq M)$ takes place during the time inter-

val $[t, t + \tau)$, given the system state $\mathbf{X}(t)$ at time t. The value of τ is selected to satisfy the so-called "leap condition", which requires that none of the propensity functions a_j $(1 \leq j \leq M)$ suffers a noticeable change in its value. Then $K_j(\tau|\mathbf{X}, t)$ can be approximated with a Poisson random variable $P\{a_j(\mathbf{X}), \tau\}$ whose mean and variance are $a_j \tau$. The system state is now updated according to

$$\mathbf{X}(t + \tau) = \mathbf{X}(t) + \sum_{j=1}^{M} \boldsymbol{\nu}_j P\{a_j(\mathbf{X}), \tau\}. \tag{A.8}$$

As the time interval τ becomes smaller, it allows for few reactions to take place simultaneously, eventually reaching the limit of one reaction per τ. In this limit, $P\{a_j(\mathbf{x}), \tau\} \to 1$ and we get the Gillespie algorithm.

Algorithmic consistency requires that, in addition to satisfying the leap condition, τ be selected in a way that prevents number of any species from becoming negative. The binomial tau-leap algorithm [90, 23] imposes this constraint by introducing a new control parameter n_c (typically a small positive integer), which defines "critical reactions" as those having at least one species with the number of molecules less than n_c. If there are one or more critical reactions then τ is chosen so that no critical reaction fires more than once. The binomial tau-leap algorithm [90, 23] also expresses the leap condition in terms of a bound on the change rate of $a_j[\mathbf{X}(t)]$ as $|\Delta a_j(\mathbf{X}(t))| \leq \epsilon a_j(\mathbf{X}(t))$, where $0 < \epsilon \ll 1$.

A.3   Chemical Langevin equation

To increase the computational efficiency further, the leap time τ can be increased so that $a_j(\mathbf{X})\tau$ becomes large enough to ensure that it contains a large number of reactions for each reaction channel. Now the Poisson random variable $P\{a_j(\mathbf{X}), \tau\}$ can be approximated with a normal random variable [61] with the same mean and variance: $a_j[\mathbf{X}(t)]\tau + \sqrt{a_j[\mathbf{X}(t)]\tau}Z_j$, where Z_j are independent normal random variables on the interval $(0,1)$. This approximation replaces Eq. (A.8) with a chemical Langevin equation (CLE)

$$\mathbf{Y}(t + \tau) = \mathbf{Y}(t) + \tau \sum_{j=1}^{M} \boldsymbol{\nu}_j a_j[\mathbf{Y}(t)] + \sqrt{\tau} \sum_{j=1}^{M} \sqrt{\boldsymbol{\nu}_j a_j[\mathbf{Y}(t)]}Z_j, \qquad (A.9)$$

where $\mathbf{Y}(t)$ is a continuous counterpart of the discrete random variable $\mathbf{X}(t)$, replacing the number of molecules of the j-th species, X_j, with the respective concentrations Y_j $(j = 1, \ldots, N)$.

Appendix B

Diffusion processes and GMP

algorithm

B.1 Diffusion process: Brownian dynamics

In cells, molecules such as proteins and metabolites, have a non-zero instantaneous speed at room temperature or at the temperature of the human body. A typical protein molecule is immersed in the aqueous medium of a living cell. It collides with other molecules in the solution, exhibiting a random walk or Brownian motion.

Let $\mathbf{X}(t) \in \mathbb{R}^3$ denote the position of a diffusing molecule at time t. Diffusive spreading of molecules of the i-th species $(i = 1, \ldots, M)$ is characterized by a molecular diffusion coefficient D_i, whose value depends on the molecule size, absolute temperature and the viscosity of a solution. The molecule's position at

the end of the time interval $\triangle t$ is computed as follows [77].

1. Generate three normally distributed random numbers ξ_1, ξ_2, and ξ_3 that serve as components of the random displacement vector $\boldsymbol{\xi} = (\xi_1, \xi_2, \xi_3)^T$.

2. Compute the molecule's position at time $t + \triangle t$ as

$$\mathbf{X}(t + \triangle t) = \mathbf{X}(t) + \sqrt{2D_i \triangle t}\, \boldsymbol{\xi}. \tag{B.1}$$

3. Set $t = t + \triangle t$ and go to step 1.

B.2 Diffusion process: Cellular automata

In general, cellular automata depend on mesh size and diffusion constant. Simulation accuracy and computational time vary according to neighborhood types [20]. For the two-dimensional example in Figs. 1B-C (main manuscript), molecules can diffuse to four adjacent cells (voxel) or stay in the original voxel in the von Neumann neighborhood, whereas in the Moore neighborhood they can diffuse to eight adjacent cells or stay in the original voxel. If $(0, 0)$ denotes the original voxel, the von Neumann neighborhood is a set $N_N = \{(-1, 0), (0, -1), (0, 0), (0, 1), (1, 0)\}$. The Moore neighborhood is a set $N_M = N_N \cup \{(-1, -1), (-1, 1), (1, -1), (1, 1)\}$.

The Gillespie multi-particle (GMP) algorithm [17] employs cellular automata to simulate diffusion. A diffusion-time constant τ_{D_i}, the time during which a molecule of the i-th species remains in one cell of a mesh, is given by [18]

$$\tau_{D_i} = \frac{1}{2d} \frac{(\Delta x)^2}{D_i}, \tag{B.2}$$

where D_i is the diffusion coefficient for the i-th species. Moreover, a reaction-time constant $\bar{\tau}_R$ is defined as the ensemble average of the equivalent time constants for all reactions related to diffusing molecules.

B.3   Gillespie multi-particle (GMP) method

We implemented the following GMP algorithm based on [18].

---

1. Set $t_S = \Delta t = \min_i\{\tau_{D_i}\}$ for all diffusing species i.

2. Initialize $t = 0$ and $n_i = 1$ for all diffusing species.

3. While $t \leq t_{\text{final}}$

   - Reset $t_S = \min_i\{\tau_{D_i} \cdot n_i\}$ for all diffusing species.

   - Reset $t_{\text{old}} = t$.

   - For each cell, use the Gillespie algorithm to simulate reactions.

   (a) While $t \leq t_S$

   Calculate τ_R using Eq. 3.3.

   - If $t \leq t_S$, find which reaction takes place within τ_R using Eq. 3.3.

     Update number of species and time:

   $$\mathbf{x} \leftarrow \mathbf{x} + \boldsymbol{\nu}_j, \qquad t \leftarrow t + \tau_R \qquad \text{(B.3)}$$

     where $\boldsymbol{\nu}_j$ is defined as the change of number of molecules.

   - Else; do not update the state vector $\mathbf{x}$ since no reaction has occured.

end *while*

(b) Reset $t = t_{\text{old}}$ for the next cell.

end *for*

- Use the cellular automata to diffuse the species.

- Reset $n_i \leftarrow n_i + 1$ for the diffused species.

- Set $t = t_S$.

end *while*

Bibliography

[1] Erwin SCHRÖDINGER, *What is Life?: With Mind and Matter and Autobiographical Sketches*, Cambridge University Press, 1992.

[2] Michael B. Elowitz, Arnold J. Levine, Eric D. Siggia, and Peter S. Swain, "Stochastic gene expression in a single cell," *Science*, vol. 297, no. 5584, pp. 1183–1186, 08 2002.

[3] Harley H. McAdams and Adam Arkin, "Stochastic mechanisms in gene expression," *Proceedings of the National Academy of Sciences*, vol. 94, no. 3, pp. 814–819, 02 1997.

[4] Oleg Yarchuk, Nathalie Jacques, Jean Guillerez, and Marc Dreyfus, "Interdependence of translation, transcription and mrna degradation in the lacz gene," *Journal of Molecular Biology*, vol. 226, no. 3, pp. 581–596, 8 1992.

[5] Yoon-Dong Yu, Yoonjoo Choi, Yik-Ying Teo, and Andrew R. Dalby, "Developing stochastic models for spatial inference: Bacterial chemotaxis," *PLoS ONE*, vol. 5, no. 5, 05 2010.

[6] Gabriel Amselem, Matthias Theves, Albert Bae, Eberhard Bodenschatz, and Carsten Beta, "A stochastic description of ¡italic¿dictyostelium¡/italic¿ chemotaxis," *PLoS ONE*, vol. 7, no. 5, 05 2012.

[7] D. A. Lauffenburger, "Measurement of phenomenological parameters for leukocyte motility and chemotaxis.," *Agents Actions Suppl*, vol. 12, pp. 34–53, 1983.

[8] U. S. Bhalla, P. T. Ram, and R. Iyengar, "Map kinase phosphatase as a locus of flexibility in a mitogen-activated protein kinase signaling network," *Science*, vol. 297, no. 5583, pp. 1018–1023, 2002.

[9] M. R. Maurya and S. Subramaniam, "A kinetic model for calcium dynamics in raw 264.7 cells: 1. mechanisms, parameters, and subpopulational variability," *Biophysical Journal*, vol. 93, pp. 709–728, 2007.

[10] Jürgen Pahle Ursula Kummer, Borut Krajnc, Anne K. Green, C. Jane Dixon, and Marko Marhl, "Transition from stochastic to deterministic behavior in calcium oscillations," *Biophysical Journal*, vol. 89, pp. 1603–1611, 2005.

[11] Daniel T. Gillespie, "Exact stochastic simulation of coupled chemical reactions," *The Journal of Physical Chemistry*, vol. 81, no. 25, pp. 2340–2361, 12 1977.

[12] Yang Cao, Daniel T. Gillespie, and Linda R. Petzold, "Efficient step size selection for the tau-leaping simulation method," *The Journal of Chemical Physics*, vol. 124, no. 4, pp. 044109–11, 01 2006.

[13] Vassilios Sotiropoulos, Marrie-Nathalie Contou-Carrere, Prodromos Daoutidis, and Yiannis N. Kaznessis, "Model reduction of multiscale chemical langevin equations: A numerical case study," *IEEE/ACM Trans. Comput. Biol. Bioinformatics*, vol. 6, no. 3, pp. 470–482, 2009.

[14] Jeroen S. van Zon and Pieter Rein ten Wolde, "Green's-function reaction dynamics: A particle-based approach for simulating biochemical networks in time and space," *J. Chem. Phys.*, vol. 123, no. 23, pp. 234910, 2005.

[15] Steven S Andrews and Dennis Bray, "Stochastic simulation of chemical reactions with spatial resolution and single molecule detail," *Phys. Biol.*, vol. 1, no. 3, pp. 137–151, 2004.

[16] Johan Hattne, David Fange, and Johan Elf, "Stochastic reaction-diffusion simulation with mesord," *Bioinformatics*, vol. 21, no. 12, pp. 2923–2924, 2005.

[17] Maciej Dobrzyński, Jordi Vidal Rodríguez, Jaap A. Kaandorp, and Joke G. Blom, "Computational methods for diffusion-influenced biochemical reactions," *Bioinformatics*, vol. 23, no. 15, pp. 1969–1977, 2007.

[18] J. Vidal Rodríguez, Jaap A. Kaandorp, Maciej Dobrzyński, and Joke G. Blom, "Spatial stochastic modelling of the phosphoenolpyruvate-dependent phosphotransferase (pts) pathway in escherichia coli," *Bioinformatics*, vol. 22, no. 15, pp. 1895–1901, 2006.

[19] F. Baras and M. Malek Mansour, "Reaction-diffusion master equation: A comparison with microscopic simulations," *Phys. Rev. E*, vol. 54, no. 6, pp. 6139–6148, 12 1996.

[20] B. Chopard and M. Droz, *Cellular automata modeling of physical systems*, Cambrige University Press, New York, 1998.

[21] B. Novak, J. J. Tyson, B. Gyorffy, and A. Csikasz-Nagy, "Irreversible cell-cycle transitions are due to systems-level feedback," *Nat Cell Biol*, vol. 9, no. 7, pp. 724–8, 2007.

[22] Christopher V. Rao and Adam P. Arkin, "Stochastic chemical kinetics and the quasi-steady-state assumption: Application to the gillespie algorithm," *The Journal of Chemical Physics*, vol. 118, no. 11, pp. 4999–5010, 03 2003.

[23] A. Chatterjee, D. G. Vlachos, and M. A. Katsoulakis, "Binomial distribution based tau-leap accelerated stochastic simulation," *The Journal of Chemical Physics*, vol. 122, no. 2, pp. 024112, 2005.

[24] X. Cai and J. Wen, "Efficient exact and k-skip methods for stochastic simulation of coupled chemical reactions," *The Journal of Chemical Physics*, vol. 131, no. 6, pp. 064108, 2009.

[25] M. W. Chevalier and H. El-Samad, "A rigorous framework for multiscale simulation of stochastic cellular networks," *The Journal of Chemical Physics*, vol. 131, no. 5, pp. 054102, 2009.

[26] Howard Salis and Yiannis Kaznessis, "Accurate hybrid stochastic simulation of a system of coupled chemical or biochemical reactions," *The Journal of Chemical Physics*, vol. 122, no. 5, pp. 054103, 2005.

[27] D. Barik, M. R. Paul, W. T. Baumann, Y. Cao, and J. J. Tyson, "Stochastic simulation of enzyme-catalyzed reactions with disparate timescales," *BIO-PHYSICAL JOURNAL*, vol. 95, no. 8, pp. 3563–3574, OCT 15 2008.

[28] R. Ramaswamy and I. F. Sbalzarini, "A partial-propensity variant of the composition-rejection stochastic simulation algorithm for chemical reaction networks," *The Journal of Chemical Physics*, vol. 132, no. 4, pp. 044102, 2010.

[29] B. Mélykúti, K. Burrage, and K. C. Zygalakis, "Fast stochastic simulation of biochemical reaction systems by alternative formulations of the chemical langevin equation," *The Journal of Chemical Physics*, vol. 132, no. 16, pp. 164109, 2010.

[30] Xiaodong Cai, "Exact stochastic simulation of coupled chemical reactions with delays," *The Journal of Chemical Physics*, vol. 126, no. 12, pp. 124108, 2007.

[31] J. M. Frazier, Y. Chushak, and B. Foy, "Stochastic simulation and analysis of biomolecular reaction networks," *BMC SYSTEMS BIOLOGY*, vol. 3, pp. 64, JUN 17 2009.

[32] D. T. Gillespie, Y. Cao, K. R. Sanft, and L. R. Petzold, "The subtle business of model reduction for stochastic chemical kinetics," *The Journal of Chemical Physics*, vol. 130, no. 6, pp. 064103, 2009.

[33] M. Rathinam, P. W. Sheppard, and M. Khammash, "Efficient computation of parameter sensitivities of discrete stochastic chemical reaction networks," *The Journal of Chemical Physics*, vol. 132, no. 3, pp. 034103, 2010.

[34] R. L. Karcher, J. T. Roland, F. Zappacosta, M. J. Huddleston, R. S. Annan, S. A. Carr, and V. I. Gelfand, "Cell cycle regulation of myosin-v by calcium/calmodulin-dependent protein kinase ii," *Science*, vol. 293, no. 5533, pp. 1317–20, 2001.

[35] K. Bedard, E. Szabo, M. Michalak, and M. Opas, "Cellular functions of endoplasmic reticulum chaperones calreticulin, calnexin, and erp57," *Int Rev Cytol*, vol. 245, pp. 91–121, 2005.

[36] S. Johnson, M. Michalak, M. Opas, and P. Eggleton, "The ins and outs of calreticulin: from the er lumen to the extracellular space," *Trends Cell Biol*, vol. 11, no. 3, pp. 122–9, 2001.

[37] M. Michalak, P. Mariani, and M. Opas, "Calreticulin, a multifunctional ca2+ binding chaperone of the endoplasmic reticulum," *Biochem Cell Biol*, vol. 76, no. 5, pp. 779–85, 1998.

[38] Y. Kim, J. S. Moon, K. S. Lee, S. Y. Park, J. Cheong, H. S. Kang, H. Y. Lee, and H. D. Kim, "Ca2+/calmodulin-dependent protein phosphatase calcineurin mediates the expression of inos through ikk and nf-kappab activity in lps-stimulated mouse peritoneal macrophages and raw 264.7 cells," *Biochem Biophys Res Commun*, vol. 314, no. 3, pp. 695–703, 2004.

[39] T. Hofer, L. Venance, and C. Giaume, "Control and plasticity of intercellular calcium waves in astrocytes: a modeling approach," *J Neurosci*, vol. 22, no. 12, pp. 4850–9, 2002.

[40] J. Wagner and J. Keizer, "Effects of rapid buffers on ca2+ diffusion and ca2+ oscillations," *Biophys J*, vol. 67, no. 1, pp. 447–56, 1994.

[41] G. Lemon, W. G. Gibson, and M. R. Bennett, "Metabotropic receptor activation, desensitization and sequestration-i: modelling calcium and inositol 1,4,5-trisphosphate dynamics following receptor activation," *Journal of Theoretical Biology*, vol. 223, no. 1, pp. 93–111, 2003.

[42] M. R. Maurya and S. Subramaniam, "A kinetic model for calcium dynamics in raw 264.7 cells: 2. knockdown response and long-term response," *Biophys J*, vol. 93, pp. 729–740, 2007.

[43] M. J. Berridge, M. D. Bootman, and H. L. Roderick, "Calcium signalling: dynamics, homeostasis and remodeling," *Nat Rev Mol Cell Biol*, vol. 4, no. 7, pp. 517–29, 2003.

[44] M. J. Berridge, "Calcium oscillations," *J Biol Chem*, vol. 265, no. 17, pp. 9583–6, 1990.

[45] T. Meyer and L. Stryer, "Molecular model for receptor-stimulated calcium spiking," *Proc Natl Acad Sci U S A*, vol. 85, no. 14, pp. 5051–5, 1988.

[46] E. G. Lakatta, T. M. Vinogradova, and V. A. Maltsev, "The missing link in the mystery of normal automaticity of cardiac pacemaker cells," *Ann N Y Acad Sci*, vol. 1123, pp. 41–57, 2008.

[47] J. Mishra and U. S. Bhalla, "Simulations of inositol phosphate metabolism and its interaction with insp(3)-mediated calcium release," *Biophysical Journal*, vol. 83, no. 3, pp. 1298–1316, 2002.

[48] T. Doi, S. Kuroda, T. Michikawa, and M. Kawato, "Inositol 1,4,5-trisphosphate-dependent ca2+ threshold dynamics detect spike timing in cerebellar purkinje cells," *J Neurosci*, vol. 25, no. 4, pp. 950–61, 2005.

[49] J. L. Greenstein, R. Hinch, and R. L. Winslow, "Mechanisms of excitation-contraction coupling in an integrative model of the cardiac ventricular myocyte," *Biophys J*, vol. 90, no. 1, pp. 77–91, 2006.

[50] J. L. Greenstein and R. L. Winslow, "An integrative model of the cardiac ventricular myocyte incorporating local control of ca2+ release," *Biophys J*, vol. 83, no. 6, pp. 2918–45, 2002.

[51] G. R. Van Den Brink, S. M. Bloemers, B. Van Den Blink, L. G. Tertoolen, S. J. Van Deventer, and M. P. Peppelenbosch, "Study of calcium signaling in non-excitable cells," *Microsc Res Tech*, vol. 46, no. 6, pp. 418–33, 1999.

[52] T. F. Wiesner, B. C. Berk, and R. M. Nerem, "A mathematical model of cytosolic calcium dynamics in human umbilical vein endothelial cells," *American Journal of Physiology-Cell Physiology*, vol. 39, no. 5, pp. C1556–C1569, 1996.

[53] C. C. Fink, B. Slepchenko, II Moraru, J. Watras, J. C. Schaff, and L. M. Loew, "An image-based model of calcium waves in differentiated neuroblastoma cells," *Biophys J*, vol. 79, no. 1, pp. 163–83, 2000.

[54] S. Rudiger, J. W. Shuai, W. Huisinga, C. Nagaiah, G. Warnecke, I. Parker, and M. Falcke, "Hybrid stochastic and deterministic simulations of calcium blips," *Biophys J*, vol. 93, no. 6, pp. 1847–57, 2007.

[55] M. S. Jafri, J. J. Rice, and R. L. Winslow, "Cardiac ca2+ dynamics: the roles of ryanodine receptor adaptation and sarcoplasmic reticulum load," *Biophys J*, vol. 74, no. 3, pp. 1149–68, 1998.

[56] M. Puceat and M. Jaconi, "Ca2+ signalling in cardiogenesis," *Cell Calcium*, vol. 38, no. 3-4, pp. 383–9, 2005.

[57] J. M. Raser and E. K. O'Shea, "Noise in gene expression: origins, consequences, and control," *Science*, vol. 309, no. 5743, pp. 2010–3, 2005.

[58] S. Schuster, M. Marhl, and T. Hofer, "Modeling of simple and complex calcium oscillations, from single-cell responses to intercellular signaling," *Eur. J. Biochem.*, vol. 269, pp. 1333–1355, 2002.

[59] Y. X. Li and J. Rinzel, "Equations for insp3 receptor-mediated [ca2+]i oscillations derived from a detailed kinetic model: a hodgkin-huxley like formalism," *J Theor Biol*, vol. 166, no. 4, pp. 461–73, 1994.

[60] M. R. Maurya and S. Subramaniam, "Supporting material to: A kinetic model for calcium dynamics in raw 264.7 cells: 1. mechanisms, parameters, and subpopulational variability," *Biophysical Journal*, vol. 93, pp. 709–728, 2007.

[61] Desmond J. Higham, "Modeling and simulating chemical reactions," *SIAM Rev.*, vol. 50, pp. 347–368, 2008.

[62] Eric L. Haseltine and James B. Rawlings, "Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics," *The Journal of Chemical Physics*, vol. 117, no. 15, pp. 6959–6969, 10 2002.

[63] E. L. Haseltine and J. B. Rawlings, "On the origins of approximations for stochastic chemical kinetics," *Journal of Chemical Physics*, vol. 123, no. 16, OCT 22 2005.

[64] M. A. Gibson and J. Bruck, "Efficient exact stochastic simulation of chemical systems with many species and many channels," *Journal of Physical Chemistry A*, vol. 104, no. 9, pp. 1876–1889, MAR 9 2000.

[65] K. Vasudeva and U.S. Bhalla, "Adaptive stochastic-deterministic chemical kinetic simulations," *Bioinformatics*, vol. 20, no. 1, pp. 78–84, JAN 1 2004.

[66] U. S. Bhalla, "Signaling in small subcellular volumes. i. stochastic and diffusion effects on individual pathways," *Biophysical Journal*, vol. 87, no. 2, pp. 733–744, AUG 2004.

[67] G. W. DeYoung and J. Keizer, "A single-pool inositol 1,4,5-trisphosphate-receptor-based model for agonist-stimulated oscillations in ca2+ concentration," *Proc. Natl. Acad. Sci. USA*, vol. 89, no. 20, pp. 9895–9899, OCT 15 1992.

[68] E. A. Mastny, E. L. Haseltine, and J. B. Rawlings, "Two classes of quasi-steady-state model reductions for stochastic kinetics," *Journal of Chemical Physics*, vol. 127, no. 9, SEP 7 2007.

[69] G. Srinivasan, D. M. Tartakovsky, B. A. Robinson, and A. B. Aceves, "Quantification of uncertainty in geochemical reactions," *Water Resour. Res.*, vol. 43, pp. W12415, 2007.

[70] TaiJung Choi, Mano Ram Maurya, Daniel M. Tartakovsky, and Shankar Subramaniam, "Stochastic hybrid modeling of intracellular calcium dynamics," *J. Chem. Phys.*, vol. 133, no. 16, pp. 165101, 2010.

[71] Juan M. Pedraza and Alexander van Oudenaarden, "Noise propagation in gene networks," *Science*, vol. 307, no. 5717, pp. 1965–1969, 2005.

[72] Michael A. Gibson and Jehoshua Bruck, "Efficient exact stochastic simulation of chemical systems with many species and many channels," *J. Phys. Chem.*, vol. 104, no. 9, pp. 1876–1889, 02 2000.

[73] Karen Lipkow, Steven S. Andrews, and Dennis Bray, "Simulated diffusion of phosphorylated chey through the cytoplasm of escherichia coli," *J. Bacteriology*, vol. 187, no. 1, pp. 45–53, 1 2005.

[74] Jim Douglas and James E. Gunn, "A general formulation of alternating direction methods," *Numerische Mathematik*, vol. 6, no. 1, pp. 428–453, 1964.

[75] William E. Schiesser, *The Numerical Method of Lines: Integration of Partial Differential Equations*, Academic Press, San Diego, 1991.

[76] Willem Hundsdorfer and Jan G. Verwer, *Numerical Solution of Time-Dependent Advection-Diffusion-Reaction Equations*, Springer, New York, 2003.

[77] N. G. Van Kampen, *Stochastic processes in physics and chemistry*, Elsevier, San Diego, 2007.

[78] Julius Adler and Wung-Wai Tso, "Decision-making in bacteria: Chemotactic response of escherichia coli to conflicting stimuli," *Science*, vol. 184, no. 4143, pp. 1292–1294, 06 1974.

[79] TaiJung Choi, Mano Ram Maurya, Daniel M. Tartakovsky, and Shankar Subramaniam, "Stochastic operator-splitting method for reaction-diffusion systems," *J. Chem. Phys.*, vol. 137, 2012.

[80] Wayne K Stadelmann, Alexander G Digenis, and Gordon R Tobin, "Physiology and healing dynamics of chronic cutaneous wounds," *The American Journal of Surgery*, vol. 176, no. 2, Supplement 1, pp. 26S–38S, 8 1998.

[81] Kim S. Midwood, Leyla Valenick Williams, and Jean E. Schwarzbauer, "Tissue repair and the dynamics of the extracellular matrix," *The International Journal of Biochemistry & Cell Biology*, vol. 36, no. 6, pp. 1031–1037, 6 2004.

[82] "inflammation: The leukocyte adhesion cascade," .

[83] Eduard Ryschich, Vachtang Kerkadze, Paulius Lizdenis, Saulius Paskauskas, Hanns-Peter Knaebel, Wolfgang Gross, Martha Maria Gebhard, Markus W. Büchler, and Jan Schmidt, "Active leukocyte crawling in microvessels assessed by digital time-lapse intravital microscopy," *Journal of Surgical Research*, vol. 135, no. 2, pp. 291–296, 10 2006.

[84] Shana Marmon, Michael Cammer, Cedric S. Raine, and Michael P. Lisanti, "Transcellular migration of neutrophils is a quantitatively significant pathway across dermal microvascular endothelial cells," *Experimental Dermatology*, vol. 18, no. 1, pp. 88–90, 2009.

[85] Muller William A., "Leukocyte–endothelial-cell interactions in leukocyte transmigration and the inflammatory response," *Trends in Immunology*, vol. 24, no. 6, pp. 326–333, 6 2003.

[86] S. Natarajan, Diane Williamson, Amanda J. Stiltz, and Keith Harding, "Advances in wound care and healing technology," *American Journal of Clinical Dermatology*, vol. 1, no. 5, 2000.

[87] Sneyd James Keener, James, *Mathematical Physiology, I: Cellular Physiology*, Springer, 2009.

[88] Brian E. Farrell, Ronald P. Daniele, and Douglas A. Lauffenburger, "Quantitative relationships between single-cell and cell-population model parameters for chemosensory migration responses of alveolar macrophages to c5a," *Cell Motility and the Cytoskeleton*, vol. 16, no. 4, pp. 279–293, 1990.

[89] John Goutsias, "Quasiequilibrium approximation of fast reaction kinetics in stochastic biochemical systems," *The Journal of Chemical Physics*, vol. 122, no. 18, pp. 184102–15, 05 2005.

[90] Yang Cao, Daniel T. Gillespie, and Linda R. Petzold, "Avoiding negative populations in explicit poisson tau-leaping," *The Journal of Chemical Physics*, vol. 123, no. 5, pp. 054104, 2005.