Biomolecular Feedback Systems

Domitilla Del Vecchio
MIT

Richard M. Murray
Caltech

Classroom Copy v0.6c, July 11, 2012

Chapter 4
Stochastic Modeling and Analysis

In this chapter we explore stochastic behavior in biomolecular systems, building on our preliminary discussion of stochastic modeling in Section 2.1. We begin by reviewing the various methods for modeling stochastic processes, including the chemical master equation (CME), the chemical Langevin equation (CLE) and the Fokker-Planck equation (FPE). Given a stochastic description, we can then analyze the behavior of the system using a variety of stochastic simulation and analysis tools.

Prerequisites. This chapter makes use of a variety of topics in stochastic processes that are not covered in AM08. Readers should have a good working knowledge of basic probability and some exposure to simple stochastic processes (e.g., Brownian motion), at the level of the material presented in Appendix B (drawn from [70]).

4.1 Stochastic Modeling of Biochemical Systems

Biomolecular systems are inherently noisy due to the random nature of molecular reactions. When the concentrations of molecules are high, the deterministic models we have used in the previous chapters provide a good description of the dynamics of the system. However, if the molecular counts are low then it is often necessary to explictly account for the random nature of events. In this case, th chemical reactions in the cell can be modeled as a collection of stochastic events corresponding to chemical reactions between species, including binding and unbinding of molecules (such as RNA polymerase and DNA), conversion of one set of species into another, and enzymatically controlled covalent modifications such as phosphorylation. In this section we will briefly survey some of the different representations that can be used for stochastic models of biochemical systems, following the material in the textbooks by Phillips *et al*. [76], Gillespie [33] and Van Kampen [52].

Statistical mechanics

At the core of many of the reactions and multi-molecular interactions that take place inside of cells is the chemical physics associated with binding between two molecules. One way to capture some of the properties of these interactions is through the use of statistical mechanics and thermodynamics.

As described briefly already in Chapter 2, the underlying representation for both statistical mechanics and chemical kinetics is to identify the appropriate microstates of the system. A microstate corresponds to a given configuration of the components (species) in the system relative to each other and we must enumerate all possible configurations between the molecules that are being modeled.

In statistical mechanics, we model the configuration of the cell by the probability that system is in a given microstate. This probability can be calculated based on the energy levels of the different microstates. Consider a setting in which our system is contained within a reservoir. Let E_r represent the energy in the resevoir, E_s the energy in the system and $E_{\text{tot}} = E_r + Es$ the total (conserved) energy. Given two different energy levels $E_s^{(1)}$ and $E_s^{(2)}$ for the system of interest, let $W_r(E_{\text{tot}} - E_s^{(i)})$ be the number of possible microstates of the reservoir with energy $E_r = E_{\text{tot}} - E_s^{(i)}$, $i = 1, 2$. The laws of statistical mechanics state that the ratio of probabilities of being at the energy levels $E_s^{(1)}$ and $E_s^{(2)}$ is given by the ratio of number of possible states of the reservoir:

$$\frac{\mathbb{P}(E_s^{(1)})}{\mathbb{P}(E_s^{(2)})} = \frac{W_r(E_{\text{tot}} - E_s^{(1)})}{W_r(E_{\text{tot}} - E_s^{(2)})}. \tag{4.1}$$

Defining the entropy of the system as $S = k_B \ln W$, where k_B is Boltmann's constant, we can rewrite equation (4.1) as

$$\frac{W_r(E_{\text{tot}} - E_s^{(1)})}{W_r(E_{\text{tot}} - E_s^{(2)})} = \frac{e^{S_r(E_{\text{tot}} - E_s^{(1)})/k_B}}{e^{S_r(E_{\text{tot}} - E_s^{(2)})/k_B}}.$$

We now approximate $S_r(E_{\text{tot}} - E_s)$ in a Taylor series expansion around E_{tot}, under the assumption that $E_r \gg E_s$:

$$S_r(E_{\text{tot}} - E_s) \approx S_r(E_{\text{tot}}) - \frac{\partial S_r}{\partial E} E_s.$$

From the properties of thermodynamics, if we hold the volume and number of molecules constant, then we can define the temperature as

$$\left.\frac{\partial S}{\partial E}\right|_{V,N} = \frac{1}{T}$$

and we obtain

$$\frac{\mathbb{P}(E_s^{(1)})}{\mathbb{P}(E_s^{(2)})} = \frac{e^{-E_s^{(1)}/k_B T}}{e^{-E_s^{(2)}/k_B T}}.$$

This implies that

$$\mathbb{P}(E_s^{(q)}) \propto e^{-E_s^{(q)}/(k_B T)}$$

and hence the probability of being in a microstate q is given by

$$\mathbb{P}(q) = \frac{1}{Z} e^{-E_q/(k_B T)}, \tag{4.2}$$

where we have written E_q for the energy of the microstate and Z is a normalizing factor, known as the *partition function*, defined by

$$Z = \sum_{q \in Q} e^{-E_q/(k_B T)}.$$

By keeping track of those microstates that correspond to a given system state (also called a macrostate), we can compute the overall probability that a given macrostate is reached.

In order to determine the energy levels associated with different microstates, we will often make use of the *free energy* of the system. Consider an elementary reaction $\mathrm{A} + \mathrm{B} \rightleftharpoons \mathrm{AB}$. Let E be the energy of the system, taken to be operating at pressure P in a volume V. The *enthalpy* of the system is defined as $H = E + PV$ and the *Gibbs free energy* is defined as $G = H - TS$ where T is the temperature of the system and S is its entropy (defined above). The change in bond energy due to the reaction is given by

$$\Delta H = \Delta G + T \Delta S,$$

where the Δ represents the change in the respective quantity. $-\Delta H$ represents the amount of heat that is absorbed from the reservoir, which then affects the entropy of the reservoir.

Derivation to be added later. Review

The resulting formula for the probability of being in a microstate q is given by

$$\mathbb{P}(q) = \frac{1}{Z} e^{-\Delta G/k_B T}.$$

Example 4.1 (Transcription factor binding)**.** Suppose that we have a transcription factor R that binds to a specific target region on a DNA strand (such as the promoter region upstream of a gene). We wish to find the probability P_{bound} that the transcription factor will be bound to this location as a function of the number of transcription factor molecules n_{R} in the system. If the transcription factor is a repressor, for example, knowing $P_{\text{bound}}(n_{\mathrm{R}})$ will allow us to calculate the likelihood of transcription occurring.

To compute the probability of binding, we assume that the transcription factor can bind non-specifically to other sections of the DNA (or other locations in the cell) and we let N_{ns} represent the number of such sites. We let E_{bound} represent the free energy associated with R bound to its specified target region and E_{ns} represent the free energy for R in any other non-specific location, where we assume that $E_{\text{bound}} < E_{\text{ns}}$. The microstates of the system consist of all possible assignments of the n_{R} transcription factors to either a non-specific location or the target region of the DNA. Since there is only one target site, there can be at most one transcription factor attached there and hence we must count all of the ways in which either zero or one molecule of R are attached to the target site.

If none of the n_R copies of R are bound to the target region then these must be distributed between the N_ns non-specific locations. Each bound protein has energy E_ns, so the total energy for any such configuration is $n_\mathrm{R}E_\mathrm{ns}$. The number of such combinations is $\binom{N_\mathrm{ns}}{n_\mathrm{R}}$ and so the contribution to the partition function from these microstates is

$$Z_\mathrm{ns} = \binom{N_\mathrm{ns}}{n_\mathrm{R}} e^{-n_\mathrm{R}E_\mathrm{ns}/(k_BT)} = \frac{N_\mathrm{ns}!}{n_\mathrm{R}!(N_\mathrm{ns}-n_\mathrm{R})!} e^{-n_\mathrm{R}E_\mathrm{ns}/(k_BT)}$$

For the microstates in which one molecule of R is bound at a target site and the other $n_\mathrm{R}-1$ molecules are at the non-specific locations, we have a total energy of $E_\mathrm{bound}+(n_\mathrm{R}-1)E_\mathrm{ns}$ and $\binom{N_\mathrm{ns}}{(n_\mathrm{R}-1)}$ possible such states. The resulting contribution to the partition function is

$$Z_\mathrm{bound} = \frac{N_\mathrm{ns}!}{(n_\mathrm{R}-1)!(N_\mathrm{ns}-n_\mathrm{R}+1)!} e^{-(E_\mathrm{bound}-(n_\mathrm{R}-1)E_\mathrm{ns})/(k_BT)}.$$

The probability that the target site is occupied is now computed by looking at the ratio of the Z_bound to $Z = Z_\mathrm{ns} + Z_\mathrm{bound}$. After some basic algebraic manipulations, it can be shown that

$$P_\mathrm{bound}(n_\mathrm{R}) = \frac{\left(\frac{n_\mathrm{R}}{N_\mathrm{ns}-n_\mathrm{R}+1}\right)\exp[-(E_\mathrm{bound}+E_\mathrm{ns})/(k_BT)]}{1+\left(\frac{n_\mathrm{R}}{N_\mathrm{ns}-n_\mathrm{R}+1}\right)\exp[-(E_\mathrm{bound}+E_\mathrm{ns})/(k_BT)]}.$$

If we assume that $N_\mathrm{ns} \gg n_R$ then $N_\mathrm{ns}-n_R+1 \approx N_\mathrm{ns}$, and we can write

$$P_\mathrm{bound}(n_\mathrm{R}) \approx \frac{kn_\mathrm{R}}{1+kn_\mathrm{R}}, \quad \text{where} \quad k = \frac{1}{N_\mathrm{ns}}\exp[-(E_\mathrm{bound}-E_\mathrm{ns})/(k_BT)].$$

As we would expect, this says that for very small numbers of repressors, P_bound is close to zero, while for large numbers of repressors, $P_\mathrm{bound} \to 1$. The point at which we get a binding probability of 0.5 is when $n_\mathrm{R} = 1/k$, which depends on the relative binding energies and the number of non-specific binding sites. ∇

Example 4.2 (Combinatorial promoter)**.** A combinatorial promoter is a region of DNA in which multiple transcription factors can bind and influence the subsequent binding of RNA polymerase. Combinatorial promoters appear in a number of natural and engineered circuits and represent a mechanism for creating switch-like behavior, for example by having a gene that controls expression of its own transcription factors.

One method to model a combinatorial promoter is to use the binding energies of the different combinations of proteins to the operator region, and then compute the probability of being in a given promoter state given the concentration of each of the transcription factors. Table 4.1 shows the possible states of a notional promoter that has two operator regions—one that binds a repressor protein R and another

Table 4.1: Configurations for a combinatorial promoter with an activator and a repressor. Each row corresponds to a specific macrostate of the promoter in which the listed molecules are bound to the target region. The relative energy of state compared with the ground state provides a measure of the likelihood of that state occurring, with more negative numbers corresponding to more energetically favorable configurations.

State	OR1	OR2	Prom	E_q (ΔG)	Comment
S_1	–	–	–	0	No binding (ground state)
S_2	–	–	RNAP	−5	RNA polymerase bound
S_3	R	–	–	−10	Repressor bound
S_4	–	A	–	−12	Activator bound
S_5	–	A	RNAP	−15	Activator and RNA polymerase

that binds an activator protein A. As indicated in the table, the promoter has three (possibly overlapping) regions of DNA: OR1 and OR2 are binding sites for the repressor and activator proteins, and Prom is the location where RNA polymerase binds. (The individual labels are primarily for bookkeeping purposes and may not correspond to physically separate regions of DNA.)

To determine the probabilities of being in a given macrostate, we must compute the individual microstates that occur at a given concentrations of repressor, activator and RNA polymerase. Each microstate corresponds to an individual set of molecules binding in a specific configuration. So if we have n_R repressor molecules, then there is one microstate corresponding to *each* different repressor molecule that is bound, resulting in n_R individual microstates. In the case of configuration S_5, where two different molecules are bound, the number of combinations is given by the product of the numbers of individual molecules, $n_\mathrm{A} \cdot n_\mathrm{RNAP}$, reflecting the possible combinations of molecules that can occupy the promoter sites. The overall partition function is given by summing up the contributions from each microstate:

$$Z = e^{-E_0/(k_B T)} + n_\mathrm{RNAP}\, e^{-E_\mathrm{RNAP}/(k_B T)} + n_\mathrm{R}\, e^{-E_\mathrm{R}/(k_B T)} + n_\mathrm{A}\, e^{-E_\mathrm{A}/(k_B T)} + n_\mathrm{A} n_\mathrm{RNAP}\, e^{-E_\mathrm{A:RNAP}/(k_B T)}. \quad (4.3)$$

The probability of a given macrostate is determined using equation (2.2). For example, if we define the promoter to be "active" if RNA polymerase is bound to the DNA, then the probability of being in this macrostate as a function of the various molecular counts is given by

$$\begin{aligned} P_\mathrm{active}(n_\mathrm{R}, n_\mathrm{A}, n_\mathrm{RNAP}) &= \frac{1}{Z}\left(n_\mathrm{RNAP}\, e^{-E_\mathrm{RNAP}/(k_B T)} + n_\mathrm{A}\, n_\mathrm{RNAP} e^{-E_\mathrm{A:RNAP}/(k_B T)}\right) \\ &= \frac{k_\mathrm{A:RNAP}\, n_\mathrm{A} + k_\mathrm{RNAP}}{1 + k_\mathrm{RNAP} + k_\mathrm{R}\, n_\mathrm{R} + (k_\mathrm{A} + k_\mathrm{A:RNAP}) n_\mathrm{A}}, \end{aligned}$$

where

$$k_X = e^{-(E_X - E_0)/(k_B T)}.$$

From this expression we see that if $n_R \gg n_A$ then P_{active} tends to 0 while if $n_A \gg n_R$ then P_{active} tends to 1, as expected. ∇

Chemical master equation (CME)

The statistical physics model we have just considered gives a description of the *steady state* properties of the system. In many cases, it is clear that the system reaches this steady state quickly and hence we can reason about the behavior of the system just by modeling the free energy of the system. In other situations, however, we care about the transient behavior of a system or the dynamics of a system that does not have an equilibrium configuration. In these instances, we must extend our formulation to keep track of how quickly the system transitions from one microstate to another, known as the *chemical kinetics* of the system.

To model these dynamics, we return to our enumeration of all possible microstates of the system. Let $P(q,t)$ represent the probability that the system is in microstate q at a given time t. Here q can be any of the very large number of possible microstates for the system, which for chemical reaction systems we can represent in terms of a vector consisting of the number of molecules of each species that is present. We wish to write an explicit expression for how $P(q,t)$ varies as a function of time, from which we can study the stochastic dynamics of the system.

We begin by assuming we have a set of M reactions R_j, $j = 1,\ldots,M$, with ξ_j representing the change in state associated with reaction R_j. Specifically, ξ_j is given by the jth column of the stoichiometry matrix N. The *propensity function* defines the probability that a given reaction occurs in a sufficiently small time step dt:

$$a_j(q,t)dt \quad = \quad \text{Probability that reaction } R_j \text{ will occur between time } t \text{ and time } t+dt \text{ given that the microstate is } q.$$

The linear dependence on dt relies on the fact that dt is chosen sufficiently small. We will typically assume that a_j does not depend on the time t and write $a_j(q)dt$ for the probability that reaction j occurs in state q.

Using the propensity function, we can compute the distribution of states at time $t+dt$ given the distribution at time t:

$$\begin{aligned} P(q,t+dt) &= P(q,t)\Big(1 - \sum_{j=1}^{M} a_j(q)dt\Big) + \sum_{j=1}^{M} P(q-\xi_j)a_j(q-\xi_j)dt \\ &= P(q,t) + \sum_{j=1}^{M}\Big(a_j(q-\xi_j)P(q-\xi_j,t) - a_j(q)P(q,t)\Big)dt. \end{aligned} \tag{4.4}$$

Since dt is small, we can take the limit as $dt \to 0$ and we obtain the *chemical master equation* (CME):

$$\frac{\partial P}{\partial t}(q,t) = \sum_{j=1}^{M} \Big(a_j(q-\xi_j) P(q-\xi_j,t) - a_j(q) P(q,t) \Big) \tag{4.5}$$

This equation is also referred to as the *forward Kolmogorov equation* for a discrete state, continuous time random process.

Despite its complexity, the master equation does capture many of the important details of the chemical physics of the system and we shall use it as our basic representation of the underlying dynamics. As we shall see, starting from this equation we can then derive a variety of alternative approximations that allow us to answer specific equations of interest.

The key element of the master equation is the propensity function $a_\xi(q,t)$, which governs the rate of transition between microstates. Although the detailed value of the propensity function can be quite complex, its functional form is often relatively simple. In particular, for a unimolecular reaction of the form $\mathrm{A} \to B$, the propensity function is proportional to the number of molecules of A that are present:

$$a_i(q,t) = k_i n_A. \tag{4.6}$$

This follows from the fact that each reaction is independent and hence the likelihood of a reaction happening depends directly on the number of copies of A that are present.

Similarly, for a bimolecular reaction, we have that the likelihood of a reaction occurring is proportional to the product of the number of molecules of each type that are present (since this is the number of independent reactions that can occur) and inversely proportional to the volume Ω. Hence, for a reaction of the form $\mathrm{A} + \mathrm{B} \longrightarrow \mathrm{C}$ we have

$$a_i(q,t) = \frac{k_i}{\Omega} n_A n_B. \tag{4.7}$$

The rigorous verification of this functional form is beyond the scope of this text, but roughly we keep track of the likelihood of a single reaction occurring between A and B and then multiply by the total number of combinations of the two molecules that can react ($n_\mathrm{A} \cdot n_\mathrm{B}$).

A special case of a bimolecular reaction occurs when $\mathrm{A} = \mathrm{B}$, so that our reaction is given by $2\,\mathrm{A} \to B$. In this case we must take into account that a molecule cannot react with itself, and so the propensity function is of the form

$$a_i(q,t) = \frac{k_i}{\Omega} n_A (n_A - 1). \tag{4.8}$$

The term $n_A(n_A - 1)$ reprents the number of ways that two molecules can be chosen from a collection of n_A identical molecules.

Table 4.2: Examples of propensity functions for some common cases [35]. Here we take r_a and r_b to be the effective radii of the molecules, $m^* = m_a m_b/(m_a + m_b)$ is the reduced mass of the two molecules, Ω is the volume over which the reaction occurs, T is temperature, k_B is Boltzmann's constant and n_A, n_B are the numbers of molecules of A and B present.

Reaction type	Propensity function coefficient, k_i
Reaction occurs if molecules "touch"	$\left(\frac{8k_BT}{\pi m^*}\right)^{1/2} \pi(r_a + r_b)^2$
Reaction occurs if molecules collide with energy ϵ	$\left(\frac{8k_BT}{\pi m^*}\right)^{1/2} \pi(r_a + r_b)^2 \cdot e^{-\epsilon/k_BT}$
Steady state transcription factor	$P_{\text{bound}} k_{\text{oc}} n_{\text{RNAP}}$

Note that the use of the parameter k_i in the propensity functions above is intentional: it corresponds to the reaction rate parameter that is present in the reaction rate equation model. The factor of Ω for biomolecular reactions models the fact that the propensity of a biomolecular reaction occuring depends explicitly on the volume in which the reaction takes place.

Although it is tempting to extend the formula for a biomolecular reaction to the case of more than two species being involved in a reaction, usually such reactions actually involve combinations of bimolecular reactions, e.g.:

$$\mathrm{A} + \mathrm{B} + \mathrm{C} \longrightarrow \mathrm{D} \qquad \Longrightarrow \qquad \mathrm{A} + \mathrm{B} \longrightarrow \mathrm{AB} \quad \mathrm{AB} + \mathrm{C} \longrightarrow \mathrm{D}$$

This more detailed description reflects that fact that it is extremely unlikely that three molecules will all come together at precisely the same instant, versus the much more likely possibility that two molecules will initially react, followed be a second reaction involving the third molecule.

The propensity functions for these cases and some others are given in Table 4.2.

Example 4.3 (Repression of gene expression)**.** We consider a simple model of repression in which we have a promoter that contains binding sites for RNA polymerase and a repressor protein R. RNA polymerase only binds when the repressor is absent, after which it can undergo an isomerization reaction to form an open complex and initiate transcription. Once the RNA polymerase begins to create mRNA, we assume the promoter region is uncovered, allowing another repressor or RNA polymerase to bind.

The following reactions describe this process:

$$\begin{aligned}
&\texttt{R1}: && \mathrm{R} + \mathrm{DNA} \rightleftharpoons \mathrm{R{:}DNA} \\
&\texttt{R2}: && \mathrm{RNAP} + \mathrm{DNA} \rightleftharpoons \mathrm{RNAP{:}DNA^c} \\
&\texttt{R3}: && \mathrm{RNAP{:}DNA^c} \longrightarrow \mathrm{RNAP{:}DNA^o} \\
&\texttt{R4}: && \mathrm{RNAP{:}DNA^o} \longrightarrow \mathrm{RNAP} + \mathrm{DNA} \quad (+\mathrm{mRNA}),
\end{aligned}$$

where $\mathrm{RNAP{:}DNA^c}$ represents the closed complex and $\mathrm{RNAP{:}DNA^o}$ represents the open complex. The states for the system depend on the number of molecules

of each species and complex that are present. If we assume that we start with n_R repressors and n_{RNAP} RNA polymerases, then the possible states for our system are given by

State	DNA	R	RNAP	R:DNA	RNAP:DNA^c	RNAP:DNA^o
q_1	1	n_R	n_{RNAP}	0	0	0
q_2	0	$n_R - 1$	n_{RNAP}	1	0	0
q_3	0	n_R	$n_{RNAP} - 1$	0	1	0
q_4	0	n_R	$n_{RNAP} - 1$	0	0	1

Note that we do not keep track of each individual repressor or RNA polymerase molecule that binds to the DNA, but simply keep track of whether they are bound or not.

We can now rewrite the chemical reactions as a set of transitions between the possible microstates of the system. Assuming that all reactions take place in a volume Ω, we use the propensity functions for unimolecular and bimolecular reactions to obtain:

$$\begin{aligned}
&\xi_1^f: \quad \mathrm{q}_1 \longrightarrow \mathrm{q}_2; \quad a(\xi_1^f) = (k_1^f/\Omega)\, n_R && \xi_1^r: \quad \mathrm{q}_2 \longrightarrow \mathrm{q}_1; \quad a(\xi_1^r) = k_1^r \\
&\xi_2^f: \quad \mathrm{q}_1 \longrightarrow \mathrm{q}_3; \quad a(\xi_2^f) = (k_2^f/\Omega)\, n_{RNAP} && \xi_2^r: \quad \mathrm{q}_3 \longrightarrow \mathrm{q}_1; \quad a(\xi_2^r) = k_2^r \\
&\xi_3: \quad \mathrm{q}_3 \longrightarrow \mathrm{q}_4; \quad a(\xi_3) = k_3 && \xi_4: \quad \mathrm{q}_4 \longrightarrow \mathrm{q}_1; \quad a(\xi_4^r) = k_4
\end{aligned}$$

The chemical master equation can now be written down using the propensity functions for each reaction:

$$\frac{d}{dt}\begin{pmatrix} P(q_1,t) \\ P(q_2,t) \\ P(q_3,t) \\ P(q_4,t) \end{pmatrix} = \begin{pmatrix} -(k_1^f/\Omega)n_R - (k_2^f/\Omega)n_{RNAP} & k_1^r & k_2^r & k_4 \\ (k_1^f/\Omega)n_R & -k_1^r & 0 & 0 \\ (k_2^f/\Omega)n_{RNAP} & 0 & -k_2^r - k_3 & 0 \\ 0 & 0 & k_3 & -k_4 \end{pmatrix} \begin{pmatrix} P(q_1,t) \\ P(q_2,t) \\ P(q_3,t) \\ P(q_4,t) \end{pmatrix}.$$

The initial condition for the system can be taken as $P(q,0) = (1,0,0,0)$, corresponding to the state q_1. A simulation showing the evolution of the probabilities is shown in Figure 4.1.

The equilibrium solution for the probabilities can be solved by setting $\dot{P} = 0$,

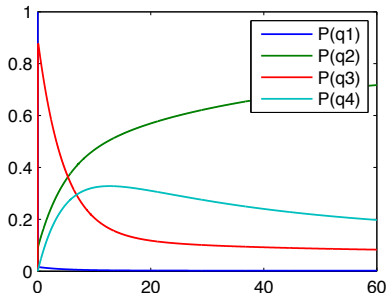

Figure 4.1: Numerical solution of chemical master equation for simple repression model.

which yields:

$$P_e(q_1) = \frac{k_1^r k_4 \Omega (k_2^r + k_3)}{k_1^f k_4 n_{\mathrm{R}} (k_2^r + k_3) + k_1^r k_2^f n_{\mathrm{RNAP}} (k_3 + k_4) + k_1^r k_4 \Omega (k_2^r + k_3)}$$

$$P_e(q_2) = \frac{k_1^f k_4 n_{\mathrm{R}} (k_2^r + k_3)}{k_1^f k_4 n_{\mathrm{R}} (k_2^r + k_3) + k_1^r k_2^f n_{\mathrm{RNAP}} (k_3 + k_4) + k_1^r k_4 \Omega (k_2^r + k_3)}$$

$$P_e(q_3) = \frac{k_1^r k_2^f k_4 n_{\mathrm{RNAP}}}{k_1^f k_4 n_{\mathrm{R}} (k_2^r + k_3) + k_1^r k_2^f n_{\mathrm{RNAP}} (k_3 + k_4) + k_1^r k_4 \Omega (k_2^r + k_3)}$$

$$P_e(q_4) = \frac{k_1^r k_2^f k_3 n_{\mathrm{RNAP}}}{k_1^f k_4 n_{\mathrm{R}} (k_2^r + k_3) + k_1^r k_2^f n_{\mathrm{RNAP}} (k_3 + k_4) + k_1^r k_4 \Omega (k_2^r + k_3)}$$

We see that the functional dependencies are similar to the case of the combinatorial promoter of Example 4.2, but with the binding energies replaced by kinetic rate constants. ∇

Example 4.4 (Transcription of mRNA)**.** Consider the production of mRNA from a single copy of DNA. We have two basic reactions that can occur: mRNA can be produced by RNA polymerase transcribing the DNA and producing an mRNA strand, or mRNA can be degraded. We represent the microstate q of the system in terms of the number of mRNA's that are present, which we write as n for ease of notation. The reactions can now be represented as $\xi_1 = +1$, corresponding to transcription and $\xi_2 = -1$, corresponding to degradation. We choose as our propensity functions

$$a_1(n,t) = \alpha, \qquad a_2(n,t) = \gamma n,$$

by which we mean that the probability of that a gene is transcribed in time dt is αdt and the probability that a transcript is created in time dt is $\gamma n dt$ (proportional to the number of mRNA's).

We can now write down the master equation as described above. Equation (4.4) becomes

$$\begin{aligned} P(n,t+dt) &= P(n,t)\Big(1-\sum_{i=1,2} a_i(n,t)dt\Big) + \sum_{i=1,2} P(n-\xi_i,t)a_i(q-\xi_i)dt \\ &= P(n,t) - a_1(n,t)P(n,t) - a_2(n,t)P(n,t) \\ &\quad + a_1(n-1,t)P(n-1,t) + a_2(n+1,t)P(n+1) \\ &= P(n,t) + \alpha P(n-1,t)dt - (\alpha-\gamma n)P(n,t)dt + \gamma(n+1)P(n+1,t)dt. \end{aligned}$$

This formula holds for $n > 0$, with the $n = 0$ case satisfying

$$P(0,t+dt) = P(0,t) - \alpha P(0,t)dt + \gamma P(1,t)dt.$$

Notice that we have an infinite number of equations, since n can be any positive integer.

We can write the differential equation version of the master equation by subtracting the first term on the right hand side and dividing by dt:

$$\begin{aligned} \frac{d}{dt}P(n,t) &= \alpha P(n-1,t) - (\alpha+\gamma n)P(n,t) + \gamma(n+1)P(n+1,t), \qquad n>0 \\ \frac{d}{dt}P(0,t) &= -\alpha P(0,t)dt + \gamma P(1,t). \end{aligned}$$

Again, this is an infinite number of differential equations, although we could take some limit N and simply declare that $P(N,t) = 0$ to yield a finite number.

One simple type of analysis that can be done on this equation without truncating it to a finite number is to look for a steady state solution to the equation. In this case, we set $\dot{P}(n,t) = 0$ and look for a constant solution $P(n,t) = p_e(n)$. This yields an algebraic set of relations

$$\begin{array}{lll} 0 = -\alpha p_e(0) + \gamma p_e(1) & \Longrightarrow & \alpha p_e(0) = \gamma p_e(1) \\ 0 = \alpha p_e(0) - (\alpha+\gamma)p_e(1) + 2\gamma p_e(2) & & \alpha p_e(1) = 2\gamma p_e(2) \\ 0 = \alpha p_e(1) - (\alpha+2\gamma)p_e(2) + 3\gamma p_e(3) & & \alpha p_e(1) = 3\gamma p_e(3) \\ \quad\vdots & & \quad\vdots \\ & & \alpha p(n-1) = n\gamma p(n). \end{array}$$

It follows that the distribution of steady state probabilities is given by the Poisson distribution

$$p(n) = e^{\alpha/\gamma}\frac{(\alpha/\gamma)^n}{n!},$$

and the mean, variance and coefficient of variation are thus

$$\mu = \frac{\alpha}{\gamma}, \qquad \sigma^2 = \frac{\alpha}{\gamma}, \qquad CV = \frac{\mu}{\sigma} = \frac{1}{\sqrt{\mu}} = \sqrt{\frac{\gamma}{\alpha}}.$$

Note taht the coefficient of variation increases if μ decreases. ∇

Chemical Langevin equation (CLE)

The chemical master equation gives a complete description of the evolution of the distribution of a system, but it can often be quite cumbersome to work with directly. A number of approximations to the master equation are thus used to provide more tractable formulations of the dynamics. The first of these that we shall consider is known as the *chemical Langevin equation* (CLE).

To derive the chemical Langevin equation, we start by assuming that the number of molecules in the system is large and that we can therefore represent the system using a vector of real numbers X, with X_i representing the (real-valued) number of molecules in S_i. (Often X_i will be divided by the volume to give a real-valued concentration of species S_i.) In addition, we assume that we are interested in the dynamics on time scales in which individual reactions are not important and so we can look at how the system state changes over time intervals in which many reactions occur and hence the system state evolves in a smooth fashion.

Let $X(t)$ be the state vector for the system, where we assume now that the elements of X are real-valued rather than integer valued. We make the further approximation that we can lump together multiple reactions so that instead of keeping track of the individual reactions, we can average across a number of reactions over a time τ to allow the continuous state to evolve in continuous time. The resulting dynamics can be described by a stochastic process of the form

$$X_i(t+\tau) = X_i(t) + \sum_{j=1}^{M} \xi_{ij} a_j(X(t))\tau + \sum_{j=1}^{M} \xi_{ij} a_j^{1/2}(X(t)) \mathcal{N}_j(0, \sqrt{\tau}),$$

where a_j are the propensity functions for the individual reactions, ξ_{ij} are the corresponding changes in the system states X_i and $\mathcal{N}_j$ are a set of independent Gaussian random variables with zero mean and variance τ.

If we assume that τ is small enough that we can use the derivative to approximate the previous equation (but still large enough that we can average over multiple reactions), then we can write

$$\frac{dX_i(t)}{dt} = \sum_{j=1}^{M} \xi_{ji} a_j(X(t)) + \sum_{j=1}^{M} \xi_{ji} a_j^{1/2}(X(t)) \Gamma_j(t) \quad =: A_i(X(t)) + \sum_{j=1}^{M} B_{ij}(X(t)) \Gamma_j(t), \tag{4.9}$$

where Γ_j are white noise processes (see Appendix B.2). This equation is called the *chemical Langevin equation* (CLE).

Example 4.5 (Protein production)**.** Consider a simplified model of protein production in which mRNAs are produced by transcription and proteins by translation. We also include degradation of both mRNAs and proteins, but we do not model the detailed processes of elongation of the mRNA and polypeptide chains.

We can capture the state of the system by keeping track of the number of copies of mRNA and proteins. We further approximate this by assuming that the number

of each of these is sufficiently large that we can keep track of its concentration, and hence $X = (x_m, x_p)$ where $x_m \in \mathbb{R}$ is the amount of mRNA and $x_p \in \mathbb{R}$ is the concentration of protein. Letting Ω represent the volume, the reactions that govern the dynamics of the system are given by:

$$
\begin{aligned}
&R_1 : \phi \xrightarrow{\alpha} \text{mRNA} && \xi_1 = (1,0) && a_1(X) = \alpha \\
&R_2 : \text{mRNA} \xrightarrow{\gamma} \phi && \xi_2 = (-1,0) && a_2(X) = \gamma\, x_m \\
&R_3 : \text{mRNA} \xrightarrow{\beta} \text{mRNA} + \text{protein} && \xi_3 = (0,1) && a_3(X) = \beta\, x_m \\
&R_4 : \text{protein} \xrightarrow{\delta} \phi && \xi_4 = (0,-1) && a_4(X) = \delta\, x_p.
\end{aligned}
$$

Substituting these expressions into equation (4.9), we obtain a stochastic differential equation of the form

$$
\frac{d}{dt}\begin{pmatrix} x_m \\ x_p \end{pmatrix} = \begin{pmatrix} -\gamma & 0 \\ \beta & -\delta \end{pmatrix} \begin{pmatrix} x_m \\ x_p \end{pmatrix} + \begin{pmatrix} \alpha \\ 0 \end{pmatrix} + \sqrt{\Omega} \begin{pmatrix} \left(\sqrt{\alpha + \gamma x_m}\right)\Gamma_m \\ \left(\sqrt{\beta x_m + \delta x_p}\right)\Gamma_p \end{pmatrix},
$$

where Γ_m and Γ_p are independent white noise processes with unit variance. (Note that in deriving this equation we have used the fact that the sum of two independent Gaussian processes is a Gaussian process.) ∇

Fokker-Planck equations (FPE)

The chemical Langevin equation provides a stochastic ordinary differential equation that describes the evolution of the system state. A slightly different (but completely equivalent) representation of the dynamics is to model how the probability distribution $P(x,t)$ evolves in time. As in the case of the chemical Langevin equation, we will assume that the system state is continuous and write down a formula for the evolution of the density function $p(x,t)$. This formula is known as the *Fokker-Planck equations* (FPE) and is essentially an approximation on the chemical master equation.

Consider first the case of a random process in one dimension. We assume that the random process is in the same form as the previous section:

$$
\frac{dX(t)}{dt} = A(X(t)) + B(X(t))\Gamma(t). \tag{4.10}
$$

The function $A(X)$ is called the *drift term* and $B(X)$ is the *diffusion term*. It can be shown that the probability density function for X, $p(x,t)$, satisfies the partial differential equation

$$
\frac{\partial p}{\partial t}(x,t) = -\frac{\partial}{\partial x}(A(x,t)p(x,t)) + \frac{1}{2}\frac{\partial^2}{\partial x^2}(B^2(x,t)p(x,t)) \tag{4.11}
$$

Note that here we have shifted to the probability density function since we are considering X to be a continuous state random process.

In the multivariate case, a bit more care is required. Using the chemical Langevin equation (4.9), we define

$$D_i(x,t) = \sum_{j=1}^{M} B_{ij}^2(x,t), \qquad C_{ij}(x,t) = \sum_{k=1}^{M} B_{ik}(x,t)B_{jk}(x,t),\ i < j = 1,\ldots,M.$$

The Fokker-Planck equation now becomes

$$\frac{\partial p}{\partial t}(x,t) = -\sum_{i=1}^{M} \frac{\partial}{\partial x_i}(A_i(x,t)p(x,t)) + \frac{1}{2}\sum_{i=1}^{M} \frac{\partial}{\partial x_i}\frac{\partial^2}{\partial x^2}(D_i(x,t)p(x,t)) + \sum_{\substack{i,j=1 \\ i<j}}^{M} \frac{\partial^2}{\partial x_i \partial x_j}(C_{ij}(x,t)p(x,t)). \quad (4.12)$$

Note that the Fokker-Planck equation is very similar to the chemical master equation: both provide a description of how the probability distribution varies as a function of time. In the case of the Fokker-Planck equation, we regard the state as a continuous set of variables and we write a partial differential equation for how the probability density function evolves in time. In the case of the chemical master equation, we have a discrete state (microstates) and we write an ordinary differential equation for how the probability distribution (formally the probability mass function) evolves in time. Both formulations contain the same basic information, just using slightly different representations of the system and the probability of being in a given state.

Linear noise approximation (LNA)

The chemical Langevin equation and the Fokker-Planck equation provide approximations to the chemical master equation. A slightly different approximation can be obtained by expanding the density function in terms of a size parameter Ω. This approximation is know as the *linear noise approximation* (LNA) or the Ω *expansion* [52].

We begin with the master equation for a *continuous* random variable X. Formally deriving this requires a considerable effort since we have to extend our previous discussions to the case where the random variable has a continuous set of values. To do this, we rewrite the propensity function $a_i(q,t)$ as $a_\xi(q,t;\Omega)$, where $q \in \mathbb{R}^n$ is a vector of continuous states and $\xi \in \mathbb{R}^n$ is a vector of continuous "increments" (the analog of reactions). We also explicitly keep track of the dependence of the propensity function on a parameter Ω (the volume in our case).

Using this notation, we can write the master equation for the random variable X as

$$\frac{\partial P}{\partial t}(x,t) = \int \left(a_\xi(x-\xi,t;\Omega)P(x-\xi,t) - a_\xi(x,t;\Omega)P(x,t)\right) d\xi.$$

Since we are working with continuous variables, we now have an integral in place of our previous sum. In addition, if we take the derivative of $P(x,t)$ with respect to the continuous variable x, we can obtain the pdf of the distribution $p(x,t)$ and this satisfies the equation

$$\frac{\partial p}{\partial t}(x,t) = \int \left(a_\xi(x-\xi,t;\Omega)p(x-\xi,t) - a_\xi(x,t;\Omega)p(x,t)\right) d\xi.$$

Although we are skipping important theoretical details, the basic idea of this formulation is the same as the discrete chemical master equation: we keep track of how the probability density changes by "summing" (integrating) over all (incremental) reactions going into and out of that particular state.

We now assume that the mean of X can be written as $\Omega\phi(t)$ where $\phi(t)$ is a continuous function of time that represents the evolution of the mean of X/Ω. To understand the fluctuations of the system about this mean, we write

$$X = \Omega\phi + \Omega^{\frac{1}{2}} Z,$$

where Z is a new variable representing the perturbations of the system about its mean. We can write the distribution for Z as

$$p_Z(z,t) = p_X(\Omega\phi(t) + \Omega^{\frac{1}{2}} z, t)$$

and it follows that the derivatives of p_Z can be written as

$$\frac{\partial^\nu p_Z}{z^\nu} = \Omega^{\frac{1}{2}\nu} \frac{\partial^\nu p_X}{x^\nu}$$
$$\frac{\partial p_Z}{\partial t} = \frac{\partial p_X}{\partial t} + \Omega \frac{d\phi}{dt}\frac{\partial p_X}{\partial x} = \frac{\partial p_X}{\partial t} + \Omega^{\frac{1}{2}} \frac{d\phi}{dt}\frac{\partial p_Z}{\partial z}.$$

We further assume that the Ω dependence of the propensity function is such that

$$a_\xi(\Omega\phi, t; \Omega) = f(\Omega)\tilde{a}_\xi(\phi),$$

where $\tilde{a}$ is not dependent on the parameter Ω or the time t. From these relations, we can now derive the master equation for p_Z in terms of powers of Ω (derivation omitted).

The $\Omega^{1/2}$ term in the expansion turns out to yield

$$\frac{d\phi}{dt} = \int \xi\, a_\xi(\Omega\phi)\, d\xi, \qquad \phi(0) = \frac{X(0)}{\Omega},$$

which is precisely the equation for the mean of the concentration. It can further be shown that the terms in Ω^0 are given by

$$\frac{\partial p_Z(z,\tau)}{\partial \tau} = -\alpha_1'(\phi)\frac{\partial}{\partial z}(z p_Z(z,t)) + \frac{1}{2}\alpha_2(\phi)\frac{\partial^2 p_Z(z,t)}{\partial z^2}, \tag{4.13}$$

where

$$\alpha_v(x) = \int \xi^v \tilde{a}_\xi(x)\, d\xi, \qquad \tau = \Omega^{-1} f(\Omega) t.$$

Notice that in the case that $\phi(t) = \phi_0$ (a constant), this equation becomes the Fokker-Planck equation derived previously.

Higher order approximations to this equation can also be carried out by keeping track of the expansion terms in higher order powers of Ω. In the case where Ω represents the volume of the system, the next term in the expansion is Ω^{-1} and this represents fluctuations that are on the order of a single molecule, which can usually be ignored.

Reaction rate equations (RRE)

As we already saw in Chapter 2, the reaction rate equations can be used to describe the dynamics of a chemical system in the case where there are a large number of molecules whose state can be approximated using just the concentrations of the molecules. We re-derive the results from Section 2.1 here, being more careful to point out what approximations are being made.

We start with the chemical Langevin equations (4.9), from which we can write the dynamics for the average quantity of the each species at each point in time:

$$\frac{d\langle X_i(t)\rangle}{dt} = \sum_{j=1}^{M} \xi_{ji} \langle a_j(X(t))\rangle, \tag{4.14}$$

where the second order term drops out under the assumption that the Γ_j's are independent processes with zero mean. We see that the reaction rate equations follow by defining $x_i = \langle X_i\rangle/\Omega$ and *assuming* that $\langle a_j(X(t))\rangle = a_j(\langle X(t)\rangle)$. This relationship is true when a_j is linear (e.g., in the case of a unimolecular reaction), but is an approximation otherwise.

4.2 Simulation of Stochastic Systems

Suppose that we want to generate a collection of sample trajectories for a stochastic system whose evolution is described by the chemical master equation (4.5):

$$\frac{d}{dt}P(q,t) = \sum_i a_i(q-\xi_i)P(q-\xi_i,t) - \sum_i a_i(q)P(q,t),$$

where $P(q,t)$ is the probability of being in a microstate q at time t (starting from q_0 at time t_0) and $a_i(q)$ is the propensity function for a reaction i starting at a microstate q and ending at microstate $q+\xi_i$. Instead of simulating the distribution function $P(q,t)$, we wish to simulate a specific instance $q(t)$ starting from some initial condition $q_0(t_0)$. If we simulate many such instances of $q(t)$, their distribution at time t should match $P(q,t)$.

To illustrate the basic ideas that we will use, consider first a simple birth process in which the microstate is given by an integer $q \in \{0,1,2,\dots\}$ and we assume that the propensity function is given by

$$a(q)\,dt = \lambda\,dt, \qquad \xi = +1.$$

Thus the probability of transition increases linearly with the time increment dt (so birth events occur at rate λ, on average). If we assume that the birth events are independent of each other, then it can be shown (see Appendix B) that this process has Poisson distribution with parameter $\lambda\tau$:

$$P(q(t+\tau)-q(t)=\ell) = \frac{(\lambda\tau)^\ell}{\ell!}e^{-\lambda\tau},$$

where τ is the difference in time and ℓ is the difference in count q. In fact, this distribution is a joint distribution in time τ and count ℓ, and by setting $\ell = 1$ it can be seen that the time to the next reaction T follows an exponential distribution and has density function

$$p_T(\tau) = \lambda\, e^{-\lambda\tau}.$$

The exponential distribution has expectation $1/\lambda$ and so we see that the average time between events is inversely proportional to the reaction rate λ.

Consider next a more general case in which we have a countable number of microstates $q \in \{0,1,2,\dots\}$ and we let k_{ji} represent the transition probability between a microstate i and microstate j. The birth process is a special case given by $k_{i+1,i} = \lambda$ and all other $k_{ji} = 0$. The chemical master equation describes the joint probability that we are in state $q = i$ at a particular time t. We would like to know the probability that we transition to a new state $q = j$ at time $t+dt$. Given this probability, we can attempt to generate an instance of the variable $q(t)$ by first determining which reaction occurs and then when the reaction occurs.

Let $P(j,\tau) := P(j,t+\tau+d\tau \mid i,t+\tau)$ represent the probability that we transition from the state i to the state j in the time interval $[t+\tau, t+\tau+d\tau]$. For simplicity and ease of notation, we will take $t = 0$. Let $T := T_{j,i}$ be the time at which the reaction first occurs. We can write the probability that we transition to state j in the interval $[\tau,\tau+d\tau]$ as

$$P(j,\tau) = P(T>\tau)\,k_{ji}\,d\tau, \tag{4.15}$$

where $P(T>\tau)$ is the probability that no reaction occurs in the time interval $[0,\tau]$ and $k_{ji}d\tau$ is the probability that the reaction taking state i to state j occurs in the

next $d\tau$ seconds (assumed to be independent events, giving the product of these probabilities).

To compute $P(T > \tau)$, define

$$k_i = \sum_j k_{ji}$$

so that $(1 - k_i)d\tau$ is the probability that no transition occurs from state i in the next $d\tau$ seconds. Then, the probability that no reaction occurs in the interval $[\tau, \tau + d\tau]$ can be written as

$$P(T > \tau + d\tau) = P(T > \tau)(1 - k_i)\, d\tau. \tag{4.16}$$

It follows that

$$\frac{d}{d\tau} P(T > \tau) = \lim_{d\tau \to 0} \frac{P(T > \tau + d\tau) - P(T > \tau)}{d\tau} = -P(T > \tau)\, k_i.$$

Solving this differential equation, we obtain

$$P(T > \tau) = e^{-k_i \tau}, \tag{4.17}$$

so that the probability that no reaction occurs in time τ decreases exponentially with the amount of time that we wait, with rate given by the sum of all the reactions that can occur from state i.

We can now combine equation (4.17) with equation (4.15) to obtain

$$P(j, \tau) = P(j, \tau + d\tau \mid i, 0) = k_{ji}\, e^{-k_i \tau}\, d\tau.$$

We see that this has the form of a density function in time and hence the probability that the next reaction is reaction j, independent of the time in which it occurs, is

$$P_{ji} = \int_0^\infty k_{ji} e^{-k_i \tau}\, d\tau = \frac{k_{ji}}{k_i}. \tag{4.18}$$

Thus, to choose the next reaction to occur from a state i, we choose between N possible reactions, with the probability of each reaction weighted by k_{ji}/k_i.

To determine the time that the next reaction occurs, we sum over all possible reactions j to get the density function for the reaction time:

$$p_T(\tau) = \sum_j k_{ji} e^{-k_i \tau} = k_i e^{-k_i \tau}.$$

This is the density function associated with a Poisson distribution. To compute a time of reaction Δt that draws from this distribution, we note that the cumulative distribution function for T is given by

$$\int_0^{\Delta t} f_T(\tau)\, d\tau = \int_0^{\Delta t} k_i e^{-k_i \tau}\, d\tau = 1 - e^{-k_i \Delta t}.$$

The cumulative distribution function is always in the range $[0,1]$ and hence we can compute Δt by choosing a (uniformly distributed) random number r in $[0,1]$ and then computing

$$\Delta t = \frac{1}{k_i} \ln \frac{1}{1-r}. \tag{4.19}$$

(This equation can be simplified somewhat by replacing $1-r$ with r' and noting that r' can also be drawn from a uniform distribution on $[0,1]$.)

Note that in the case of a birth process, this computation agrees with our earlier analysis. Namely, $k_i = \lambda$ and hence the (only) reaction occurs according to an exponential distribution with parameter λ.

This set of calculations gives the following algorithm for computing an instance of the chemical master equation:

1. Choose an initial condition q at time $t = 0$.
2. Calculate the propensity functions $a_\xi(q)$ for each possible reaction q.
3. Choose the time for the reaction according to equation (4.19), where $r \in [0,1]$ is chosen from a uniform distribution.
4. Use a weighted random number generator to identify which reaction will take place next, using the weights in equation (4.18).
5. Update q by implementing the reaction ξ and update the time t by δt
6. If $T < T_{\text{stop}}$, goto step 2.

This method is sometimes called "Gillespie's direct method" [33, 34], but we shall refer to it here as the "stochastic simulation algorithm" (SSA). We note that the reaction number in step 4 can be computed by calculating a uniform random number on $[0,1]$, scaling this by the total propensity $\sum_i a(\xi_i, q)$, and then finding the first reaction i such that $\sum_0^i a(\xi_i, q)$ is larger than this scaled random number.

Example 4.6 (Transcription)**.** To be completed. ∇ Review

4.3 Input/Output Linear Stochastic Systems

In many situations, we wish to noise how noise propogates through a biomolecular system. For example, we may wish to understand how stochastic variations in RNA polymerase concentraton affect gene expression. In order to analyze these cases, we specialize to the case of a biomolecular system operating around a fixed operating point.

We now consider the problem of how to compute the response of a linear system to a random process. We assume we have a linear system described in state space as

$$\dot{X} = AX + FW, \qquad Y = CX \tag{4.20}$$

Given an "input" W, which is itself a random process with mean $\mu(t)$, variance $\sigma^2(t)$ and correlation $r(t, t+\tau)$, what is the description of the random process Y?

Let W be a white noise process, with zero mean and noise intensity Q:

$$r(\tau) = Q\delta(\tau).$$

We can write the output of the system in terms of the convolution integral

$$Y(t) = \int_0^t h(t-\tau)W(\tau)\,d\tau,$$

where $h(t-\tau)$ is the impulse response for the system

$$h(t-\tau) = Ce^{A(t-\tau)}B + D\delta(t-\tau).$$

We now compute the statistics of the output, starting with the mean:

$$\begin{aligned}\mathbb{E}(Y(t)) &= \mathbb{E}\Big(\int_0^t h(t-\eta)W(\eta)\,d\eta\Big)\\ &= \int_0^t h(t-\eta)\mathbb{E}(W(\eta))\,d\eta = 0.\end{aligned}$$

Note here that we have relied on the linearity of the convolution integral to pull the expectation inside the integral.

We can compute the covariance of the output by computing the correlation $r_Y(\tau)$ and setting $\sigma_Y^2 = r_Y(0)$. The correlation function for y is

$$\begin{aligned}r_Y(t,s) = \mathbb{E}(Y(t)Y(s)) &= \mathbb{E}\Big(\int_0^t h(t-\eta)W(\eta)\,d\eta \cdot \int_0^s h(s-\xi)W(\xi)\,d\xi\Big)\\ &= \mathbb{E}\Big(\int_0^t\int_0^s h(t-\eta)W(\eta)W(\xi)h(s-\xi)\,d\eta d\xi\Big)\end{aligned}$$

Once again linearity allows us to exchange expectation and integration

$$\begin{aligned}r_Y(t,s) &= \int_0^t\int_0^s h(t-\eta)\mathbb{E}(W(\eta)W(\xi))h(s-\xi)\,d\eta d\xi\\ &= \int_0^t\int_0^s h(t-\eta)Q\delta(\eta-\xi)h(s-\xi)\,d\eta d\xi\\ &= \int_0^t h(t-\eta)Qh(s-\eta)\,d\eta\end{aligned}$$

Now let $\tau = s-t$ and write

$$\begin{aligned}r_Y(\tau) = r_Y(t,t+\tau) &= \int_0^t h(t-\eta)Qh(t+\tau-\eta)\,d\eta\\ &= \int_0^t h(\xi)Qh(\xi+\tau)\,d\xi \qquad \text{(setting } \xi = t-\eta\text{)}\end{aligned}$$

Finally, we let $t \to \infty$ (steady state)

$$\lim_{t\to\infty} r_Y(t, t+\tau) = r_Y(\tau) = \int_0^\infty h(\xi) Q h(\xi+\tau) d\xi \tag{4.21}$$

If this integral exists, then we can compute the second order statistics for the output Y.

We can provide a more explicit formula for the correlation function r in terms of the matrices A, F and C by expanding equation (4.21). We will consider the general case where $W \in \mathbb{R}^p$ and $Y \in \mathbb{R}^q$ and use the correlation matrix $R(t,s)$ instead of the correlation function $r(t,s)$. Define the *state transition matrix* $\Phi(t,t_0) = e^{A(t-t_0)}$ so that the solution of system (4.20) is given by

$$x(t) = \Phi(t,t_0)x(t_0) + \int_{t_0}^t \Phi(t,\lambda) F w(\lambda) d\lambda$$

Proposition 4.1 (Stochastic response to white noise)**.** *Let* $\mathbb{E}(X(t_0)X^T(t_0)) = P(t_0)$ *and* W *be white noise with* $\mathbb{E}(W(\lambda)W^T(\xi)) = R_W\delta(\lambda-\xi)$. *Then the correlation matrix for* X *is given by*

$$R_X(t,s) = P(t)\Phi^T(s,t)$$

where $P(t)$ *satisfies the linear matrix differential equation*

$$\dot{P}(t) = AP + PA^T + FR_WF, \qquad P(0) \quad = P_0.$$

Proof. Using the definition of the correlation matrix, we have

$$\begin{aligned}
\mathbb{E}(X(t)X^T(s)) &= \mathbb{E}\Big(\Phi(t,0)X(0)X^T(0)\Phi^T(t,0) + \text{cross terms} \\
&\qquad + \int_0^t \Phi(t,\xi)FW(\xi)\,d\xi \int_0^s W^t(\lambda)F^T\Phi(s,\lambda)\,d\lambda\Big) \\
&= \Phi(t,0)\mathbb{E}(X(0)X^T(0))\Phi(s,0) \\
&\qquad + \int_0^t\int_0^s \Phi(t,\xi)F\mathbb{E}(W(\xi)W^T(\lambda))F^T\Phi(s,\lambda)\,d\xi\,d\lambda \\
&= \Phi(t,0)P(0)\phi^T(s,0) + \int_0^t \Phi(t,\lambda)FR_W(\lambda)F^T\Phi(s,\lambda)\,d\lambda.
\end{aligned}$$

Now use the fact that $\Phi(s,0) = \Phi(s,t)\Phi(t,0)$ (and similar relations) to obtain

$$R_X(t,s) = P(t)\Phi^T(s,t)$$

where

$$P(t) = \Phi(t,0)P(0)\Phi^T(t,0) + \int_0^T \Phi(t,\lambda)FR_WF^T(\lambda)\Phi^T(t,\lambda)d\lambda$$

Finally, differentiate to obtain

$$\dot{P}(t) = AP + PA^T + FR_WF, \qquad P(0) \quad = P_0$$

(see Friedland [30] for details). □

The correlation matrix for the output Y can be computed using the fact that $Y = CX$ and hence $R_Y = C^T R_X C$. We will often be interested in the steady state properties of the output, which are given by the following proposition.

Proposition 4.2 (Steady state response to white noise)**.** *For a time-invariant linear system driven by white noise, the correlation matrices for the state and output converge in steady state to*

$$R_X(\tau) = R_X(t, t+\tau) = Pe^{A^T\tau}, \qquad R_Y(\tau) = CR_X(\tau)C^T$$

where P satisfies the algebraic equation

$$AP + PA^T + FR_W F^T = 0 \qquad P > 0. \tag{4.22}$$

Equation (4.22) is called the *Lyapunov equation* and can be solved in MATLAB using the function `lyap`.

Example 4.7 (First-order system)**.** Consider a scalar linear process

$$\dot{X} = -aX + W, \qquad Y = cX,$$

where W is a white, Gaussian random process with noise intensity σ^2. Using the results of Proposition 4.1, the correlation function for X is given by

$$R_X(t, t+\tau) = p(t)e^{-a\tau}$$

where $p(t) > 0$ satisfies

$$p(t) = -2ap + \sigma^2.$$

We can solve explicitly for $p(t)$ since it is a (non-homogeneous) linear differential equation:

$$p(t) = e^{-2at}p(0) + (1 - e^{-2at})\frac{\sigma^2}{2a}.$$

Finally, making use of the fact that $Y = cX$ we have

$$r(t, t+\tau) = c^2(e^{-2at}p(0) + (1 - e^{-2at})\frac{\sigma^2}{2a})e^{-a\tau}.$$

In steady state, the correlation function for the output becomes

$$r(\tau) = \frac{c^2\sigma^2}{2a}e^{-a\tau}.$$

Note correlation function has the same form as the Ornstein-Uhlenbeck process in Example B.7 (with $Q = c^2\sigma^2$). ∇

As in the case of deterministic linear systems, we can analyze a stochastic linear system either in the state space or the frequency domain. The frequency domain approach provides a very rich set of tools for modeling and analysis of interconnected systems, relying on the frequency response and transfer functions to represent the flow of signals around the system.

Given a random process $X(t)$, we can look at the frequency content of the properties of the response. In particular, if we let $\rho(\tau)$ be the correlation function for a (scalar) random process, then we define the *power spectral density function* as the Fourier transform of ρ:

$$S(\omega) = \int_{-\infty}^{\infty} \rho(\tau) e^{-j\omega\tau}\, d\tau, \qquad \rho(\tau) \quad = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) e^{j\omega\tau}\, d\tau.$$

The power spectral density provides an indication of how quickly the values of a random process can change through the frequency content: if there is high frequency content in the power spectral density, the values of the random variable can change quickly in time.

Example 4.8 (Ornstein-Uhlenbeck process)**.** To illustrate the use of these measures, consider a first-order Markov process where the correlation function is

$$\rho(\tau) = \frac{Q}{2\omega_0} e^{-\omega_0(\tau)}.$$

This correspnds to Example 4.7 (also called an *Ornstein-Uhlenbeck process*). The power spectral density becomes

$$\begin{aligned} S(\omega) &= \int_{-\infty}^{\infty} \frac{Q}{2\omega_0} e^{-\omega|\tau|} e^{-j\omega\tau}\, d\tau \\ &= \int_{-\infty}^{0} \frac{Q}{2\omega_0} e^{(\omega - j\omega)\tau}\, d\tau + \int_{0}^{\infty} \frac{Q}{2\omega_0} e^{(-\omega - j\omega)\tau}\, d\tau = \frac{Q}{\omega^2 + \omega_0^2}. \end{aligned}$$

We see that the power spectral density is similar to a transfer function and we can plot $S(\omega)$ as a function of ω in a manner similar to a Bode plot, as shown in Figure 4.2. Note that although $S(\omega)$ has a form similar to a transfer function, it is a real-valued function and is not defined for complex s. ∇

Using the power spectral density, we can more formally define "white noise": a *white noise process* is a zero-mean, random process with power spectral density $S(\omega) = W$ = constant for all ω. If $X(t) \in \mathbb{R}^n$ (a random vector), then $W \in \mathbb{R}^{n \times n}$. We see that a random process is white if all frequencies are equally represented in its power spectral density; this spectral property is the reason for the terminology "white".

Given a linear system

$$\dot{X} = AX + FW, \qquad Y = CX,$$

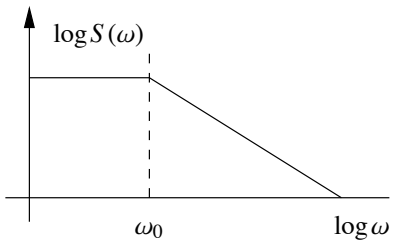

Figure 4.2: Spectral power density for a first-order Markov process.

.

with W given by white noise, we can compute the spectral density function corresponding to the output Y. We start by computing the Fourier transform of the steady state correlation function (4.21):

$$\begin{aligned} S_Y(\omega) &= \int_{-\infty}^{\infty} \left[\int_0^{\infty} h(\xi) Q h(\xi+\tau) d\xi \right] e^{-j\omega\tau} d\tau \\ &= \int_0^{\infty} h(\xi) Q \left[\int_{-\infty}^{\infty} h(\xi+\tau) e^{-j\omega\tau} d\tau \right] d\xi \\ &= \int_0^{\infty} h(\xi) Q \left[\int_0^{\infty} h(\lambda) e^{-j\omega(\lambda-\xi)} d\lambda \right] d\xi \\ &= \int_0^{\infty} h(\xi) e^{j\omega\xi} d\xi \cdot QH(j\omega) = H(-j\omega) Q H(j\omega). \end{aligned}$$

This is then the (steady state) response of a linear system to white noise.

As with transfer functions, one of the advantages of computations in the frequency domain is that the composition of two linear systems can be represented by multiplication. In the case of the power spectral density, if we pass white noise through a system with transfer function $H_1(s)$ followed by transfer function $H_2(s)$, the resulting power spectral density of the output is given by

$$S_Y(\omega) = H_1(-j\omega) H_2(-j\omega) Q_u H_2(j\omega) H_1(j\omega).$$

As stated earlier, white noise is an idealized signal that is not seen in practice. One of the ways to produced more realistic models of noise and disturbances is to apply a filter to white noise that matches a measured power spectral density function. Thus, we wish to find a covariance W and filter $H(s)$ such that we match the statistics $S(\omega)$ of a measured noise or disturbance signal. In other words, given $S(\omega)$, find $W > 0$ and $H(s)$ such that $S(\omega) = H(-j\omega) W H(j\omega)$. This problem is know as the *spectral factorization problem*.

Figure 4.3 summarizes the relationship between the time and frequency domains.

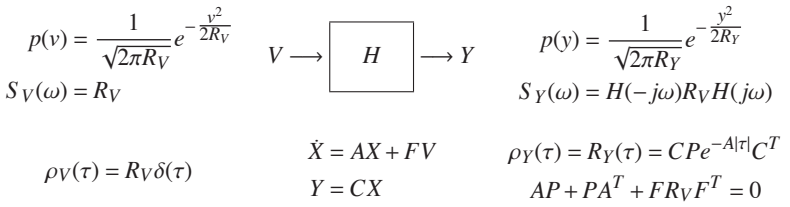

Figure 4.3: Summary of steady state stochastic response.

Exercises

4.1 (BE 150, Winter 2011) For this problem, we return to our standard model of transcription and transcription process with probabilistic creation and degradation of discrete mRNA and protein molecules. The *propensity functions* for each reaction are as follows:
Probability of transcribing 1 mRNA molecule: $0.2dt$
Probability of degrading 1 mRNA molecule: $0.5dt$ and is proportional to the number of mRNA molecules.
Probability of translating 1 protein: $5dt$ and is proportional to the number of mRNA molecules.
Probability of degrading 1 protein molecule: $0.5dt$ and is proportional to the number of protein molecules.
dt is the time step chosen for your simulation. Here we choose $dt = 0.05$.

(a) Simulate the stochastic system above until time $T = 100$. Plot the resulting number of mRNA and protein over time.

(b) Now assume that the proteins are degraded much more slowly than mRNA and the propensity function of protein degradation is now $0.05dt$. To maintain similar protein levels, the translation probability is now $0.5dt$ (and still proportional to the number of mRNA molecules). Simulate this system as above. What difference do you see in protein level? Comment on the effect of protein degradation rates on noise.

4.2 (BE 150, Winter 2011) Compare a simple model of negative autoregulation with one without autoregulation:

$$\frac{dX}{dt} = \beta_0 - \gamma X$$

and

$$\frac{dX}{dt} = \frac{\beta}{1 + \frac{X}{K}} - \gamma X$$

(a) Assume that the basal transcription rates β and β_0 vary between cells, following a Gaussian distribution with $\frac{\sigma^2}{<X>} = 0.1$. Simulate time courses of both models for

100 different "cells" using the following parameters: $\beta = 2, \beta_0 = 1, \gamma = 1, K = 1$. Plot the nonregulated and autoregulated systems in two separate plots. Comment on the variation you see in the time courses.

(b) Calculate the deterministic steady state for both models above. How does variation in the basal transcription rate β or β_0 enter into the steady state and relate it to what you see in part (a).

4.3 Consider gene expression: $\phi \xrightarrow{k} \mathrm{m}$, $\mathrm{m} \xrightarrow{\beta} \mathrm{m} + \mathrm{P}$, $\mathrm{m} \xrightarrow{\gamma} \phi$, and $\mathrm{P} \xrightarrow{\delta} \phi$. Answer the following questions:

(a) Use the stochastic simulation algorithm (SSA) to obtain realizations of the stochastic process of gene expression and numerically compare with the deterministic ODE solution. Explore how the realizations become close to or apart from the ODE solution when the volume is changed. Determine the stationary probability distribution for the protein (you can do this numerically, but note that this process is linear, so you can compute the probability distribution analytically in closed form).

(b) Now consider the additional binding reaction of protein P with downstream DNA binding sites D: $\mathrm{P} + \mathrm{D} \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} \mathrm{C}$. Note that the system no longer linear due to the presence of a bi-molecular reaction. Use the SSA algorithm to obtain sample realizations and numerically compute the probability distribution of the protein and compare it to what you obtained in part (a). Explore how this probability distribution and the one of C change as the rates k_{on} and k_{off} become larger and larger with respect to δ, k, β, γ. Do you think we can use a QSS approximation similar to what we have done for ODE models?

(c) Determine the Langevin equation for the system in part (b) and obtain sample realizations. Explore numerically how good this approximation is when the volume decreases/increases.

4.4 Consider the bi-molecular reaction $\mathrm{A} + \mathrm{B} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \mathrm{C}$, in which A and B are in total amounts A_T and B_T, respectively. Compare the steady state value of C obtained from the deterministic model to the mean value of C obtained from the stochastic model as the volume is changed in the stochastic model. What do you observe? You can perform this investigation through numerical simulation.

4.5 Consider the simple birth and death process: $\mathrm{Z} \underset{k_1 G}{\overset{k_2 G}{\rightleftharpoons}} \emptyset$, in which G is a "gain". Assume that the reactions are catalyzed by enzymes and that the gain G can be tuned by changing the amounts of these enzymes. A deterministic ODE model for this system incorporating noise and disturbances due to the stochasticity of the cellular environment is given by

$$\dot{Z} = k_1 G - k_2 G Z + d(t),$$

in which $d(t)$ incorporates noise, as seen in the previous homework. Determine the Langevin equation for this birth and death process and compare its form to the deterministic one. Also, determine the frequency response of Z to noise for both the deterministic model and for the Langevin model. Does increasing the gain G has the same effect in both models? Explain.

4.6 Consider a second order system with dynamics

$$\begin{pmatrix} \dot{X}_1 \\ \dot{X}_2 \end{pmatrix} = \begin{pmatrix} -a & 0 \\ 0 & -b \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} v, \qquad Y = \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$$

that is forced by Gaussian white noise with zero mean and variance σ^2. Assume $a, b > 0$.

(a) Compute the correlation function $\rho(\tau)$ for the output of the system. Your answer should be an explicit formula in terms of a, b and σ.

(b) Assuming that the input transients have died out, compute the mean and variance of the output.