Analysis of Noise Induced Stochastic Fluctuations in Gene Regulatory Networks

Brian Munsky* and Mustafa Khammash†

April 12, 2007

Abstract

Stochasticity is well recognized to be of crucial importance in the analysis of gene regulatory problems. This importance stems from the fact that extremely rare but important regulatory molecules often cause a great amount of intrinsic noise within a cell. Such systems are frequently modeled at the mesoscopic level as jump Markov processes, whose probability distributions evolve according to the chemical master equation (CME). In this paper we review a number of attempts that have been made to solve the CME. These include various kinetic Monte Carlo approaches, such as the Stochastic Simulation Algorithm (SSA) and its deviates, as well as systems theory based analytical solutions to the CME, such as the Finite State Projection (FSP) method and various moment closure techniques.

1 Introduction

The cellular environment is abuzz with noise. The origin of this noise is attributed to the random events that govern the motion of cellular constituents at the molecular level. Cellular noise not only results in random fluctuations within individual cells, but it is also a source of phenotypic variability among clonal cellular populations. In some instances these fluctuations are suppressed downstream through an intricate dynamical network that acts to filter the noise, much like a low pass filter attenuates high frequency signals. Yet in other instances, noise induced fluctuations are exploited to the cell's advantage. Researchers are only now beginning to understand that the richness of stochastic phenomena in biology depends directly upon these interactions of dynamics and noise and upon the mechanisms through which these interactions occur. Intriguing examples of mechanisms that rely on noise include stochastic switches, coherence resonance in oscillators, and stochastic focusing for the amplification of signals [1].

Given the importance of noise induced stochastic fluctuations in the cell, the quantitative modeling and analysis of these fluctuations is of paramount importance for the understanding and synthesis of biological networks. While mathematical models of genetic networks often represent gene expression and regulation as deterministic processes with continuous variables, the stochastic nature of cellular noise necessitates an approach that models these variables as discrete and stochastic. The continuous and deterministic approach makes sense when large numbers of molecules justify a continuous valued concentration description using mass-action kinetics. In this case, chemical reactions are modeled as reaction diffusion processes, and their dynamics can be found with partial differential equations (PDEs). When the reacting chemical solutions are well-mixed, these PDEs can then be well approximated with ordinary differential equations (ODEs). On the other hand, the cellular milieu is often home to key molecules that can be found in very small integer populations. Indeed in a typical living cell, it is not uncommon for some of the key molecules have ten or fewer copies. Clearly, in these instances the concentration description is meaningless, and a discrete stochastic model of the chemical species is essential. The choice between the two modeling approaches is not always clear. What is clear, however, is that as the size of the system of interacting species decreases, intrinsic noise becomes increasingly important (a relative change of one molecule is very important when there are only ten to begin with). At the sub-cellular level where gene regulatory networks reside, crucial chemical species such as DNA, RNA, and regulatory proteins may be present in only one or two copies per cell [2]. In these networks, which affect all aspects of life, stochastic effects have been found to play a significant and often a detrimental role in various aspects of cell function.

As a simple example, Fig. 1 represents a generic gene regulatory network comprised of only three mechanisms: transcription, translation, and regulatory feedback. With intrinsic noise, even this simple system can exhibit a rich variety of behaviors. For example, consider an open-loop system where transcription is slow, but translation is very fast. Such a strategy, which may be used to conserve energy [3], can result in systems where the transcripts may be entirely absent from the cell most of the time. However, because of efficient translation, one of these rare transcripts may occasionally result large bursts of proteins [3, 4]. Because such events can happen in some cells and not in others, they may account for huge variation in phenotype despite isogenic populations [4]. Conversely, if transcription were much faster and translation slower, the same average amount of protein may be found, but the variation could be far less [3].

Chemical regulators may also induce phenotypical variation despite homogenous genotypes. One excellent example of this can be found in the *pap* switch in *E. coli* [5, 6, 7]. In that system, DNA adenine methylase (DAM) applies irremovable methyl groups at some key regulatory regions of the DNA. In

*Brian Munsky is with the Department of Mechanical Engineering, University of California, Santa Barbara, CA 93106-5070 brianem@engr.ucsb.edu

†Mustafa Khammash is with the Department of Mechanical Engineering, University of California, Santa Barbara, CA 93106-5070 khammash@engr.ucsb.edu

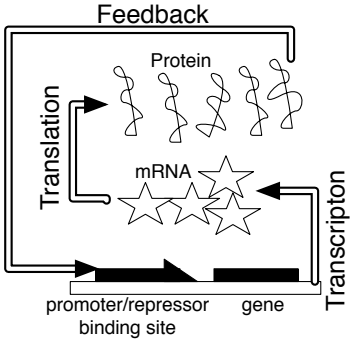

Figure 1: Schematic representation of gene transcription, translation and regulation. When in an "on" configuration the gene will transcribe mRNA molecules (stars). These, in turn, are translated to produce regulatory proteins, which can regulate the gene, turning it "off" in the case of negative feedback or "on" in the case of positive feedback.

one location, these methyl groups can help activate the *pap* gene; in another location, the methyl group will deactivate the gene [6]. The system is further affected by the intrinsic noise due to a transcriptional feedback mechanism similar to that illustrated in Fig. 1. In this case, the *pap*-encoded protein PapI works in conjunction with Leucine-responsive regulator protein (Lrp) to block Dam from methylating the sites which turn the gene expression off.

The topic of this paper is the mathematical modeling and analysis of systems of discrete stochastic chemically reacting systems, with an eye on applications to gene regulatory networks. We aim to provide a brief overview of the various approaches that have been recently proposed in this area. In the next section, we begin by providing a review of the chemical problem on the mesoscopic scale and derive what is commonly referred to as the Chemical Master Equation. This equation governs the evolution of the probability densities of the system states. In Section 3, we review a few of the recent Kinetic Monte Carlo approaches for generating sample trajectories. In Section 4 we discuss analytical techniques for the solution of the CME. In particular, we present a new direct approach for computing the relevant statistics, which involves the projection of the solution of the CME onto finite subsets. We illustrate the algorithm underlying our Finite State Projection approach and describe some system theory based modifications and enhancements that enable large reductions and increased efficiency with little to no loss in accuracy. In Section 5 we summarize and make some concluding remarks.

2 Formulation of Stochastic Chemical Kinetics

Gillespie's 1992 paper [8], provides a good background on the stochastic chemical kinetics problems and its major result: the chemical master equation (CME). For convenience, we provide a much simplified and less rigorous outline of his argument. Consider two molecules s_1 and s_2 moving around in a system of volume V. Suppose that molecule s_1 moves with the speed u, but in randomly changing directions. Suppose that a reaction $s_1 + s_2 \rightarrow s_3$ will occur when the center of molecule s_1 comes within a distance r of the center of molecule s_2. In some small fraction of time, dt, the molecule s_1 will cover a distance udt and will sweep a region dV whose volume is approximately $\pi r^2 udt$. If the center of s_2 is in dV then a reaction will occur; otherwise it will not. Since the system is well mixed, the probability that s_2 is in that region and that a reaction will occur is $\pi r^2 u V^{-1} dt$. If there were ξ_1 molecules of s_1 and ξ_2 molecules of s_2, then the probability that *any* such reaction will occur is given by $\xi_1 \xi_2 \pi r^2 u V^{-1} dt$.

For a chemical solution of N species, $\{s_1, \ldots, s_N\}$, one can define the system state as $\mathbf{x} = [\xi_1, \ldots, \xi_N]$. Each μ^{th} reaction is a transition from some state $\mathbf{x}_i$ to some other state $\mathbf{x}_j = \mathbf{x}_i + \nu_\mu$, where ν_μ is known as the *stoichiometric vector*. Following the methodology above, each reaction also has a *propensity function*, $a_\mu(\mathbf{x})dt$, which is the probability that the μ^{th} reaction will happen in a time step of length dt. For example, the reaction $s_1 + s_2 \rightarrow s_3$ discussed above has the stoichiometric vector $\nu = [-1, -1, 1]^T$, and a propensity $a(\mathbf{x})dt = \xi_1 \xi_2 \pi r^2 u V^{-1} dt$.

The stoichiometry and propensity functions for each of the M possible reactions fully define the system dynamics and are sufficient to find sample trajectories with the Monte Carlo methods of Section 3. However, for many interesting gene regulatory problems individual system trajectories are not the best description. Instead, it is desirable to analyze the dynamics in terms of probability distributions. For this it is useful to derive the chemical master equation.

Suppose that one knows the probability of all states $\mathbf{x}_i$ at time t, then the probability that the system will be in the state $\mathbf{x}_i$ at time, $t+dt$, is equal to the sum of (i) the probability that the system begins in the state $\mathbf{x}_i$ at t and remains there until $t+dt$, and (ii) the probability that the system is in a different state at time t and will transition to $\mathbf{x}_i$ in the considered time step, dt. This probability can be written as:

$$p(\mathbf{x}_i; t+dt) = p(\mathbf{x}_i; t)\left(1 - \sum_{\mu=1}^{M} a_\mu(\mathbf{x})dt\right) + \sum_{\mu=1}^{M} p(\mathbf{x}_i - \nu_\mu; t) a_\mu(\mathbf{x}_i - \nu_\mu)dt. \quad (1)$$

If one enumerates all possible $\mathbf{x}_i$ and defines the probability distribution vector $\mathbf{P}(t) = [p(\mathbf{x}_1; t), p(\mathbf{x}_2; t), \ldots]^T$, then it is relatively easy to derive the set of linear ordinary differential equations, known as the chemical master equation (CME) [9]:

$$\dot{\mathbf{P}}(t) = \mathbf{A}\mathbf{P}(t). \quad (2)$$

In the next two sections, this CME is solved using first Monte Carlo Methods and then a few analytical approaches.

3 Monte Carlo algorithms for the CME

Because the CME is often infinite dimensional, it is usually impossible to solve exactly. For this reason, the majority of analyses at the mesoscopic scale have been conducted using Monte Carlo (MC) algorithms. The most widely used

of these algorithms is Gillespie's Stochastic Simulation Algorithm (SSA) [10], but there are also a number of approximations to the SSA [11, 12, 13, 14, 15, 16, 17, 18, 19]. These are discussed in the following subsections.

3.1 Gillespie's Stochastic Simulation Algorithm

Gillespie Stochastic Simulation Algorithm (SSA) [10] is the most common tool in use for stochastic analyses at the mesoscopic level. This is to be expected, because once one defines the propensity functions and the stoichiometry for each of the M reactions, the SSA is very easy to apply. Each step of the SSA begins at a state $\mathbf{x}$ and a time t and is comprised of three tasks, (*i*) generate the time until the next reaction, (*ii*) determine which reaction happens at that time, and (*iii*) update the time and state to reflect the previous two choices. For a single reaction with propensity function, $a(\mathbf{x})$, the time of the next reaction, τ, is an exponentially distributed random variable with mean $(a(\mathbf{x}))^{-1}$. For M different possible reactions with propensities $\{a_\mu(\mathbf{x})\}$, τ is the minimum of M such random variables, or, equivalently an exponentially distributed random variable with mean equal to $\left(\sum_{\mu=1}^{M} a_\mu(\mathbf{x})\right)^{-1}$. To determine which of the M reactions occurs at $t+\tau$, one must generate a second random variable from the set $\mu = \{1, 2, \ldots, M\}$ with the probability distribution given by $P(\mu) = a_\mu(\mathbf{x}) \left(\left(\sum_{\mu=1}^{M} a_\mu(\mathbf{x})\right)^{-1}\right.$. Once τ and μ have been chosen the system can be updated to $t = t+\tau$ and $\mathbf{x} = \mathbf{x} + \nu_\mu$.

The SSA approach is exact in the sense that its result is a random variable with a probability distribution exactly equal to the solution of the corresponding CME. However, each run of the SSA provides only a single, not necessarily representative, trajectory; should one actually wish to reproduce the probability distribution, the SSA must be run many times. For this reason, researchers have proposed many accelerated approximations of the SSA.

3.2 System partitioning methods

In the first type of approximation to the SSA, the system is partitioned into slow and fast portions. This partitioning has been approached in a number of different manners. In [11] the system is separated into slow "primary" and fast "intermediate" species. This method uses three random variables at each step: first, the primary species' populations are held constant, and the population of the intermediate species is generated as a random variable from its quasi-steady-state (QSS) distribution. The dynamics of the "primary" species are then found with two more random variables, similar to the SSA above but with propensity functions depending upon the chosen populations of the intermediates species. The more recently developed Slow-Scale SSA (ssSSA) [12] is very similar in that the system is again separated into sets of slow and fast species. The ssSSA differs in that it does not explicitly generate a realization for the fast species, but instead uses the QSS distribution to scale the propensities of the slow reactions.

So-called hybrid methods such as [13] and [14] also separate the system into fast and slow reactions, but these methods do not then rely upon a QSS approximation. Instead, the fast reactions are approximated with deterministic ODEs or as continuous valued Markov processes using Langevin equations, and the slow reactions are treated in a manner similar to the SSA except now with time varying propensity functions.

3.3 τ leap methods

In a second approach to accelerating the SSA, researchers frequently assume that propensity functions are constant over small time intervals. With this "τ leap assumption" one can model each of the M reaction channels as an independent Poisson random process [15]. Beginning at time t and state $\mathbf{x}(t)$, the state at the end of a time step of length τ is approximated as $\mathbf{x}(t+\tau) = \mathbf{x}(t) + \sum_{\mu=1}^{M} k_\mu \nu_\mu$, where each k_μ is a random variable chosen from the Poisson distribution $k_\mu \in \mathcal{P}(a_\mu(\mathbf{x}(t)), \tau)$. The accuracy of τ leaping methods depends only upon how well the τ leap assumption is satisfied. Naturally, the τ leap assumption is best satisfied when all species have sufficiently large populations and all propensities functions are relatively smooth. Otherwise small changes in populations could result in large relative changes in propensities. Ignoring these changes can easily lead to unrealistic predictions of negative populations and/or numerical stiffness. One may avoid negative populations by using a Binomial τ leap strategy [16] or by adaptively choosing the size of each τ leap [17]. One can also ameliorate the problem of numerical stiffness using implicit methods such as that in [18].

When the populations are very large, and the propensity functions are very smooth, the chemical species may be more easily modeled with continuous variables using the *chemical Langevin equation* [19]. In this solution scheme, one assumes that many reactions will occur in the *macroscopic infinitesimal* times step dt without violating the τ leap assumption. One can therefore replace the Poisson distributions with Gaussian distributions, and treat the resulting process as a stochastic differential equation driven by white noise [19].

4 Analytical solutions to the CME

Each of the previous methods relied upon Monte Carlo simulations to explore the system dynamics and provides only one process realization at a time. To gain a probabilistic description of how likely the system is to exhibit certain behaviors at certain times, one would have to compile the results from simulations. Unfortunately, Monte Carlo algorithms have a very poor rate of convergence; even with an exact method such as the SSA, the error decreases with the number of runs only according to $\varepsilon = O(N^{-\frac{1}{2}})$. For this reason, an analytical solution to the CME is highly desirable.

4.1 The Finite State Projection solution

We recently presented the Finite State Projection algorithm for the solution of the CME [20]. In this approach we recognized two very important properties of the CME. First,

any system $\dot{\mathbf{P}}_{FSP}(t) = \mathbf{A}_J \mathbf{P}_{FSP}(t)$ corresponding to a truncation of (2) provides a lower bound on the solution to (2). Second, probability distributions are non-negative and sum to exactly one. As a result if the truncated solution, $\mathbf{P}_{FSP}(t) = \exp(\mathbf{A}_J t)\mathbf{P}_J(0)$, sums to $(1-\varepsilon)$, then the true solution, $\mathbf{P}(t)$ is within ε of $\mathbf{P}_{FSP}(t)$. Therefore, if one wishes to solve the CME to within a prespecified error tolerance, ε, one need only include enough rows and columns of $\mathbf{A}$ in $\mathbf{A}_J$ such that $\exp(\mathbf{A}_J t)\mathbf{P}_J(0)$ has a sum greater than $(1-\varepsilon)$. (For complete proofs and details, see [20]).

This FSP approach has effectively reduced the infinite dimensional CME to a finite dimensional set of linear time invariant ODEs. At this point, additional system theoretic tools can also be applied to evaluate stochastic gene regulatory networks. Furthermore, these tools lead to even stronger reductions of the CME and expand the realm of solvable problems.

4.1.1 Projection based reductions of the FSP

Even after projecting the CME onto a finite space, it is often useful to project the system onto an even lower dimensional space. For systems in which there is a clear separation of time-scales, one can utilize perturbation theory to average over the fast dynamics and project the system onto its slow manifold [21, 22, 23]. This approach is similar to the ssSSA [12] discussed above, in that the existence of "fast" and "slow" species do indeed result in a separation of time scales in the CME. However, time-scale based reductions to the master equation are more general in that they may be possible even in the absence of clear separations of fast and slow species.

For a second reduction, one can assume that the probability distribution varies linearly over some portions of the configuration space, solve the FSP problem on a coarse grid, and then interpolate to find the distributions at intervening points. This approach has shown great promise for certain problems as we have shown in [23]. Like the original FSP and its slow manifold approximation, this method does not explicitly make use of initial conditions or desired outputs. The advantage of this is that many different initial conditions or outputs can be considered with no extra computational cost. However, when this flexibility is not required, further reductions are possible.

Often one is not interested in the full solution to the CME. Instead, perhaps one desires only key information regarding that solution such as the likelihood of important events or statistical moments of important chemical species. In these cases one can define an output for the system $\mathbf{y}(t) = \mathbf{C}\mathbf{P}(t)$, and reduce the system to its minimal observable realization as explored in [24, 23]. In a third reduction approach one may utilize the initial condition $\mathbf{P}_{FSP}(0)$ and use a Krylov-based projection approach to solve directly for the matrix-vector product, $\exp(\mathbf{A}_J t)\mathbf{P}_{FSP}(0)$. This approach taken by Burrage and coauthors in [25] is particularly useful when taken in conjunction with a multiple time step algorithm as discussed next.

4.1.2 Multiple time step solutions of the FSP

In addition to using secondary projections to speed up the FSP solution, one can reduce computational effort with multiple time step routines. In both [25] and [26], it was independently observed that the probability distributions of some systems may travel over large portions of the state space, yet they may be only sparsely supported during any given instant in time. By splitting the full time interval into many small subintervals, one can consider much smaller portions of the state space for each time increment.

In [25] Burrage et. al observe that computing the full matrix exponential is more expensive than computing the product of that exponential with the initial distribution vector. This observation motivates the use of Roger Sidje's Krylov-based software package [27]. By adapting *expokit* to perform inexact matrix-vector products at each intermediate time step, the Krylov-based FSP algorithm of [25] only considers a portion of the state space during any one step, and is much more efficient than the original FSP. The tradeoff is that this solution is valid only for a specific initial distribution and must be recomputed for every different initial condition.

Our own Multiple Time Step FSP algorithm [26] takes a very different approach to improve the efficiency of the FSP. As in the previous method, we recognized that the solution of the full exponential is more work than necessary if one is only interested in the transition from an initial distribution at t to final distribution at $t+\tau$. However, it is also important to note that the matrix exponential contains far more information than the transition from one specific vector to another. In particular, each column of this exponential gives an estimate of a probability distribution at time $t+\tau$ conditioned on a specific state at time t. With this in mind, one can break the distribution at the beginning of each time step into many single-element initial vectors, solve these independently of one another and then apply the property of superposition. By restricting all time steps to the same length, many matrix exponentials can be reused from one time step to the next as well as for different initial conditions.

For further improvements in efficiency, either of these multiple time step solutions of the FSP can readily be combined with the projection based reductions of the previous subsection.

4.2 Moment closure techniques

One can also approximate the solution of the CME by representing the population of each species as a continuous variable and solving for the means and variances of the multi-variate distribution under closure assumptions. The first and most common such approach is the Linear Noise Approximation (LNA) [9, 28, 29]. In the LNA, one expands the solution of the master equation in a Taylor series about the macroscopic trajectory. The first order terms correspond to the macroscopic rate equations, and the second order terms approximate the system noise. The end result is a first order Fokker Planck equation, which is far more readily solved than the CME. In [30] a similar approach is taken except that the computation of the mean is coupled with that of the variances; this mass fluctuations kinetics (MFK) approach allows one to capture fluctuations where the mean deviates from the macroscopic

equation. This is particularly important for systems that exhibit stochastic focusing [1].

In a similar approach, the dynamics of each *uncentered* moment of the CME can be shown to depend linearly upon the rest to form an *infinite dimensional moment dynamics* linear ODE equivalent to the CME [31]. By assuming that the distributions are normal, lognormal, poisson, binomial, or another common form, one can approximate higher order moments in terms of the lower moments and effectively truncate the dynamics. Abhi and Hespanha review a few of these approaches for the *stochastic logistic model* in population biology [31]. In the same paper, Abhi and Hespanha also introduce an effective moment closure technique, which does not make an *a priori* assumption on the distribution shape, but instead defines a moment closure scheme in which they match the time derivatives of the truncated moment dynamics to the full moment dynamics at the initial time t_0.

Problems with a single macroscopic steady state often result in unimodal distributions and can be expressed with only the first few moments. For these, the above techniques are very well suited. However, problems that exhibit multi-modal distributions, such as switching systems, will require many higher order moments, and the applicability of these methods may quickly degrade.

5 Conclusions

Stochasticity is a very important concern in the numerical study of gene regulatory networks, and there has been significant attention given to developing the necessary mathematical tools. At the heart of these tools is the chemical master equation (CME). This infinite dimensional linear ODE is usually impossible to solve exactly, but there have been many successful attempts to simulate and approximate its solution.

This paper has reviewed a number of Monte Carlo (MC) algorithms as well as analytical approaches to solving the CME. For the MC approaches, the classic stochastic simulation algorithm (SSA) persists as the most commonly used tool, but many recent approximations to the SSA, such as τ leaping methods and system partitioning methods, are quickly gaining popularity. In terms of analytical approaches to solving the CME, our recent finite state projection (FSP) method has reduced the CME to a finite dimensional problem and has opened the door for many secondary reductions. In an entirely different direction, other researchers have recast the CME as an infinite dimensional moment dynamics problem, to which they apply moment closure techniques and reduce to a finite dimensional space before solving. At present no one method is sufficient for all systems, and the choice of which method to use in which circumstance depends primarily on the scale of the problem. However, as researchers continue to adapt the advanced tools of systems theory to these important stochastic chemical problems, more methods will become available, and computational researchers will be better able to approach a broader range of interesting biological questions.

6 Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant NSF-ITR CCF-0326576 and the Institute for Collaborative Biotechnologies through Grant DAAD19-03-D-0004 from the U.S. Army Research Office.

References

[1] J. Paulsson, O. Berg, and M. Ehrenberg. Stochastic focusing: Fluctuation-enhanced sensitivity of intracellular regulation. *PNAS*, 97(13):7148–7153, 2000.

[2] M. McAdams and A. Arkin. Its a noisy business! *Tren. Gen.*, 15(2):65–69, 1999.

[3] M. McAdams and A. Arkin. Stochastic mechanisms in gene expression. *PNAS*, 94:814–1819, 1997.

[4] E. Ozbudak, M. Thattai, I. Kurtser, A. Grossman, and A. van Oudenaarden. Regulation of noise in the expression of a single gene. *Nature Genetics*, 31:69–73, 2002.

[5] A. Hernday, M. Krabbe, B. Braaten, and D. Low. Self-perpetuating epigenetic pili switches in bacteria. *PNAS*, 99(4):16470–16476, December 2002.

[6] A. D. Hernday, B. A. Braaten, and D. A. Low. The mechanism by which dna adenine methylase and papi activate the pap epigenetic switch. *Mol. Cell*, 12:947–957, October 2003.

[7] B. Munsky, A. Hernday, D. Low, and M. Khammash. Stochastic modeling of the pap-pili epigenetic switch. *Proc. FOSBE*, pages 145–148, August 2005.

[8] D. T. Gillespie. A rigorous derivation of the chemical master equation. *Physica A*, 188:404–425, 1992.

[9] van Kampen. *Stochastic Processes in Physics and Chemistry*. Elsevier, 3 edition, 2001.

[10] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.*, 81(25):2340–2360, May 1977.

[11] C. V. Rao and A. P. Arkin. Stochastic chemical kinetics and the quasi-steady-state assumption: Application to the gillespie algorithm. *J. Chem. Phys.*, 118(11):4999–5010, Mar. 2003.

[12] Y. Cao, D. Gillespie, and L. Petzold. The slow-scale stochastic simulation algorithm. *J. Chem. Phys.*, 122(014116), Jan. 2005.

[13] E. Haseltine and J. Rawlings. Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics. *J. Chem. Phys.*, 117(15):6959–6969, Jul. 2002.

[14] H. Salis and Y. Kaznessis. Accurate hybrid stochastic simulation of a system of coupled chemical or biological reactions. *J. Chem. Phys.*, 112(054103), 2005.

[15] D. T. Gillespie. Approximate accelerated stochastic simulation of chemically reacting systems. *J. Chem. Phys.*, 115(4):1716–1733, Jul. 2001.

[16] T. Tian and K. Burrage. Binomial leap methods for simulating stochastic chemical kinetics. *J. Chem. Phys.*, 121(21):10356–10364, Dec. 2004.

[17] Y. Cao, D. T. Gillespie, and L. R. Petzold. Avoiding negative populations in explicit poisson tau-leaping. *J. Chem. Phys.*, 123(054104), 2005.

[18] M. Rathinam, L. R. Petzold, Y. Cao, and D. T. Gillespie. Stiffness in stochastic chemically reacting systems: The implicit tau-leaping method. *J. Chem. Phys.*, 119(24):12784–12794, Dec. 2003.

[19] D. T. Gillespie. The chemical langevin equation. *J. Chem. Phys.*, 113(1):297–306, Jul. 2000.

[20] B. Munsky and M. Khammash. The finite state projection algorithm for the solution of the chemical master equation. *J. Chem. Phys.*, 124(044104), 2006.

[21] S. Peles, B. Munsky, and M. Khammash. Reduction and solution of the chemical master equation using time-scale separation and finite state projection. *J. Chem. Phys.*, 125(204104), Nov. 2006.

[22] B. Munsky, S. Peles, and M. Khammash. Stochastic analysis of gene regulatory networks using finite state projections and singular perturbation. *Submitted to ACC*, Jul. 2007.

[23] B. Munsky and M. Khammash. The finite state projection approach for the analysis of stochastic noise in gene networks. *Submitted*, 2007.

[24] B. Munsky and M. Khammash. A reduced model solution for the chemical master equation arising in stochastic analyses of biological networks. *Proc. 45th IEEE CDC*, Dec. 2006.

[25] K. Burrage, M. Hegland, S. Macnamara, and R. Sidje. A krylov-based finite state projection algorithm for solving the chemical master equation arising in the discrete modelling of biological systems. *Proc. of The A.A.Markov 150th Anniversary Meeting*, 2006.

[26] B. Munsky and Khammash M. A multiple time-step finite state projection algorithm for the solution to the chemical master equation. *Submitted elsewhere*, 2007.

[27] Roger B. Sidje. EXPOKIT: Software package for computing matrix exponentials. *ACM Transactions on Mathematical Software*, 24(1):130–156, March 1998.

[28] J. Elf and M. Ehrenberg. Fast evaluations of fluctuations in biochemical networks with the linear noise approximation. *Genome Research*, 13:2475–2484, 2003.

[29] R. Tomioka, H. Kimura, T. Kobayashi, and K. Aihara. Multivariate analysis of noise in genetic regulatory networks. *J. Theoretical Biology*, 229(4):501–521, 2004.

[30] C. Gmez-Uribe and G. Verghese. Mass fluctuation kinetics: Capturing stochastic effects in systems of chemical reactions through coupled mean-variance computations. *JCP*, 126(024109), Jan. 2007.

[31] A. Singh and J. Hespanha. Moment closure techniques for stochastic models in population biology. *Proc. of the 2006 ACC*, pages 4730–4735, June 2006.

Authors' profiles

Brian Munsky received his B.S. and M.S. degrees from the Department of Aerospace Engineering at the Penn State University in 2000 and 2002 respectively. During those studies, he worked at the Penn State Rotorcraft Center of Excellence (RCOE) to develop detailed aeroelastic models with which to analyze helicopter noise and vibrations. Since 2003, he has pursued his Ph.D. degree at the University of California at Santa Barbara, where he developed a stochastic model of the Pap Pili epigenetic switch in *E. coli*, and he formulated the Finite State Projection solution for the Chemical Master Equation. His main research interests include applying systems theory to the modeling and analysis of stochastic gene regulatory networks.

Mustafa Khammash ...