drives differentiation of cells in the early embryo of the sea urchin *Strongylocentrotus purpuratus* (figure 7.37). The behavior of this network begins with maternally specified asymmetries in the egg and leads to development of the endoderm (inside layer), skeletal, and mesoderm (middle layer) components of the embryo. A full kinetic characterization of the interactions in a network of this size is daunting, and so models are typically constructed using simpler methods, such as Boolean frameworks. (Appropriate modeling frameworks are reviewed in Bolouri and Davidson (2002).)

The study of complex gene regulatory networks has revealed an important insight into their structure: they often exhibit a *modular* architecture, meaning that the network is composed of subnetworks that play their role somewhat independently of one another. Modularity is a key aspect of human-engineered systems: it allows individual components to be designed, constructed, and tested independently of the entire system. Moreover, modularity allows the re-use of components in multiple systems—a feature that is likely of use in evolutionary "design." (Modularity is reviewed in Wagner et al. (2007); the challenges and opportunities that modular design presents to synthetic biology are discussed in Purnick and Weiss (2009).)

7.6* Stochastic Modeling of Biochemical and Genetic Networks

Chemical reactions result from collisions of individual molecules. Most molecular collisions do not cause reactions. On a molecular scale, reactions are thus rare events and are difficult to predict. In many cellular processes, this molecular randomness is averaged out over large numbers of reaction events, resulting in predictable system behavior. In contrast, processes that depend on small numbers of molecules can be strongly affected by the randomness of biochemical events. Gene expression often involves molecular species that are present in low numbers, and so gene regulatory networks can be subject to this random variation. (The effects of noise on developmental gene networks reveals itself in differences between genetically identical organisms, from bacteria to humans. Stochasticity in gene expression is discussed in Raj and van Oudenaarden (2008).)

Random variation is often considered an inconvenience that must be overcome: the fact that this randomness is usually referred to as "noise" suggests it is a nuisance. However, in some biological contexts, random behavior can be exploited for improved performance. An example is provided by the phenomenon of bacterial *persistence*, in which a genetically identical population gives rise to a small number of so-called persistent cells that exhibit antibiotic resistance at the cost of a reduced growth rate. In the absence of antibiotics, slow-growing persistent cells are quickly out-competed, but the presence of a handful of these cells ensures the population's survival when antibiotics are encountered.

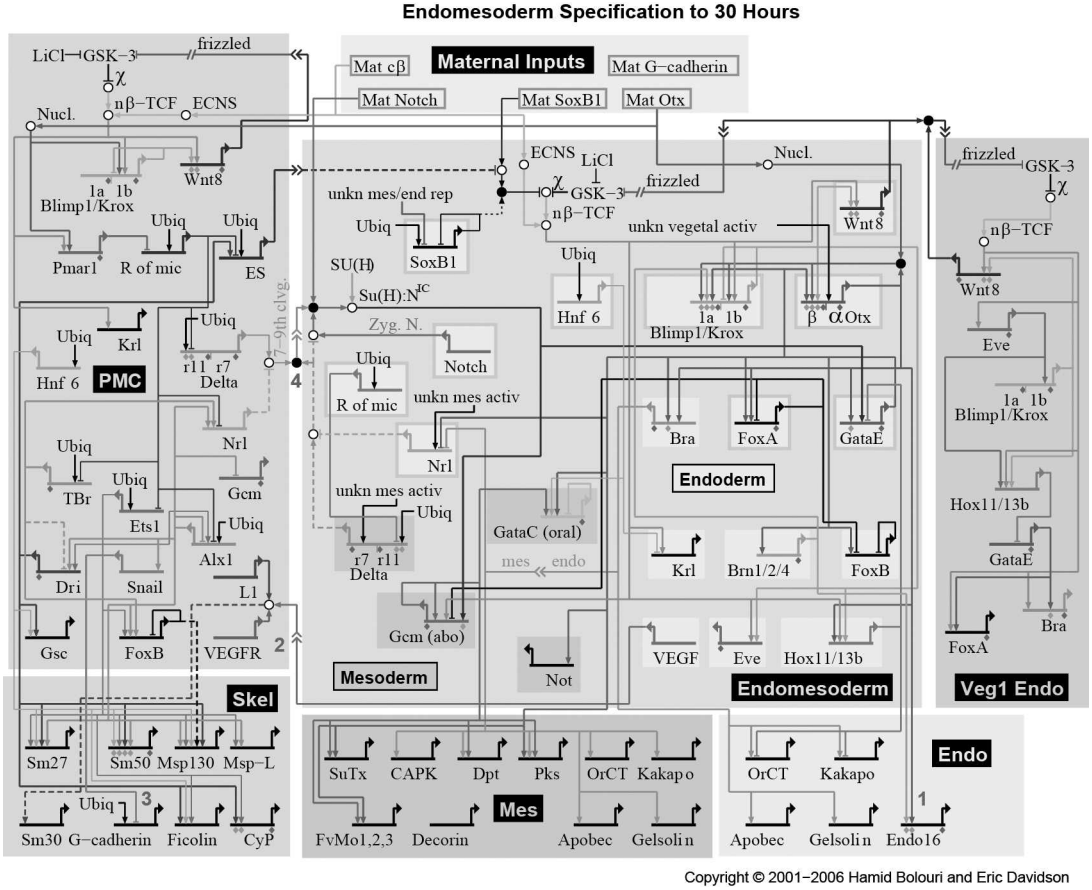

Figure 7.37
Endomesoderm specification network in the sea urchin *Strongylocentrotus purpuratus*. The genes are organized into boxes based on their function. Maternal inputs appear at the top; differentiation proteins are encoded by genes in the bottom boxes. The network describes events that occur in the 30 hours after fertilization. Reproduced, with permission, from Davidson (2006), figure 4.2.

At the cellular level, randomness can be partitioned into two categories: *extrinsic noise*, which refers to random variations that impact all processes in the cell equally, and *intrinsic noise*, which is driven by thermal fluctuations at the molecular level. In models of intracellular networks, extrinsic noise appears as randomness in the values of model parameters and so can be directly incorporated into a differential equation–based framework. In contrast, treatment of intrinsic noise demands the adoption of a modeling framework that takes into account the randomness of the biochemical events that drive reaction dynamics.

A reaction network that comprises large numbers of reactant molecules will involve many simultaneous reaction events. In such cases, network behavior corresponds to the average over these events and is well described by deterministic differential equation models. Figure 7.38, which shows the behavior of a decaying population of molecules, illustrates this averaging effect. The solid curve in each panel of the figure shows a simulation that incorporates randomness; the dashed curve shows a corresponding deterministic simulation. In panel A, the initial population size is large. In this case, individual decay events have a negligible effect on the overall pool. Averaged over many events, the random timing of the reactions is smoothed out, so the deterministic model provides a good description of system behavior. In panel B, the initial population consists of a smaller number of molecules, so the averaging effect is not as strong. Panel C shows a simulation that starts with just 10 molecules. Each decay event has an appreciable effect on the overall

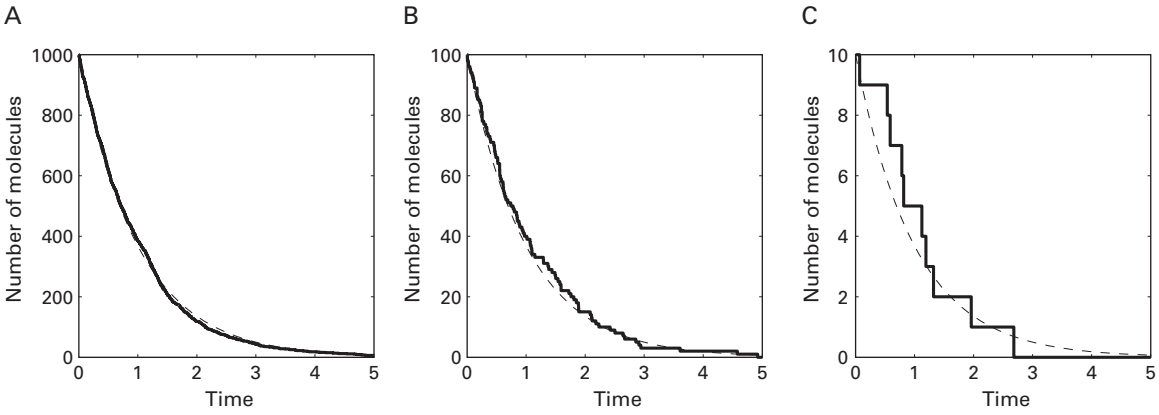

Figure 7.38
Simulations of constitutive decay. The solid curves show simulations that incorporate randomness (stochastic simulations). The dashed curves show the corresponding deterministic (differential equation–based) simulations. The initial pool sizes are 1000 (A), 100 (B), and 10 (C) molecules. For large pool size (A), the simulations agree. However, as the molecule count decreases (B, then C), random effects become more pronounced and are not well described by the deterministic model.

abundance. In this case, the system's discrete, random behavior is not well described by the deterministic simulation.

In this section, we introduce a *stochastic* modeling framework that is suitable for describing systems that involve small numbers of molecules. The term *stochastic* means "random": it is used to describe dynamic processes that have some element of randomness in their progression. (Appendix B contains a brief introduction to some basic concepts from probability.)

7.6.1 A Discrete Modeling Framework

In developing a stochastic modeling framework for chemical reaction networks, we will continue to assume spatial homogeneity and a fixed volume. The abundance of each chemical species will be described by the *number* of molecules in the reaction volume. The state of the system is then the vector $\mathbf{N}$ of molecule counts. (In contrast, the state of a differential equation model is the vector $\boldsymbol{s}$ of species *concentrations*, which change smoothly over time.) As the stochastic dynamics proceed, the molecule counts will change their values in discrete jumps.

We will characterize each reaction in the network by a *stoichiometry vector* $\boldsymbol{s}$ and a *propensity function* a. For each reaction, the stoichiometry vector indicates the identity and number of reactants and products: the j-th component of this vector is the net number of molecules of species j produced or consumed in the reaction. The propensity is a description of reaction rate.

To illustrate these ideas, consider the network composed of the two reactions

$$R_1: \quad A + B \xrightarrow{k_1} C \qquad\qquad R_2: \quad C \xrightarrow{k_2} .$$

The state of this system describes the numbers of molecules of species A, B, and C present at any given time. The stoichiometry vectors are

$$\mathbf{s}_1 = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \begin{matrix} \leftarrow A \\ \leftarrow B \\ \leftarrow C \end{matrix} \qquad \text{and} \qquad \mathbf{s}_2 = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} \begin{matrix} \leftarrow A \\ \leftarrow B \\ \leftarrow C \end{matrix}.$$

When a reaction occurs, the state vector $\mathbf{N}$ is updated by addition of the corresponding stoichiometry vector. For example, suppose that at a given time the state is $\mathbf{N} = (N_A, N_B, N_C) = (12, 3, 4)$. If reaction R_1 were to occur, we would update the state by replacing $\mathbf{N}$ with $\mathbf{N} + \mathbf{s}_1 = (11, 2, 5)$.

The reaction propensities are functions of reactant abundance. We will assume that the probability of a reaction event is proportional to the product of the abundance of each reactant species (as in mass action). The propensities for this example are then

$$a_1(\mathbf{N}) = k_1 N_A N_B \qquad\qquad a_2(\mathbf{N}) = k_2 N_C.$$

Reaction propensities take the same form as mass-action rate laws, but differences appears when multiple copies of an individual reactant are involved.[2]

7.6.2 The Chemical Master Equation

We will build a stochastic modeling framework on the assumption that there are small time increments dt for which:

- At most one reaction event can occur during any time interval of length dt.
- The probability that reaction R_k occurs in any time interval $[t, t+dt]$ is the product of the reaction propensity at time t and the length of the interval: $a_k(\mathbf{N}(t))dt$.

Under these assumptions, the probability that no reactions occur during a time interval $[t, t+dt]$ is $1-\sum_k a_k(\mathbf{N}(t))dt$, where the sum is taken over all reactions in the system.

Let $P(\mathbf{N}, t)$ denote the probability that the system is in state $\mathbf{N}$ at time t. This is called the *probability distribution* of the state (and is dependent on the initial condition—it is a conditional probability distribution). If the distribution $P(\mathbf{N}, t)$ is known at time t, we can use the assumptions above to describe the distribution at time $t+dt$:

$$P(\mathbf{N},t+dt) = P(\mathbf{N},t)\cdot \underbrace{\left(1-\sum_k a_k(\mathbf{N})dt\right)}_{\text{Probability of no reactions firing}} + \sum_k \underbrace{P(\mathbf{N}-\mathbf{s}_k,t)a_k(\mathbf{N}-\mathbf{s}_k)dt}_{\text{Probability of reaction } R_k \text{ occurring while in state } \mathbf{N}-\mathbf{s}_k}. \tag{7.27}$$

This equation is called a *probability balance*. The first term is the probability of being in state $\mathbf{N}$ at time t and remaining in that state until time $t+dt$ (because no reaction events occur). The second term is the sum of the probabilities of transitioning into state $\mathbf{N}$ from another state (because reaction R_k causes a transition from $\mathbf{N}-\mathbf{s}_k$ to $(\mathbf{N}-\mathbf{s}_k)+\mathbf{s}_k = \mathbf{N}$).

As an example, consider the simple reaction chain in which species A is produced at zero order and degrades at first order:

$$R_1:\ \xrightarrow{k_1} A \qquad\qquad R_2:\ A \xrightarrow{k_2}.$$

The state of the system is the number of molecules of A (i.e., $N = N_A$). The reaction stoichiometries are $s_1 = [1]$, $s_2 = [-1]$. The reaction propensities are $a_1 = k_1$ and $a_2 = k_2 N_A$. The transitions between states follow the scheme in figure 7.39.

2. For instance, the propensity of the bimolecular reaction $A + A \xrightarrow{k} C$ is $kN_A(N_A-1)/2$. This formula reflects the number of unique pairings of two A molecules.

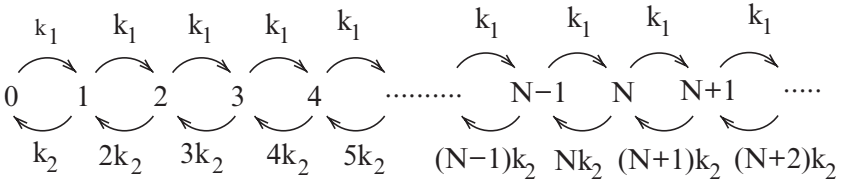

Figure 7.39
Transitions among states for the simple reaction chain $\xrightarrow{k_1} A \xrightarrow{k_2}$. The reaction propensities are indicated.

In this case, the probability balance reads:

$$P(0, t + dt) = P(0, t)(1 - k_1 dt) + P(1, t)\cdot k_2 dt$$

$$P(1, t + dt) = P(1, t)(1 - (k_1 + k_2)dt) + P(0, t)\cdot k_1 dt + P(2, t)\cdot 2k_2 dt$$

$$P(2, t + dt) = P(2, t)(1 - (k_1 + 2k_2)dt) + P(1, t)\cdot k_1 dt + P(3, t)\cdot 3k_2 dt$$

$$P(3, t + dt) = P(3, t)(1 - (k_1 + 3k_2)dt) + P(2, t)\cdot k_1 dt + P(4, t)\cdot 4k_2 dt$$

$$\vdots$$

$$P(N, t + dt) = P(N, t)(1 - (k_1 + Nk_2)dt) + P(N - 1, t)\cdot k_1 dt + P(N + 1, t)\cdot (N + 1)k_2 dt$$

$$\vdots$$

Exercise 7.6.1 Verify that the probability balance for the scheme:

R_1 : $\xrightarrow{k_1} A$

R_2 : $\xrightarrow{k_2} B$

R_3 : $A + B \xrightarrow{k_3}$

is

$$\begin{aligned} P((N_A, N_B), t + dt) = {} & P((N_A, N_B), t)(1 - (k_1 + k_2 + N_A N_B k_3)dt) \\ & + P((N_A - 1, N_B), t)\cdot k_1 dt + P((N_A, N_B - 1), t)\cdot k_2 dt \\ & + P((N_A + 1, N_B + 1), t)\cdot (N_A + 1)(N_B + 1)k_3 dt. \end{aligned}$$

□

The probability balance (7.27) can be used to derive a differential equation describing the rate of change of the probability distribution, as follows. Subtracting $P(\mathbf{N}, t)$ from each side of equation (7.27) gives

$$P(\mathbf{N}, t+dt) - P(\mathbf{N}, t) = -P(\mathbf{N}, t)\left(\sum_k a_k(\mathbf{N})dt\right) + \sum_k P(\mathbf{N} - \mathbf{s}_k, t)a_k(\mathbf{N} - \mathbf{s}_k)dt.$$

Dividing both sides by dt and taking the limit as dt tends to zero results in

$$\frac{d}{dt}P(\mathbf{N},t) = -P(\mathbf{N},t)\left(\sum_k a_k(\mathbf{N})\right) + \sum_k P(\mathbf{N}-\mathbf{s}_k,t)a_k(\mathbf{N}-\mathbf{s}_k)$$
$$= \sum_k \left(\underbrace{-P(\mathbf{N},t)a_k(\mathbf{N})}_{\text{Flow out of state } \mathbf{N}} + \underbrace{P(\mathbf{N}-\mathbf{s}_k,t)a_k(\mathbf{N}-\mathbf{s}_k)}_{\text{Flow into state } \mathbf{N}} \right)$$

This is called the *chemical master equation*. It is a system of differential equations describing the time-varying behavior of the probability distribution. The terms on the right-hand side account for probability flow out of, and into, the state $\mathbf{N}$ at time t. The master equation includes a differential equation for every state $\mathbf{N}$ that the system can adopt, and so typically involves an infinite number of equations.

For the simple reaction chain

$$\xrightarrow{k_1} A \xrightarrow{k_2}$$

described earlier, the master equation is

$$\begin{aligned}
\frac{d}{dt}P(0,t) &= -P(0,t)k_1 + P(1,t)k_2 \\
\frac{d}{dt}P(1,t) &= -P(1,t)(k_1+k_2) + P(0,t)k_1 + P(2,t)2k_2 \\
\frac{d}{dt}P(2,t) &= -P(2,t)(k_1+2k_2) + P(1,t)k_1 + P(3,t)3k_2 \qquad (7.28) \\
&\vdots \\
\frac{d}{dt}P(N,t) &= -P(N,t)(k_1+Nk_2) + P(N-1,t)k_1 + P(N+1,t)(N+1)k_2 \\
&\vdots
\end{aligned}$$

Exercise 7.6.2 Determine the chemical master equation for the system in exercise 7.6.1. □

To illustrate the behavior of solutions of the master equation, we consider the closed reaction network:

$$R_1 : A \xrightarrow{k_1} B \qquad R_2 : B \xrightarrow{k_2} A .$$

To keep the analysis simple, we suppose that there are only two molecules present in the system. The system state $\mathbf{N} = (N_A, N_B)$ can then take only three possible values: $(2,0)$, $(1,1)$, or $(0,2)$. The master equation is a system of three differential equations:

$$\frac{d}{dt}P((2,0),t) = -P((2,0),t)2k_1 + P((1,1),t)k_2$$

$$\frac{d}{dt}P((1,1),t) = -P((1,1),t)k_2 - P((1,1),t)k_1 + P((2,0),t)2k_1 + P((0,2),t)2k_2 \qquad (7.29)$$

$$\frac{d}{dt}P((0,2),t) = -P((0,2),t)2k_2 + P((1,1),t)k_1 .$$

Note that the right-hand-sides sum to zero, as dictated by conservation of probability.

A simulation of system (7.29) is illustrated in figure 7.40, which shows plots of the probability distribution (histograms) at three time points.

For this network, the steady-state distribution $P^{ss}(N_A, N_B)$ can be found by setting the time rates of change to zero:

$$0 = -P^{ss}(2,0)2k_1 + P^{ss}(1,1)k_2$$

$$0 = -P^{ss}(1,1)k_2 - P^{ss}(1,1)k_1 + P^{ss}(2,0)2k_1 + P^{ss}(0,2)2k_2$$

$$0 = -P^{ss}(0,2)2k_2 + P^{ss}(1,1)k_1.$$

Solving these equations, along with the condition that probability is conserved ($P^{ss}(2,0) + P^{ss}(1,1) + P^{ss}(0,2) = 1$) yields the steady-state probability distribution:

$$P^{ss}(2,0) = \frac{k_2^2}{(k_1+k_2)^2}, \quad P^{ss}(1,1) = \frac{2k_1k_2}{(k_1+k_2)^2}, \quad P^{ss}(0,2) = \frac{k_1^2}{(k_1+k_2)^2}. \qquad (7.30)$$

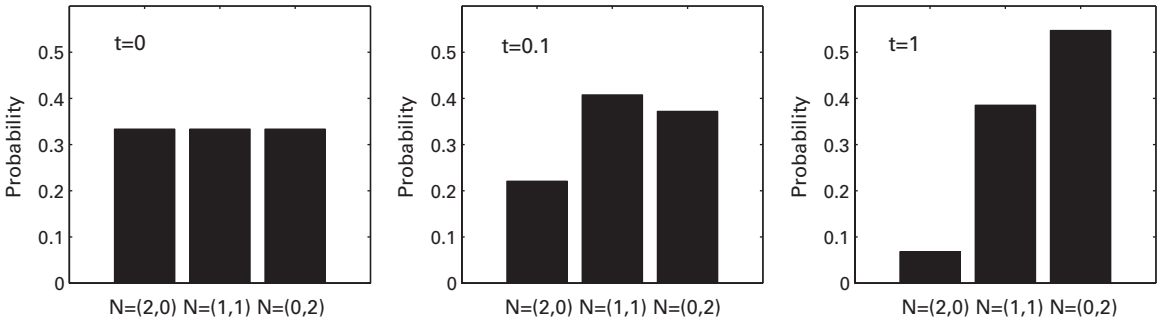

Figure 7.40
Evolution of probabilities for the closed reaction network (7.29). Probability distributions for $\mathbf{N} = (N_A, N_B)$ at times $t = 0$, $t = 0.1$, and $t = 1$ are shown. A uniform initial distribution is chosen, so that at time $t = 0$, all states are equally likely: $P((2,0),0) = P((1,1),0) = P((0,2),0) = 1/3$. Parameter values (in time^{-1}): $k_1 = 3$, $k_2 = 1$. Time units are arbitrary.

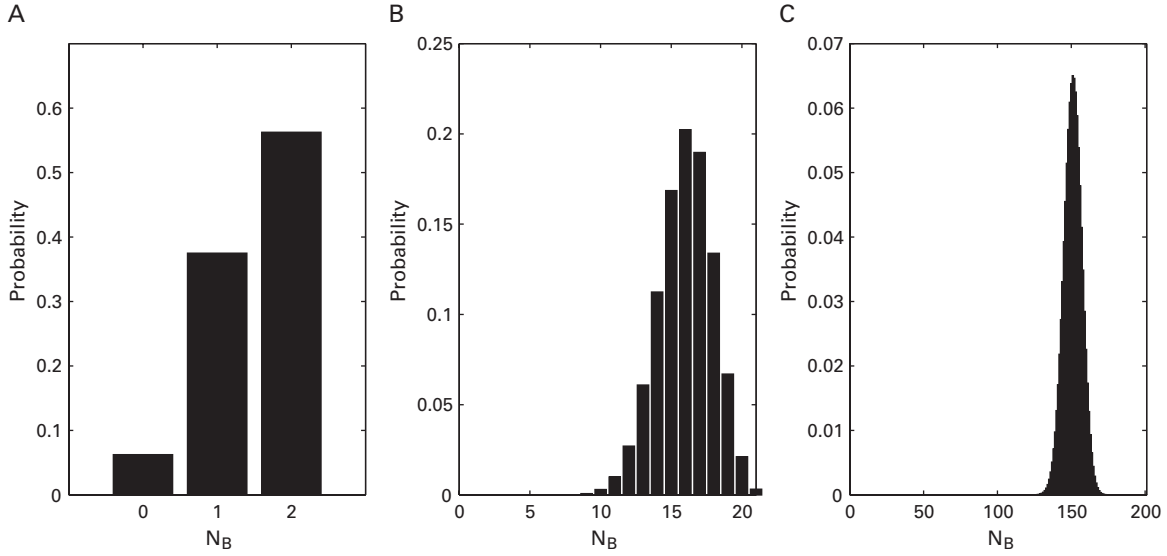

Figure 7.41
Steady-state probability distribution for N_B in reaction network (7.29) for a total molecule count of 2 (A), 20 (B), and 200 (C). As the molecule count grows, the distribution tends to a single peak for which $N_B = 3N_A$, which corresponds to the deterministic (mass-action) steady state.

Exercise 7.6.3 Verify equations (7.30). Does the simulation in figure 7.40 appear to have reached steady state by time $t = 1$? □

As the number of molecules in the system increases, the steady-state distribution of probabilities becomes smoother and more tightly peaked. Figure 7.41 shows the steady-state probability distributions for N_B in system (7.29) when there are 2, 20, and 200 molecules present. As the molecule count increases, the distribution converges to a tight peak at which three fourths of the total pool consists of molecules of B. The deterministic (mass action–based) description of the system yields a steady state that is concentrated at this single point. The probabilistic solution thus converges to the deterministic description for large molecule counts.

Exercise 7.6.4 Verify that the means (i.e., expected values) of N_A and N_B in the steady-state probability distribution (7.30) correspond to the deterministic (mass action–based) model of system (7.40). □

For most systems, the chemical master equation is intractable. (Simulations typically need to incorporate an infinite number of equations!) Consequently, a number of methods have been developed to provide alternative descriptions of stochastic behavior (reviewed in Khammash (2010)). These include the *linear noise approxi-*

mation, which generates differential equations whose solutions approximate the mean and variance of system behavior; *moment closure methods*, which allow calculation of approximate statistics for the probability distribution; and the *finite state projection*, which approximates the chemical master equation by a finite system of differential equations.

Rather than address these analytic approaches, we next consider a numerical method for generation of simulations of stochastic systems.

7.6.3 Gillespie's Stochastic Simulation Algorithm

Numerical algorithms that incorporate stochastic effects (by calling on random number generators) are called *Monte Carlo* methods. In 1977, Dan Gillespie published a Monte Carlo method for simulation of individual trajectories of chemical reaction networks characterized by the chemical master equation (reviewed in Gillespie (2007)). These trajectories, called *sample paths*, represent single elements drawn from a probability distribution generated by the system. Statistics of the trajectory distribution can be determined by generating a large collection of these sample paths (called an *ensemble*).

Gillespie's method, which he called the stochastic simulation algorithm (SSA), tracks each individual reaction event. The simulation does not proceed over a fixed time-grid but jumps forward in time from one reaction event to the next. After each reaction, the algorithm determines which reaction will occur next and how much time will elapse before it occurs.

The simulation algorithm depends on the properties of two *random variables*: the time T to the firing of the next reaction, and the reaction R that will occur next. We next consider how these two random variables are determined.

Determining the Next Reaction The probability that a particular reaction will occur is proportional to the propensity of the reaction. Consider a network that involves three reactions, R_1, R_2, and R_3, with propensities a_1, a_2, and a_3. Let $P(R = R_i)$ denote the probability that R_i will be the next reaction to occur. Probability $P(R = R_i)$ is proportional to the propensity a_i of reaction R_i. Together, these probabilities sum to one. The probability distribution is

$$\begin{aligned} P(R = R_1) &= \frac{a_1}{a_1 + a_2 + a_3} \\ P(R = R_2) &= \frac{a_2}{a_1 + a_2 + a_3} \\ P(R = R_3) &= \frac{a_3}{a_1 + a_2 + a_3}. \end{aligned} \tag{7.31}$$

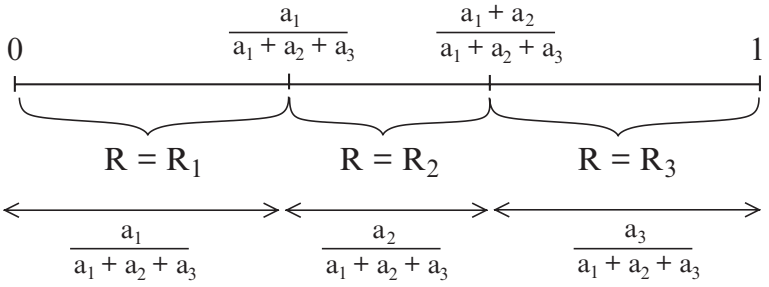

Figure 7.42
Selection of the next reaction. For a network with three reactions, the interval from zero to one is divided into three subintervals, whose lengths correspond to the probabilities of the reactions. A number sampled from the uniform zero-to-one distribution corresponds to a selection of the next reaction.

To implement a simulation of this network's behavior, we need to sample from this probability distribution. Most numerical software packages have built-in functions that generate random numbers drawn uniformly between zero and one. Samples from this uniform distribution can be converted to samples from the distribution (7.31), as follows. We divide the zero-to-one interval into three subintervals—one for each reaction—as in figure 7.42. The length of each subinterval is equal to the probability of the corresponding reaction. A number u that is drawn from the uniform distribution falls into one of these subintervals and thus corresponds to a particular reaction. This procedure can be formalized as follows:

$$\begin{aligned}
&\text{If} \quad 0 \le u \le \frac{a_1}{a_1 + a_2 + a_3}, \quad \text{then we set } R = R_1. \\
&\text{If} \quad \frac{a_1}{a_1 + a_2 + a_3} < u \le \frac{a_1 + a_2}{a_1 + a_2 + a_3}, \quad \text{then we set } R = R_2. \\
&\text{If} \quad \frac{a_1 + a_2}{a_1 + a_2 + a_3} < u \le \frac{a_1 + a_2 + a_3}{a_1 + a_2 + a_3} = 1, \quad \text{then we set } R = R_3.
\end{aligned} \tag{7.32}$$

Figure 7.43 provides a visualization of this process. Here, the uniform random number u is assigned to the vertical axis. The height of the staircase graph corresponds to the cumulative probabilities as used in algorithm (7.32). The next reaction is determined by selecting a number u from the uniform zero-to-one distribution and then extending a horizontal line to the staircase graph, as shown. This graph is called the *cumulative distribution function* for the random variable R.

Determining the Time to the Next Reaction The time T that elapses between reactions is also a random variable. Unlike R, it does not have a discrete value-set, but can

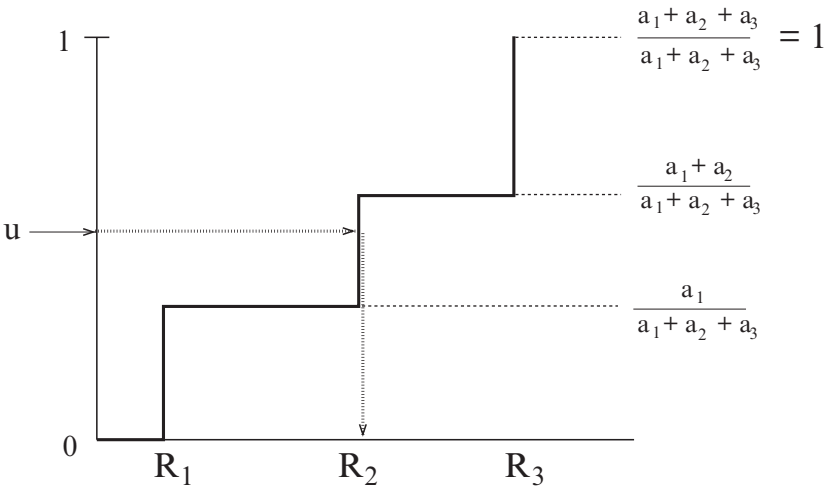

Figure 7.43
Cumulative distribution function for the random variable R. The height of the staircase graph corresponds to the cumulative probability as in algorithm (7.32). A reaction is chosen by selecting a number u from the uniform distribution on the vertical axis and then extending a horizontal line to the staircase graph.

take any nonnegative value. Because it can take infinitely many values, the probability of T having any particular value is vanishingly small. Thus, rather than frame our discussion in terms of point-wise probabilities, we will instead sample T directly from the cumulative distribution function, as we did for the random variable R in figure 7.43. The cumulative distribution function for T is given by

$$P(0 \le T \le t) = 1 - e^{-at}, \tag{7.33}$$

where a is the sum of the reaction propensities:

$$a = a_1 + a_2 + a_3.$$

Equation (7.33) characterizes T as an *exponential random variable*.

The cumulative distribution function for T is shown in figure 7.44. Most often, samples u from the uniform zero-to-one distribution will correspond to short wait-times between reactions: only occasionally (when u is chosen near 1) will a long time be selected. The steepness of the curve depends on a, the sum of the propensities. If this sum is large (many highly probable reactions), then the curve rises steeply, and waiting times are almost always short. If the sum is smaller, then the curve rises more slowly, and longer waiting times are more likely.

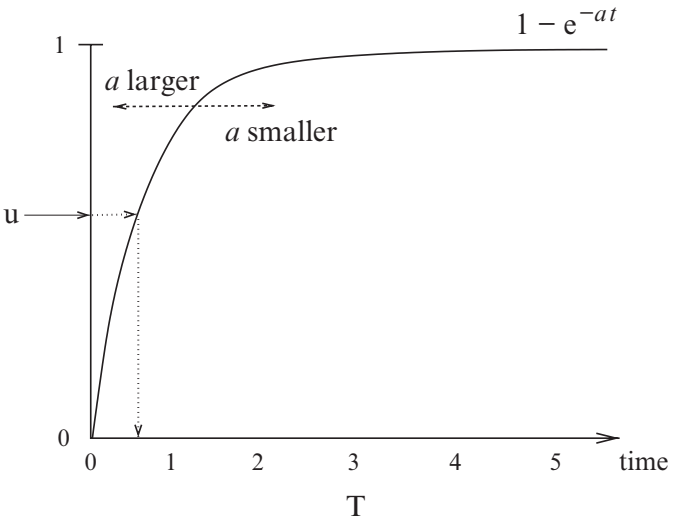

Figure 7.44
Cumulative distribution function for the waiting time T. The parameter a is the sum of the propensities. Waiting times T are determined by selecting numbers u from the uniform zero-to-one distribution and then extending a horizontal line to the graph, as shown. For large values of a, the curve rises sharply—most samples u correspond to short waiting times. For smaller a values, larger waiting times are more likely.

Gilliespie's algorithm can be summarized as follows:

Stochastic Simulation Algorithm (SSA)

1. Set the initial state $\mathbf{N}$. Initialize time t to zero.
2. Calculate the reaction propensities $a_k(\mathbf{N})$.
3. Draw a sample R_k from the random variable R (figure 7.43).
4. Draw a sample τ from the random variable T (figure 7.44).
5. Increment the simulation time $t \to t + \tau$ to account for the elapsed time.
6. Update the state vector $\mathbf{N} \to \mathbf{N} + \mathbf{s}_k$ to reflect the fact that reaction R_k has occurred.
7. Return to step 2.

The algorithm is usually continued until the simulation time t reaches the end of a specified time interval.

7.6.4 Examples

We conclude by using Gillespie's SSA to explore the behavior of some simple reaction networks.

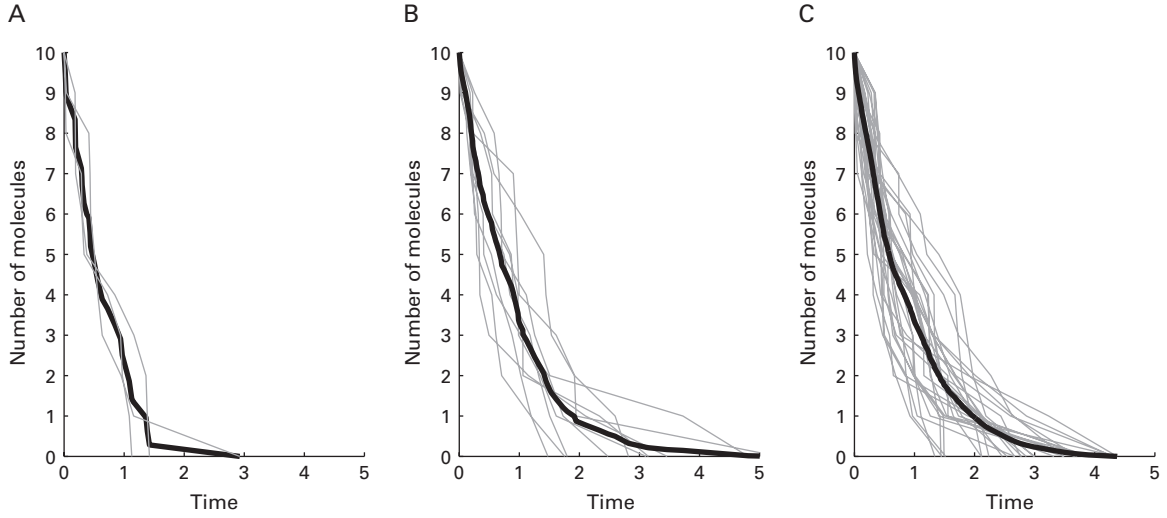

Figure 7.45
Ensembles of sample paths for the decay reaction. Each sample path begins with 10 molecules. Ensembles of 3 (A), 10 (B), and 30 (C) sample paths are shown (gray curves). The black lines show the ensemble average. This averaged behavior approaches the deterministic prediction (exponential decay) as the ensemble size grows. Parameter value: $k = 1$ (time^{-1}). Time units are arbitrary.

Constitutive Decay Consider the decay reaction

$$A \xrightarrow{k} .$$

The behavior of this system was illustrated by the stochastic simulations in figure 7.38, which show that the trajectories are highly variable when the system consists of only a small number of molecules. Figure 7.45 shows ensembles of sample paths, each starting with only 10 molecules. Three ensembles are shown, along with the average behavior (solid line). Although the individual sample paths show significant variability, the average is more consistent. As the ensemble size increases, the averaged behavior converges to the solution of the deterministic model. By generating a large ensemble, a complete description of system behavior—including measures of the variability in the distribution of trajectories—can be reached (see problem 7.8.24).

In some cases, a very large number of sample paths is needed to guarantee confidence in these ensemble-derived results: generating a sufficiently large ensemble can be a time-consuming process. A number of refinements of the SSA have been proposed that aim to reduce the computational requirements for simulation (see problem 7.8.25 for an example).

Constitutive Gene Expression We next consider a simple model of unregulated gene expression, involving mRNA, M, and protein, P. The reaction network is

$$\begin{aligned} R_1 &: (\text{transcription}) \quad \rightarrow M \quad \text{propensity} : k_r \\ R_2 &: (\text{translation}) \quad \rightarrow P \quad \text{propensity} : k_p N_M \\ R_3 &: (\text{degradation}) \quad M \rightarrow \quad \text{propensity} : \delta_r N_M \\ R_4 &: (\text{degradation}) \quad P \rightarrow \quad \text{propensity} : \delta_p N_P \, . \end{aligned} \tag{7.34}$$

Sample paths from a Gillespie simulation are shown in figure 7.46A. The mRNA traces are centered around an average of about 10 molecules. The protein count shows an average of about 60.

Experimental observations have revealed that transcription is sometimes a "bursty" process in which each transcription event leads to the production of multiple copies of mRNA (reviewed in Chubb and Liverpool (2010)).

This model can be modified to describe bursty transcription by replacing reaction R_1 with

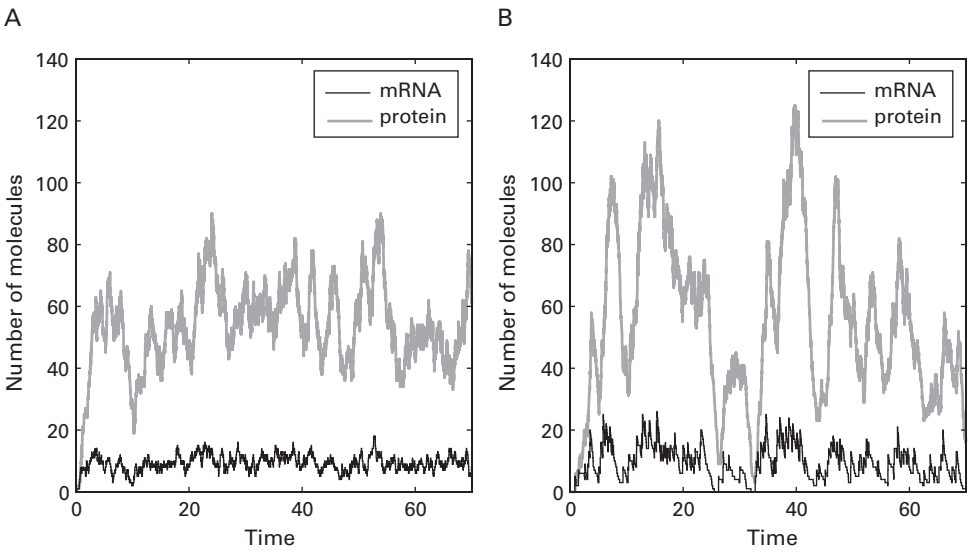

Figure 7.46
Stochastic simulations of constitutive gene expression. (A) Each transcription event produces a single mRNA transcript. (B) Transcription is modeled as "bursty": each transcription event produces five mRNA molecules. The propensity of the transcription reaction has been reduced by a factor of 5 to give the same average as in (A). Parameter values (in time^{-1}): $k_r = 10$, $k_p = 6$, $\delta_r = 1$, $\delta_p = 1$. Time units are arbitrary.

$$\tilde{R}_1 : (\text{bursty transcription}) \quad \to 5M \quad \text{propensity} : \frac{k_r}{5}.$$

In this modified model, each transcription event produces five mRNA molecules. To allow direct comparison with the original model, the propensity of this bursty transcription reaction has been reduced by a factor of 5, so that the time-averaged mRNA production rate is unchanged. Figure 7.46B shows simulations of this modified model. Although the mRNA and protein averages are the same in both models, the modified model exhibits considerably more variability. This difference in behavior could not be described by a mass action–based model: the deterministic versions of these two models are identical (in both cases, the transcription rate is k_r). Variability is an experimentally observable feature of system behavior that can only be captured in a stochastic modeling framework.

The Brusselator Our final example, the *Brusselator*, is a theoretical chemical system that exhibits sustained oscillations (see exercise 4.3.1 of chapter 4). The reaction network is

$$\begin{aligned}
R_1 &: \quad \to X \quad && \text{propensity} : k_1 \\
R_2 &: \quad X \to Y \quad && \text{propensity} : k_2 N_X \\
R_3 &: \quad 2X + Y \to 3X \quad && \text{propensity} : \frac{k_3}{2} N_X (N_X - 1) N_Y \\
R_4 &: \quad X \to \quad && \text{propensity} : k_4 N_X .
\end{aligned}$$

A sample path is shown in figure 7.47, in both the time domain (panel A) and the phase space (panel B). The trajectories are somewhat jagged, but the oscillations are fairly regular. In contrast, some oscillatory stochastic systems exhibit considerable variability in the timing and shape of the cycles (see problem 7.8.27).

7.7 Suggestions for Further Reading

• **Modeling Gene Regulatory Networks** The book *An Introduction to Systems Biology: Design Principles of Biological Circuits* (Alon, 2007) surveys a number of models of gene regulatory networks. The text *Computational Modeling of Gene Regulatory Networks—A Primer* (Bolouri, 2008) addresses a wider range of modeling approaches than discussed in this chapter.

• **Phage Lambda** The book *A Genetic Switch: Phage Lambda Revisited* (Ptashne, 2004) provides a detailed description of the molecular genetics of the decision switch, including an accessible account of the experiments that led to these discoveries.

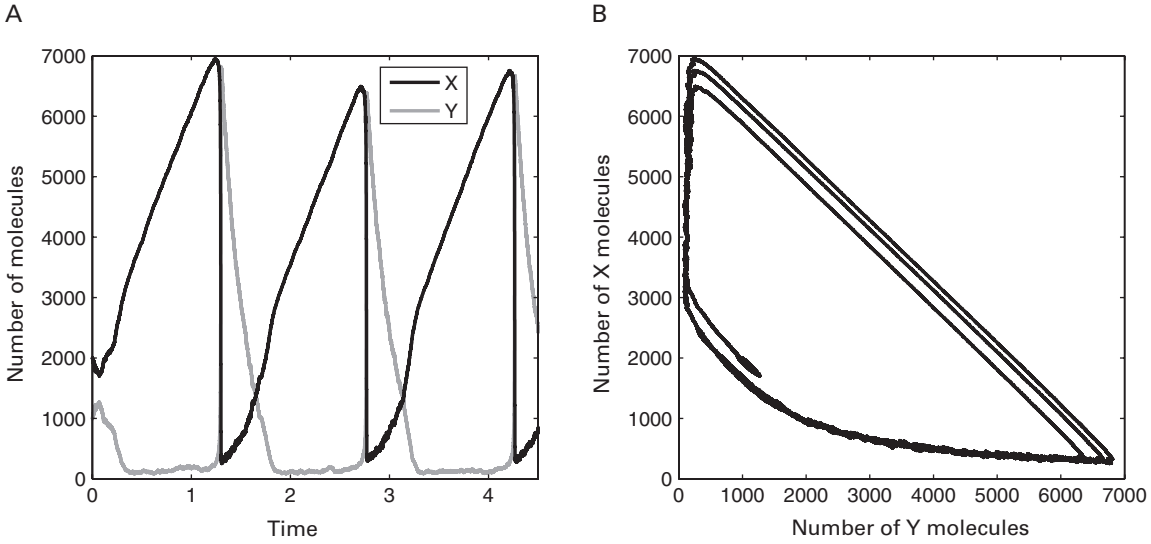

Figure 7.47
Stochastic simulation of the Brusselator. (A) Oscillations are evident in the time domain: a slow increase in Y is followed by a sudden crash when Y levels are sufficiently high. (B) This phase portrait shows the approximate limit cycle followed by the periodic trajectories. Initial conditions are $X = 1000$, $Y = 2000$. Parameter values (in time^{-1}): $k_1 = 5000$, $k_2 = 50$, $k_3 = 0.00005$, and $k_4 = 5$. Time units are arbitrary.

• **Synthetic Gene Circuits** Discussions of modeling and design in synthetic biology are provided in the book chapter "Synthetic Gene Regulatory Systems" (Weiss and Kaern, 2006) and in *Engineering Gene Circuits* (Myers, 2010). The nontechnical book *Biology is Technology: The Promise, Peril, and New Business of Engineering Life* (Carlson, 2010) provides a thoughtful discussion of the potential impact of synthetic biology.

• **Stochastic Modeling in Systems Biology** An introduction to stochastic modeling techniques in systems biology is provided in the book chapter "Modeling and Analysis of Stochastic Biochemical Networks" (Khammash, 2010). The book *Stochastic Modeling for Systems Biology* (Wilkinson, 2006) provides a detailed treatment of stochastic approaches.

7.8 Problem Set

7.8.1 Response Time of Autoinhibitory Genes

Consider expression from an unregulated gene as modeled in equation (7.2):

$$\frac{d}{dt}p(t) = \alpha_0 - \delta p(t)\ . \tag{7.35}$$

decay rates for the two autoinducers; the parameter D characterizes dilution of all species.

(i) Take parameter values $k_2 = 0.4$ hr^{-1}, $c_m = 10^5$ cells per nanoliter, $\beta = 2$, $d_1 = 1$ hr^{-1}, $d_2 = 0.3$ hr^{-1}, $K_1 = K_2 = 10$ nM, $\gamma_1 = \gamma_2 = 0.1$ nM ml hr^{-1}, $\delta_1 = 0.017$ hr^{-1}, $\delta_2 = 0.11$ hr^{-1}, and $D = 0.2$ hr^{-1}. Simulate the system for three cases: $k_1 = 0.2, 0.8$, and 1.4 hr^{-1}. Verify that in the first and last case, one population dominates over the other, whereas for $k_1 = 0.8$ hr^{-1}, the populations tend to a persistent oscillatory pattern. Explain the long-time system behavior in each case.

(ii) Using this model, Balagaddé and co-workers discovered that the system is more likely to exhibit oscillations when the predator death rate (d_1) is sufficiently large. They thus engineered a modified killer protein to reach a higher death rate. Confirm their finding by exploring the range of k_1 values over which the system oscillates at low ($d_1 = 0.2$) and high ($d_1 = 1$) predator death rates.

7.8.21 Genetic Logic Gate Design

(a) Develop a differential equation model of a regulated promoter that implements an OR logic gate, as discussed in section 7.5.1. Hint: Recall the expression rates for regulated promoters in section 7.1.2.

(b) Using genes that implement only OR, AND, and NOT gates, design a genetic circuit that implements an XOR (exclusive OR) gate. The output from an XOR gate is ON when exactly one of its two inputs is ON. Develop a differential equation model of your circuit.

7.8.22* Chemical Master Equation: Closed System

Consider the reaction network

$$A \xrightarrow{k_1} B + C \qquad\qquad B + C \xrightarrow{k_{-1}} A.$$

(a) Suppose that the system starts with two molecules of A, one molecule of B, and no molecules of C; that is, $(N_A, N_B, N_C) = (2, 1, 0)$. Determine the set of possible states the system can adopt, and write the chemical master equation that describes the corresponding probability distribution.

(b) Take $k_1 = 1$ (time^{-1}) and $k_{-1} = 1$ (time^{-1}), and solve for the steady-state probability distribution.

7.8.23* Chemical Master Equation: Open System

Consider the open system

$$\xrightarrow{k_1} A \xrightarrow{k_2}.$$

(a) Take $k_1 = 1$ (concentration $\cdot$ time^{-1}) and $k_2 = 1$ (time^{-1}). Referring to the corresponding master equation (7.28), verify that in steady state, the probabilities satisfy

$$P(N_A = n) = \frac{1}{n} P(N_A = n-1).$$

(b) Use the fact that $\sum_{n=0}^{\infty}(1/n!) = e$ (where the factorial $n! = n(n-1)(n-2)\cdots 3\cdot 2\cdot 1$, and e is Euler's number $e \approx 2.71828$) to derive the steady-state probability distribution:

$$P(N_A = n) = \frac{1/e}{n!},$$

for each $n = 0, 1, 2, \ldots$ (by convention $0! = 1$).

7.8.24* Statistics of an Ensemble of Sample Paths

Consider the simple model of unregulated gene expression in section 7.6.4:

R_1 : (transcription) $\rightarrow M$ propensity : k_r

R_2 : (translation) $\rightarrow P$ propensity : $k_p N_M$

R_3 : (degradation/dilution) $M \rightarrow$ propensity : $\delta_r N_M$

R_4 : (degradation/dilution) $P \rightarrow$ propensity : $\delta_p N_P$.

Take parameter values as in figure 7.46A. Simulate sample paths using the stochastic simulation algorithm. Analyze the statistics of your ensemble to verify that in steady state (so-called *stationary* behavior), the coefficient of variation (standard deviation divided by mean) for each species is

$$mRNA: \; C_r = \left(\frac{\delta_r}{k_r}\right)^{1/2}, \qquad \text{protein}: \; C_p = \left(\frac{\delta_r \delta_p}{k_r k_p}\right)^{1/2} \left(1 + \frac{k_p}{\delta_r + \delta_p}\right)^{1/2}.$$

(For a derivation of these formulas, see Khammash (2010).) Note that statistics can be gathered from an ensemble of sample paths or a single long simulation. XPPAUT users will have to export the data from the simulation and calculate the mean and variance using another program.

7.8.25* Stochastic Simulation: The First-Reaction Method

An alternative to the stochastic simulation algorithm presented in section 7.6.3 is the *first-reaction method*, which also involves stepping from reaction event to reaction event. However, rather than sample the next reaction and the waiting time separately (as in the SSA), the first-reaction algorithm samples a waiting time for each reaction in the network and then selects the shortest of this collection of times.

This selection specifies the identity of the next reaction and the elapsed time. Because these waiting times can often be re-used from one time-step to the next, this algorithm can be significantly more efficient than the SSA. (An implementation of the first-reaction method was presented in Gibson and Bruck (2000).)

Recall that the waiting time $T = T_{\text{wait}}$ in the SSA has a cumulative distribution function given by $P(0 \le T_{\text{wait}} \le t) = 1 - e^{-at}$, where a is the sum of the propensities for all of the reactions in the network. In the first-reaction algorithm, the waiting time T_{first} is the minimum of a collection of reaction-specific waiting times T_i, each of which is characterized by $P(0 \le T_i \le t) = 1 - e^{-a_i t}$, where a_i is the reaction propensity. Confirm that the cumulative distribution function for the first-reaction waiting time ($T_{\text{first}} = \min(T_1, T_2, \ldots, T_m)$) agrees with the distribution of $T = T_{\text{wait}}$ in the SSA. Hint: Verify that $P(T_{\text{wait}} > t) = e^{-at}$ and $P(T_i > t) = e^{-a_i t}$. Then use the fact that $P(T_{\text{first}} > t) = P((T_1 > t) \text{ and } (T_2 > t) \ldots \text{and } (T_{\text{m}} > t))$ where the T_i are independent of one another.

7.8.26* Noisy Toggle Switch

Stochastic systems can exhibit a range of bistable-like behaviors, ranging from "true" bistability to frequent noise-induced transitions between two nominally stable states. To explore this behavior, consider a stochastic system that recapitulates the bistable toggle switch discussed in section 7.2.3:

$$R_1 : (\text{synthesis}) \quad \rightarrow P_1 \quad \text{propensity} : \frac{\alpha}{1 + N_2^{\beta}}$$

$$R_2 : (\text{synthesis}) \quad \rightarrow P_2 \quad \text{propensity} : \frac{\alpha}{1 + N_1^{\beta}}$$

$$R_3 : (\text{decay}) \quad P_1 \rightarrow \quad \text{propensity} : \delta N_1$$

$$R_4 : (\text{decay}) \quad P_2 \rightarrow \quad \text{propensity} : \delta N_2 .$$

Here, N_1 and N_2 are the molecular counts for the two repressors. The Hill-type propensities for the synthesis reactions are not well-justified at the molecular level, but these expressions nevertheless provide a simple formulation of a bistable stochastic system. Take parameter values $\delta = 1$ and $\beta = 4$. The corresponding deterministic system (i.e., $dp_i / dt = \alpha / (1 + p_j^4) - p_i$) is bistable for any $\alpha > 1$. Run simulations of the stochastic system for $\alpha = 5$, 50, 500, and 5000. Be sure to run the simulations sufficiently long so that the steady trend is clear (i.e., at least 10,000 reaction steps). Verify that for $\alpha = 5000$, the system exhibits bistability (with about 5000 molecules of the dominant species, in the long term). In contrast, verify that with $\alpha = 5$, noise dominates and the system shows no signs of bistability. What about at $\alpha = 50$ and 500? Comment on how the steady-state molecule abundance affects system behav-

ior. (Note that it may be necessary to run multiple simulations to confirm your findings.)

7.8.27* Noise-Induced Oscillations

Stochastic systems can exhibit a range of oscillatory behaviors, ranging from near-perfect periodicity to erratic cycles. To explore this behavior, consider a stochastic relaxation oscillator studied by José Vilar and colleagues (Vilar et al., 2002). The system involves an activator and a repressor. The activator enhances expression of both proteins. The repressor acts by binding the activator, forming an inert complex. A simple model of the system is

$$R_1 : \text{(activator synthesis)} \qquad \to b_A A \qquad \text{propensity} : \frac{\gamma_A}{b_A}\frac{\alpha_0 + N_A / K_A}{1 + N_A / K_A}$$

$$R_2 : \text{(repressor synthesis)} \qquad \to b_R R \qquad \text{propensity} : \frac{\gamma_R}{b_R}\frac{N_A / K_R}{1 + N_A / K_R}$$

$$R_3 : \text{(activator decay)} \qquad A \to \qquad \text{propensity} : \delta_A N_A$$

$$R_4 : \text{(repressor decay)} \qquad R \to \qquad \text{propensity} : \delta_R N_R$$

$$R_5 : \text{(association)} \qquad A + R \to C \qquad \text{propensity} : k_C N_A N_R$$

$$R_6 : \text{(dissociation and decay)} \qquad C \to R \qquad \text{propensity} : \delta_A N_C$$

Here, N_A, N_R, and N_C are the molecular counts for the activator, repressor, and activator–repressor complex. The parameter b_A and b_R characterize the expression burst size. The Hill-type propensities for the synthesis reactions are not well-justified at the molecular level, but these expressions nevertheless provide a simple formulation of a stochastic relaxation oscillator.

(a) Take parameter values $\gamma_A = 250$, $b_A = 5$, $K_A = 0.5$, $\alpha_0 = 0.1$, $\delta_A = 1$, $\gamma_R = 50$, $b_R = 10$, $K_R = 1$, $k_C = 200$, and $\delta_R = 0.1$. Run simulations of this model and verify its quasi-periodic behavior.

(b) The deterministic version of this model is

$$\frac{d}{dt}a(t) = \gamma_A \frac{\alpha_0 + a(t)/K_A}{1 + a(t)/K_A} - k_C a(t) r(t) - \delta_A a(t)$$

$$\frac{d}{dt}r(t) = \gamma_R \frac{a(t)/K_R}{1 + a(t)/K_R} - k_C a(t) r(t) + \delta_A c(t) - \delta_R r(t)$$

$$\frac{d}{dt}c(t) = k_C a(t) r(t) - \delta_A c(t),$$

where a, r, and c are the concentrations of activator, repressor, and complex. Run a simulation with the same parameter values as in part (a). Does the system exhibit oscillations? How is the behavior different if you set $\delta_R = 0.2$?

(c) The contrast between the behavior of the models in parts (a) and (b), for $\delta_R = 0.1$, can be explained by the excitability of this relaxation oscillator. Run two simulations of the deterministic model ($\delta_R = 0.1$), one from initial conditions (a, r, c) = (0, 10, 35) and another from initial conditions (a, r, c) = (5, 10, 35). Verify that in the first case, the activator is quenched by the repressor, and the system remains at a low-activator steady state, whereas in the second case, this small quantity of activator is able to break free from the repressor and invoke a (single) spike in expression. Explain how noise in the activator abundance could cause repeated excitations by allowing the activator abundance to regularly cross this threshold. This is referred to as *noise-induced oscillation*.