

RESEARCH STATEMENT

CHE TAN D. PAHLAJANI

My primary research interests lie at the intersection of *probability theory* and *mathematical biology*. More precisely, I am interested in stochastic processes that arise in cellular and systems biology. I have recently been working on limit theorems (via averaging) for certain stochastic chemical kinetics models with multiple time scales. These questions have been motivated by a desire to understand stochastic phenomena in gene regulatory networks - the complex web of interactions in cells through which genes regulate the production of proteins. I am also interested in some theoretical questions involving random perturbations of dynamical systems.

1. STOCHASTIC PROCESSES IN SYSTEMS BIOLOGY

1.1. Introduction. Recent years have seen an explosion of interest in the nascent field of systems biology. This has been an interdisciplinary endeavor that has attracted the interest of engineers, applied mathematicians, computer scientists, statisticians, and of course, biologists. Although the term “systems biology” eludes clear definition, a common viewpoint is to regard the field as the application of a systems perspective to biology while taking an integrative approach to understanding complex interactions in biological systems.

At the heart of biological systems analysis is the analysis of *biochemical reaction networks*. While the dynamics of chemical reactions at the macroscopic scale (think of a test-tube) are fairly well understood, analyzing chemical kinetics at the cellular and sub-cellular scale poses a unique set of challenges. Perhaps the most significant of these is the fact that reactions at the cellular level often involve small numbers of molecules of reactants and products; the result is an inherent, observable *stochasticity* that cannot be accounted for by the classical (deterministic) ODE/PDE models for chemical reactions (see [BKPR, RWA02, SKPG]). Also, cellular reaction networks often have reactions occurring on *multiple time scales*. The presence of complicated feedback loops further adds to the complexity.

Stochasticity is especially pronounced in gene regulatory networks. For the present discussion, it will suffice to think of a gene as a specific stretch of DNA that encodes for the production of a specific protein. A cell typically contains thousands of different types of proteins and certain special proteins (called transcription factors) regulate the rate at which certain target genes are read, thereby regulating the rate at which the corresponding protein is synthesized. The result is a network structure (see [A07]) with genes as nodes and a directed edge taking Gene_i to Gene_k signifying that the protein product of Gene_i controls the rate of synthesis of the protein product of Gene_k .

The issue of randomness in cellular reaction networks can be addressed by working with stochastic models - typically continuous-time Markov chains that keep track of populations of various chemical species¹. The fundamental theory for continuous-time Markov chains is very well-developed. In the chemical kinetics context, Markov chains are most commonly characterized using either random

¹We work with the assumption that the cell is well-mixed. Physically this corresponds to assuming that the transport processes inside cells occur on a faster time scale than the chemical reactions. Mathematically, we can average out the spatial component and work with populations/concentrations without reference to the spatial variables.

time changes ([EK86]), or using the forward equation (called the chemical master equation in the physics literature) which describes the time evolution of probabilities. Also, trajectories of the Markov chain can be simulated using the Stochastic Simulation Algorithm, also known as the Gillespie Algorithm.

A different, but extremely useful mathematical framework for studying chemical kinetics is that of stochastic differential equations (SDE). One such SDE model, which is frequently used, is the *linear noise approximation* (LNA), also called the *van Kampen approximation* (see [EE03],[vK]). The LNA is a mesoscopic model - intermediate between the microscopic Markov chain and the macroscopic ODE. The LNA arises as follows: in the limit of large reaction volume, a form of the Law of Large Numbers (LLN) asserts that the rescaled Markov chain converges to the deterministic (macroscopic) ODE. Moreover, a form of the Central Limit Theorem (CLT) states that the rescaled fluctuations of the Markov chain from the ODE converge (in distribution) to a time-inhomogeneous Gaussian process which solves a linear stochastic differential equation (SDE). The approximation of the Markov chain by the system consisting of the deterministic ODE together with the SDE for the Gaussian fluctuations constitutes the linear noise approximation.

My research over the last two years has focused on deriving reduced (averaged) models for general chemical kinetics systems with multiple time scales, working in the framework of the linear noise approximation. My interest in the application area of systems biology was sparked through discussions and collaboration with Paul J. Atzberger (UCSB Mathematics) and Mustafa Khammash (UCSB Mechanical Engineering); indeed, much of the ensuing is joint work.

1.2. Multiscale chemical kinetics systems: model reduction. Working with the LNA, we describe in [PKA] a method to effect a stochastic model reduction for a general chemical kinetics system with two distinct time scales. Our results and techniques are very much along the lines of the stochastic mode reduction of [MTV] (see also [Pap76], [PS08]). In addition to providing a proof of concept, we apply our technique to a simple yet frequently encountered gene regulatory network - a single gene with negative feedback. Our results are substantiated by numerical computation.

Our general result can be described as follows: we have N chemically reacting species interacting through M reaction channels. We assume that of these M reactions, M_s are slow and M_f are fast (the fast reactions occur at “rate” $1/\varepsilon$ with $0 < \varepsilon \ll 1$); $M_s + M_f = M$. Assuming further that ℓ species change numbers only through slow reactions, while the remaining $N - \ell$ species change numbers only through fast reactions, we get, with $\mathbf{x}^\varepsilon(t) = (\mathbf{x}_s^\varepsilon(t), \mathbf{x}_f^\varepsilon(t)) \in (\mathbb{R}^+)^{\ell} \times (\mathbb{R}^+)^{N-\ell}$, $\mathbf{V}^\varepsilon(t) = (\mathbf{V}_s^\varepsilon(t), \mathbf{V}_f^\varepsilon(t)) \in \mathbb{R}^{\ell} \times \mathbb{R}^{N-\ell}$,

$$\begin{aligned} \dot{\mathbf{x}}_s^\varepsilon(t) &= F_s(\mathbf{x}_s^\varepsilon(t), \mathbf{x}_f^\varepsilon(t)) \\ \dot{\mathbf{x}}_f^\varepsilon(t) &= \frac{1}{\varepsilon} F_f(\mathbf{x}_s^\varepsilon(t), \mathbf{x}_f^\varepsilon(t)) \\ (1) \quad d\mathbf{V}_s^\varepsilon(t) &= [D_s F_s(\mathbf{x}^\varepsilon(t)) \mathbf{V}_s^\varepsilon(t) + D_f F_s(\mathbf{x}^\varepsilon(t)) \mathbf{V}_f^\varepsilon(t)] dt + \sum_{k=1}^{M_s} \sigma_k^s(\mathbf{x}^\varepsilon(t)) dW_k(t) \\ d\mathbf{V}_f^\varepsilon(t) &= \frac{1}{\varepsilon} [D_s F_f(\mathbf{x}^\varepsilon(t)) \mathbf{V}_s^\varepsilon(t) + D_f F_f(\mathbf{x}^\varepsilon(t)) \mathbf{V}_f^\varepsilon(t)] dt + \frac{1}{\sqrt{\varepsilon}} \sum_{k=M_s+1}^M \sigma_k^f(\mathbf{x}^\varepsilon(t)) dW_k(t). \end{aligned}$$

Here, $(W_1(t), W_2(t), \dots, W_M(t))$ is an M -dimensional Brownian motion and D_s (respectively D_f) denotes Jacobian with respect to the variables in the vector \mathbf{x}_s (respectively \mathbf{x}_f). The significance

of $\mathbf{x}^\varepsilon(t)$, $\mathbf{V}^\varepsilon(t)$ is captured by the assertion that for large reaction volume

$$(2) \quad \frac{\text{Population vector at time } t}{\text{Volume}} \approx \mathbf{x}^\varepsilon(t) + \frac{1}{\sqrt{\text{Volume}}} \mathbf{V}^\varepsilon(t).$$

In other words, $\mathbf{x}^\varepsilon(t)$ is the macroscopic concentration and $\mathbf{V}^\varepsilon(t)$ is the (rescaled) fluctuation process from the ODE limit. Equation (2) is, of course, the statement of the linear noise approximation and is made precise through the LLN and CLT arguments alluded to above.

Equation (1) displays a separation of time scales; the variables $\mathbf{x}_s^\varepsilon(t)$ and $\mathbf{V}_s^\varepsilon(t)$ vary slowly, while the variables $\mathbf{x}_f^\varepsilon(t)$ and $\mathbf{V}_f^\varepsilon(t)$ vary quickly. Standard arguments from the theory of singular perturbations imply that, as $\varepsilon \rightarrow 0$, $\mathbf{x}_s^\varepsilon(t)$ can be approximated by the ODE

$$(3) \quad \dot{\mathbf{x}}_s^0(t) = \bar{F}_s(\mathbf{x}_s^0(t))$$

where the effective vector field \bar{F}_s can be explicitly computed and the approximation justified using perturbation expansions.

In [PKA], we explore a similar model reduction for the stochastic process $\mathbf{V}_s^\varepsilon(t)$. For $\varepsilon > 0$ small, the dynamics of $\mathbf{V}_s^\varepsilon(t)$ (the fluctuation process for the slow species) depends on $\mathbf{x}_s^\varepsilon(t)$, $\mathbf{x}_f^\varepsilon(t)$ and $\mathbf{V}_f^\varepsilon(t)$. The intuition is, that as $\varepsilon \rightarrow 0$, one should make use of equation (3) to eliminate dependence on $\mathbf{x}_f^\varepsilon(t)$ and *average* out the fast fluctuations of $\mathbf{V}_f^\varepsilon(t)$. Under certain technical assumptions, we establish that as $\varepsilon \rightarrow 0$, $\mathbf{V}_s^\varepsilon(t)$ can be approximated by the stochastic process $\mathbf{V}_s^0(t)$ with $\mathbf{V}_s^0(t)$ solving the effective SDE

$$(4) \quad d\mathbf{V}_s^0(t) = B(t)\mathbf{V}_s^0(t)dt + \sum_{k=1}^{M_s} \bar{\sigma}_k(\mathbf{x}_s^0(t))dW_k(t)$$

where $B(t)$ and $\bar{\sigma}_k$ can be explicitly computed. The upshot of our result is that the macroscopic concentration and the fluctuation process for the slow species can be described, as $\varepsilon \rightarrow 0$, by the closed set of equations (3), (4), without reference to any of the fast variables. In essence, we have reduced the $2N$ -dimensional ODE/SDE system (1) to the 2ℓ -dimensional ODE/SDE system (3), (4).

1.3. Ongoing work.

- **Weak convergence - sharpening the results of [PKA].**

The asymptotic results of [PKA] were obtained using perturbation expansions for the backward equation (as in [MTV]) corresponding to $\mathbf{V}^\varepsilon(t)$. These formal arguments rely on a theorem of Kurtz ([Kur73]) and justify approximation of $\mathbf{V}_s^\varepsilon(t)$ by $\mathbf{V}_s^0(t)$ for *each fixed time* t as $\varepsilon \rightarrow 0$. A natural question to ask is whether the probability law of $\mathbf{V}_s^\varepsilon(t)$ converges weakly ([EK86]) to the law of $\mathbf{V}_s^0(t)$ on the space of paths $C_{\mathbb{R}^\ell}[0, \infty)$.²

I am currently working on this problem ([P1]) within the LNA framework (equations (1), (2)) for certain specific reaction networks. From the vantage point of stochastic averaging, this problem poses some interesting technical challenges. If one views equation (1) as a single SDE, then the degeneracy of the SDE interferes with establishing the ergodic properties central to averaging calculations. Instead, the approach taken in [PKA], [P1] is to view (1) as a slow-fast ODE (for $\mathbf{x}^\varepsilon(t)$) *driving* a slow-fast SDE (for $\mathbf{V}^\varepsilon(t)$). This latter approach sidesteps the issue of degeneracy, but one now has an SDE for $\mathbf{V}^\varepsilon(t)$ whose coefficients depend on both time t and ε (via their dependence on $\mathbf{x}^\varepsilon(t)$) - a feature that requires extra care (as $\varepsilon \rightarrow 0$) and that certainly seems new in the averaging literature.

² $C_{\mathbb{R}^\ell}[0, \infty)$ is the space of continuous functions mapping $[0, \infty)$ into \mathbb{R}^ℓ , equipped with the topology of uniform convergence on compact intervals.

1.4. Future work.

- **Stochastic filtering in gene regulatory networks**

Recent scientific and technological advances have made it possible to experimentally measure the level of expression of specific genes inside single cells (see, for instance [EJSS]). These studies often use genes that encode for *fluorescent proteins*; a gene that encodes for Green Fluorescent Protein (GFP) is present in jellyfish. A remarkable aspect about biology is that a gene taken from one organism (say, a jellyfish) can be expressed in a different organism (say, a bacterium). This has tremendous ramifications for measuring gene expression. Suppose one is interested in the level of expression of Gene_i in a certain cell, i.e. the level of Protein_i in the cell. By cleverly introducing a gene for fluorescent protein into the genome of the cell under study, it is possible to determine the level of Protein_i in the cell on the basis of observed fluorescence³. If one has a genetically identical population of such cells, one can sample the distribution of the protein level.

This leads to a whole class of naturally posed stochastic filtering questions. As noted earlier, genes are part of large interacting networks - genes produce proteins which regulate the expression of other genes and so on. If Gene_i above is part of a network, a natural question is: given the distribution of Protein_i for times $0 \leq s \leq t$, what is the best guess of the distribution of Protein_k at time t ? How does this guess depend on the network topology and the nature of the kinetics?

More abstractly, the problem can be stated as follows: one has a stochastic process of interest $X(t)$, one can observe a related stochastic process $Y(t)$. One wants to find the best guess of $X(t)$ given observations of $Y(s)$ for times $0 \leq s \leq t$, i.e. the conditional distribution $\pi(t)$ of $X(t)$ given the σ -algebra $\mathcal{F}_t^Y \stackrel{\text{def}}{=} \sigma(Y(s) : 0 \leq s \leq t)$. This is, of course, the central problem of stochastic filtering - a field that has been extensively studied since the 1960's (see [Kal80, KO88, Kus90]), especially in the context of diffusion processes. In contrast, the case where $X(t)$ and $Y(t)$ are both pure jump processes with countable state space seems to have received somewhat less attention (see [CG01, CG02, KKM90] and the references therein). I am very interested in exploring filtering questions for the continuous-time Markov chains of stochastic chemical kinetics - both in developing new techniques and in applying existing results to specific biochemical reaction networks.

- **Analysis of rare events**

Some of the most interesting and mathematically rich questions in many probabilistic investigations concern the probabilities of *rare events*. For the Markov chains of stochastic chemical kinetics, understanding rare events by direct stochastic simulation - generating several realizations and computing probabilities - becomes prohibitively expensive from a computational standpoint, thereby creating the need for *analytical* tools to address the problem. The mathematical framework for estimating probabilities of rare events is provided by the theory of *large deviations* (see [Var84, DZ98, FW98, FK06]). A typical problem in large deviations is the following: one has a sequence of random variables $\{X_n : n \in \mathbb{N}\}$ taking values in some Polish⁴ space \mathcal{X} with $X_n \rightarrow x_0$ as $n \rightarrow \infty$, x_0 deterministic. Establishing a large deviations principle (LDP) for the X_n 's (see [Var84] for a precise formulation) is

³Using these techniques, differences in gene expression have been observed in isogenic (genetically identical) populations, even when other factors are controlled. This provides experimental validation for the assumption of stochasticity in theoretical models.

⁴complete, separable, metric space.

tantamount to identifying a so-called *rate function* $I : \mathcal{X} \rightarrow [0, \infty]$ such that

$$(5) \quad \mathbb{P}(X_n \in A) \asymp e^{-n \inf_{x \in A} I(x)} \quad \text{as } n \rightarrow \infty.$$

In other words, I gives the exponential rate of decay of probabilities; of course, $I(x_0) = 0$.

At this point, it is instructive to note the common underlying mathematical structure between, on the one hand, continuous-time Markov chains of stochastic chemical kinetics converging to the large volume ODE limit, and on the other, queueing networks converging to the fluid limit. In both cases, one has a pure jump stochastic process (often Markovian) on $(\mathbb{Z}^+)^d$ converging path-wise to an ODE. Questions of large deviations for fluid limits of queueing networks have been extensively investigated (see [AD99] and references therein). A natural question that follows is: how far does the LDP analysis from the queueing context carry over to the chemical kinetics context? How does the differing nature of Markov chain transition intensities and network topology in the chemical kinetics context affect the existence of LDP's and identification of rate functions? This seems to be a promising source of fascinating problems, with enough common ground to get started, but sufficiently many differences to require new insights.

- **Explorations in biology**

I am also interested in learning more about biology. There is a wide range of problems in biology (other than intracellular reaction networks) where noise plays a significant role. Examples include ion-channel gating, neuronal dynamics, cytoskeleton dynamics and motors. I would like to further my knowledge about some of these areas to be better equipped to pose and solve probabilistic questions that naturally arise in their study.

Also of interest are some *biological* questions which necessarily must be addressed using mathematics, such as the role of noise in biological function. The consequences of intracellular noise are varied. There are instances where genetically identical bacteria exploit noise (manifested as population heterogeneity) to “hedge their bets” in matters of survival. In other instances, noise is suppressed and cellular operations proceed with great regularity - a consequence of *robustness* in biological networks. Sometimes, noise is used in the amplification of signals. A better understanding of the emergent properties of the biological system (as a whole) involves an intricate interplay of ideas from systems and control theory, biology and probability. One of the challenges here is finding the right mathematical question that can shed light on the biology.

2. RANDOM PERTURBATIONS OF DYNAMICAL SYSTEMS

2.1. Stochastic averaging for Hamiltonian systems. In my doctoral dissertation (under the direction of Rich Sowers at UIUC), I developed certain corrector functions for stochastic averaging of a planar noisy Hamiltonian system with discontinuous statistics. The problem can be described as follows: one starts with a randomly perturbed planar Hamiltonian system $\mathbf{X}^\varepsilon(t) \stackrel{\text{def}}{=} (X_1^\varepsilon(t), X_2^\varepsilon(t))$ given by

$$(6) \quad \begin{aligned} d\mathbf{X}^\varepsilon(t) &= \frac{1}{\varepsilon^2} (\nabla^\perp \mathbf{H})(\mathbf{X}^\varepsilon(t)) dt + dW(t) \\ \mathbf{X}^\varepsilon(0) &= \mathbf{x}_0 \end{aligned}$$

with a smooth double-well Hamiltonian function $\mathbf{H} : \mathbb{R}^2 \rightarrow \mathbb{R}$, $W = (W_1, W_2)$ a two-dimensional Brownian motion and $\nabla^\perp \mathbf{H} \stackrel{\text{def}}{=} (\mathbf{H}_{x_2}, -\mathbf{H}_{x_1})$ the symplectic gradient of \mathbf{H} . Under the assumption that the unperturbed Hamiltonian system (i.e. without noise) has a homoclinic orbit (shape of a figure-eight), Freidlin and Wentzell ([FW94]) established that as $\varepsilon \rightarrow 0$, the dynamics of the

slowly-varying quantity $H(\mathbf{X}^\varepsilon(t))$ can be approximated by averaging over the fast Hamiltonian flow to get a Markov process on a three-legged graph with *glueing conditions* at the vertex.

Sowers ([Sow03, Sow05]) investigated similar questions in terms of a singular perturbations problem near the homoclinic orbit and developed a methodology for constructing correctors to be used as *perturbed test functions* ([Kus84]) in averaging. In my dissertation, I developed the corresponding calculations when one has *skewness* (discontinuous statistics) at the homoclinic orbit; in other words, on hitting a loop of the figure-eight, a coin is flipped to decide whether the next excursion will be outside or inside the loop. I am currently in the process of streamlining these results for publication ([P2]).

2.2. Diffusion limits for noisy discrete-time dynamical systems. A large number of problems in (deterministic) dynamical systems take the form of mappings; this may be the natural model for the problem at hand or a reduction of the actual model (e.g. a Poincaré map). I am interested in exploring scaling limits for discrete-time dynamical systems perturbed by small noise. A prototype of such problems follows.

Start with the discrete-time dynamical system $x \mapsto b(x, \alpha)$ where $x \in \mathbb{R}$ and $\alpha \in \mathbb{R}$ is a parameter. Suppose this dynamical system exhibits a period-doubling bifurcation at $\alpha = \alpha^*$, i.e. for $\alpha < \alpha^*$, there is a unique stable fixed point and for $\alpha > \alpha^*$, there is a stable periodic orbit of period two. With $0 < \varepsilon \ll 1$ a small parameter and $\{\xi_n : n \geq 1\}$ a family of independent identically distributed random variables, we consider the noisy planar system

$$(7) \quad \begin{aligned} x_{n+1}^\varepsilon &= b(x_n^\varepsilon, \alpha_n^\varepsilon) + \varepsilon f(x_n^\varepsilon, \alpha_n^\varepsilon) \xi_{n+1} \\ \alpha_{n+1}^\varepsilon &= \alpha_n^\varepsilon + \varepsilon g(x_n^\varepsilon, \alpha_n^\varepsilon) \xi_{n+1}. \end{aligned}$$

Here, f and g are suitable functions that describe the dependence of noise strength on the state. Define the speeded-up continuous-time processes

$$(8) \quad \hat{x}_t^\varepsilon \stackrel{\text{def}}{=} x_{\lfloor t/\varepsilon^2 \rfloor}^\varepsilon \quad \text{and} \quad \hat{\alpha}_t^\varepsilon \stackrel{\text{def}}{=} \alpha_{\lfloor t/\varepsilon^2 \rfloor}^\varepsilon$$

where $\lfloor \cdot \rfloor$ denotes the integer floor. Owing to the separation of time scales, \hat{x}_t^ε should quickly converge to the stable fixed point (respectively, stable orbit of period two) for $\alpha < \alpha^*$ (respectively, $\alpha > \alpha^*$), leading to an *effective* $\bar{g}(\alpha)$ for the limiting dynamics (as $\varepsilon \rightarrow 0$) of $\hat{\alpha}_t^\varepsilon$. The natural question is: what happens at the bifurcation $\alpha = \alpha^*$?

In analogy with continuous-time results, one expects that as $\varepsilon \rightarrow 0$, $\hat{\alpha}_t^\varepsilon$ can be approximated in distribution by a limiting diffusion given by an effective generator (involving $\bar{g}(\alpha)$) away from the bifurcation $\alpha = \alpha^*$, and *glueing conditions* at the bifurcation. I am interested in such problems involving an interplay of discrete and continuous-time descriptions, bifurcations and weak convergence.

REFERENCES

- [A07] Uri Alon. *An Introduction to Systems Biology. Design Principles of Biological Circuits*. Chapman & Hall/CRC.
- [AD99] Rami Atar and Paul Dupuis. Large deviations and queueing networks: Methods for rate function identification. *Stoch. Proc. Appl.*, 84, 255 - 296, 1999.
- [BKPR] Karen Ball, Thomas G. Kurtz, Lea Popovic and Greg Rempala. Asymptotic analysis of multiscale approximations to reaction networks. *Ann. Appl. Prob.*, vol. 16, no. 4, 1925 - 1961, 2006.
- [CG01] Claudia Ceci and Anna Gerardi. Nonlinear filtering equation of a jump process with counting observations. *Acta Appl. Math.*, 66, no. 2, 139 - 154, 2001.
- [CG02] Claudia Ceci and Anna Gerardi. Conditional law of a branching process observing a subpopulation. *J. Appl. Prob.*, 39, 112 - 122, 2002.

- [DZ98] Amir Dembo and Ofer Zeitouni. *Large Deviations Techniques and Applications*. Second Edition, Springer-Verlag, New York, 1998.
- [EE03] Johan Elf and Mans Ehrenberg. Fast Evaluation of Fluctuations in Biochemical Networks with the Linear Noise Approximation. *Genome Research*, 13:2475-2484, 2003.
- [EJSS] Michael B. Elowitz, Arnold J. Levine, Eric D. Siggia and Peter S. Swain. Stochastic gene expression in a single cell. *Science*, volume 297, 1183-1186, 2002.
- [EK86] Stewart N. Ethier and Thomas G. Kurtz. *Markov Processes: Characterization and Convergence*. John Wiley & Sons Inc., New York, 1986.
- [FK06] Jin Feng and Thomas G. Kurtz. *Large deviations for stochastic processes*. American Mathematical Society, Providence, RI, 2006.
- [FW94] Mark I. Freidlin and Alexander D. Wentzell. Random Perturbations of Hamiltonian Systems. *Mem. Amer. Math. Soc.*, 109(523):viii+82, 1994.
- [FW98] Mark I. Freidlin and Alexander D. Wentzell. *Random Perturbations of Dynamical Systems*. Springer-Verlag, New York, 1998.
- [Kal80] Gopinath Kallianpur. *Stochastic filtering theory*. Springer-Verlag, New York-Berlin, 1980.
- [KKM90] Wolfgang H. Kliemann, Giorgio Koch and Federico Marchetti. On the unnormalized solution of the filtering problem with counting process observations. *IEEE Transactions on Information Theory*, vol. 36, no. 6, 1990.
- [KO88] Thomas G. Kurtz and Daniel L. Ocone. Unique characterization of conditional distributions in nonlinear filtering. *Ann. Prob.*, volume 16, no. 1, 80 - 107, 1988.
- [Kur73] Thomas G. Kurtz. A limit theorem for perturbed operator semigroups with applications to random evolutions. *J. Functional Analysis*, 12:55-67, 1973.
- [Kus84] Harold J. Kushner. *Approximation and Weak Convergence Methods for Random Processes, with Applications to Stochastic Systems Theory*, volume 6 of *MIT Press Series in Signal Processing, Optimization, and Control*. MIT Press, Cambridge, MA, 1984.
- [Kus90] Harold J. Kushner. *Weak convergence methods and singularly perturbed stochastic control and filtering problems*. Birkhauser Boston, Inc., Boston, MA 1990.
- [MTV] Andrew J. Majda, Ilya Timofeyev, Eric Vanden Eijnden. A mathematical framework for stochastic climate models. *Comm. Pure Appl. Math.*, 54, no.8, 891-974, 2001.
- [PKA] Chetan D. Pahlajani, Mustafa Khammash and Paul J. Atzberger. Stochastic reduction method for biological chemical kinetics using time-scale separation. (*in preparation*)
- [P1] Chetan D. Pahlajani. Averaging of fluctuations in a multiscale gene transcription network. (*in preparation*)
- [P2] Chetan D. Pahlajani. Stochastic averaging correctors for a noisy Hamiltonian system with discontinuous statistics. (*in preparation*)
- [Pap76] George C. Papanicolaou. Some probabilistic problems and methods in singular perturbations. *Rocky Mountain Journal of Mathematics*, Volume 6, Number 4, 653-674, 1976.
- [PS08] Grigorios A. Pavliotis and Andrew M. Stuart. *Multiscale Methods: Averaging and Homogenization*. Springer, 2008.
- [RWA02] Christopher V. Rao, Denise M. Wolf and Adam P. Arkin. Control, exploitation and tolerance of intracellular noise. *Nature*, vol. 420, 231-237, 2002.
- [SKPG] H. E. Samad, M. Khammash, L. Petzold and D. T. Gillespie. Stochastic modeling of gene regulatory networks. *Internat. J. Robust Nonlinear Control*, 15, 691-711, 2005.
- [Sow03] Richard B. Sowers. Stochastic averaging near a homoclinic orbit with multiplicative noise. *Stoch. Dyn.*, 3(3):299-391, 2003.
- [Sow05] Richard B. Sowers. A boundary layer theory for diffusively perturbed transport around a heteroclinic cycle. *Comm. Pure Appl. Math.*, 58:30-84, 2005.
- [Var84] S. R. S. Varadhan. *Large Deviations and Applications*. SIAM, Philadelphia, PA, 1984.
- [vK] N. G. van Kampen. *Stochastic Processes in Physics and Chemistry*, 2001.