Multi-scale Stochastic Modeling of Cell Dynamics

Jonathan Mattingly (Duke University),
Lea Popovic (Concordia University),
Peter Swain (University of Edinburgh)

January 17–22, 2010

1 Overview of the Field

A cellular network can be modelled mathematically using principles of chemical kinetics. These are particularly useful for studying the dynamic aspects of cells such as gene transcription, translation, regulation, and protein-protein interactions. Due to a large number of parameters, variables and constraints in cellular networks, numerical and computational techniques are often necessary. The development of computational approaches and analytical results to understand these dynamic processes is essential for the elucidation of cellular mechanisms. An increasing number of scientists are working to improve these approaches and to create, refine and test dynamical models in order to accurately reflect observations, and acquire predictive explanatory power.

At the cellular level, chemical dynamics are often dominated by the action of regulatory molecules present at levels of only a few copies per cell. Intrinsic noise due to random fluctuations of these components appears to have significant consequences: the observed large variation in morphology, rates of development, physiological responses in a cell often lead to a randomization of phenotypic outcomes and non-genetic population heterogeneity. Hence, stochastic modeling of the molecular dynamics within a cell is necessary in order to fully describe a set of expected outcomes. In many cases of biological interest some of the chemical species in the network are present in much lower abundance than others and the reaction rate constants can vary over several orders of magnitude. This implies that standard concentration scaling of stochastic models in assessing chemical dynamics does not provide a good representation of the behavior of the system, and that development of stochastic models on multiple scales is necessary.

2 Recent Developments and Open Problems

Cellular pathways involve many different molecular species, which are interconnected by an even larger number of chemical reactions, which poses a complex analytical problem. For prediction and simulation purposes, it is essential to reduce both the modeling and computational complexity of the problem, while still capturing all the essential characteristics and behavior of such a network. This has recently stimulated the development and analysis of stochastic models for biochemical networks and dynamics with multiple scales. A rigorous approach to this problem poses many interesting and challenging mathematical questions. Mathematical approaches can contribute directly to solve some of these problems in the following ways:

a. Numerous simulation algorithms are currently being developed for stochastic reaction systems on different time scales. The error of such simulations has not yet been fully explored, nor have the results from stochastic theory for systems on multiple scales been fully utilized to improve the development of such computational efforts.

b. Experimentalists are exploring the effect of stochastic mechanisms on the stability and robustness of components of these complex pathways. They are hoping to observe the effects that the dynamical properties of the system have on noise transmission, its amplification or damping, as well as whether some molecular systems may have evolved to use the stochasticity to its advantage. Fluctuation methods and stochastic dynamics may be combined to yield analytical answers to some of these questions.

c. Recent advances in single cell measurement techniques have yielded a wealth of new data which are being mined for important biological content. The temporal data are observations from stochastic dynamical systems, and statistical methods for stochastic processes can help extract some relevant parameters and make predictions for the behavior of the system.

3 Presentation Highlights

Presentations covered a wide range of mathematical, theoretical and computational issues in systems biology modeling, focusing in particular on stochastic and multiscale issues. In order of discussion -

Des Higham: Discrete versus Continuous in Simple Gene Regulation Models

Markov jump processes can provide accurate models in many applications, notably chemical and biochemical kinetics, and population dynamics. Stochastic differential equations offer a computationally effcient way to approximate these processes. It is therefore of interest to establish results that shed light on the extent to which the jump and diffusion models agree. This talk focused on mean hitting time behaviour in a thermodynamic limit, with examples of three simple types of reaction where analytical results can be derived, and where we found that the match between mean hitting time behavior of the two models is vastly different in each case. Furthermore, we stress that in many examples there is no guarantee that the diffusion model stays non-negative.Thus, care must be exercised when using diffusion models for reaction systems. ([1])

Jin Wang: Potential and Flux Landscape Framework for Understanding Stability and Robustness of Cellular Network

Studying the cell cycle process is crucial for understanding cell growth, proliferation, development, and death. We uncovered some key factors in determining the global robustness and function of the budding yeast cell cycle by exploring the underlying landscape and flux of this nonequilibrium network. The dynamics of the system is determined by both the landscape which attracts the system down to the oscillation orbit and the curl flux which drives the periodic motion on the ring. This global structure of landscape is crucial for the coherent cell cycle dynamics and function. The topography of the underlying landscape, specifically the barrier height separating basins of attractions, characterizes the capability of changing from one part of the system to another. This quantifies the stability and robustness of the system. We studied how barrier height is influenced by environmental fluctuations and perturbations on specific wirings of the cell cycle network. When the fluctuations increase, the barrier height decreases and the period and amplitude of cell cycle oscillation is more dispersed and less coherent. The corresponding dissipation of the system quantitatively measured by the entropy production rate increases. This implies that the system is less stable under fluctuations. In this talk we identify some key structural elements for wirings of the cell cycle network responsible for the change of the barrier height and therefore the global stability of the system through the sensitivity analysis. We show results are in agreement with recent experiments and also provide new predictions. ([2])

Ted Perkins: Trajectory inference for stochastic chemical kinetic models

Continuous-time Markov chains are used to model systems in which transitions between states as well as the time the system spends in each state are random. Many computational problems related to such chains have been solved, including determining state distributions as a function of time, parameter estimation, and control. However, the problem of inferring most likely trajectories, where a trajectory is a sequence of states as well as the amount of time spent in each state, appears unsolved. We studied three versions of this problem: (i) an initial value problem, in which an initial state is given and we seek the most likely trajectory until a given final time, (ii) a boundary value problem, in which initial and final states and times are given, and we seek the most likely trajectory connecting them, and (iii) trajectory inference under partial observability, analogous to finding maximum likelihood trajectories for hidden Markov models. In this talk we show that maximum likelihood trajectories are not always well-defined, and describe a polynomial time test for well-definedness. When well-definedness holds, we show that each of the three problems can be solved in polynomial time, and we develop efficient dynamic programming algorithms for doing so. ([3])

Ruth Williams: Coupled enzymatic degradation of proteins

A major challenge for systems biology is to deduce the molecular interactions that underly correlations observed between concentrations of different intracellular molecules. While direct explanations such as coupled transcription/translation or direct protein-protein interactions are often considered, potential indirect sources of coupling have received much less attention. In this talk, I will report on an investigation involving both theory and experiment of how correlations can arise generically from a post-translational coupling mechanism involving the processing of multiple protein species by a common enzyme. In this talk we show how the model can be posed in framework of multiclass queueing systems, and how we obtained the correlation of the stationary distribution for the protein species.

Samuel Kou: Multi-resolution inference of stochastic models from partially observed data

Inferring parameter values for diffusion models from data is often complicated by the fact that the underlying stochastic process is only partially observed. Likelihood based inference faces the difficulty that likelihood is usually not available even numerically. Conventional approach discretizes the stochastic model to approximate the likelihood. In order to have desirable accuracy, one has to use highly dense discretization. However, this usually imposes unbearable computation burden. In this talk we will introduce the framework of Bayesian multi-resolution inference to address this difficulty. By working on different resolution (discretization) levels simultaneously and by letting the resolutions talk to each other, we improve not only the computational efficiency, but also estimation accuracy. We illustrate our approach by examples.

Matthew Scott: Modeling intrinsic noise in continuous systems

In this talk I will consider systems that involve both reactions and spatial transport, treating chemical species at different locations as different species types. I will discuss recent work using an analytic approximation method to derive moments and spatiotemporal correlations in systems modeled using a reaction transport master equation. Our approach obtains moments via Van Kampen expansion, and uses a Fourier transform of factorial cumulants in order to obtain spatial correlation spectrum.

Di Liu: Numerical methods for stochastic bio-chemical reacting networks with multiple time scales

Multiscale and stochastic approaches play a crucial role in faithfully capturing the dynamical features and making insightful predictions of cellular reacting systems involving gene expression. Despite their accuracy, the standard stochastic simulation algorithms are necessarily inefficient for most of the realistic problems with a multiscale nature characterized by multiple time scales induced by widely disparate reactions rates. In this talk, I will discuss some recent progress on using asymptotic techniques for probability theory to simplify the complex networks and help to design efficient numerical schemes. We compare Nested SSA, multiscale SSA, and the slow-scale SSA algorithms, and we discuss methods of dynamic adaptation (re-evaluation of fast/slow classification) in multiscale algorithms. ([4])

Darren Wilkinson: Bayesian inference for stochastic networks

In this talk l give an overview of statistical methods for parameter inference for stochastic kinetic models, with emphasis on Bayesian approaches and sequential likelihood free MCMC, and discuss software. I will show an example application to stochastic kinetic modelling of p53/Mdm2 oscillations. ([5])

Moises Santillan: Evolution of the distributions for stochastic gene expression subject to negative feedback regulation

In this talk we discuss a simplification of the master equation (via an adiabatic approximation) and the numerical solution of the reduced master equation. The accuracy of this procedure is tested by comparing its results with analytic solutions (when available) and with Gillespie stochastic simulations. We employ our approach to study the stochastic expression of a simple gene network, which is subject to negative feedback regulation at the transcriptional level. We consider the influence of negative feedback on the amplitude of

intrinsic noise, and the relaxation rate of the system probability distribution function to the steady solution. Our results suggest the existence of an optimal feedback strength that maximizes this relaxation rate. ([6])

Paul Tupper: An Apparent Paradox of State-Dependent Diffusion

We consider an experiment of molecular diffusion on a two-dimensional rectangular lattice periodic boundary conditions. This rectangular lattice is partitioned into panels of equal size that alternate between two diffusion coefficients, one being twice the value of the other. Two contrary theoretical predictions of the motion of the diffusing particle are available. The first prediction (from stochastic calculus) observes that the diffusion coefficient influences how much time is spent on each side of the lattice. The second prediction (from statistical mechanics) states that the particle should spend the same amount of time on both sides of the partition. In this talk we discuss how these predictions depend on our interpretation of the diffusion coefficient as we take the limit of discrete systems.

Sayan Mukherjee: Multiscale factor models for molecular networks

In this talk a factor modeling framework is developed that is both predictive of phenotypic or response variation and the inferred factors offer insight with respect to underlying physical or biological processes. The method is general and can be applied to a variety of scientific problems. We focus on modeling complex disease phenotypes (etiology of cancer) as a motivating example. In this setting, the factors capture gene or protein interaction networks at different scales – breadth of the interaction network. The method integrates multiscale analysis on graphs and manifolds developed in applied harmonic analysis with sparse factor models, a mainstay of applied statistics. ([7])

Rachel Kuske: Model choice for mixed mode oscillations: coherence resonance and delay bifurcations

Many neuronal systems and models display a certain class of mixed mode oscillations (MMOs) consisting of periods of small amplitude oscillations interspersed with spikes. Various models with different underlying mechanisms have been proposed to generate this type of behavior. Stochastic versions of these models can produce similarly looking time series, often with noise-driven mechanisms different from those of the deterministic models. We present a suite of measures which, when applied to the time series, serves to distinguish models and classify routes to producing MMOs, such as noise-induced oscillations or delay bifurcation. By focusing on the subthreshold oscillations, we analyze the interspike interval density, trends in the amplitude and a coherence measure. We develop these measures on a biophysical model for stellate cells and a phenomenological FitzHugh-Nagumo-type model and apply them on related models. The analysis highlights the influence of model parameters and reset and return mechanisms in the context of a novel approach using noise level to distinguish model types and MMO mechanisms. Ultimately, we indicate how the suite of measures can be applied to experimental time series to reveal the underlying dynamical structure, while exploiting either the intrinsic noise of the system or tunable extrinsic noise. ([8])

David Anderson: Simulation methods for stochastically modeled population processes

While exact simulation methods exist for discrete-stochastic models of biochemical reaction networks, they are often times too inefficient for use because the number of computations scales linearly with the number of reaction events; thus, approximate algorithms have been developed. However, stochastically modeled reaction networks often have "natural scales" and it is crucial that these be accounted for when developing and analyzing numerical approximation methods. I will show that conducting such a non-standard error analysis leads to fundamentally different conclusions than previous analyses. Another option for approximating discrete-stochastic models of chemical reaction networks is to use a diffusion approximation. However, even in the regimes where such an approximation is preferable to the discrete numerical approximation methods, it is now necessary to approximate the diffusion process. In the second portion of my talk I will show how the special structure of chemical reaction networks can be utilized to develop an efficient and easy to implement method that is second order accurate in the weak sense for such diffusion processes. ([9],[10])

Eldon Emberly: A mechanism for polar protein localization in bacteria

We discuss a model for the cell cycle in the asymmetrically dividing bacteria Caulobacter Crescentus. We are interested in how polar localization of PopZ protein influences the biochemical calculations that bacteria need to perform cell differentiation.

John Fricks: Modeling Neck Linker Extension in Kinesin Molecular Motors

The kinesin molecular motor family takes a single 8 nanometer step forward for each ATP hydrolyzed except in rare cases. Recent experiments have demonstrated multiple steps including frequent back steps may be possible if the necklinker connecting the heads of the kinesin are extended. In this talk I will present a detailed intra-step model of kinesin stepping which allows for multiple steps and show that asymptotic quantities can be calculated usinga combination of limit theorems for semi-Markov processes and matrix analytic techniques for Markov chains. ([11])

Hye-Won Kang: The optimal size for space discretization for chemical reaction-diffusion networks

In this talk, I will discuss how to discretize space for the stochastic spatially-discrete model for chemical reaction-diffusion networks. A system with reaction and diffusion is modeled using a continuous time Markov jump process. Diffusion is described as a jump to the neighboring compartments with proper spatial discretization. Considering stationary mean and variance of each species in each compartment, the optimal size for spatial discretization will be suggested. I will show conditions for the exponential convergence to the uniform solution in the corresponding deterministic spatially-continuous model for chemical reaction-diffusion networks. Conditions obtained from the deterministic model approximate criteria for the optimal size for space discretization from the stochastic model well.

Peter Pfaffelhuber: Spatial aspects of multiscale chemical reaction networks

In modeling interactions between different types of molecular species in a population one often makes the assumption that the system is "well mixed". This is reflected in the fact that the rate at which reactions between species occur is proportional the overall number of each of the species types that are needed as inputs for the reaction. Intuitively this assumption is correct if molecular transport is "much faster" than the interactions. When molecular transport is not fast enough to insure spatial homogeneity of the system, one needs to address the role of space in the evolution of the total amount of each species in the system. In this talk we will present a model for a spatially inhomogeneous system. By making different assumptions on how fast the molecular transport is relative to the interactions, we will derive results for the evolution of the total amount of each species in the system, and discuss how they differ from results in a homogeneous system.

Mads Kaern: A framework for stochastic simulations of gene expression within evolving heterogeneous cell populations

The stochastic dynamics of individual cells typically take place within larger populations that are heterogeneous with respect to each individual and evolve with their own intricate dynamics. Noticeably, phenotypic heterogeneity associated with intrinsic or extrinsic noise in single cells may greatly impact population-level dynamics. Such couplings can be demonstrated experimentally using engineered gene regulatory networks, and is expected to have implications in numerous areas ranging from stem cell differentiation and development to drug-resistance within microbial or cancer cell populations. To facilitate the study of such multi-scale dynamical problems, we have developed an algorithm that combines an exact method to simulate molecular-level fluctations in single cells and a constant-number Monte Carlo method to simulate time-dependent statistical characteristics of growing cell populations. To benchmark performance, we compare simulation results with steady-state and time-dependent analytical solutions for several scenarios, including steady-state and time-dependent gene expression and the effects of cell growth, division, and DNA replication. This comparison demonstrates that the algorithm provides an efficient and accurate approach to model the effects of complex biological features on gene expression dynamics. Additionally, we show that the algorithm can quantitatively reproduce expression dynamics within bet-hedging cell populations during their adaption to

environmental stress, indicating that it provides the framework suitable for simulating and analyzing realistic models of heterogeneous population dynamics combining molecular-level stochastic reaction kinetics, relevant physiological details and phenotypic variability. ([12])

Hans Othmer: A Multi-Scale Analysis of Reacting Systems

We discuss theoretical and experimental approaches to three distinct developmental systems that illustrate how theory can influence experimental work and vice-versa. The chosen systems - Drosophila melanogaster, bacterial pattern formation, and pigmentation patterns - illustrate the fundamental physical processes of signaling, growth and cell division, and cell movement involved in pattern formation and development. These systems exemplify the current state of theoretical and experimental understanding of how these processes produce the observed patterns, and illustrate how theoretical and experimental approaches can interact to lead to a better understanding of development. We use multiscale analysis of the chemical reaction systems to obtain analytic approximations for amounts of different chemical species. ([13])

Greg Rempala: Statistical and Algebraic Methods for Analyzing Stochastic Mass Action Kinetics

With the development of new sequencing technologies of modern molecular biology, it is increasingly common to collect time-series data on the abundance of molecular species encoded within the genomes. This presentation shall illustrate how such data may be used to infer the parameters as well as the structure of the biochemical network under mass-action kinetics. We use algebraic methods as an alternative to conventional hierarchical statistical methods, and carry out network inference by deciding which rate constants are significantly different from zero. ([14])

David McMillen: Bacterial gene expression: modelling and (some) experiments

Plasmid-borne gene expression systems have found wide application in the emerging fields of systems biology and synthetic biology, where plasmids are used to implement simple network architectures, either to test systems biology hypotheses about issues such as gene expression noise or as a means of exerting artificial control over a cell's dynamics. In both these cases, fluorescent proteins are commonly applied as a means of monitoring the expression of genes in the living cell, and efforts have been made to quantify protein expression levels through fluorescence intensity calibration and by monitoring the partitioning of proteins among the two daughter cells after division; such quantification is important in formulating the predictive models desired in systems and synthetic biology research. A potential pitfall of using plasmid-based gene expression systems is that the high protein levels associated with expression from plasmids can lead to the formation of inclusion bodies, insoluble aggregates of misfolded, nonfunctional proteins that will not generate fluorescence output; proteins caught in these inclusion bodies are thus dark to fluorescence-based detection methods. Our results suggest that computational models using protein numbers derived from fluorescence measurements should take these into account, especially when working with rapidly growing cells. ([15])

Peter Swain: Modelling stochasticity in gene expression

Gene expression is a stochastic, or noisy, process. This noise comes about in two ways. The inherent stochasticity of biochemical processes such as transcription and translation generates intrinsic noise. In addition, uctuations in the amounts or states of other cellular components lead indirectly to variation in the expression of a particular gene and thus represent extrinsic noise. We show that simultaneous measurement of two identical genes per cell enables discrimination of these two types of noise. Although the intrinsic stochasticity inherent in biochemistry is relatively well understood, cellular variation, or 'noise', is predominantly generated by interactions of the system of interest with other stochastic systems in the cell or its environment. Such extrinsic fluctuations are nonspecific, affecting many system components, and have a substantial lifetime, comparable to the cell cycle (they are 'colored'). Here, we extend the standard stochastic simulation algorithm to include extrinsic fluctuations. On a model that involves all the major steps for transcription and translation, we show that these fluctuations affect mean protein numbers and intrinsic noise, can speed up typical network response times, and can explain trends in high-throughput measurements of variation. If extrinsic fluctuations in two

components of the network are correlated, they may combine constructively (amplifying each other) or destructively (attenuating each other). Our results demonstrate that both the timescales of extrinsic fluctuations and their nonspecificity affect the function and performance of biochemical networks. ([16], [17])

Konstantin Mischaikow: Developing a Database for the global dynamics of multiparameter systems

My most recent work has been on developing topological methods for the analysis of the global dynamics of multiparameter nonlinear systems. I will discuss new computational tools based on topological methods that extracts coarse, but rigorous, combinatorial descriptions of global dynamics of multiparameter nonlinear systems. These techniques are motivated by several observations: 1) In many applications there are models for the dynamics, but specific parameters are unknown or not directly computable. To identify the parameters one needs to be able to match dynamics produced by the model against that which is observed experimentally; 2) It is well established that nonlinear dynamical systems can produce extremely complicated dynamics, e.g. chaos, that is not structurally stable. However experimental measurements are often too crude to identify such fine structure in the dynamics or to establish the parameter values to sufficient precision even at points that are structurally stable; 3) Often the models themselves are based on heuristics as opposed to being derived from first principles and thus the fine structure of the dynamics produced by the models may be of little interest for the applications in mind. To make the above mentioned comments concrete I will use a simple model arising in population biology. I am very interested in combining these methods with stochastic dynamics. ([18])

Qian, Hong: Nonequilibrium phase transition Emerging landscape, time scales, and the chemical basis for epigenetic-inheritance

We consider a small driven biochemical network, the phosphorylation-dephosphorylation cycle (or GTPase) with a positive feedback. We investigate its bistability, with fluctuations, in terms of a nonequilibrium phase transition based on ideas from large-deviation theory. We show that the nonequilibrium phase transition has many of the characteristics of classic equilibrium phase transition: Maxwell construction, discontinuous first-derivative of the "free energy function", Lee-Yang's zero for the generating function, and a tricritical point that matches the cusp in nonlinear bifurcation theory. As for the biochemical system, we establish mathematically an emergent "landscape" for the system. The landscape suggests three different time scales in the dynamics: (i) molecular signaling, (ii) biochemical network dynamics, and (iii) cellular evolution. For finite mesoscopic systems such as a cell, motions associated with (i) and (iii) are stochastic while that with (ii) is deterministic. We suggest that the mesoscopic signature of the nonequilibrium phase transition is the biochemical basis of epigenetic inheritance. ([19])

Lev Tsimring: Dynamics and synchronization of synthetic gene oscillators

We designed and constructed a novel two-component gene oscillator in bacteria E. coli, based on principles observed to be critical for the core of many circadian clock networks. The design of the oscillator was based on a common motif of two inter-connecting positive and negative feedback loops. A small transcriptional delay in the negative feedback loop leads to stochastic relaxation oscillations which are further amplified and stabilized by the positive feedback loop. We use computational modeling to develop design criteria for achieving oscillations in this system. Drawing on analogy to integrate-and-fire dynamics in neuroscience, we have coined the term degradeand fire oscillations to describe the essence of the dynamics. In our subsequent work, we engineered gene network with global intercellular coupling that is capable of generating synchronized oscillations in a growing population of cells. Using microfluidic devices tailored for cellular populations at differing length scales, we investigated the collective synchronization properties along with spatiotemporal waves occurring at millimetre scales. ([20])

Tom Kurtz: Diffusion Approximation for Multiscale Reaction Network Models

Ćlassical stochastic models for chemical reaction networks are given by continuous time Markov chains. Methods for characterizing these models will be reviewed focusing primarily on obtaining the models as solutions of stochastic equations. The primary focus of the talk will be on employing stochastic analytic methods for these equations to understand the multiscale nature of complex networks and to exploit the

multiscale properties to simplify the network models. A diffusion approximation will be described for the slow changing components of a multi-scale system. ([21])

Katharina Best: Is anybody out there? Modelling spatial scaling in quorum sensing

Intercellular communication by means of small signal molecules synchronizes gene expression and coordinates functions among bacteria. This population density-dependent regulation is known as quorum sensing. Quorum sensing is frequently mediated by N-acylhomoserine lactone autoinducers. We investigate the molecular mechanism and regulation of quorum sensing in order to establish a predictive mathematical model of autoinducer signalling in this organism. To this end, we describe the dynamical system of the reactions responsible for the autoinducing behaviour within the bacterial cell and extend it to include coupling via a common environment. In a further step, we investigate the spatial organisation of the communicating cells.

Tomas Gedeon: Somitogenesis clock-wave initiation requires differential decay and multiple binding sites for clock protein

Somitogenesis is a process common to all vertebrate embroys in which repeated blocks of cells arise from presomatic mesoderm (PSM) to lay the foundational pattern for trunk and tail development. Somites form in the wake of passing waves of periodic gene expression that originate in the tailbud and sweep posteriorly across the PSM. Previous work has suggested that the waves result from a spatiotemporally graded control protein that affects the oscillation rate of the clock-gene expression. With a minimally constructed mathematical model, we study the contribution of two control mechanisms to the formation of this gene-expression wave. We test four biologically motivated model scenarios with either one or two clock protein transcription binding sites, and with or without differential decay rates for clock protein monomers and dimers. We examine the sensitivity of wave formation with respect to multiple model parameters and robustness to heterogeneity in cell population. We find that only a model with both multiple binding sites and differential decay rates is able to reproduce experimentally observed waveforms. The wave formation is robust to heterogeneous parameters in the cell population. Our results show that the experimentally observed characteristics of somitogenesis wave initiation constrain the underlying genetic control mechanisms.

David Cottrell: Incorporating diffusion in stochastic models of gene expression

A natural subject to explore in the study of biomolecular reaction networks, is the effect of spatial diffusion and molecular motility on the quantitative and qualitative properties of the biological systems. When there are no bimolecular reactions present, as in the case of the standard model of gene expression, the evolution equations are linear, and an analytical treatment becomes feasible. Our aim is to consider a simple branching-process model for gene expression and extend this to a branching-diffusion model where molecules diffuse in free-space. We derive differential equations related to several statistics and analytically compute a number of quantities related to the spatial correlations of the system.The mathematical approach that we take relies on the fact that the system can be modeled by a branching process, i.e. a process for which there is no interaction between the particles. Branching processes permit a useful decomposition due to the linearity of the process with respect to the initial condition. We take advantage of this property by considering a system of moment generating functions with xed initial conditions and construct the corresponding evolution equations. ([23])

4 Scientific Progress Made

The workshop explored new research directions and the ways in which researchers in stochastic dynamics, statistics and systems biology can join efforts in advancing our understanding of complex networks and stochastic dynamics in cells. Presentations stimulated many productive conversations between participants, resulting in a number of new project ideas. In particular, as a result of meeting for the first time at the workshop, Hong Qian (University of Washington) and Matthew Scott (Waterloo) have started up a collaboration to analyze nonlinear effects in fluorescence correlation spectroscopy. Also, David Anderson (University of Wisconsin), Des Higham (Strathclyde, Glasgow) and Ruth Williams (UC San Diego) spent some time at BIRS

discussing possible diffusion approximations to biochemical reaction network models that capture the natural positivity constraints on concentrations. This has led to a nice joint collaboration that they are currently working on. In both cases the groups did not really know each other before and this BIRS workshop was fundamental to bringing members of the groups together. Inspired by the talk of Mads Kaern (Ottawa), Jonathan Mattingly (Duke), Lea Popovic (Concordia) and postdoc John McSweeney (SAMSI) started a project on the effect that randomness from cell division plays in intracellular chemical reaction systems. A number of other projects between workshop participants previously begun reached significant progress during the workshop. Specifically Paul Tupper (Simon Fraser) finished a paper with Peter Swain (Edinburgh) and made significant progress on a project with Jonathan Mattingly (Duke). Likewise, Tom Kurtz (University of Wisconsin) and Lea Popovic (Concordia) finished a paper on diffusion approximations for reaction networks on multiple scales. Many workshop attendees stressed that the excellent facilities at BIRS (in particular the smaller rooms with lots of board space) were really helpful for their in depth discussions.

5 Outcome of the Meeting

The workshop brought together mathematical scientists working in disparate scientific communities: probability, stochastic analysis, stochastic numerics, applied mathematics, statistics, bioinformatics, bioengineering. There was also a significant presence of young researchers: 2 graduate students, 3 postdocs, 5 assistant professors. Important problems for future research directions were identified and discussed, starting a number of new collaborations in the process.

References

[1] L. Szpruch and D. J. Higham, Comparing hitting time behaviour of Markov jump processes and their diffusion approximations, *Multiscale Modeling and Simulation*, to appear.

[2] J. Wang, L. Xu, and E. K. Wang, Potential Landscape and Flux Framework of Non-Equilibrium Networks: Robustness, Dissipation and Coherence of Biochemical Oscillations, *Proc. Natl. Acad. Sci.* **105**, 12271–12276, 2008.

[3] T. J. Perkins, Maximum likelihood trajectories for continuous-time Markov chains, *Advances in Neural Information Processing Systems*, **22**, NIPS-2009.

[4] D. Liu, Analysis of multiscale methods for stochastic dynamical systems with multiple time scales, *SIAM Multiscale Modeling and Simulation*, **8**, 944–964, 2010

[5] D. A. Henderson, R. J. Boys, C. J. Proctor, and D. J. Wilkinson, Linking systems biology models to data: a stochastic kinetic model of p53 oscillations, *Oxford Handbook of Applied Bayesian Analysis*, A. O'Hagan and M. West (eds.), 155–187, Oxford University Press, 2010.

[6] E. S. Zeron and M. Santilln, Distributions for negative-feedback-regulated stochastic gene expression: dimension reduction and numerical solution of the chemical master equation, *J. Theor. Biol.*, **264**, 377–385, 2010.

[7] J. Guinney, P. Febbo, M. Maggioni, and S. Mukherjee, Multiscale factor models for molecular networks, submitted.

[8] P. Borowski, R. Kuske, Y.-X. Li and J. L. Cabrera, Characterizing mixed mode oscillations shaped by noise and bifurcation structure, submitted.

[9] D. F. Anderson, A. Ganguly, and T. G. Kurtz, Error analysis of tau-leap simulation methods, *Annals of Applied Probability*, accepted.

[10] D. F. Anderson and J. C. Mattingly, A weak trapezoidal method for a class of stochastic differential equations, *Communications in Mathematical Sciences*, to appear.

[11] J. Hughes, W. Hancock, and J. Fricks, A Matrix Computa- tional Approach to Kinesin Neck Linker Extension, submitted.

[12] D. A. Charlebois, J. Intosalmi, D. Fraser, M. Kaern, An Algorithm for the Stochastic Simulation of Gene Expression and Heterogeneous Population Dynamics,*Commun. Comput. Phys.*, accepted.

[13] H. Othmer, K. Painter, D. Umulis, and C. Xue, The intersection of theory and application in elucidating pattern formation in developmental biology, *Math. Mod. Nat. Phenom.*, **4**, 3–79, 2009.

[14] G. Craciun, C. Pantea, and G. Rempala, Algebraic Methods for Inferring Biochemical Networks: a Maximum Likelihood Approach, *Computational Biology and Chemistry*,**33:5**, 361–367, 2009.

[15] Dark proteins: Effect of inclusion body formation on quantification of protein expression. Marco A. J. Iafolla, Mostafizur Mazumder, Vandit Sardana, Tharsan Velauthapillai, Karanbir Pannu, David R. McMillen. Proteins: Structure, Function, and Bioinformatics Volume 72 Issue 4, Pages 1233 - 1242, 2008

[16] P. S. Swain, M. B. Elowitz, and E. D. Siggia, Intrinsic and extrinsic contributions to stochasticity in gene expression, *Proc. Nat. Acad. Sci.*,**99:20**, 12795–12800, 2002.

[17] V. Shahrezaei, J. F. Ollivier and P. S Swain, Colored extrinsic fluctuations and stochastic gene expression, *Molecular Systems Biology*, **4: 196**, 2008.

[18] Z. Arai, W. Kalies, H. Kokubu, K. Mischaikow, H. Oka, and Pl. Pilarczyk, A Database Schema for the Analysis of Global Dynamics of Multiparameter Systems, *SIAM J. Applied Dyn. Syst.*, **8** 757, 2009.

[19] Ge, H. and Qian, H., Nonequilibrium phase transition in a mesoscoipic biochemical system: From stochastic to nonlinear dynamics and beyond. *Journal of the Royal Society Interface*, 2010.

[20] T. Danino, O. Mondragon-Palomina, L. S. Tsimring, and J. Hasty, A synchronized quorum of genetic clocks, *Nature*,**463**, 326–330, 2010.

[21] H.-W. Kang, T. G. Kurtz, and L. Popovic, Diffusion approximations for multiscale chemical reaction models, submitted.

[22] M. Campanelli and T. Gedeon, Somitogenesis clock-wave initiation requires differential decay and multiple binding sites for clock protein, *PLoS Comput. Biol.*,**6:4**, 2010.

[23] David Cottrell, Peter Swain and Paul Tupper, A stochastic branching-diffusion model for gene expression, preprint.