CHAPTER 5

Stochastic Modeling of Cellular Networks

Jacob Stewart-Ornstein and Hana El-Samad

Department of Biochemistry and Biophysics, California Institute for Quantitative Biosciences, University of California, San Francisco, CA

Abstract

Noise and stochasticity are fundamental to biology because they derive from the nature of biochemical reactions. Thermal motions of molecules translate into randomness in the sequence and timing of reactions, which leads to cell–cell variability ("noise") in mRNA and protein levels even in clonal populations of genetically

METHODS IN CELL BIOLOGY, VOL 110

0091-679X/10 $35.00
DOI 10.1016/B978-0-12-394620-1.00005-9

identical cells. This is a quantitative phenotype that has important functional repercussions, including persistence in bacterial subpopulations challenged with antibiotics, and variability in the response of cancer cells to drugs. In this chapter, we present the modeling of such stochastic cellular behaviors using the formalism of jump Markov processes, whose probability distributions evolve according to the chemical master equation (CME). We also discuss the techniques used to solve the CME. These include kinetic Monte Carlo simulations techniques such as the stochastic simulation algorithm (SSA) and method closure techniques such as the linear noise approximation (LNA).

I. Introduction

Cells are microscopic reactors where multitudes of chemical reactions occur. Biochemical reactions are probabilistic collisions between randomly moving molecules, with each event resulting in the increment or decrement of molecular species by integer amounts (Hasty and Collins, 2002; McAdams and Arkin, 1999; Rao *et al*., 2002; Raser and O'Shea, 2005). As many crucial biological species including RNA and DNA are present in small quantities (ones or tens) per cell, these stochastic events can have measurable effects. The amplified effect of fluctuations in a molecular reactant, or the compounded of fluctuations across many molecular reactants, referred to as "molecular noise," often can accumulate as an observable phenotype, endowing the cell with individuality and generating nongenetic cell-to-cell variability in a population.

Observations of such nongenetic variation date back to the 1940s when it was determined that bacterial cultures were not completely killed by antibiotic treatment—a small fraction of cells "persist" (Bigger, 1944). The insensitivity to antibiotics exhibited by these persister cells was nonheritable (Moyed and Broderick, 1986), and persister cells spontaneously switched back to the nonpersistent state, regaining sensitivity to antibiotics. The advent of optical measurement methods, which monitor fluorescent reporter expression in single cells using flow cytometry or fluorescence microscopy, further illustrated that isogenic populations of cells can show great variability or "noise" in their gene expression (Cai *et al*., 2006; Thattai and van Oudenaarden, 2001). By measuring the fluorescence intensity of single cells, probability distributions representing variability in a process across a population of cells can be constructed **(**Fig. 1). A broad distribution indicates a large dispersion of expression levels across the population. Recently, genome scale assays of variability in gene expression revealed that specific types of genes—those involved in energy metabolism and stress response—showed heightened variability (Bar-Even *et al*., 2006; Newman *et al*., 2006b). These data were used to lend support to the hypothesis that variability in protein content among cells might be a regulated trait that confers a selective advantage through a "bet-hedging" strategy with respect to future environmental shifts (Avery *et al*., 2007; Blake *et al*., 2006). Such stochastic fate specification has also been postulated in other contexts. For example, each cell in the mouse olfactory bulb must select only one olfactory receptor to express, and is thought to implement this decision by stocastically selecting to express a gene which then mediates global repression of the other ∼1300

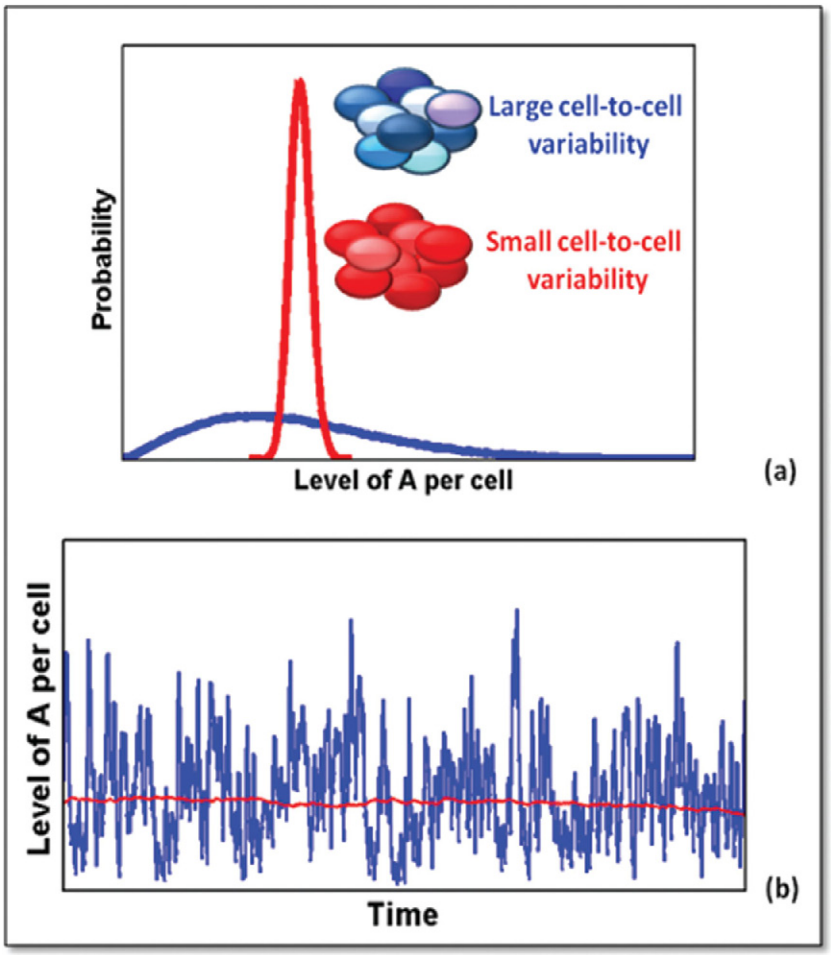

Fig. 1 Biochemical noise. (a) Distribution of a cellular component A for large cell–cell variability (blue, noisy) and small cell-to-cell variability (red). (b) Fluctuations of A as a function of time in one cell. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this book.)

receptors (Serizawa *et al*., 2003). A similar model exists in two precursor cells in *Caenorhabditis elegans*, called Z1.ppp and Z4.aaa. In 50% of embryos, Z1.ppp differentiates into the AC cell, whereas Z4.aaa adopts the VU cell fate. In the other 50% of embryos, the opposite occurs. Through a random process, one cell adopts the VU cell fate, and then inhibits that choice in the other through a Notch signaling mechanism (Karp and Greenwald, 2003).

Variability, however, is not always beneficial. In the cell cycle for example, numerous feedback loops exist to ensure a tightly regulated and orderly transition through DNA replication and cell division (Tsai *et al*., 2008). Similarly, in the *Drosophila* embryos, variability in the pattern of the bicoid protein results in

undesirable developmental alterations, and studies suggest that the system is poised at the fundamental limit of the precision it can achieve (Gregor *et al*., 2007a, b). In all these cases, understanding the roots and consequences of variability in the cell through careful measurements and quantitative modeling was of paramount importance for understanding the functioning of the underlying biological networks.

II. The Need for a Stochastic Modeling Framework

Most often than not, mathematical models represent the dynamic operation of cellular networks as deterministic processes with continuous variables. This continuous and deterministic approach may be warranted when large numbers of molecules justify a continuous valued concentration description. This is, for example, the case in metabolic networks where the concentration of reactants is in the millimolar range. There, chemical reactions can be modeled as reaction diffusion processes, and their dynamics described by partial differential equations (PDEs). When the reacting chemical solutions are well-mixed, these PDEs can then be well approximated with ordinary differential equations (ODEs).

There are many situations where this continuous deterministic modeling fails and stochastic models are necessary to capture biologically relevant properties of the systems under study. One such scenario is one where continuous models fail to describe quantitatively the behavior of a system because key regulatory molecules are found in very small integer populations. For example, the Lac operon in *Escherichia coli* is regulated by lactose binding to the repressor *Lac*I, which needs to be inactivated to allow for transcription of the operon. In this system, the key regulatory event in sensing lactose is the stochastic expression of a very small number of copies of the lactose permease lacY. As a result, the switching rate of *E. coli* to a lactose metabolizing state is governed by small number fluctuations of lacY (Choi *et al*., 2008), necessitating a discrete stochastic model of the chemical species involved in this regulation.

A second situation where stochastic models are needed arises when fluctuations induce dynamical behaviors, which cannot be captured even qualitatively using deterministic models. For example, stochastic fluctuations in excitable systems cause them to undergo large excursions away from their equilibrium point. Such excitable behavior occurs in the prokaryote *Bacillus subtilis* when it transitions between low- and high-"competence" states that have differential abilities to absorb DNA from their environments. Periods of high competence occur stochastically when the master regulators comK and comS exceed a certain concentration. After an individual cell has passed the threshold, strong positive feedback loops drive the cell toward competence, followed by a slower negative feedback loop, which switches the system off after a defined period (Suel *et al*., 2006, 2009). These dynamics occur nonsynchonously in a small population of cells, and therefore cannot be recapitulated using a deterministic mode, which, in this instance, can only settle into its only stable equilibrium. In contrast, accounting quantitatively for stochastic variation in protein concentrations is needed to reproduce this behavior.

III. Overview of Computational Approach

A quantitative modeling framework that takes into account the inherent stochasticity of biochemical interactions occurring inside a cell should handle discrete systems, should be adaptable to many different problems, and should be computationally tractable. A widely used such approach that we review in this chapter is one developed to address the chemical kinetics of well-mixed homogeneous systems. In this approach the cell is treated as a well-mixed bag of chemical species (Gillespie, 1977; Mcquarri, 1967). A model then probabilistically describes the chemical interactions of a subset of these species as a Markov (memoryless) jump process. After such a model is initiated from a defined state (in terms of the number of molecules of different species), reactions are allowed to occur between the chemical species. These reactions are represented by state transitions in a Markov chain, and transitions occur in discrete steps after a random time period, with the change and the time both depending only on the previous state. In this way, the transitions model the change in the number of each type of biological molecule in accordance with the stoichiometry of the chemical reaction (Fig. 2).

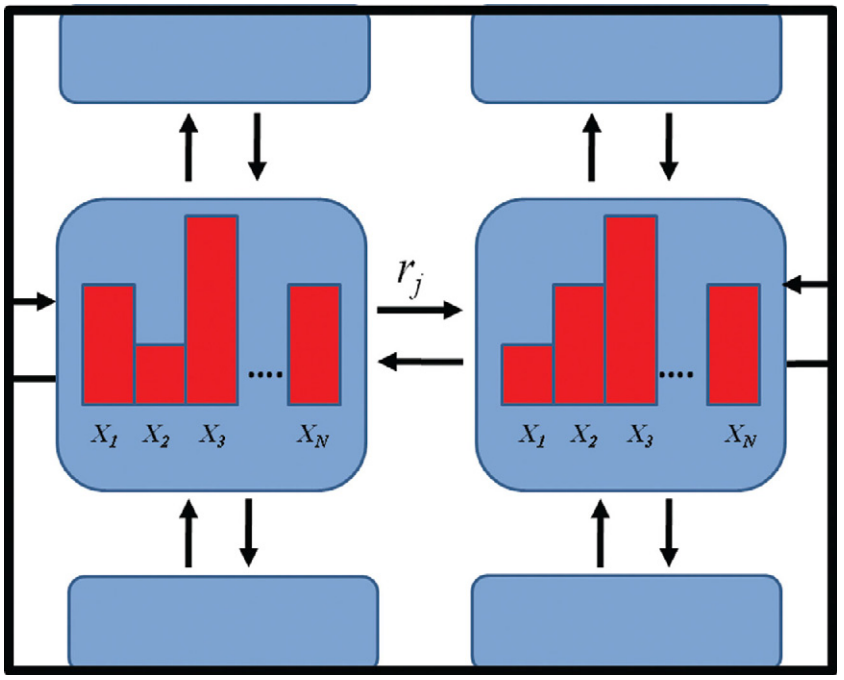

Fig. 2 Markov chain model for chemical kinetics. The states of the Markov chain are defined by the numbers of biological molecules of each chemical species, labeled X_1, X_2,. . ., X_N. Transitions between these states model the individual chemical reactions which may occur in the system. The transition corresponding to the chemical reaction of type k is labeled by R_k. (For color version of this figure, the reader is referred to the web version of this book.)

The chemical master equation (CME) is a differential equation that governs the time evolution of the probability for observing the Markov chain in a given state at a given time. The CME is generally derived using the Markov property, by writing the Chapman–Kolmogorov equation, an identity that must be obeyed by the transition probability of any Markov process (Gillespie, 1992; Mcquarri, 1967). Although the CME is straightforward to write, it cannot be analytically solved for any but the simplest problems. Therefore, numerical simulations on a computer are the key tool used for understanding the behavior of a system described by a CME. Monte Carlo simulation techniques are routinely used. Specifically, in this context, an algorithm known as the stochastic simulation algorithm (SSA, but more commonly known as the Gillespie algorithm) is used to generate exact realizations (or "runs") of the Markov jump process (Gillespie, 1977). The algorithm generates time course trajectories of the system states over a given time window, starting from a given initial system state. Each such run is "exact" in the sense that it is an independent realization from the true underlying process. However, each realization is also stochastic and is therefore different for each simulation run. A construction of the probability distributions of the underlying stochastic processes can then be done by executing and compiling a sufficient number of such runs.

IV. Biological Insights from Computational Approaches

Cell-to-cell variability (molecular noise) is ubiquitous in the cellular world where typical transcription factors can exist in as a few as 10 copies per cell and bind to promoters of individual genes, which produce bursts of a few mRNAs. Although the functional repercussions of this variability were observed in bacterial persistence as early as 1944 (Bigger, 1944), it is only recently that this aspect of cellular physiology has captured the imagination of both theorists and biologists. As a result, the last few decades have witnessed many discoveries about how cells and organisms attenuate or exploit their molecular fluctuations, and what implications these bear on cellular phenotypes. Computational methods based on the formalism presented in this chapter continue to play a central role in these investigations.

Cellular decision making has been one area where stochastic models have made a crucial contribution. One of the earlier landmark works to apply the Gillespie algorithm (Gillespie, 1977) for modeling a natural gene network included a comprehensive model of the Lambda switch (McAdams *et al.*, 1998). This seminal work described how the Lamba phage balanced lytic and lysogenic outcomes of bacterial infection and illustrated how stochastic molecular events, originating from the random movement of molecules, can trigger decisions on a much larger scale leading to divergent cellular fates. A flurry of subsequent work used the same approach to investigate stochastic cellular switching and decision making in a number of biological contexts. For example, theoretical work illustrated that a population of cells capable of random phenotypic switching can have an advantage in a fluctuating environment (Kussell *et al.*, 2005; Thattai and van Oudenaarden, 2001; Wolf *et al.*, 2005). Some of these predictions

have subsequently been confirmed, showing that noise can aid survival in severe stress (Blake *et al*., 2006) and can optimize the efficiency of resource uptake during starvation (Suel *et al*., 2009) and survival in fluctuating environments (van Oudenaarden *et al*., 2008).

In addition to their role in unraveling the functional repercussions of molecular noise, computational methods that capture biological fluctuations have been instrumental in pinpointing their origins and the cellular mechanisms that modulate them. Stochasticity in gene expression received special attention. There, the synergy between quantitative measurements at the single cell or single molecule level and appropriate quantitative models deepened our understanding of the processes involved in transcription and translation and yielded some unexpected observations (Cai *et al*., 2006; Chubb *et al*., 2006; Cluzel *et al*., 2005; Golding *et al*., 2005; Raj *et al*., 2006; Yu *et al*., 2006). For example, it was demonstrated that transcription of genes in *E. coli* is not as simple as RNA polymerases transcribing with a constant flux. Instead, the process is highly variable and proceeds in bursts rather than continuously. The origin of this behavior is still unknown, although possible candidates include global fluctuations of chromosome supercoiling states and RNA polymerase availability. A discrete stochastic framework accounting for all possible promoter states was also necessary to interpret experimental measurements of stochastic expression from eukaryotic promoters (Murphy *et al*., 2007). Quantitative computational approaches of the type we discuss in this chapter and high-resolution measurement technologies are poised to further reveal the workings of these fundamental cellular processes.

Synthetic biology is a nascent branch of biological investigation where accurate predictive modeling is of crucial importance. The aim of synthetic biology is to bring together ideas from biology and engineering to design and build biological networks that can achieve novel functions inside cells. It is now appreciated that the robust operation of synthetic cellular networks requires an understanding of molecular fluctuations, and that this understanding stems from rigorous probing of their stochastic dynamics. Analysis of stochastic models of the type we will tackle in this chapter has, for example, enabled the design and construction of synthetic oscillators that are robust to expected cellular variability (Tigges *et al*., 2009).

V. Computational Methods

A. A Simple Example

In a simple model of transcription, a gene is transcribed to generate a mRNA at a constant rate k, and each mRNA molecule is independently degraded at a rate γ. The mRNA copy number is then a random variable $M(t)$, which can assume positive integer values m. These interactions can be written using chemical reaction notation as:

$$\emptyset \xrightarrow{k} M$$

$$M \xrightarrow{\gamma m} \emptyset$$

From a deterministic perspective, the mean mRNA copy number per cell across a population can be described with the differential equation:

$$\frac{dM}{dt} = k - \gamma M$$

At steady state, $dM/dt = 0$ and hence the mean mRNA copy number is then given by:

$$M^{ss} = k/\gamma$$

This result gives the mean mRNA per cell as a ratio of synthesis and decay rates. Note that this mean value does not necessarily represent the number of mRNA in any given cell. It is just the average expected value of mRNA at steady state across the population.

In a stochastic context, we are concerned with finding the distribution of mRNA numbers across a population of cells. That is, we want to document the number of mRNA molecules in individual cells, and use this information to determine how many cells in a population are expected to contain a given number of mRNA molecules. To do this, we begin by writing an equation governing the time evolution of $p(m, t)$, the probability that $M(t) = m$. We can start with $p(m, t + dt)$, the probability that the system achieves m mRNA molecules at time $t + dt$. This probability is intuitively computed by enumerating the number of scenarios through which this outcome could be achieved. For example, the system could achieve m molecules at time $t + dt$ if it had $m - 1$ molecules at time t, and then one molecule is transcribed during time interval dt. This probability is simply given by $P(m - 1, t)kdt$. Similarly, the probability that the system has $m + 1$ molecules and loses one by degradation in time dt is given by $P(m + 1, t)\ (m + 1)\gamma dt$, whereas the probability of the system to have exactly m mRNA molecules at time t and not lose or gain any additional molecules in the time interval dt is given by $P(m, t)(1-kdt)(1-m\gamma dt)$. As a result, $P(m, t + dt)$ can be written as:

$$\begin{aligned} P(m, t + dt) = {} & P(m - 1, t)kdt + P(m + 1, t)(m + 1)\gamma dt + P(m, t) \\ & (1 - kdt)(1 - m\gamma dt) \end{aligned} \tag{1}$$

Multiplying out and rearranging terms in Eq. (1), we get:

$$\begin{aligned} P(m, t + dt) - P(m, t) = {} & P(m - 1, t)kdt + P(m + 1, t)(m + 1)\gamma dt \\ & -P(m, t)(K + m\gamma)dt + \varphi(dt^2) \end{aligned} \tag{2}$$

Dividing Eq. (2) by dt and taking the limit as $dt \rightarrow 0$, we get:

$$\frac{d}{dt}P(m, t) = kP(m - 1, t) + (m + 1)\gamma P(m + 1, t)dt - (K + m\gamma)P(m, t)dt \tag{3}$$

Eq. (3) is known as the CME. Although the derivation of the CME was illustrated for this specific example, similar derivations can be done for any biomolecular network described by a system of chemical reactions. Below, we provide a general formulation of the CME.

B. The General Formulation for Building Discrete Stochastic Models for Biomolecular Networks Using the Chemical Master Equation

In this section we describe the discrete state, continuous time Markov process model for well-stirred chemical reaction systems. First, we consider a system of chemical reactions with N molecular species (S_1, S_2, .., S_N) occurring in a volume V. We make two key assumptions. The first is that the system is well-mixed, that is the probability of finding any molecule in the volume V is given by dV/V. In many biological systems this is a reasonable assumption. For example, the length of a bacterial cell is around 1 μm and the diffusion coefficient of a protein *in vivo* has been measured to be on the order of 10 μm/s. Therefore, complete mixing of the bacterial cytosolic protein pool can possibly occur on the milisecond to second time scale (Konopka *et al*., 2006). However, the diffusion constant of many proteins moving in 2D on membranes may be much less than the area over which reactions occur, causing local depletion or enrichment of chemical species that renders the well-mixed assumption invalid (Vrljic *et al*., 2002). The second assumption we make is that the system is at thermal equilibrium. As a result, the velocity v of a molecule moving due to thermal energy is given by the Boltzman distribution:

$$f = \sqrt{\frac{m}{2\pi k_B T}} e^{-(m/2k_B T)v^2}$$

where T is the constant system temperature. We use the state $X(t) \in Z_+^N$ to denote the vector whose elements $X_i(t)$ are the number of molecules of the ith species at time t. If there are M elementary chemical reactions that can occur among these N species, then we associate with each reaction r_j ($j = 1, \ldots M$) a nonnegative *propensity function* a_j defined such that $a_j(X(t))\tau + o(\tau^2)$ is the probability that reaction r_j will happen in the next small time interval $(t, t + \tau)$ as $\tau \to 0$. The polynomial form of the propensities $a_j(x)$ may be derived from fundamental physical principles under certain assumptions (Gillespie, 1977). If r_j is the unimolecular reaction $S_1 \to product$, then a quantum mechanical argument dictates the existence of some constant c_j such that $c_j dt$ gives the probability that any particular S_1 molecule will transform into product in the next infinitesimal time dt. If there are currently n_1 such S_1 molecules in the system, then the probability that one of them will undergo the reaction in the next dt is $n_1 c_j dt$. Therefore, the propensity function of this unimolecular reaction is $a_j = n_1 c_j$. By contrast, if r_j is a bimolecular reaction of the form $S_1 + S_2 \to product$, then kinetic arguments can be used to assert the presence of a constant c_j such that $c_j dt$ is the probability that a randomly chosen pair of molecules S_1 and S_2 will react in the next infinitesimal time interval dt. Therefore, if n_1 molecules of S_1 and n_2 molecules of S_2 exist in volume V, then a reaction r_j will occur in the next dt with a probability $a_j dt = n_1 n_2 c_j dt$ (a_j is again called the propensity function of this reaction). Propensity functions for different types of reactions are summarized in Table I.

Table I
Stochastic Reaction Propensities

Reaction	Propensity $a_j(x)$
$\emptyset \xrightarrow{c_j}$ product	c_j
$S_i \xrightarrow{c_j}$ product	$c_j n_i$
$S_i + S_j \xrightarrow{c_j}$ product	$c_j n_i n_j$
$S_i + S_i \xrightarrow{c_j}$ product	$c_j \frac{n_i(n_i-1)}{2}$

The occurrence of a reaction r_j leads to a stoichiometric change of ϑ_j for the state X of the reactants involved. ϑ_j is therefore a stoichiometric vector that reflects the integer change in reactant species due to a reaction r_j.

It is useful to define these quantities:

Probability that reaction r_j fires one in $[t, t + dt] = a_j(x)dt + O(dt^2)$

Probability that no reaction in the system fires in $[t, t + dt] = 1 - \sum_{j=1}^{M} a_j(x)dt + O(dt^2)$

Probability that more than one reaction fires in $[t, t + dt] = O(dt^2)$

As in the simple example above, the CME for this system can be written by inspection using these quantities. Specifically, the probability of achieving state $X = x$ at time $t + dt$, $p(x, t + dt)$, is the sum of the following terms:

$$p(x,t+dt) = p(x,t)\left[1 - \sum_{j=1}^{M} a_j(x)dt + O(dt^2)\right] + \sum_{j=1}^{M}[p(x-\vartheta_j,t)a_j(x-\vartheta_j)dt + O(dt^2)] + O(dt^2) \tag{4}$$

The first term in Eq. (4) is simply the probability that the system was already in state x in terms of the number of its molecules for different species, and remained in that state with no reactions occurring during dt. The second term is the probability that the system was a ϑ_j step away from state x, and then was brought into that state by the occurrence of a reaction. Obviously, one has to account for all the reactions that can drive the system into that state, hence the summation.

Rearranging Eq. (4) we obtain:

$$p(x,t+dt) - p(x,t) = -p(x,t)\sum_{j=1}^{M} a_j(x)dt + \sum_{j=1}^{M}[p(x-\vartheta_j,t)a_j(x-\vartheta_j)dt] + O(dt^2) \tag{5}$$

Dividing Eq. (5) by dt and taking the limit as $dt \to 0$ gives the differential form

$$\frac{dP(x,t)}{dt} = \sum_{j=1}^{M} a_j(x-\vartheta_j)P(x-\vartheta_j) - a_j(x)P(x,t) \tag{6}$$

Eq. (6) is the CME for a general set of chemically reacting species in a constant, well-stirred volume.

C. Stationary Solutions of the CME

The stationary (steady state) distribution of the CME is solved for by setting $dP(x, t)/dt = 0$. For the simple model of transcription described by the CME in Eq. (3), this translates to: $kp(m-1) + (m+1)\gamma p(m+1) = (K + m\gamma)p(m)$

Solution of this balance equation can be done by induction. We observe that:

$$kp(0) = \gamma\, p(1)$$

$$kp(1) = 2\gamma p(2)$$

$$kp(m-1) = m\gamma p(m)$$

As a result, $p(m)$ can be expressed as a function of $p(0)$ as:

$$p(m) = \left(\frac{k}{\gamma}\right)^m \frac{1}{m!} p(0) \tag{7}$$

We can solve for $p(0)$ from Eq. (7) using the fact that $\sum_m p(m) = 1$. Therefore, $1 = \sum_m (k/\gamma)^m \frac{1}{m!} p(0) = e^{k/\gamma} p(0)$. As a result, $p(0) = e^{-(k/\gamma)}$ and $p(n) = e^{-a(a^m/m!)}$ with $= k/\gamma$. This corresponds to a Poisson distribution with equal mean and variance $\mu = \sigma^2 = a$.

This model has recently been validated using RNA fluorescence *in situ* hybridization (FISH) for ~100 well-expressed bacterial genes. These measurements conformed reasonably well to the predicted Poisson distribution, showing a relationship $\mu = 1.6\sigma^2$ (in contrast, protein expression in *Saccharomyces cerevisiae* scales as $\mu = 1200\sigma^2$ (Bar-Even *et al.*, 2006)). However, the subtle quantitative deviation from the Poisson relationship also suggested that other processes beyond simple production/degradation model might be at play to account for all the variability occurring in bacterial gene expression (Taniguchi *et al.*, 2010).

In general for a typical biological problem with several species and parameters neither the time evolution nor the stationary distribution described by the CME are analytically solvable. Therefore, one has to resort to numerical techniques to determine these quantities through sample path computations.

1. The Stochastic Simulation Algorithm: Generating Sample Paths

The approach here is to run a simulation describing the fluctuating behavior of a set of interacting chemical reactions in a single cell over time, and then to repeat this procedure multiple times to build an ensemble of behaviors across a population of cells.

To each of the chemical reactions $r_j (j = 1, \ldots, M)$ occurring among species $(S_1, S_2, \ldots, S_N)$ in a well-stirred volume, we attribute a random variable τ_j defined as the time to the firing of the next reaction r_j. Based on this formulation, τ_j is exponentially distributed with parameter $a_j(x)$ (a_j is the propensity function of this reaction). It can be shown that the time to the next reaction, defined as the random variable $\tau = \min\{\tau_j\}$, is exponentially distributed with parameter $\sum_{j=1}^{M} a_j(x)$. The random variable representing the index of the next reaction to occur $\mu = \text{argmin}\{\tau_j\}$ can also be shown to be uniformly distributed with $(\mu = j) = a_j(x) / \sum_{j=1}^{M} a_j(x)$. Using these quantities, one can then simulate the system with the four simple steps:

Initialize time t_0 and state x_0

Draw a sample $\hat{\tau}$ from $P(\tau)$, the distribution of τ

Draw a sample $\hat{\mu}$ from $P(\mu)$, the distribution of μ

Update time $t\, t + \hat{\tau}$ and state $x\, x + \hat{\mu}$ and repeat if final time is not reached.

This method is known as the SSA, and belongs to a wider class of numerical techniques known as Kinetic Monte Carlo algorithms. Every run of the algorithm above will generate a sample path of the stochastic process described by the CME (see for example Fig. 1(b)). To generate the probability distributions, one can run a large number of such sample paths.

D. Moment Computations

The CME is an equation for the probability distribution and can therefore be used in a straightforward manner to derive an expression for the evolution of the mean and higher order moments of these distributions. Simply put, for the first-order moment, $E(X_i)$, we can multiply the CME by x_i and then sum over all values of x. That is, $E[X_i] = \Sigma x_i p(x, t)$, and $(d\Sigma x_i p(x, t)/dt) = (dE[X_i]/dt)$. Similarly, for the second moment $E[X_i X_j]$, we can multiply the CME by $x_i x_j$ and sum over values of x. If we define $A(X) = [a_1(X), a_1(X), \ldots . a_M(X)]^T$ as the vector of propensity functions, and $S = [\vartheta_1 \vartheta_2 \ldots . \vartheta_M]$ as the stoichiometry matrix, then we can derive (using some straightforward algebraic manipulations that we will omit here) the following equations for the mean and second-order moments:

$$\frac{dE[X]}{dt} = SE[A(X)] \tag{8}$$

$$\frac{dE[XX^T]}{dt} = SE[A(X)X^T] + E[XA^T(X)]S^T + S\, diag(E[A(X)])S^T \tag{9}$$

1. Moment Equations for a System With Affine Propensities

An especially tractable form of the moment equations derived above arises when the propensity functions are affine, that is $A(X) = WX + w_0$, where W is an $N \times N$ matrix and w_0 is an $N \times 1$ vector. In this case, $E[A(X)] = W\,E[X] + w_o$ and

$E[A(X)X^T] = W\,E[XX^T] + w_o E[X^T]$. Replacing these expressions in Eqs. (8) and (9) above gives the moments equations:

$$\frac{dE[X]}{dt} = SWE[X] + Sw_o \tag{10}$$

$$\begin{aligned}\frac{dE[XX^T]}{dt} = {} & SWE[XX^T] + E[XX^T]W^T S^T + S\,diag(WE[X] + w_o)S^T \\ & + Sw_o E[X^T] + E[X]{w_o}^T S^T\end{aligned} \tag{11}$$

Eq. (11) is for the uncentered second moment. The covariance matrix (containing the centered second-order moments) is defined as $C = E[(X - E[X])(X - E[X])^T]$. Therefore, an expression for its time evolution can be derived by manipulation of Eqs. (11) and (12) to give:

$$\frac{dC}{dt} = SWC + CW^T S^T + S\,diag(WE[X] + w_o)S^T$$

The steady state means and covariances can be obtained by solving the linear algebraic equations corresponding to setting $(dE[X]/dt) = 0$ and $(dC/dt) = 0$. Let $\overline{X} = lim_{t\to\infty} E[X(t)]$ and $\overline{C} = lim_{t\to\infty} C(t)$. Then,

$$SW\,\overline{X} = -Sw_o \tag{12}$$

$$SW\,\overline{C} + \overline{C}W^T S^T + S\,diag(W\,\overline{X} + w_o)S^T = 0 \tag{13}$$

Now, if we define $M = SW$, $B = S\sqrt{diag(W\,\overline{X} + w_o)}$, and $D = BB^T$, then the steady state covariance given by Eq. (13) becomes

$$M\overline{C} + \overline{C}M^T + D = 0$$

This is the well-known Lyapunov equation, which characterizes the steady state covariance of the output of the linear dynamical system

$$\frac{dY}{dt} = MY + Bw$$

where w is the unit intensity white Gaussian noise.

E. An Example Where Calculations of Means and Covariances Generated Rich Biological Insight

Consider as extension of our initial model of transcription to include translation of a protein product from an mRNA (Figure 3). mRNA and protein can also decay with first-order kinetics. The simplest representation of this module contains four biochemical reactions:

$$R_1 : \emptyset \xrightarrow{k_r} mRNA$$

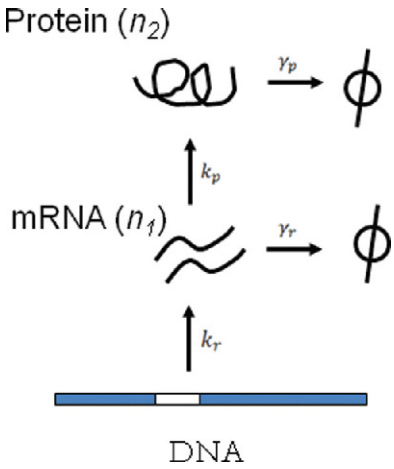

Fig. 3 Simple transcription and translation module. (For color version of this figure, the reader is referred to the web version of this book.)

$$R_2 : mRNA \xrightarrow{\gamma_r} \emptyset$$

$$R_3 : mRNA \xrightarrow{k_p} protein + mRNA$$

$$R_4 : protein \xrightarrow{\gamma_p} \emptyset$$

If we denote the number of molecules of mRNA by $X_1(t)$ and that of the protein by $X_2(t)$, then $X(t) = [X_1(t)X_2(t)]^T$. Also, the stoichiometry matrix is given by:

$$S = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}$$

Although the propensity vector is given by:

$$A(X) = \begin{bmatrix} k_r \\ \gamma_r X_1 \\ k_p X_1 \\ \gamma_p X_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ \gamma_r & 0 \\ k_p & 0 \\ 0 & \gamma_p \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} + \begin{bmatrix} k_r \\ 0 \\ 0 \\ 0 \end{bmatrix} = WX + w_0$$

Therefore, $M = SW = \begin{bmatrix} -\gamma_r & 0 \\ k_p & -\gamma_p \end{bmatrix}$ and $Sw_0 = \begin{bmatrix} k_r \\ 0 \end{bmatrix}$. As a result, the steady state as given by Eq. (12) is: $\overline{X} = -M^{-1}Sw_0 = \begin{bmatrix} \frac{k_r}{\gamma_r} \\ \frac{k_p k_r}{\gamma_p \gamma_r} \end{bmatrix}$.

The steady-covariance matrix can also be computed using Eq. (13). Specifically,

$$BB^T = S\,diag(W\overline{X} + w_o)S^T = \begin{bmatrix} 2k_r & 0 \\ 0 & \frac{2k_pk_r}{\gamma_r} \end{bmatrix}$$

As a result, the steady state covariance matrix $\overline{C}$ is given by:

$$\overline{C} = \begin{bmatrix} \frac{k_r}{\gamma_r} & \frac{k_pk_r}{\gamma_r(\gamma_r+\gamma_p)} \\ \frac{k_pk_r}{\gamma_r(\gamma_r+\gamma_p)} & \frac{k_pk_r}{\gamma_p\gamma_r}(1+\frac{k_p}{\gamma_r+\gamma_p}) \end{bmatrix} \tag{14}$$

Notice that for the mRNA in Eq. (14), we have exactly recapitulated the result derived based on the exact solution of the CME above, namely that its stationary distribution has an equal mean and variance given by k_r/γ_r. The mean of the protein is given by: $\overline{X}_2 = (k_pk_r/\gamma_p\gamma_r)$, while its variance is $\overline{C}_{22} = (k_pk_r/\gamma_p\gamma_r)(1 + (k_p/(\gamma_r + \gamma_p)))$. Therefore, the coefficient of variation for the protein (a unitless quantity to be intuitively thought of as a normalized standard deviation) is given by:

$$CV = \frac{\sqrt{\overline{C}_{22}}}{\overline{X}_2} = \frac{1}{\sqrt{\overline{X}_2}}\left(1+\frac{k_p}{\gamma_r+\gamma_p}\right)^{1/2} \tag{15}$$

This equation confirms our intuition that as the number of molecules increases, the CV ("noise") of the system would decrease. Most importantly, it assigns a very specific pattern for this decrease in that it should follow an inverse square-root function of the mean with a scaling constant dependant on the translation rate of the mRNA and decay rates of the protein and mRNA. Experimental investigations of noise in gene expression of a large set of genes in the yeast *S. cerevisae* and bacterium *E. coli* subsequently confirmed this prediction (Newman *et al*., 2006a). However, does a large $\overline{X}_2$ necessarily imply a small CV?

Notice that:

$$\begin{aligned} CV^2 &= \frac{1}{\overline{X}_2}\left(1+\frac{k_p}{\gamma_r+\gamma_p}\right) = \frac{1}{\frac{k_pk_r}{\gamma_p\gamma_r}}\left(1+\frac{k_p}{\gamma_r+\gamma_p}\right) \geq \frac{1}{\frac{k_pk_r}{\gamma_p\gamma_r}}\left(\frac{k_p}{\gamma_r+\gamma_p}\right) \\ &= \frac{\gamma_p\gamma_r}{k_r}\cdot\frac{1}{\gamma_r+\gamma_p} \end{aligned} \tag{16}$$

Therefore, for some values of γ_r, γ_p, and γ_r, CV^2 in Eq. (16) can be arbitrarily large. Simultaneously, through choice of k_p, $\overline{X}_2 = (k_pk_r/\gamma_p\gamma_r)$ can be set independently of CV^2 to be arbitrarily large. Therefore, large mean does NOT necessarily imply small fluctuations. This model of gene expression predicts that decreased translation rates should decrease noise in gene expression, a result that was confirmed experimentally (Ozbudak *et al*., 2002). More generally this framework suggests cellular contexts where noise might be expected to be particularly problematic. For example, this

model predicts that when proteins are rapidly degraded and expressed at low copy number, such as the cyclins in the cell cycle, high variability would ensue. Given this insight, many recent investigations of the cell cycle focused precisely on what control strategies implemented through interlinked positive and negative feedback loops can compensate for this effect to provide robust noise free oscillations (Tsai *et al*., 2008).

Exceptions aside, Eq. (15) and some of its variations have guided many investigations that delineated fundamental properties of noise in gene expression. Researchers have used this equation to infer promoter, mRNA and protein dynamics based on snapshots of protein distributions (see Paulsson (2005) for a review). Furthermore, these analyses proved particularly useful in describing the effect of chromatin features on gene expression. In one recent such study, a viral vector was used to integrate a green fluorescent protein (GFP) reporter construct randomly in a mammalian cell line and the CV of each integrant was measured. Fitting the data to a two-state gene expression model similar to Eq. (16), with the addition that a promoter can transition between OFF and ON states, suggested that the chromatin state of the integration site affects the stability and productivity of the ON state, but not the frequency of activation (Skupsky et al., 2010). It is worth noting here that these static snapshots of noise in gene expression are not always sufficient to resolve all the parameters involved in the process. For example, in the study mentioned above, these distributions were sufficient to determine the promoter activation frequency but not its active duration. Dynamic measurements might be necessary to resolve such parameters.

F. Linearization of Macroscopic Dynamics and the Linear Noise Approximation: Computing Approximate Moments for Nonlinear Propensity Functions

Although computation of first and second moments at steady state could be done using an algebraic equation when the propensity functions that appear in the CME are affine, no such calculation is possible when these propensity functions are nonlinear as is the case for many biological reactions. The reason is rather simple; close inspection of Eqs. (10) and (11) reveals that in this case, every moment depends on higher order moments, resulting in an infinite hierarchy of ODEs to solve. The Linear Noise Approximation (LNA) is a procedure to truncate this hierarchy. Before we present the LNA, we review selected parts of the standard treatment of linearized dynamics around a steady state (Strogatz, 1994). First, we remind the reader that the system of reaction rate equations describing the macroscopic behavior of the concentration of N biochemical species interacting through a set of M biochemical reactions is given by the coupled ODEs (Cornish-Bowden, 1979):

$$\frac{dx}{dt} = SA(x) \tag{17}$$

where $x(t) = [x_1(t)x_2(t) \ldots x_N(t)]^T$ is the vector of macroscopic concentrations, and S is the $N \times M$ stoichiometry matrix. If a steady state $\overline{x}$ exists for the macroscopic dynamics, it follows from solving the algebraic system of equations:

$$0 = SA(\overline{x})$$

Linearization of Eq. (17) around the steady state vector $\overline{x} = [\overline{x}_1 \overline{x}_2 \ldots \overline{x}_n]^T$ leads to a matrix equation for the deviations $\delta x = [\delta x_1 \delta x_2 \ldots \delta x_N]^T$ from $\overline{x}$ given by:

$$\frac{d}{dt}\delta x = M\delta x$$

M is the Jacobian matrix, with the elements:

$$M_{ij} = \frac{\partial[S_i A(x)]}{\partial x_k}|_{x=\overline{x}}$$

Therefore, in compact notation $M = S\frac{\partial A(x)}{\partial x}|_{x=\overline{x}}$.

Going back to the stochastic representation, we assume that the distribution of the chemical species is tightly distributed around its mean. We also assume that $x(t) = (X(t)/V)$ (where $X(t)$ is the mean of the distribution) is identical to the solution $\varphi(t)$ of the reaction rate equations (Eq. (17)) that describe the macroscopic concentrations of molecular species in the system. Notice that $\varphi(t)$ is a vector of concentrations, while $X(t)$ is a vector containing the number of molecules, hence the need for a volume scaling factor V.

More formally, let $X(t) = V\varphi(t) + \varepsilon(t)$, where $\varepsilon(t)$ is the zero mean random variable denoting the deviation from the deterministic term $V\varphi(t)$ (Tomioka *et al.*, 2004). Expanding in Taylor series around $\varphi(t)$ in Eq. (10), we get

$$\frac{dE[X]}{dt} = \frac{dV\varphi}{dt} + \frac{dE[\varepsilon]}{dt} = VSA(\varphi) + S\frac{\partial A(Vx)}{\partial Vx}|_{x=\varphi}\, E(\varepsilon) + O(\varepsilon^2) \qquad (18)$$

The assumptions on the distributions imply that $O(\varepsilon^2)$ can be neglected in Eq. (18) above. Therefore, recovering the equation: $\frac{d\varphi}{dt} = SA(\varphi)$. Furthermore, we obtain:

$$\frac{dE[\varepsilon]}{dt} = S\frac{\partial A(Vx)}{\partial Vx}|_{x=\varphi}\, E[\varepsilon]$$

Rewriting Eq. (11) similarly in terms of Taylor series expansion and truncating the $O(\varepsilon^2)$ terms generates the following equation for the time evolution of the noise covariance matrix $C_\varepsilon = E[\varepsilon\varepsilon^T] - E[\varepsilon]E[\varepsilon^T]$:

$$\frac{dC_\varepsilon}{dt} = M(\varphi)C_\varepsilon + C_\varepsilon M^T(\varphi) + D(V\varphi) + \frac{\partial D(Vx)}{\partial Vx}|_{x=\varphi}\, E[\varepsilon] \qquad (19)$$

In Eq. (19), we defined $M = S(\partial A(Vx)/\partial Vx)|_{x=\varphi}$ as the Jacobian matrix and $D = S\,diag[A(V\varphi)]S^T$. Now, we have the closed simultaneous questions for the time evolution of mean and covariance of the random fluctuations around the macroscopic solution. We assume that the macroscopic solution is stable around $\varphi(t)$. That is, the eigenvalues of the Jacobian matrix M are negative for all t. This assumption is necessary to justify the linearization.

We also assume that the macroscopic rate equations converge to a stable steady state $\overline{\varphi}$. Under these assumptions, there exists a distribution around $\overline{\varphi}$ with mean $E[\varepsilon] = 0$ and covariance matrix C_ε that satisfies the following equation:

$$M(\overline{\varphi})\overline{C} + \overline{C}M^T(\overline{\varphi}) + D(V\overline{\varphi}) = 0 \tag{20}$$

Notice again that Eq. (20) is a Lyapunov equation, with $M(\overline{\varphi})$ being the Jacobian matrix obtained by linearizing the system around its macroscopic steady state.

In summary, one can obtain the covariance matrix of this distribution around a macroscopic steady state by taking the following simple procedure:

Find the stoichiometry matrix S and the propensity vector $A(X)$

Find a stable equilibrium of the reaction rate equations of the system

Calculate two matrices $M(\overline{\varphi})$ and $D(V\overline{\varphi})$

Solve Lyapunov equation (Eq. (20))

Above, we have presented a multivariable and compact derivation of the LNA. Multiple forms of this derivation exist under alternative names such as the system size expansion (Elf and Ehrenberg, 2003; Kampen, 1992).

Due to its minimal computation costs, the LNA makes rapid analytical investigation of noise features for different models and parameter sets possible. For example, LNA analysis of all possible three node networks over a wide range of parameter sets has recently been used to show that both positive and negative feedback motifs can buffer noise from an upstream node, but that only positive feedback loops can do so while maintaining network responsiveness. This insight was confirmed by a detailed analysis of nitrogen metabolism in yeast, which suggested that coupled positive and negative feedback in this system may indeed act to buffer noise (Hornung and Barkai, 2008).

G. Other Closure Techniques for the Moment Equations

As discussed above, the solution to the CME can be expanded in a Taylor series about the macroscopic deterministic trajectory. The first-order terms correspond to the macroscopic rate equations, and the second-order terms approximate the system noise. Variations on this procedure exist. For example, mass fluctuations kinetics (MFK) calculations take a similar approach to the LNA except that the computation of the mean is coupled with that of the variances (Gomez-Uribe and Verghese, 2007). Therefore, the MFK approach allows one to capture situations where the mean of the stochastic distributions may deviate from the solution of the macroscopic rate equations. This is particularly important for systems that exhibit emergent stochastic phenomena such as, for example, excitability (Suel *et al*., 2006, 2007) and stochastic resonance or focusing (Paulsson *et al*., 2000).

Other moment closure techniques proceed by assuming specific probability distributions for the underlying stochastic processes, and then using this assumption to express higher order moments as a function of the lower order ones to effectively truncate the dynamics. This has been done for well-known classes of distributions, such as normal (Whittle, 1957), lognormal (Keeling, 2000), Poisson binomial

(Nasell, 2003). Moment closure techniques that do not make explicit assumptions about the shape of the distribution also exist. One such moment closure approximation known as separable derivative matching (Singh and Hespanha, 2007) approximates the $(N+1)$th moment as a polynomial function of the first N moments. This approach matches time derivatives between the approximate closed system and the exact nonclosed system at the initial time t_0 and the given initial conditions. This allows the exponents (which remain constant over the simulation) in the polynomial function to be uniquely determined, and the solution turns out to be consistent with the underlying distribution probability distribution being lognormal. It is worth noting here that the derivation of the moment equations implicitly assumes the presence of a single macroscopic steady state. Hence, the distributions are unimodal and the process is well characterized by the first few moments. However, problems that exhibit multimodal distributions will require many higher order moments, and the applicability of these methods may quickly degrade. Usually, the choice between accurate numerical approaches and approximation analytical approaches (such as the LNA and moment closure techniques) is done on a case-by-case basis to balance computational cost versus accuracy.

VI. Open Challenges

Stochastic modeling of biological dynamics, especially at the cellular level, is increasingly making its way to the mainstream of quantitative biology investigation. The CME and its accompanying SSA have proven to be invaluable computational tools for such studies. There are, however, many challenges that need to be addressed in order to make stochastic modeling a widely applicable tool for realistic biological problems. Below, we discuss some of these challenges and recent developments in the literatures to address them.

A. Efficient Stochastic Simulation and Analysis for Systems Evolving at Disparate Temporal and Spatial Scales

For many cellular networks of biological importance, the chemical reactions occur at significantly different rates. As a motivating example, consider gene regulation in the bacterium *Escherichia coli.* There, a typical time scale for mRNA transcription is on the order of minutes, whereas the time scale for protein degradation/dilution is on the order of an hour (Alon, 2007). This suggests that the protein concentrations do not depend strongly on the instantaneous number of mRNAs but rather on their average over time. Even more drastically, posttranslational modifications of the protein (e.g., phosphorylation) often occur on the time scale of seconds. These disparate time scales in the chemical reactions pose great challenges for efficient numerical simulation of these processes. These challenges arise from having to resolve the stochastic dynamics on the fastest characteristic time scales of the

system. Take for example a model in which a kinase activates a transcription factor by phosphorylating it, while a phosphotase removes the phosphate. We are interested in understanding the fluctuations in the expression of the gene that is regulated by the transcription factor. It is often the case that the competing phosphorylation and dephosphorylation reactions occur rapidly (fast reactions), whereas gene expression is relatively slow. In this situation a stochastic simulation of the system will spend most of its computational time fruitlessly adding and removing phosphates from the transcription factor and relatively little time on reactions that result in gene expression, our actual interest.

Multiple approaches have emerged to address this problem. On the analytical side, the strategy is often to derive reduced models by explicitly representing the chemical species having dynamics with relatively slow characteristic time scales while eliminating representations of the chemical species having dynamics with relatively fast characteristic time scales (Atzberger *et al*., 2011; Cao *et al*., 2005; Haseltine and Rawlings, 2002; Rao and Arkin, 2003). Roughly speaking, these methods parallel quasi-steady state approximations for deterministic chemical kinetics where a subset of species is assumed to be asymptotically at steady state on the time scale of interest. One commonly used example is the Hill function (a[TF/(TF + Kd)]), which describes the expression of a gene for a given concentration of a transcription factor (*TF*), affinity of the transcription factor for the promoter (K_d), and maximal activation (a). This expression is derived using the assumption that transcription factor binding and unbinding events are rapid relative to the rate of gene expression, and so one can approximate them as an average occupancy rather than explicitly model every individual event (Nemenman *et al*., 2009).

On the numerical side, several approximate methods have been developed to speed up simulations while sacrificing some of the exactness of the SSA. The basic idea behind these approximate methods is that instead of simulating a single reaction per step, a number of reactions can occur in each simulation step. These approximate methods are known as leap methods including the τ-leap method (Gillespie, 2001; Gillespie and Petzold, 2003), the binomial τ-leap method (Chatterjee *et al*., 2005; Rathinam and El Samad, 2007), and the K-leap method (Cai and Xu, 2007).

Despite such productive work on the subject, the efficient analysis and simulation of stochastic cellular dynamics for realistic problems is still very difficult. For example, there is little theory that can provide reassurance about the accuracy of the approximate SSAs in challenging scenarios. Furthermore, quasi-steady state approximations of stochastic fast scales are done based on intuition and assumptions derived from deterministic chemical kinetics. For these methods to be broadly applicable, they need to be placed on more solid theoretical footing in terms of the assumptions that can and cannot be made in a stochastic context and rigorous proofs need to be generated for their accuracy in different realistic contexts.

The holistic understanding of biological systems often involves the probing of cellular biochemical networks in the context of the cell, of cells in the context of a tissue, and of a tissue in the context of the organism. How to account for and move between these spatial scales remains an open problem for stochastic modeling. This

"multiscale" problem is of poignant relevance to pharmacological studies, which need to integrate effects of small molecules therapies at the single cell level with global metabolic processes within the body such as prodrug activation, degradation of the active molecules, and off-target toxicities (Eissing *et al*., 2011).

B. Efficient Spatiotemporal Simulations

Previous sections cover the stochastic algorithms for modeling biological pathways with no spatial information. However, biological networks in practice consist of components that interact in a three-dimensional space and are not necessarily distributed homogeneously as they diffuse between different cellular compartments. For example, even within *E. coli* (the prototypical cell-as-a-bag modeling system) membrane invaginations can dramatically alter the diffusive properties of molecules (Weisshaar *et al*., 2006). In eukaryotic neuronal cells, axons can be meters long raising immense barriers to diffusive mixing. Thus, the basic assumption of spatial homogeneity and large concentration diffusion may be challenged in some biological systems. In this context, stochastic spatiotemporal representations are required.

Roughly speaking, discrete spatial stochastic simulations can be separated into lattice and off-lattice particle based methods. In off-lattice methods, the Brownian movements of the individual molecules are accounted for and all particles in the system have explicit spatial coordinates (Bartol, 2002). At each time step, molecules with nonzero diffusion coefficients are able to move, in a random walk fashion, to new positions. In this case, the motion and direction of the molecules are determined by using random numbers during the simulation. Similarly, collisions with potential binding sites and surfaces are detected and handled by using only random numbers with a computed binding probability. Particle methods can provide very detailed simulations of highly complex systems at the cost of exceedingly large amounts of computational effort.

For lattice methods, the two- or three-dimensional volume used to represent a cellular compartment (organelles or membranes) is covered by a computational mesh (Morton-Firth and Bray, 1998; Schnell *et al*., 2004). The lattice is then "populated" with particles of the different molecular species that comprise the system. Particles with nonzero diffusion coefficient are able to diffuse by jumping to an empty neighboring domain. If the domain is assumed to accommodate only one molecule, chemical reactions can take place with a certain probability among molecules in adjacent domains. Another scenario is one in which subvolumes can host many molecules, with well-mixedness assumed in each subvolume. In both cases, diffusion steps are treated as treated first-order reactions, with a reaction rate constant proportional to the diffusion coefficient (Ander *et al*., 2004; Baras and Mansour, 1996; Elf *et al*., 2010; Stundzia and Lumsden, 1996). As a result, diffusion can be treated as an additional chemical reaction, and one is back to the SSA formalism.

Many caveats of these methods exist. For example, the artificial nature of the lattice may introduce lattice anisotropy (Ridgway *et al*., 2009). Furthermore, in many physiologically relevant situations, molecular crowding can prevent reacting molecules from reaching regions of the domain due to the high concentration of macromolecules impeding their passage (Ridgway *et al*., 2009). A particularly striking example of this is diffusive motion in the context of the eukaryotic nucleus where densely packed nucleoli and hetrochromatin structures greatly reduce diffusive rates, suggesting one mechanism whereby heterchormatin prevents active transcription (Bancaud et al., 2009). Therefore, despite their conceptual appeal, these spatiotemporal algorithms need to be updated to capture the full scope of biological reality. Furthermore, even in their current approximate forms, these algorithms require substantial and sometimes prohibitive computational power and have only been successfully applied to small systems with finite number of molecular species. As a result, many computational innovations are still needed to enable the quantitative probing of the spatial stochastic dynamics of biological systems.

C. Parametrization and Sensitivity Analysis of Stochastic Models

Stochastic models of biological systems typically depend on a set of kinetic parameters whose values are often unknown or fluctuate due to an uncertain environment. These parameters determine the dynamic behavior of the model, and changes in them may alter the system's output in nonintuitive ways. Typically, many of the parameters in a biological system have not been measured or are unmeasurable. For example, a typical assay for measuring the affinity of a transcription factor for its promoter by gel shift will describe this interaction in terms of a disassociation constant (K_d), which gives the ratio of binding and unbinding rates. A stochastic model, however, requires explicit ON and OFF rates that are rarely available. In this case, one strategy would be to estimate the ON and OFF rates under the assumption that binding of two molecules is "diffusion limited." However, a more commonly encountered situation is one in which no direct measurement exists from which to base a choice of parameters. In this case, it becomes imperative to establish that specific choices for the value of these parameters do not substantially change the model behavior of interest.

Assessing the change in a system output pursuant to perturbations in its kinetic parameters is carried out using sensitivity analysis. Traditionally, the concept of sensitivity analysis has been applied largely to continuous deterministic systems, for example, systems described by differential (or differential-algebraic) equations. Much of these analyses have focused on the effects of infinitesimal perturbations of certain parameters. In deterministic chemical kinetics, the infinitesimal sensitivities are represented using the first-order sensitivity coefficients, given by (Varma *et al*., 2005):

$$S_{ij}(t) = \frac{\partial x_i(t)}{\partial \theta_j} \tag{21}$$

where x_i denotes that i^{th} output of the system at time t (e.g., the concentration of chemical species as given by Eq. (17)) and θ_j is the jth parameter. This equation assumes implicitly that the output $\boldsymbol{x}_i$ is continuous with respect to the parameter θ_j. Using the definition in Eq. (21), dynamic evolution equations can be derived for $S_{ij}(t)$ and solved along with the original system equation. In the context of biological systems modeling, sensitivity analysis has been indispensable to deduce important system properties, such as robustness in an uncertain environment (Stelling *et al*., 2004). In large networks, sensitivity analysis can pinpoint critical or rate limiting pathways and aid in reduced order modeling. Despite their usefulness, these sensitivities report on changes of model behavior changes as parameters change *locally*, but do not address the outcome of large changes to parameters or simultaneous perturbations to multiple parameters. Assessing the effect of large perturbations is typically carried out numerically by recomputing the reaction rate equations for the perturbed parameter values and comparing these to the nominal parameter values.

The most common approach for sensitivity analysis in stochastic systems resembles the simulation-based strategy. Monte Carlo (SSA) simulations are run for various values of the parameter whose sensitivity is of interest, and the variation in the outcome of these simulations for a variable of interest, such as mean, quantified. The sensitivity at time T to a finite perturbation h of a parameter θ about its nominal value $\theta = \theta_0$ can be computed via a finite difference of the expected value, such as

$$S = \frac{E[X(T, \theta_0 + h)] - E[X(T, \theta_0)]}{h}$$

Basically, one uses SSA to compute these expected values by generating many samples of $X(T, \theta_0 + h)$ and $X(T, \theta_0)$, usually using two independent streams of random numbers to generate samples of $X(T, \theta_0 + h)$ and $X(T, \theta_0)$. This is called the independent random number (IRN) approach and has been recently used in combination with the Fisher information matrix to generate several different sensitivity measures (Gunawan *et al*., 2005). Evidently, Monte Carlo simulations need to be carried out for the nominal and perturbed parameter value making this approach often computationally expensive. Furthermore, the use of IRNs usually results in a statistical estimator with large variance, thereby increasing the computational effort as large samples may be required. Recent work has shown that using the same stream of common random numbers (CRNs) to generate samples of $X(T, \theta_0 + h)$ and $X(T, \theta_0)$ can typically result in an estimator with low variance and thus requires far fewer samples (Rathinam *et al*., 2010). Approaches based on the Girasnov measure have also been proposed to smooth the sensitivity estimates and reduce their bias (Plyasunov and Arkin, 2006). Finally, more tractable but approximate approaches to computing sensitivities of stochastic models have also been formulated based on the LNA (Hornung and Barkai, 2008).

The application of sensitivity analysis, nonetheless, is still prohibitive for most realistic models of stochastic cellular networks. This problem is further compounded by the aforementioned challenge posed by large numbers of unknown model parameters, which need to be identified from data. Many parameter identifiability

analyses use the concept of sensitivity to determine *a priori* whether certain parameters can be estimated from experimental data and to search for these parameters using iterative algorithms. Efficient computation of parameter sensitivities is therefore a topic of great interest and bearing on the applicability of stochastic methods, and one where many challenges still lie ahead.

VII. Conclusions

Stochastic modeling methods are generating many important insights into the operation and organizational principles of cellular networks. Challenges remain before the full power of these methods can be unleashed in the study of many complex biological dynamics. This is an area of great promise, and one where progress will greatly deepen our understanding of the stochastic underpinnings of life.

References

Alon, U. (2007). *An Introduction to Systems Biology: Design Principles of Biological Circuits.* Chapman & Hall/CRC, Boca Raton, FL.

Ander, M., Beltrao, P., and Di Ventura, B., *et al*. (2004). SmartCell, a framework to simulate cellular processes that combines stochastic approximation with diffusion and localisation: analysis of simple networks. *Syst. Biol.* **1**, 129–138.

Atzberger, P. J., Pahlajani, C. D., and Khammash, M. (2011). Stochastic reduction method for biological chemical kinetics using time-scale separation. *J. Theor. Biol.* **272**, 96–112.

Avery, S. V., Smith, M. C. A., and Sumner, E. R. (2007). Glutathione and Gts1p drive beneficial variability in the cadmium resistances of individual yeast cells. *Mol. Microbiol.* **66**, 699–712.

Bancaud, A., Huet, S., Daigle, N., Mozziconacci, J., Beaudouin, J., and Ellenberg, J. (2009). Molecular crowding affects diffusion and binding of nuclear proteins in heterochromatin and reveals the fractal organization of chromatin. *EMBO J.* **28**(24), 3785–3798.

Bar-Even, A., Paulsson, J., Maheshri, N., Carmi, M., O'Shea, E., Pilpel, Y., and Barkai, N. (2006). Noise in protein expression scales with natural protein abundance. *Nat. Genet.* **38**, 636–643.

Baras, F., and Mansour, M. M. (1996). Reaction-diffusion master equation: a comparison with microscopic simulations. *Phys. Rev. E* **54**, 6139–6148.

Bartol, TM., Stiles, JR. (2002). MCell: A Monte Carlo Simulation of Cellular Physiology.

Bigger, W. B. (1944). Treatment of staphylococcal infections with penicillin by intermittent sterilization. *Lancet* **2**, 497–500.

Blake, W. J., Balazsi, G., Kohanski, M. A., Isaacs, F. J., Murphy, K. F., Kuang, Y., Cantor, C. R., Walt, D. R., and Collins, J. J. (2006). Phenotypic consequences of promoter-mediated transcriptional noise. *Mol. Cell* **24**, 853–865.

Cai, L., Friedman, N., and Xie, X. S. (2006). Stochastic protein expression in individual cells at the single molecule level. *Nature* **440**, 358–362.

Cai, X. D., and Xu, Z. Y. (2007). K-leap method for accelerating stochastic simulation of coupled chemical reactions. *J. Chem. Phys.* **126**.

Cao, Y., Gillespie, D. T., and Petzold, L. R. (2005). The slow-scale stochastic simulation algorithm. *J. Chem. Phys.* **122**, 14116.

Chatterjee, A., Vlachos, D. G., and Katsoulakis, M. A. (2005). Binomial distribution based tau-leap accelerated stochastic simulation. *J. Chem. Phys.* **122**.

Choi, P. J., Cai, L., Frieda, K., and Xie, S. (2008). A stochastic single-molecule event triggers phenotype switching of a bacterial cell. *Science* **322**, 442–446.

Chubb, J. R., Trcek, T., Shenoy, S. M., and Singer, R. H. (2006). Transcriptional pulsing of a developmental gene. *Curr. Biol.* **16**, 1018–1025.

Cluzel, P., Le, T. T., Harlepp, S., Guet, C. C., Dittmar, K., Emonet, T., and Pan, T. (2005). Real-time RNA profiling within a single bacterium. *Proc. Natl. Acad. Sci. U S A* **102**, 9160–9164.

Cornish-Bowden, A. (1979). *Fundamentals of Enzyme Kinetics.* Butterworths, London, Boston.

Eissing, T., Kuepfer, L., Becker, C., Block, M., Coboeken, K., Gaub, T., Goerlitz, L., Jaeger, J., Loosen, R., and Ludewig, B., *et al*. (2011). A computational systems biology software platform for multiscale modeling and simulation: integrating whole-body physiology, disease biology, and molecular reaction networks. *Front. Physiol.* **2**, 4.

Elf, J., and Ehrenberg, M. (2003). Fast evaluation of fluctuations in biochemical networks with the linear noise approximation. *Genome Res.* **13**, 2475–2484.

Elf, J., Fange, D., Berg, O. G., and Sjoberg, P. (2010). Stochastic reaction-diffusion kinetics in the microscopic limit. *Proc. Natl. Acad. Sci. U S A* **107**, 19820–19825.

Gillespie, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* **81**, 2340–2361.

Gillespie, D. T. (1992). A rigorous derivation of the chemical master equation. *Physica A* **188**, 404–425.

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

Gillespie, D. T., and Petzold, L. R. (2003). Improved leap-size selection for accelerated stochastic simulation. *J. Chem. Phys.* **119**, 8229–8234.

Golding, I., Paulsson, J., Zawilski, S. M., and Cox, E. C. (2005). Real-time kinetics of gene activity in individual bacteria. *Cell* **123**, 1025–1036.

Gomez-Uribe, C. A., and Verghese, G. C. (2007). Mass fluctuation kinetics: capturing stochastic effects in systems of chemical reactions through coupled mean-variance computations. *J. Chem. Phys.* **126**, .

Gregor, T., Tank, D. W., Wieschaus, E. F., and Bialek, W. (2007a). Probing the limits to positional information. *Cell* **130**, 153–164.

Gregor, T., Wieschaus, E. F., McGregor, A. P., Bialek, W., and Tank, D. W. (2007b). Stability and nuclear dynamics of the bicoid morphogen gradient. *Cell* **130**, 141–152.

Gunawan, R., Cao, Y., Petzold, L., and Doyle, F. J. (2005). Sensitivity analysis of discrete stochastic systems. *Biophys. J.* **88**, 2530–2540.

Haseltine, E. L., and Rawlings, J. B. (2002). Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics. *J. Chem. Phys.* **117**, 6959–6969.

Hasty, J., and Collins, J. J. (2002). Translating the noise. *Nat. Genet.* **31**, 13–14.

Hornung, G., and Barkai, N. (2008). Noise propagation and signaling sensitivity in biological networks: a role for positive feedback. *PLOS Comput. Biol.* **4**.

Kampen, N. (1992). *Stochastic Processes in Chemistry and Physics.* Elsevier, .

Karp, X., and Greenwald, I. (2003). Post-transcriptional regulation of the E/Daughterless ortholog HLH-2, negative feedback, and birth order bias during the AC/VU decision in *C. elegans*. *Genes Dev.* **17**, 3100–3111.

Keeling, M. J. (2000). Multiplicative moments and measures of persistence in ecology. *J. Theor. Biol.* **205**, 269–281.

Konopka, M. C., Shkel, I. A., Cayley, S., Record, M. T., and Weisshaar, J. C. (Sep 2006). Crowding and confinement effects on protein diffusion in vivo. *J Bacteriol.* **188**(17), 6115–6123.

Kussell, E., Kishony, R., Balaban, N. Q., and Leibler, S. (2005). Bacterial persistence: a model of survival in changing environments. *Genetics* **169**, 1807–1814.

McAdams, H. H., and Arkin, A. (1999). It's a noisy business! Genetic regulation at the nanomolar scale. *Trends Genet.* **15**, 65–69.

McAdams, H. H., Arkin, A., and Ross, J. (1998). Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected *Escherichia coli* cells. *Genetics* **149**, 1633–1648.

Mcquarri, D. A. (1967). Stochastic approach to chemical kinetics. *J. Appl. Probability* **4**, 413–477.

Morton-Firth, C. J., and Bray, D. (1998). Predicting temporal fluctuations in an intracellular signalling pathway. *J. Theor. Biol.* **192**, 117–128.

Moyed, H. S., and Broderick, S. H. (1986). Molecular-cloning and expression of *HIPA*, a gene of *Escherichia coli* K-12 that affects frequency of persistence after inhibition of murein synthesis. *J. Bacteriol.* **166**, 399–403.

Murphy, K. F., Balazsi, G., and Collins, J. J. (2007). Combinatorial promoter design for engineering noisy gene expression. *Proc. Natl. Acad. Sci. U S A* **104**, 12726–12731.

Nasell, I. (2003). An extension of the moment closure method. *Theor. Popul. Biol.* **64**, 233–239.

Nemenman, I., Sinitsyn, N. A., and Hengartner, N. (2009). Adiabatic coarse-graining and simulations of stochastic biochemical networks. *Proc. Natl. Acad. Sci. U S A* **106**, 10546–10551.

Newman, J. R., Ghaemmaghami, S., Ihmels, J., Breslow, D. K., Noble, M., DeRisi, J. L., and Weissman, J. S. (2006a). Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. *Nature* **441**, 840–846.

Newman, J. R. S., Ghaemmaghami, S., Ihmels, J., Breslow, D. K., Noble, M., DeRisi, J. L., and Weissman, J. S. (2006b). Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. *Nature* **441**, 840–846.

Ozbudak, E. M., Thattai, M., Kurtser, I., Grossman, A. D., and van Oudenaarden, A. (2002). Regulation of noise in the expression of a single gene. *Nat. Genet.* **31**, 69–73.

Paulsson, J. (2005). Models of stochastic gene expression. *Phys. Life Rev.* **2**, 157–175.

Paulsson, J., Berg, O. G., and Ehrenberg, M. (2000). Stochastic focusing: Fluctuation-enhanced sensitivity of intracellular regulation. *Proc. Natl. Acad. Sci. U S A* **97**, 7148–7153.

Plyasunov, S., and Arkin, A. P. (2006). Averaging methods for stochastic dynamics of complex reaction networks: description of multiscale couplings. *Multiscale Modeling Simulation* **5**, 497–513.

Raj, A., Peskin, C. S., Tranchina, D., Vargas, D. Y., and Tyagi, S. (2006). Stochastic mRNA synthesis in mammalian cells. *PLOS Biol.* **4**, 1707–1719.

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

Rao, C. V., Wolf, D. M., and Arkin, A. P. (2002). Control, exploitation and tolerance of intracellular noise. *Nature* **420**, 231–237.

Raser, J. M., and O'Shea, E. K. (2005). Noise in gene expression: origins, consequences, and control. *Science* **309**, 2010–2013.

Rathinam, M., and El Samad, H. (2007). Reversible-equivalent-monomolecular tau: a leaping method for "small number and stiff" stochastic chemical systems. *J. Comput. Phys.* **224**, 897–923.

Rathinam, M., Sheppard, P. W., and Khammash, M. (2010). Efficient computation of parameter sensitivities of discrete stochastic chemical reaction networks. *J. Chem. Phys.* **132**.

Ridgway, D., Broderick, G., Lopez-Campistrous, A., Ru'aini, M., Winter, P., Hamilton, M., Boulanger, P., Kovalenko, A., and Ellison, M. J. (2009). Coarse-grained molecular simulation of diffusion and reaction kinetics in a crowded virtual cytoplasm. *Biophys. J.* **96**, 2548.

Schnell, S., Turner, T. E., and Burrage, K. (2004). Stochastic approaches for modelling in vivo reactions. *Comput. Biol. Chem.* **28**, 165–178.

Serizawa, S., Miyamichi, K., Nakatani, H., Suzuki, M., Saito, M., Yoshihara, Y., and Sakano, H. (2003). Negative feedback regulation ensures the one receptor-one olfactory neuron rule in mouse. *Science* **302**, 2088–2094.

Singh, A., and Hespanha, J. P. (2007). A derivative matching approach to moment closure for the stochastic logistic model. *Bull. Math. Biol.* **69**, 1909–1925.

Skupsky, R., Burnett, J. C., Foley, J. E., Schaffer, D. V., and Arkin, A. P. (2010). HIV promoter integration site primarily modulates transcriptional burst size rather than frequency. *PLOS Comput. Biol.* **6**.

Stelling, J., Doyle, F. J., and Gilles, E. D. (2004). Robustness properties of circadian clock architectures. *Proc. Natl. Acad. Sci. U S A* **101**, 13210–13215.

Strogatz, S. H. (1994). *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering.* Addison-Wesley, Reading, MA.

Stundzia, A. B., and Lumsden, C. J. (1996). Stochastic simulation of coupled reaction-diffusion processes. *J. Comput. Phys.* **127**, 196–207.

Suel, G. M., Cagatay, T., Turcotte, M., Elowitz, M. B., and Garcia-Ojalvo, J. (2009). Architecture-dependent noise discriminates functionally analogous differentiation circuits. *Cell* **139**, 512–522.

Suel, G. M., Garcia-Ojalvo, J., Liberman, L. M., and Elowitz, M. B. (2006). An excitable gene regulatory circuit induces transient cellular differentiation. *Nature* **440**, 545–550.

Suel, G. M., Kulkarni, R. P., Dworkin, J., Garcia-Ojalvo, J., and Elowitz, M. B. (2007). Tunability and noise dependence in differentiation dynamics. *Science* **315**, 1716–1719.

Taniguchi, Y., Choi, P. J., Li, G. W., Chen, H. Y., Babu, M., Hearn, J., Emili, A., and Xie, X. S. (2010). Quantifying *E. coli* proteome and transcriptome with single-molecule sensitivity in single cells. *Science* **329**, 533–538.

Thattai, M. T., and van Oudenaarden, A. (2001). Intrinsic noise in gene regulatory networks. *Biophys. J.* **80**, 479a.

Tigges, M., Marquez-Lago, T. T., Stelling, J., and Fussenegger, M. (2009). A tunable synthetic mammalian oscillator. *Nature* **457**, 309–312.

Tomioka, R., Kimura, H., Kobayashi, T. J., and Aihara, K. (2004). Multivariate analysis of noise in genetic regulatory networks. *J. Theor. Biol.* **229**, 501–521.

Tsai, T. Y. C., Choi, Y. S., Ma, W. Z., Pomerening, J. R., Tang, C., and Ferrell, J. E. (2008). Robust, tunable biological oscillations from interlinked positive and negative feedback loops. *Science* **321**, 126–129.

van Oudenaarden, A., Acar, M., and Mettetal, J. T. (2008). Stochastic switching as a survival strategy in fluctuating environments. *Nat. Genet.* **40**, 471–475.

Varma, A., Morbidelli, M., and Wu, H. (2005). Parametric Sensitivity in Chemical Systems. Cambridge University Press, Cambridge, New York.

Vrljic, M., Nishimura, S. Y., Brasselet, S., Moerner, W. E., and McConnell, H. M. (2002). Uncorrelated diffusion of MHC class II proteins in the plasma membrane. *Biophys. J.* **82**, 523a.

Weisshaar, J. C., Konopka, M. C., Shkel, I. A., Cayley, S., and Record, M. T. (2006). Crowding and confinement effects on protein diffusion in vivo. *J. Bacteriol.* **188**, 6115–6123.

Whittle, P. (1957). On the use of the normal approximation in the treatment of stochastic-processes. *J. R. Stat. Soc. Ser. B-Stat. Methodol.* **19**, 268–281.

Wolf, D. M., Vazirani, V. V., and Arkin, A. P. (2005). Diversity in times of adversity: probabilistic strategies in microbial survival games. *J. Theor. Biol.* **234**, 227–253.

Yu, J., Xiao, J., Ren, X. J., Lao, K. Q., and Xie, X. S. (2006). Probing gene expression in live cells, one protein molecule at a time. *Science* **311**, 1600–1603.

Further reading

Avery, S. V. (2006). Microbial cell individuality and the underlying sources of heterogeneity. *Nat. Rev. Microbiol.* **4**, 577–587.

Gillespie, D. T. (1977a). Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* **81**, 2340–2361.

Kampen, N. (1992a). *Stochastic Processes in Chemistry and Physics.* Elsevier.

Maheshri, N., and O'Shea, E. K. (2007). Living with noisy genes: how cells function reliably with inherent variability in gene expression. *Annu. Rev. Biophys. Biomol. Struct.* **36**, 413–434.