

# Particle-Based Stochastic Reaction-Diffusion Models in Biology

Daniel Coombs (University of British Columbia),  
Mark Flegg (Monash University),  
Samuel Isaacson (Boston University),  
Per Lötstedt (Uppsala University),  
Linda Petzold (University of California Santa Barbara)

November 9–14, 2014

## 1 Overview of the Field

Throughout the twentieth century, biological and medical research was centered within the laboratory. The development of detailed mathematical models of cellular systems was hindered by a need for quantitative measurements of cellular processes and insufficient computer resources. In recent years, the combination of new experimental techniques and exponential improvements in computer technology has made the construction and study of such models possible. These mathematical models have the potential to provide biological and medical research with predictive power. A quantitative scientific paradigm is developing in which the theoretical predictions of mathematical models provide complementary tools that can help guide laboratory experiments. Such is the importance of simulation techniques to the future of fields including molecular biology, cellular biology and systems biology that a full quantitative numerical model of a cell has been labeled as one of the “grand challenges of the 21st century”[24].

Mathematical models have been used during the last decade to model a number of biological processes, including gene regulation, molecular signaling, cell division and molecular transport. The development of these models has been motivated by the aforementioned advances in experimental techniques, which have revealed that many processes in cell and molecular biology are inherently stochastic [6]. This stochasticity can provide a challenge that must be compensated for to allow proper function, such as the inherent noisiness of gene expression [19]. Stochasticity can also drive biological function, facilitating phenotypic diversity of cellular populations and driving evolution.

At the scale of a single cell, stochasticity becomes important due to low copy numbers of biological molecules, such as mRNA and proteins, that take part in bio-chemical reactions driving cellular processes [19]. When trying to describe such biological processes traditional (mean-field, coarse-grained) deterministic models for the concentration of chemical species, given by systems of ordinary differential equations (ODEs) or partial differential equations (PDEs), are often inadequate, exactly because of these low copy numbers. Stochastic models are necessary to account for small particle numbers (intrinsic noise) and extrinsic noise sources. Often the complexity of these models depends crucially on several key bio-chemical reactions, and whether they are diffusion-limited or reaction-limited. In the latter case, one can describe the processes by adopting the framework of Markov jump processes [10] and stochastic differential equations (SDEs) [1] (chemical master and Fokker Planck equations), while in the former one needs to adopt the framework of

stochastic reaction-diffusion models, including the lattice reaction-diffusion master equation (RDME) [14], stochastic partial differential equations (SPDEs) and particle-based Brownian dynamics models [7].

Many fundamental biological processes involve spatially distributed components in which diffusion-limited reactions play a key role. Examples include the diffusion of proteins within microdomains and lipid bilayers; the propagation of trans-membrane potentials across cellular membranes; the movement of cells within tumors; and the transmission of chemical signals from the cell surface to regulatory sites in the nucleus. As such, models that explicitly resolve the spatial transport of molecules within cells can help to understand the influence of subcellular structure, spatially distributed signaling compartments, and cell shape on such processes [13, 16, 17]. In recent years, particle-based stochastic reaction-diffusion models have been used to develop mathematical models which both account for noise in the chemical reaction process, and incorporate the explicit diffusion and reaction of proteins and mRNAs within cells. There are two primary particle-based models that have been used: variants of the diffusion limited reaction model due to Smoluchowski [21, 15], and the reaction-diffusion master equation (RDME) [9].

In the former molecules are modeled as points or spheres undergoing Brownian motion. First order reactions of the form  $A \rightarrow B$ , are modeled as basic Poisson processes, with each molecule of type A having a probability per time to convert into a molecule of species B. When considering second order reactions of the form  $A + B \rightarrow C$  several variants of the Smoluchowski model have been proposed. In the original, pure-adsorption model [21] two molecules react instantly upon reaching a fixed separation, a model parameter called the reaction-radius. Depending on the underlying biological system being studied, this parameter will be the sum of the physical “radii” of the two molecules, or can be an empirical parameter chosen to fit certain experimentally determined statistics of the reaction process [2]. A more general partial-adsorption model then allows for the possibility of non-reactive encounters, where two reactants may sometimes reflect off each other when their separation equals the reaction radius. Finally, a third variant popularized by Doi posits that the two reactants may react with a specified probability per unit time when their separation is less than a reaction radius. Mathematically, the two Smoluchowski reaction models lead to Dirichlet and Robin boundary conditions in the associated Forward Kolmogorov equations describing the evolution of the probability density for a chemical system to be in a specified state. In contrast, the Doi model leads to interaction functions within the associated Forward Kolmogorov equation.

In contrast to the Smoluchowski type models, in the RDME space is discretized into a lattice, with molecular diffusion approximated by continuous-time random walks on the lattice. First order reactions are modeled identically to the Smoluchowski models, while bimolecular reactions occur with fixed probabilities per unit time for two reactants located at the same lattice site. The Smoluchowski model is often considered a more microscopic model, with the RDME interpreted as a mesoscopic scale coarse graining of Smoluchowski type models.

There has been a large body of work in the last decade to develop numerical methods capable of accurately and efficiently simulating the stochastic processes described by these models. The development of these numerical methods has been motivated largely by specific biological systems studied within individual research groups. As a result, a wide variety of numerical methods have been developed to simulate specific biological systems, with a wide variety of extensions of the base models. These extensions allow for more realistic biological features, including complex domain geometries; spatial transport by advection-diffusion or drift-diffusion; non-elementary chemical reaction mechanisms; and molecular crowding.

Broadly speaking, there are three general classes of numerical methods that have been widely studied. The First Passage Kinetic Monte Carlo (FPKMC) [18, 4] and Green’s Function Reaction Dynamics (GFRD) [25, 23] methods can generate exact realizations of the stochastic process described by the Smoluchowski model, while time-step based Brownian Dynamics (BD) methods approximate this process [2, 8]. The Gillespie method [10] can be used to generate exact realizations of the stochastic process described by the RDME. Whilst a great deal of literature exists on the analysis of these methods, there are few detailed comparisons of their relative merits, and no ‘gold standard’ has been established. Part of the reason why no gold standard approach has been identified in the literature is the disjointed, interdisciplinary nature of the research. Understanding the usefulness, efficiency and accuracy of the various algorithms requires careful studies by computer scientists, mathematicians, physicists, chemists and biologists, who have had limited opportunities for collaboration in this space. Another reason why no gold standard has been established is because the various approaches are dichotomous and difficult to compare objectively.

Typically, numerical methods based on the Smoluchowski model allow for high precision since each

molecule/particle is represented separately and traced exactly in space. The cost of such high precision is the requirement that each reaction/interaction must be simulated separately. Thus, the computational cost of these methods grow rapidly as the number of molecules/particles (and individual trajectories) increase, and also as the number of possible reactant combinations increase. Once the number of molecules within a domain become sufficiently large (or concentrated), the computational cost of Smoluchowski-based numerical methods can be prohibitively high. RDME-based Gillespie methods, on the other hand, are generally more computationally efficient at high particle counts as only the copy number of molecules within each compartment is tracked; the position of individual molecules within each compartment are not distinguished. Despite their computational efficiency, Gillespie methods may introduce unphysical computational artifacts. This is due to the discretization of the domain into a lattice, random walk approximations of the underlying diffusion processes, and choice of mechanism for modeling bimolecular reactions. It is difficult to objectively value accuracy over efficiency or vice versa to model problems as diverse as those in biology. Some biological systems may fluctuate between low concentrations and high concentrations in both space and time and to achieve acceptable levels of accuracy using current technology and computational resources it is important to improve our understanding of how these approaches may be hybridized, taking the best of both approach.

True to the rather fractured/diverse nature of the community, a wide variety of publicly available simulation packages have been developed to facilitate the study of cellular processes, each based on a different variant of the aforementioned methods. Some of the more popular programs include Smoldyn [2], SpatioCyte [3], URDME [5], MesoRD [11], STEPS [12], Lattice Microbes [20], MCell [22] and GFRD [25]. At the BIRS workshop, each of these simulation methodologies were represented and open problems were discussed.

## 2 Recent Developments and Open Problems

We highlight four broad areas, corresponding to the four main session topics of the workshop:

1. **Biological Models and Analytical Approximations:** In the last 10-20 years, a broad body of work has been developed to understand the dynamics of non-spatial models for biological systems. Here the underlying mathematical models are typically given by ODEs, SDEs, or integer-valued jump processes. For such systems we have gained a deep understanding of the types of dynamical behavior that can occur and how these dynamics depend on critical system parameters and network topology. For example, the possibility of multi-stability in deterministic ODE models for cellular processes can now be predicted in many cases from network topology. More recently such results have begun to be adapted to the stochastic regime, where deterministic chemical concentrations are replaced by jump processes for the number of molecules of each chemical species.

In contrast to these well-mixed models, in spatially extended models the types of dynamical behaviors that can arise, and methods for predicting them are much less developed. Moreover, there are few general results on how explicitly accounting for spatial transport in biochemical systems might influence the predicted behavior of previously non-spatial models. One reason such questions are less well-developed in the context of spatial models is due to a much reduced set of well-developed tools for the general analysis of these systems. For stochastic jump-process models of chemical systems we have a large collection of exact and approximate techniques for analytically investigating system behavior. These include exact transform and solution methods for simple systems, network-theory based approaches, and asymptotic methods based on separation of timescales. In particular, a rigorous hierarchy of mathematical models have been established to show how “large volume” expansions can lead from stochastic jump process models, to SDE models, and ultimately to deterministic ODE models. In contrast, no similar complete and rigorous theory has been developed for spatially extended particle-based reaction-diffusion models.

One area where analysis of particle-based reaction-diffusion models has made substantial progress in recent years is in estimating timescales for various reaction processes. A large body of asymptotic methods have been developed to estimate mean first passage times across a variety of particle-based reaction-diffusion models. One general theme in the development of these methods is that the “reaction-radius” at which two proteins may interact is generally a very small parameter relative to the

typical size of cellular domains, and the typical distance between reactive molecules. These methods have provided insight into a variety of biological problems, from how proteins search for DNA binding sites through to how molecules are trafficked to receptors in synaptic spines.

2. **Mesosopic Methods and Modeling:** While quite popular for modeling biological systems, there are several practical challenges in using lattice reaction-diffusion master equation (RDME) models. One difficulty arises from the underlying discretization of space, which introduces a purely numerical parameter, the lattice spacing, into RDME-based models. While the RDME can be shown to converge back to the Smoluchowski continuous-space particle model for systems with only linear reactions, bimolecular reactions are lost in the RDME as the lattice spacing is taken to zero. This difficulty arises as molecules are treated as point particles, that may only react when located at the same lattice site within the RDME. As the lattice spacing is taken to zero one obtains point particles moving by Brownian motion, which can only react when located at the same point. Since the probability for two diffusing molecules to ever have the same position is zero, bimolecular reactions are lost. To overcome this challenge the RDME is usually postulated to only be an appropriate model for lattice spacings that are sufficiently large that bimolecular reactions can be adequately resolved, while also sufficiently small that the approximation of Brownian as a continuous-time random walk on the lattice is accurate. As discussed in the workshop, a number of groups have recently investigated over what range of lattice spacings this approximation is robust, how to broaden the range of lattice spacings over which the RDME provides an accurate approximation to continuous-space particle models, and how to modify the RDME model to provide a similar lattice-based stochastic reaction-diffusion model that actually converges back to continuous-space models as the lattice spacing is taken to zero.

A second major challenge in the use of RDME-based models is in how to efficiently generate realizations of the associated stochastic process. The well-known Gillespie method (i.e. stochastic simulation algorithm or kinetic Monte Carlo method) can generate exact realizations of the stochastic process associated with the RDME [14]. While there is an enormous body of work developing optimized versions of the Gillespie method for well-mixed chemical systems, there are only a few methods that have been optimized for the RDME. Here the simulation of spatial movement of molecules as continuous-time random walks is usually the dominant computational cost, often requiring more than 90% of simulation time. In recent years several methods have been proposed to help optimize this component of RDME simulations, including methods that approximate the random walks of individual molecules by only resolving the net motion of molecules between two lattice sites, by replacing the random walks with various continuous diffusion methods, by grouping together many diffusive hops into one event, and by optimized data structures to efficiently sample and update the location and time of the next diffusive hopping event.

3. **Particle-Based Methods and Modeling:** Timestep based Brownian Dynamics (BD) methods are one of the most popular stochastic reaction-diffusion methods for studying biological systems. In their most basic formulations these methods provide first-order (in the timestep) approximations to the underlying particle dynamics that are relatively easy to implement, but can be computationally demanding when taking small timesteps to adequately resolve the dynamics of cellular processes. To overcome this challenge a number of improved BD methods have been proposed in recent years to facilitate the use of large timesteps, while still accurately resolving important statistics of the underlying stochastic reaction-diffusion processes [8].

In sufficiently dilute systems the First Passage Kinetic Monte Carlo Method was recently developed to provide exact realizations of the stochastic process associated with the Smoluchowski model [4]. While this method provides substantial computational savings for dilute systems in simplified geometries, it is not competitive with BD methods for more dense systems in which molecules are tightly packed, or when the domain geometry is tortuous. As presented at the workshop, a number of groups have recently developed variants of the method that allow for more general methods of spatial transport than diffusion, while sacrificing the exactness of the underlying simulation algorithm.

4. **Multiscale Methods and Modeling:** Current particle-based stochastic reaction-diffusion methods can accurately simulate hundreds of thousands to perhaps millions of reacting and diffusing molecules over timescales of several hours. To achieve the simulation of cellular processes that occur over longer

timescales, such as the cell cycle, and over length scales of mammalian cells, multiscale methods are needed to automatically resolve model components at the coarsest acceptable level of physical resolution. In recent years researchers have begun to investigate how to achieve such couplings, proposing a number of methods to allow for multiscale couplings across interfaces separating different regions of space, or allow for resolving different reactions or chemical species at different physical scales. Examples of the former include coupling across an interface a particle model in one region of a domain to deterministic reaction-diffusion PDE models in a second region, or similar couplings between microscopic Smoluchowski models and mesoscopic RDME models. For the latter, several groups have investigated how to accurately resolve different components of a given physical system at different scales in a manner that is consistent with a fully particle-based model.

### 3 Presentation Highlights

The workshop brought together theorists and experimentalists investigating spatially distributed stochastic processes in cell and molecular biology, with researchers developing techniques for the analysis of mathematical models and related numerical methods. One particular emphasis was models involving spatial transport and chemical reactions, which often span multiple time and length scales. Applications to biological processes such as gene expression, immunological receptor signaling, and signal transduction were discussed to illustrate the types of biological problems for which accurate and efficient stochastic spatio-temporal methods are needed.

Below we highlight, in the presenter's own words, several talks from each of the workshop's four sessions:

#### **Session on Biological Models and Analytical Approximations:**

Speaker: **Jun Allard** (University of California Irvine)

Title: *Mechanics of diffusing surface molecules modulates T cell receptor sensitivity*

Abstract: Receptors on the surface of cells control many cellular processes. An important class of receptors, e.g., T-cell receptors, attach to molecules that are anchored to other cells or surfaces, and remain poorly understood. The T-cell receptor complex spans 15 nanometers, while other nearby molecules spans  $\sim 40$  nanometers. Since all these molecules are mobile on the two-dimensional cell surface, the size differential has been proposed to lead to spatial segregation (mediated by the mechanical properties of the cell membranes) that triggers immune signaling. I will present a nanometer-scale mathematical model that couples membrane elasticity with compressional resistance and lateral mobility of molecules. We find robust supradiffusive segregation. The model predicts a time-dependent tension on the receptor leading to a nonlinearity which could enhance the receptors ability to make precise immune decisions. Understanding the full life-cycle of receptor dynamics raises questions involving surface diffusion of a population of molecules, and I will present open problems along with computational estimates and their biological importance.

Speaker: **Daniel Coombs** (University of British Columbia)

Title: *Random violence: A stochastic approach to cell cytotoxicity*

Abstract: This talk is about our recent work on the delivery of effector molecules from immune cells such as T and Natural Killer cells. These cells release fairly small numbers of molecules that induce cell death, into the tightly defined region of contact with a target cell (the "immunological synapse"). I will explain the background biology and the leading hypothesis of how this process works. I will then show how we used a spatial stochastic algorithm to analyze whether the hypothesis is correct and outline future experimental work. This research was done jointly with Daniel Woodsworth (BC Cancer Research Centre).

Speaker: **Andreas Hellander** (Department of Information Technology, Uppsala University)

Title: *Accuracy of the Michaelis-Menten approximation when analyzing effects of molecular noise*

Quantitative biology relies on the construction of accurate mathematical models, yet the effectiveness of these models is often predicated on making simplifying approximations that allow for direct comparisons with available experimental data. The Michaelis-Menten approximation is widely used in both deterministic and discrete stochastic models of intracellular reaction networks, due to the ubiquity of enzymatic activity in cellular processes and the clear biochemical interpretation of its parameters. However, it is not well understood how the approximation applies to the discrete stochastic case or how it extends to spatially inhomogeneous

systems. We study the behavior of the discrete stochastic Michaelis-Menten approximation as a function of system size and show that significant errors can occur for small volumes, in comparison with a corresponding mass action system. We then explore some consequences of these results for quantitative modeling. One consequence is that fluctuation-induced sensitivity, or stochastic focusing, can become highly exaggerated in models that make use of Michaelis-Menten kinetics even if the approximations are excellent in a deterministic model. Another consequence is that spatial stochastic simulations based on the reaction-diffusion master equation can become highly inaccurate if the model contains Michaelis-Menten terms.

#### **Session on Mesoscopic Methods and Modeling:**

Speaker: **Stefan Engblom** (Uppsala university)

Title: *Mesoscopic Stochastic Modeling: Diffusion Operators, Multiphysics Couplings, and Convergence*

Abstract: In this talk I will discuss stochastic modeling in the reaction-transport framework from various viewpoints. I shall initially be concerned with diffusion-controlled reactions targeting applications mainly in molecular cell biology. I will briefly review the basic setup and conditions for the validity of this type of modeling. In particular I will discuss the properties of the diffusion transport operator.

I will next discuss an application example from outside the diffusion-controlled domain, namely an approach towards multiphysics modeling of neuronal spiking activity affected by stochastic channel fluctuations. This example serves as a reminding illustration that questions of convergence are not that straightforward to answer.

Speaker: **Anastasios Matzavinos** (Brown University)

Title: *Dissipative particle dynamics simulations of polymer networks*

Abstract: Networks of entangled or cross-linked polymers, such as the actin cytoskeleton, are ubiquitous in phenomena pertaining to cellular and molecular biology. In many cases, the structure of these networks is dynamically altered by the mechanical feedback of biological lipid membranes and cytoplasmic flows. However, current modeling and computational approaches neglect such mechanical feedbacks for the sake of computational tractability.

In this talk, we present a *dissipative particle dynamics* approach to simulating the meso-scale dynamics of polymer networks. Our simulations explicitly include mechanical interactions with other meso-scale structures (e.g., lipid membranes) and cytoplasmic flows. We compare the results of our approach to those of Brownian dynamics simulations. We also discuss ongoing work on stochastic homogenization, bridging the gap between the meso-scale description and macroscopic models of bulk mechanical properties.

Speaker: **Kevin Sanft** (University of Minnesota)

Title: *Scaling properties of exact simulation algorithms for spatially discretized stochastic reaction-diffusion processes*

Abstract: Stochastic reaction-diffusion processes are widely used to model biochemical systems. Discretizing the spatial domain leads to a discrete state, continuous time Markov jump process that can be described by the reaction-diffusion master equation. Solutions to the master equation are approximated by generating exact trajectories using variants of the Gillespie algorithm. The choice of simulation formulation and underlying data structures has a dramatic effect on computational efficiency. In this talk, I will show how the optimal algorithm choice depends on the number and relative timescales of the transition channels. For very large problems, memory hierarchy effects lead to scaling properties that differ from the asymptotic analysis, which influences the optimal simulation algorithm parameters.

#### **Session on Particle-Based Methods and Modeling:**

Speaker: **Steven Andrews** (Fred Hutchinson Cancer Research Center)

Title: *The Smoldyn simulator: overview, applications, and hybrid simulation*

Abstract: Smoldyn is a particle-based cell biology simulator which represents proteins or other molecules of interest as individual spheres. These particles diffuse, undergo chemical reactions with each other, and interact with membranes and other surfaces in ways that closely mimic reality. In particular, all interaction rates are quite accurate. Smoldyn is easy to use and supports a wide variety of features. Several colleagues and I recently used Smoldyn to investigate transcription factor dynamics in cell nuclei to determine what processes enable transcription factors to locate their target genes quickly. In agreement with prior results, we found that non-specific binding and then diffusion along DNA accelerates target gene finding through a

process called the antenna effect. Additionally, we found that intersegmental transfer also accelerates target gene finding; here, a transcription factor transfers directly from being non-specifically bound on one DNA segment to being non-specifically bound on an adjacent DNA segment. In separate work, Martin Robinson and I recently added adjacent-volume hybrid simulation capability to Smoldyn. Here, space is partitioned into adjacent continuum and lattice regions, which are simulated with particle-based and spatial Gillespie type methods, respectively. These enable simulations to represent high levels of detail where required but lower detail (and faster computation) elsewhere.

Speaker: **Samuel Isaacson** (Boston University)

Title: *Lattice Approximation of Spatially-Continuous Particle-Based Stochastic Reaction-Diffusion Models*

Abstract: We derive a lattice, continuous-time jump process approximation of a spatially-continuous, interaction-function based stochastic reaction-diffusion model. The new model has the benefit of treating diffusion and linear reactions in exactly the same manner as the reaction-diffusion master equation (RDME). Moreover, in the limit of coarse meshes it can be shown that the RDME approximates this model. Unlike the RDME, this new, convergent reaction-diffusion master equation model (CRDME) retains bimolecular reactions in the limit that the mesh spacing approaches zero, converging to the underlying interaction-function model. The CRDME therefore offers alternatives to both Brownian Dynamics (BD) methods for solving the interaction-function model, and coupled RDME-BD methods that attempt to overcome the loss of bimolecular reactions in the RDME for small mesh sizes.

Speaker: **Frank Noe** (Free University of Berlin)

Title: *interacting-Particle Reaction-Diffusion (iPRD) dynamics*

Abstract: In cellular signal transduction, what happens where and when? Addressing this question requires to deal with protein interactions that involve low copy numbers, precise stoichiometry, the spatiotemporal arrangement within molecular machines. While modern experimental techniques such as super-resolution microscopy are taking giant leaps towards watching cells in action with molecular resolution, computer simulation is still facing the challenge of combining physical detail with computational efficiency. Here we propose the interacting-Particle Reaction-Diffusion (iPRD) approach. iPRD is a fusion of particle-based reaction-kinetics and molecular dynamics including particle-interactions aiming at simulating cellular signal transduction with rigorous physical approach. I will present the theory and methodology, briefly sketch our ReaDDy implementation of iPRD and hint to some biological applications.

### **Session on Multiscale Methods and Modeling:**

Speaker: **Ruth Baker** (Oxford University)

Title: *Adaptive multi-level Monte Carlo methods*

Abstract: Discrete-state, continuous-time Markov models are widely used in the modelling of biochemical reaction networks. Their complexity generally precludes analytic solution, and so we rely on Monte Carlo simulation to estimate system statistics of interest. Perhaps the most widely used method is the Gillespie algorithm. This algorithm is exact but computationally complex. As such, approximate stochastic simulation algorithms such as the tau-leap algorithm are often used. Sample paths are generated by taking leaps of length tau through time and using an approximate method to generate reactions within leaps. However, tau must be held relatively small to avoid significant estimator bias and this significantly impacts on potential computational advantages of the method.

The multi-level method of Anderson and Higham tackles this problem by cleverly generating a suite of sample paths with different accuracy in order to estimate statistics. A base estimator is computed using many (cheap) paths at low accuracy. The bias inherent in this estimator is then reduced using a number of correction estimators. Each correction term is estimated using a collection of (increasingly expensive) paired sample paths where one path of each pair is generated at a higher accuracy compared to the other. By sharing randomness between these paired sample paths only a relatively small number of pairs are required to calculate each correction term.

In the original multi-level method, paths are simulated using the tau-leap technique with a fixed value of tau. This approach can result in poor performance where the reaction activity of a system changes substantially over the timescale of interest. By introducing a novel, adaptive time-stepping approach we extend the applicability of the multi-level method to such cases. In our algorithm, tau is chosen according to the stochas-

tic behaviour of each sample path. We present an implementation of our adaptive time-stepping multi-level method that, despite its simplicity, performs well across a wide range of sample problems.

Speaker: **Radek Erban** (University of Oxford)

Title: *From Molecular Dynamics to Particle-based Stochastic Reaction-Diffusion Models*

Abstract: I will discuss all-atom and coarse-grained molecular dynamics (MD) models with the aim of developing and analysing multiscale methods which use MD simulations in parts of the computational domain and (less detailed) particle-based stochastic reaction-diffusion models in the remainder of the domain. Applications using all-atom MD include intracellular dynamics of ions and ion channels. Applications using coarse-grained MD include protein binding to receptors on the cellular membrane, where modern stochastic reaction-diffusion simulators of intracellular processes can be used in the bulk and accurately coupled with a (more detailed) MD model of protein binding which is used close to the membrane.

Speaker: **Christian Yates** (University of Bath)

Title: *A PDE/compartment-based hybrid method for simulating stochastic reaction-diffusion systems*

Abstract: Spatial reaction-diffusion models have been employed to describe many emergent phenomena in biological systems. The modelling technique for reaction-diffusion systems that has predominated due to its analytical tractability and ease of simulation has been the use of partial differential equations (PDEs). However, due to recent advances in computational power, the simulation, and therefore postulation, of computationally intensive individual-based models has become a popular way to investigate the effects of noise in reaction-diffusion systems.

The specific stochastic model with which we shall concern ourselves are known as ‘compartment-based’. These models are characterised by a discretisation of the computational domain into a grid/lattice of discrete voxels between which molecules can jump. Molecules are considered to be well-mixed in each one of these voxels and can react stochastically with other molecules in their voxel with prescribed rates.

In a wide variety of biological situations, stochasticity due to low copy numbers is relevant only in particular regions of the domain. In other regions, copy numbers are sufficiently high that mean field models suffice to capture the important dynamics. Such conditions necessitate the development of hybrid models in which some areas of the domain are modeled using a continuum representation and others using an individual-based representation.

In this talk we develop hybrid algorithms which couple a PDE in one region of the domain to a compartment-based model in the other. Rather than balancing flux at the interface, we use a method which is similar to the ghost-cell method. Characteristic of this method is the individual treatment of particles as they cross the interface. A small region of the PDE domain adjacent to the compartment-based region is allowed to contribute particles to the compartment-based regime. When particles cross over the interface into this pseudo-compartment from the compartment-based regime a step-function with the mass of a single particle is added. We test our hybrid method in a variety of different scenarios and analyse the error introduced in each case.

## 4 Scientific Progress Made

Discussions at the workshop identified a number of future research directions and important problems to be investigated.

After the speaker presentations outlined in Section 3, there was a panel discussion where a number of important issues were discussed specific to the research and software development community present. The panel discussions were lead by leaders from the major simulation software groups. Simulation approaches to modelling stochastic reaction-diffusion processes are quite varied and each approach has advantages and disadvantages from a numerical standpoint. Some issues that need to be taken into account by modellers when using these numerical approaches are accuracy, computational complexity and data management, and the ease with which the approach may be implemented in parallel. There are currently no standard test problems for assessing these qualities, which is just part of the reason it is very difficult to compare different approaches and choose the most appropriate one for a given model system. In a discussion lead by E. de Schutter a number of important problems were identified which prohibit the establishment of a community standard against which algorithms may be compared and tested. Whilst some of these problems had simple

solutions others remain as open problems. One of the problems that was identified was the need for the standards to be valued by the community. This discussion was the first organised discussion of this problem by the community and an important step to establishing a common scientific language to report findings. An online forum was set up to encourage future communication within the community on important issues such as this.

The structure and order of the workshop speakers played an important role in stimulating critical discussion and identifying important problems for future work. Open problems were easily identifiable by focusing on specific topics such as biological models and open problems in the feasibility of the methods to simulating real biological systems, and then specifically focusing on modelling techniques. The limitations of both particle-based methods and mesoscopic methods were summarised well in the final section which focused on hybrid modelling approaches. These limitations need to be addressed by the community before the modelling techniques available to researchers are capable of simulating very complex biological systems such as a human cell. Segmenting the workshop subject content in this way led to critical discussions and the development of new collaborations. For example, organizer S. Isaacson has two projects that were stimulated by discussions at the workshop: 1. Determining how reversible reactions should be modeled in general particle-based models to preserve key physical properties such as detailed balance, and 2. Developing asymptotic methods to provide analytic approximations to the solutions of particle-based stochastic reaction-diffusion models (in collaboration with conference participant Jay Newby). For both these projects manuscripts are now in preparation for submission.

## 5 Outcome of the Meeting

The workshop offered a unique opportunity for the wider scientific community working on mathematical biology, systems biology, numerical analysis, computational methods, and stochastic analysis to exchange ideas and work collaboratively in order to tackle the many challenging open problems of this field. These include researchers from mathematics (probabilists, numerical analysts, and applied mathematicians), statistics, computational science, physics, biology, engineering, and systems biology. The interactions between these researchers are expected to lead to the development of better mathematical and computational methods with which to study biological processes, and new insights into specific systems arising in cell and population biology. The workshop also introduced a number of more junior researchers to the broader international community of scientists interested in particle-based stochastic reaction-diffusion models.

The meeting also identified the growing importance for the community to congregate and discuss future directions on a more regular basis. A workshop was organised for the community in Cambridge in 2016 and discussions have already started for holding another meeting in Australia in 2018. Prior to the Particle-Based Stochastic Reaction-Diffusion Models in Biology workshop at BIRS, research in this area was fragmented and progress was reported in small groups separately by mathematicians, computer scientists, computational biologists etc. The importance of meeting as an interdisciplinary community to establish shared research goals was widely recognized as a key outcome from the workshop. This in turn stimulated ongoing efforts to continue holding bi-annual meetings on Particle-Based Stochastic Reaction Diffusion models.

## References

- [1] D. Anderson and J. Mattingly, A weak trapezoidal method for a class of stochastic differential equations, *Communications in Mathematical Sciences* **9** (2011), no. 1, 301–318.
- [2] S. Andrews and D. Bray, Stochastic simulation of chemical reactions with spatial resolution and single molecule detail. *Phys Biol* **1** (2004), 137–151.
- [3] S. Arjunan and M. Tomita, A new multicompartmental reaction-diffusion modeling method links transient membrane attachment of E. coli MinE to E-ring formation. *Systems and Synthetic Biol* **4** (2010), 35–53.
- [4] A. Donev, V. Bulatov, T. Oppelstrup, G. Gilmer, B. Sadigh, and M. Kalos, A first-passage kinetic Monte Carlo algorithm for complex diffusion–reaction systems, *J. Comp. Phys.* **229** (2010), 3214–3236.

- [5] S. Engblom, L. Ferm, A. Hellander and P. Lötstedt, Simulation of stochastic reaction-diffusion processes on unstructured meshes. *SIAM J Sci Comput* **31** (2009), 1774–1797.
- [6] M. Elowitz, A. Levine, E. Siggia, and P. Swain, Stochastic gene expression in a single cell, *Science* **297** (2002), 1183–1186.
- [7] R. Erban and S. J. Chapman, Stochastic modelling of reaction-diffusion processes: algorithms for bimolecular reactions, *Physical Biology* **6** (2009), no. 4, 046001.
- [8] J. Lipkova, K. Zygalakis, S. J. Chapman, and R. Erban, Analysis of Brownian Dynamics simulations of reversible bimolecular reactions, *SIAM J. Appl. Math.* **71** (2011), 714.
- [9] C. W. Gardiner, K. J. McNeil, D. F. Walls, and I. S. Matheson, Correlations in stochastic models of chemical reactions, *J. Stat. Phys.* **14** (1976), 307.
- [10] D. T. Gillespie, Exact stochastic simulation of coupled chemical-reactions, *J. Phys. Chem.* **81** (1977), 2340–2361.
- [11] J. Hattne, D. Fange and J. Elf, Stochastic reaction-diffusion simulation with MesoRD, *Bioinformatics* **21** (2005), 2923–2924.
- [12] I. Hepburn, W. Chen, S. Wils and E. De Schutter, STEPS: efficient simulation of stochastic reaction-diffusion models in realistic morphologies. *BMC Syst Biol* **6** (2012), 36.
- [13] S. A. Isaacson, D. M. McQueen and C. S. Peskin, The influence of volume exclusion by chromatin on the time required to find specific DNA binding sites by diffusion, *PNAS* **108** (2011), 3815–3820.
- [14] S. Isaacson and C. Peskin, Incorporating diffusion in complex geometries into stochastic chemical kinetics simulations, *SIAM Journal on Scientific Computing* **28** (2006), no. 1, 47–74.
- [15] J. Keizer, Nonequilibrium statistical thermodynamics and the effect of diffusion on chemical reaction rates, *J. Phys. Chem.* **86** (1982), 5052–5067.
- [16] Muñoz-García J., Neufeld Z., Kholodenko B. N., and Sauro H. M. (2009), Positional information generated by spatially distributed signaling cascades, *PLoS Comp. Biol.* **5**, no. 3:e1000330.
- [17] S. Neves, P. Tsokas, A. Sarkar, E. Grace, P. Rangamani, S. Taubenfeld, C. Alberini, J. Schaff, R. Blitzer, I. Moraru, R. and Iyengar, Cell shape and negative links in regulatory motifs together control spatial information flow in signaling networks, *Cell* **133** (2008), 666–680.
- [18] T. Opplestrup, V. Bulatov, A. Donev, M. Kalos, G. Gilmer and B. Sadigh, First-passage kinetic Monte Carlo method. *Phys Rev E* **80** (2009), 066701.
- [19] J. Paulsson, Summing up the noise in gene networks, *Nature* **427** (2004), 415–418.
- [20] E. Roberts, J. E. Stone, and Z. Luthey-Schulten, Lattice Microbes: high-performance stochastic simulation method for the reaction-diffusion master equation, *J. Comput. Chem.* **34** (2013).
- [21] M. V. Smoluchowski, Mathematical theory of the kinetics of the coagulation of colloidal solutions, *Z. Phys. Chem.* **92** (1917), 129–168.
- [22] J. Stiles and T. Bartol, Monte Carlo methods for simulating realistic synaptic microphysiology using MCell. *Computational Neuroscience: Realistic Modeling for Experimentalists* (E. Schutter, ed.), CRC Press, (2001) pp: 87–127.
- [23] K. Takahashi and S. Tănase-Nicola and P. R. ten Wolde, Spatio-temporal correlations can drastically change the response of a MAPK pathway, *PNAS* **107** (2010), 2473–2478,
- [24] M. Tomita, Whole-cell simulation: a grand challenge of the 21st century. *Trends Biotechnology* **19** (2001), 205–210.
- [25] J. van Zon and P. R. ten Wolde, Green’s-function reaction dynamics: a particle-based approach for simulating biochemical networks in time and space. *J Chem Phys* **123** (2005), 234910.