Chapter 2: Stochastic Processes on Complex Networks

W. Garrett Jenkinson and John Goutsias

November 2, 2011

1 Statistical Modeling

We now discuss the issue of modeling dynamic processes on complex networks. We start with a general presentation in terms of necessary mathematical foundations on stochastic modeling. Later, we provide specific examples to solidify these concepts. Finally, we conclude with a discussion about exact simulation techniques.

In the following, each interacting component $X_n$, $n = 1, 2, \ldots, N$, in a complex networked system assumes a state $x_n(t) \in \mathbb{N}$ at each time point $t \geq 0$, where $\mathbb{N}$ is the set of all natural numbers $\{0, 1, 2, \ldots\}$. The $N \times 1$ state vector $\mathbf{x}(t)$ completely characterizes the system at time $t$. Here, $\mathbf{x}(t)$ evolves with time, but not deterministically. Hence, if we run multiple simulations of the same system from identical initial conditions, we cannot expect to obtain the same dynamics in each simulation. Therefore, $\{\mathbf{x}(t), t \geq 0\}$ is a multivariate stochastic process, since it is a random vector whose values change as a function of time [i.e., $\mathbf{x}(t)$ is a random vector for any $t$]. To distinguish between a random vector and its realizations, we use capital letters, such as $\mathbf{X}$, to denote random vectors and small letters, such as $\mathbf{x}$, to denote realizations. For determining the statistical properties of a stochastic process $\{\mathbf{X}(t), t \geq 0\}$, we need to specify the joint probabilities $p(x_1, t_1; x_2, t_2; \ldots; x_q, t_q) := \Pr[\mathbf{X}(t_1) = x_1, \mathbf{X}(t_2) = x_2, \ldots, \mathbf{X}(t_q) = x_q]$, for every $0 \leq t_1 < t_2 < \cdots < t_q$, $x_1, x_2, \ldots, x_q$, and $q$. However, for our applications we are mainly interested in calculating the probability $\pi(\mathbf{x}, t) =$
\( \Pr[X(t) = x] \), for every \( t \geq 0 \). Note that when we know the initial condition of our system to be \( X(0) = x_0 \), then \( \pi(x, 0) = \delta(x - x_0) \), where \( \delta(\cdot) \) is the indicator function \([\text{i.e., } \delta(x - x_0) = 1, \text{ if } x = x_0, \text{ and } 0, \text{ otherwise}]\). Note, however, that is more common to assume that we know the initial state and specify the conditional probability \( p(x, t) := \Pr[X(t) = x \mid X(0) = x_0] \), for \( t > 0 \), which of course depends on \( x_0 \).

### 1.1 Markov Processes

A Markov process is a particular case of a stochastic process that satisfies an important property: the dynamic evolution of a Markov process does not depend on all past values but only on the immediate past value. More precisely, this means that, for any set of \( q \) successive times \( t_1 < t_2 < \cdots < t_q \),

\[
p(x_q, t_q \mid x_1, t_1; \ldots; x_{q-1}, t_{q-1}) = p(x_q, t_q \mid x_{q-1}, t_{q-1}),
\]

where

\[
p(x_q, t_q \mid x_1, t_1; \ldots; x_{q-1}, t_{q-1}) := \Pr[X(t_q) = x_q \mid X(t_1) = x_1, \ldots, X(t_{q-1}) = x_{q-1}].
\]

The conditional probability \( p(x_q, t_q \mid x_{q-1}, t_{q-1}) := \Pr[X(t_q) = x_q \mid X(t_{q-1}) = x_{q-1}] \) is known as the transition probability of the Markov process \( \{X(t), t \geq 0\} \). Thus, knowledge of the state at time \( t_{q-1} \) contains the same information as every state that came before \( t_{q-1} \) when it comes to predicting the future state at \( t_q \). This property simplifies specification of the joint probabilities, since

\[
p(x_1, t_1; x_2, t_2; \ldots; x_q, t_q) = p(x_1, t_1) \prod_{i=1}^{q-1} p(x_{i+1}, t_{i+1} \mid x_i, t_i).
\]  \((1)\)

Therefore, only the initial and transition probabilities are required for specifying the previous joint probabilities.

#### 1.1.1 The Chapman-Kolmogorov equation

Let us take the special case of Equation 1 for \( q = 3 \). By summing over the intermediate state \( x_2 \), we have that

\[
p(x_1, t_1; x_3, t_3) = p(x_1, t_1) \sum_{x_2} p(x_2, t_2 \mid x_1, t_1)p(x_3, t_3 \mid x_2, t_2).
\]

---

1
Then, by dividing both sides with \( p(x_1, t_1) \), we obtain

\[
p(x_3, t_4 \mid x_1, t_1) = \sum_{x_2} p(x_3, t_3 \mid x_2, t_2)p(x_2, t_2 \mid x_1, t_1).
\]

In general,

\[
p(x_{q+1}, t_{q+1} \mid x_{q-1}, t_{q-1}) = \sum_{x_q} p(x_{q+1}, t_{q+1} \mid x_q, t_q)p(x_q, t_q \mid x_{q-1}, t_{q-1}),
\]

which is known as the Chapman-Kolmogorov equation. This equation provides constraints on the transition probabilities of a Markov process [1]. An alternative form of this identity, known as the Master Equation, is crucial when modeling stochastic dynamics on complex networks.

### 1.1.2 The Master Equation

To simplify analysis, we will assume that the Markov process under consideration is homogeneous. This means that the transition probabilities depend only on the difference between time points, in which case,

\[
p(x_q, t_q \mid x_{q-1}, t_{q-1}) = p(x_q, t_q + c \mid x_{q-1}, t_{q-1} + c),
\]

for any constant \( c \). As a consequence, we can denote \( p(x_q, t_q \mid x_{q-1}, t_{q-1}) \) by \( p_r(x_q \mid x_{q-1}) \), where \( \tau := t_q - t_{q-1} \).

Let us now consider the behavior of the transition probability for arbitrarily small time steps. Specifically, it can be shown that [1]

\[
p_{d\tau}(x_{q+1} \mid x_q) = [1 - a_0(x_q)d\tau]\delta(x_{q+1} - x_q) + T(x_{q+1} \mid x_q)d\tau + o(d\tau),
\]

where \( d\tau := t_{q+1} - t_q \), \( \lim_{d\tau \to 0} o(d\tau)/d\tau = 0 \), and \( T(x_{q+1} \mid x_q) \geq 0 \) is the transition probability per unit time from \( x_q \) to \( x_{q+1} \) [by convention, we set \( T(x \mid x) = 0 \), for all \( x \)]. Thus, for a sufficiently small \( d\tau \), the probability that we transition from state \( x_q \) to state \( x_{q+1} \) in the next \( d\tau \) time step is approximately given by \( T(x_{q+1} \mid x_q)d\tau \). Meanwhile, \( a_0(x_q)d\tau \) is the probability that a transition takes place during the next \( d\tau \) time step, in which case, \( 1 - a_0(x_q)d\tau \) is the probability that no transition occurs during that step. Therefore,

\[
a_0(x_q) = \sum_x T(x \mid x_q).
\]
If we now insert Equations 3 & 4 into the Chapman-Kolmogorov Equation 2, we can show that

\[
\frac{p_{q+d\tau}(\mathbf{x}_{q+1} \mid \mathbf{x}_{q-1}) - p_{q}(\mathbf{x}_{q+1} \mid \mathbf{x}_{q-1})}{d\tau} = \sum_{\mathbf{x}_q} T(\mathbf{x}_{q+1} \mid \mathbf{x}_q)p_{q}(\mathbf{x}_q \mid \mathbf{x}_{q-1}) - \sum_{\mathbf{x}_q} T(\mathbf{x}_q \mid \mathbf{x}_{q+1})p_{q}(\mathbf{x}_{q+1} \mid \mathbf{x}_{q-1}) + \frac{o(d\tau)}{d\tau}.
\]

By taking the limit as \(d\tau \to 0\), we finally obtain the following partial differential equation:

\[
\frac{\partial p_{q}(\mathbf{x}_{q+1} \mid \mathbf{x}_{q-1})}{\partial \tau} = \sum_{\mathbf{x}_q} \left[ T(\mathbf{x}_{q+1} \mid \mathbf{x}_q)p_{q}(\mathbf{x}_q \mid \mathbf{x}_{q-1}) - T(\mathbf{x}_q \mid \mathbf{x}_{q+1})p_{q}(\mathbf{x}_{q+1} \mid \mathbf{x}_{q-1}) \right].
\] (5)

This differential form of the Chapman-Kolmogorov equation is called the Master Equation (ME). The equation is usually written in the following simplified and more intuitive form [1]:

\[
\frac{\partial p(\mathbf{x}, t)}{\partial t} = \sum_{\mathbf{x}'} \left[ T(\mathbf{x} \mid \mathbf{x}')p(\mathbf{x}', t) - T(\mathbf{x}' \mid \mathbf{x})p(\mathbf{x}, t) \right],
\] (6)

which is obtained from Equation 5 by setting \(t_{q-1} = 0\) and \(t_q = t\).

The ME is a gain-loss equation for the probabilities of state \(\mathbf{x}\). The first term is the gain of \(\mathbf{x}\) due to transitions from all other states \(\mathbf{x}'\), whereas, the second term is the loss due to transitions of \(\mathbf{x}\) into other states. So, given an initial condition \(\mathbf{X}(0) = \mathbf{x}_0\), all information about the process [i.e., \(p(\mathbf{x}, t)\), for \(t > 0\)] can be found by solving Equation 6, assuming that we know the transition probabilities per unit time \(T(\mathbf{x}' \mid \mathbf{x})\), for every \(\mathbf{x}'\) and \(\mathbf{x}\).

### 1.2 Master Equations in Complex Networks

We have discussed ME’s for general Markov processes.\(^1\) However, we can exploit the structure of a given complex network in order to derive a ME governing that network. Specifically, in a complex network, state changes

\(^1\)The discussion has been for states that take discrete values. If needed, similar results can be derived for real valued states; see [1] for details.
are only due to firing of processing units and this reduces Equation 6 to a simpler form. First, we present the simplified form of the ME, and then we proceed to discuss some specific examples.

1.2.1 Master Equations for General Complex Networks

The state of a general complex network can only be updated based upon the firing of a processing unit. When a processing unit fires, the state instantaneously updates according to the underlying stoichiometry of the network. Specifically, if the \( m \)th processing unit is the only unit that fires within the infinitesimal time interval \( [t, t + dt) \), and if \( s_m \) is the \( m \)th column of the stoichiometry matrix \( S \), then we would have

\[
(x(t + dt) = x(t) + s_m. \]

If we can assign a firing rule \( \phi_m(\cdot) \), characteristic to the \( m \)th processing unit such that

\[
T(x + s_m \mid x)dt = \phi_m(x)dt \]

is the probability that the \( m \)th processing unit will fire during the time interval \( [t, t + dt) \) when \( x(t) = x \), then we can arrive at the following ME:

\[
\frac{\partial p(x, t)}{\partial t} = \sum_{m \in M} \left\{ \phi_m(x - s_m)p(x - s_m, t) - \phi_m(x)p(x, t) \right\}. \tag{7}
\]

This is due to the fact that \( T(x' \mid x) = 0 \) for any \( x' \neq x + s_m \) for some \( m \), since there is no processing unit whose firing can take \( x \) to \( x' \). We usually refer to the firing rule \( \phi_m(\cdot) \) as the propensity function of the \( m \)th processing unit.

We have shown in the previous discussion that Equation 7 is capable of describing the stochastic dynamic evolution of our Markov process \( \{X(t), t \geq 0\} \). Recall from Chapter 1 that the evolution of \( X(t) \) depends explicitly on another stochastic quantity, namely the degree of advancement \( Z(t) \). As a matter of fact

\[
X(t) = x_0 + S Z(t), \quad \text{for} \quad t \geq 0, \tag{8}
\]

where \( \{Z(t) \geq 0\} \) is an \( M \times 1 \) multivariate stochastic process whose components \( Z_m(t) \) tally the number of times that the \( m \)th processing unit has fired within the time interval \( [0, t) \) [we always have that \( Z(0) = 0 \)]. Note

\(^2\)Note that we model here a reversible processing unit as two irreversible processing units. Thus, the forward processing unit gets a column in the stoichiometry matrix \( S \) and the reverse processing unit also gets its own column in \( S \). The column associated with the reverse processing unit equals the negative of the column associated with the forward processing unit.
that the stochastic process \( \{Z(t), t \geq 0\} \) uniquely determines the stochastic process \( \{X(t), t \geq 0\} \). However, the opposite is not true in general, due to the fact that the matrix \( S^T S \) is not invertible; if it is, then \( Z(t) = (S^T S)^{-1} S^T [X(t) - x(0)] \), in which case, the stochastic process \( \{Z(t), t \geq 0\} \) will be uniquely determined from the stochastic process \( \{X(t), t \geq 0\} \). Therefore, we can consider \( \{Z(t), t \geq 0\} \) to be more informative than \( \{X(t), t \geq 0\} \).

In many complex networked systems, it may be difficult to observe \( \{Z(t), t \geq 0\} \) and more natural to observe \( \{X(t), t \geq 0\} \). For example, in a chemical reaction network, \( X(t) \) is the number of each chemical species present at time \( t \), which can be measured, while \( Z(t) \) is the number of times each reaction has occurred within \( [0, t) \), which is usually very difficult or impossible to measure. Thus, we can consider \( \{Z(t), t \geq 0\} \) to be the more informative hidden state variables of the system under consideration, and \( \{X(t), t \geq 0\} \) as the observed state variables. If we choose to model \( Z(t) \) directly and use Equation 8 to obtain \( X(t) \), we would be using what is known as a hidden Markov model (HMM) for our system. Note that we can build additional complexity into this HMM by assuming that our observations are not \( \{X(t), t \geq 0\} \) but a noise corrupted version \( \{Y(t), t \geq 0\} \). In this case, \( X(t) \) is also a hidden variable, and we would require a noise model describing how our observable \( Y(t) \) is generated from \( X(t) \). For now, we will stick to the noise-free HMM.

Fortunately, if we know how to generate Equation 7 [i.e., if we know the propensity functions \( \phi_m(x) \), for \( m \in M \)], then it is straightforward to construct a ME that governs the stochastic process \( \{Z(t), t \geq 0\} \). As we said before, \( \phi_m(x) dt \) is the probability that the \( m \)th processing unit will fire within the time interval \([t, t + dt]\). Therefore, \( \phi_m(x) dt \) is the probability that \( Z_m(t) \) will be increased by 1 within \([t, t + dt]\). Hence, if we define \( \alpha_m(z) := \phi_m(x_0 + Sz) \), then we have constructed all transition probabilities that tell us how the state \( Z(t) \) may change [note that \( Z(t) \) can only change by incrementing its values by one]. Using this information, it is straightforward to derive the following ME for the stochastic process \( \{Z(t), t \geq 0\} \):

\[
\frac{\partial p(z, t)}{\partial t} = \sum_{m \in M} \left\{ \alpha_m(z - e_m)p(z - e_m, t) - \alpha_m(z)p(z, t) \right\}, \quad (9)
\]

which of course is a Markov process, where \( p(z, t) := \Pr[Z(t) = z \mid Z(0) = 0] \), and \( e_m \) is the \( m \)th column of the \( M \times M \) identity matrix.
1.2.2 Chemical reaction networks

Recall that our chemical system is characterized by an $N \times 1$ state vector $\mathbf{x}$, which corresponds to the molecular numbers of the $N$ different chemical species that interact through $M$ chemical reactions. The following discussion closely follows the presentation in [2].

Let us assume we have a well-stirred chemical system with constant temperature $T$ and volume $V$. Using chemical collision theory, we can show that the probability of a randomly selected combination of reactant molecules of the $m$th reaction to react within an infinitesimal time interval $[t, t + dt)$ is given by $c_m dt$, where

$$
c_m := \frac{1}{V} \left[ \frac{8k_B T (\mu_1 + \mu_2)}{\pi \mu_1 \mu_2} \right]^{1/2} \pi (r_1 + r_2)^2 e^{-E_m/k_B T}.
$$

In this expression, $k_B = 1.3806504 \times 10^{-23}$ J/K is the Boltzmann constant, $\mu_1$ and $\mu_2$ are the masses of the two reactant molecules, $r_1$ and $r_2$ are their corresponding radii, and $E_m$ is the activation energy of the reaction. We usually refer to $c_m$ as the specific probability rate constant of the $m$th reaction.

The important fact to note here is that $c_m$ is a constant characteristic to the $m$th chemical reaction. Moreover, a similar constant can be derived for any elementary reaction (i.e., a reaction that occurs in a single step – usually only monomolecular and bimolecular reactions are elementary).

Let us now define a function $\gamma_m(\mathbf{x})$ that tells us how many possible reactive pairs exist for the $m$th chemical reaction. For example, if the $m$th reaction is bimolecular with $X_1$ and $X_2$ as reactants, and if $x_1(t)$ and $x_2(t)$ are the molecular numbers of these reactants at time $t$, then $\gamma_m(\mathbf{x}(t)) = x_1(t)x_2(t)$. Likewise, if the $m$th reaction is bimolecular with $X_1$ reacting with itself, then $\gamma_m(\mathbf{x}(t)) = x_1(t)[x_1(t) - 1]/2$. Finally, if the $m$th reaction is monomolecular with $X_1$ as a reactant, then $\gamma_m(\mathbf{x}(t)) = x_1(t)$. Using the previous definitions, it can be shown (see [2]) that the probability that only reaction $m$ occurs within the time interval $[t, t + dt)$ equals to $c_m \varphi_m(\mathbf{x}(t))dt + o(dt)$. Moreover, the probability that more than one reaction occurs within $[t, t + dt)$ is $o(dt)$. Thus, the propensity function of the $m$th reaction is given by

$$
\phi_m(\mathbf{x}) = c_m \gamma_m(\mathbf{x}), \text{ for all } m \in \mathcal{M},
$$
where

\[
\gamma_m(x(t)) = \begin{cases} 
  x_n(t), & \text{for a monomolecular reaction with reactant } X_n \\
  x_n(t)[x_n(t) - 1]/2, & \text{for a bimolecular reaction with identical reactants } X_n \\
  x_m(t)x_{n'}(t), & \text{for a bimolecular reaction with different reactants } X_n, X_{n'}
\end{cases}
\]  

(10)

Putting these propensities into Equation 7, we arrive at the ME:

\[
\frac{\partial p(x, t)}{\partial t} = \sum_{m \in \mathcal{M}} \left[ c_m \gamma_m(x - s_m)p(x - s_m, t) - c_m \gamma_m(x)p(x, t) \right],
\]

which is usually referred to as the Chemical Master Equation (CME).

Alternatively, we may define \(\alpha_m(z) := \phi(x_0 + S z)\) and use Equation 9 to arrive at:

\[
\frac{\partial p(z, t)}{\partial t} = \sum_{m \in \mathcal{M}} \left[ c_m \gamma_m(x_0 + S(z - e_m))p(z - e_m, t) - c_m \gamma_m(x_0 + Sz)p(z, t) \right],
\]

which is the CME associated with the degrees of advancement.

### 1.2.3 Epidemiological networks

Here, we will discuss the simplest epidemiological model, namely the SIR model (for a gentle introduction to deterministic epidemiological models, see [3], while a more detailed account of what follows may be found in [4]). Many more considerations and relaxations of assumptions are possible to create more accurate models. Later, we will present some of these models.

At any given time, individuals are in one of three classes, \(S, I,\) or \(R\). At time \(t\), there are \(x_1(t)\) susceptible individuals in class \(S\) who can catch a disease, but are not yet infected. Likewise, the infected class \(I\) has \(x_2(t)\) individuals who are sick and capable of transmitting the disease to susceptible individuals. Finally, the resistant class \(R\) has \(x_3(t)\) individuals who have died, or have recovered and are now immune to infection. Thus, there are \(N = 3\) interacting components (\(S, I,\) and \(R\)) and \(M = 2\) processing units. The first processing unit corresponds to the infection of a susceptible individual (i.e., \(S + I \rightarrow 2I\)), whereas, the second processing unit corresponds to an infected
individual becoming resistant to the infection (i.e., \( I \to R \)). This yields the following 3 \times 2 stoichiometry matrix:

\[ S = \begin{bmatrix} -1 & 0 \\ 1 & -1 \\ 0 & 1 \end{bmatrix}. \]  

\( \text{(11)} \)

In the following, we assume that the population of individuals is homogeneously mixed. For a susceptible person to become ill, two events must occur: first, she must come in contact with an infected individual, and second, given that she has been in contact with an infected individual, the infection must be successfully transmitted to her. Since we assume a constant mixing rate, the probability of a susceptible individual coming in contact with an infected individual is proportional to the number of infected individuals; i.e., we have that

\[ \Pr[\text{a given susceptible individual comes in contact with an infected individual within } [t, t + dt)] = \lambda_1 x_2(t) dt. \]

Now, we assume that the probability that the infection is transmitted given that the susceptible individual has been in contact with an infected individual is constant equal to \( \lambda_2 \). Thus, the probability that a randomly selected susceptible individual becomes infected within the time interval \([t, t + dt)\) is \( \lambda_1 \lambda_2 x_2(t) dt \). Since there are \( x_1(t) \) susceptible individuals at time \( t \), the probability that any susceptible individual becomes infected within \([t, t + dt)\) is given by \( \phi_1(x(t)) dt = \lambda_1 \lambda_2 x_1(t) x_2(t) dt \), where

\[ x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix}. \]

Likewise, we assume that individuals recover from the infected class at a constant rate \( \gamma \), where \( 1/\gamma \) is the average time an individual spends being infected. If \( \mu \) is the rate at which an infection kills an individual, then the removal rate from the infected class is \( \gamma + \mu \). Hence, the probability that some infected individual becomes resistant within \([t, t + dt)\) is given by
\[ \phi_2(x(t))dt = (\gamma + \mu)x_2(t)dt \]. Therefore, we arrive at the following ME:

\[
\frac{\partial p(x, t)}{\partial t} = \phi_1(x - s_1)p(x - s_1, t) - \phi_1(x)p(x, t) \\
+ \phi_2(x - s_2)p(x - s_2, t) - \phi_2(x)p(x, t) \\
= \lambda_1\lambda_2(x_1 + 1)(x_2 - 1)p(x - s_1, t) - \lambda_1\lambda_2 x_1 x_2 p(x, t) \\
+ (\gamma + \mu)(x_2 + 1)p(x - s_2, t) - (\gamma + \mu)x_2 p(x, t),
\]

where \(s_1\) is the first column of the stoichiometry matrix \(S\) given by Equation 11 and \(s_2\) is the second column of \(S\); i.e.,

\[
s_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \quad \text{and} \quad s_2 = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}.
\]

### 1.2.4 Pharmacokinetic networks

We consider here a previously developed model for studying the effect of tetrachloroethylene, a widely used solvent, on carcinogenesis [5]. This model assumes a division of the human body into the lungs, which serve as the central compartment, and four peripheral compartments, namely fat tissue, poorly perfused tissue (muscles and skin), richly perfused tissue (central nervous system and viscera, except liver), and liver. To model this system, we will denote by \(X_n\) the solvent present in the \(n\)th compartment. Then, we can represent the system by \(N = 5\) species interacting by the following \(M = 10\) reactions:

\[
\emptyset \rightarrow X_1 \\
X_1 \rightleftharpoons X_2 \\
X_1 \rightleftharpoons X_3 \\
X_1 \rightleftharpoons X_4 \\
X_1 \rightleftharpoons X_5 \\
X_5 \rightarrow \emptyset.
\]

The underlying reactions model the injection of solvent into lung blood (reaction 1), the exchange of one molecule of solvent between the lung blood and fat tissue (reactions 2 & 3), poorly perfused tissue (reactions 4 & 5), richly perfused tissue (reactions 6 & 7), and liver tissue (reactions 8 & 9), as well as the metabolic clearance of the solvent by the liver (reaction 10).
If we assume that all compartments are homogeneous, that the injection of solvent into the lung blood takes place at a constant rate $\kappa_1$, and that the probability of a randomly selected solvent molecule to move from compartment $n$ to compartment $n'$ within the infinitesimal time interval $[t, t + dt)$ is proportional to $dt$, with proportionality constant $\kappa_{nn'}$, then we can model the previous pharmacokinetic system as a Markovian reaction network with mass-action propensity functions

$$
\phi_1(x) = \kappa_1, \quad \phi_2(x) = \kappa_{12}x_1, \quad \phi_3(x) = \kappa_{21}x_2, \quad \phi_4(x) = \kappa_{13}x_1, \quad \phi_5(x) = \kappa_{31}x_3,
$$
$$
\phi_6(x) = \kappa_{14}x_1, \quad \phi_7(x) = \kappa_{41}x_4, \quad \phi_8(x) = \kappa_{15}x_1, \quad \phi_9(x) = \kappa_{51}x_5, \quad (13)
$$

where the $n^{th}$ element $x_n$ of vector $x$ denotes the population of tetrachloroethylene in the $n^{th}$ compartment. Moreover, if we assume that tetrachloroethylene metabolism in the liver is saturable according to the Michaelis–Menten relationship of enzyme kinetics [5], then the propensity function of the last reaction will be given by the following hyperbolic expression [6]

$$
\phi_{10}(x) = \frac{Vx_5}{K + x_5}, \quad (14)
$$

where $V, K$ are two parameters associated with the underlying metabolic mechanism.

1.2.5 Trophic networks

Note the similarity in the propensity functions for chemical and epidemiological networks. These are known as mass action systems, meaning that the functional forms of their propensity functions simply depend on products of “masses” of interacting components (reactants, infected individuals, etc.) entering the processing units [e.g., $\phi_m(x(t)) = k_m\gamma_m(x(t))$, where $k_m$ is a rate parameter and $\gamma_m(\cdot)$ is given by Equation 10]. This is known as the law of mass action. A trophic network is likewise a mass action system. As a consequence, its ME is similar to those of chemical and epidemiological systems [7].

Let us explore this for our simple example from Chapter 1. Recall, that we are dealing only with grass (G), rabbits (R) and wolves (W). The grass supply is approximately constant, since the rabbits fertilize the grass after they eat it. Moreover, rabbits eat grass and produce more rabbits and wolves...
eat rabbits to produce more wolves. Finally, wolves die at a constant rate. This yields the following three “reactions”

\[
\begin{align*}
G + R & \rightarrow G + (1 + a_1)R \\
W + R & \rightarrow (1 + a_2)W \\
W & \rightarrow \emptyset
\end{align*}
\]  

(15)

where \(0 < a_1, a_2 < 1\) are constants that quantify the percent conversion of biomass and \(\emptyset\) is the empty set.

Now, if we assign \(X_1 = G, X_2 = R,\) and \(X_3 = W,\) then we can define the propensity function of each “reaction” in Equation 15:

\[
\begin{align*}
\phi_1(x) &= k_1x_1x_2 \\
\phi_2(x) &= k_2x_2x_3 \\
\phi_3(x) &= k_3x_3,
\end{align*}
\]

where \(k_1, k_2,\) and \(k_3\) are constants characteristic to each trophic “reaction” (e.g., \(k_3\) quantifies the rate by which wolves perish). Our stoichiometry matrix can be found in the usual way from Equation 15, yielding

\[
S = \begin{bmatrix}
0 & 0 & 0 \\
a_1 & -1 & 0 \\
0 & a_2 & -1
\end{bmatrix}.
\]  

(16)

With the stoichiometry specified and the propensity functions defined, we can simply use Equation 7 to arrive at our final ME:

\[
\frac{\partial p(x, t)}{\partial t} = \phi_1(x - s_1)p(x - s_1, t) - \phi_1(x)p(x, t) + \phi_2(x - s_2)p(x - s_2, t) - \phi_2(x)p(x, t) + \phi_3(x - s_3)p(x - s_3, t) - \phi_3(x)p(x, t)
\]

\[
= k_1x_1(x_2 - a_1)p(x - s_1, t) - k_1x_1x_2p(x, t) + k_2(x_2 + 1)(x_3 - a_2)p(x - s_2, t) - k_2x_2x_3p(x, t) + k_3(x_3 + 1)p(x - s_3, t) - k_3x_3p(x, t),
\]

where \(s_1\) is the first column of the stoichiometry matrix \(S\) given by Equation 16, \(s_2\) is the second column, and \(s_3\) is the third column; i.e.,

\[
s_1 = \begin{bmatrix}
0 \\
a_1 \\
0
\end{bmatrix}, \quad s_2 = \begin{bmatrix}
0 \\
-1 \\
a_2
\end{bmatrix}, \quad \text{and} \quad s_3 = \begin{bmatrix}
0 \\
0 \\
-1
\end{bmatrix}.
\]
Many systems can be adequately described by the law of mass action. For systems with interacting elements diffusing and interacting randomly (such as molecules bouncing around and colliding in solution, sick and healthy individuals moving about a city and meeting randomly, or predators and prey randomly moving through an ecosystem), there is ample opportunity to justify the law of mass action from probabilistic first principles. In the next section, we consider a complex networked system where the components (neurons) do not randomly diffuse to undergo interactions with other neurons. This will be our first example of a system which is not aptly described by the law of mass action.

### 1.2.6 Neural networks

We present here the model and the ME associated with neural networks (see [8]). The model consists of $L$ neurons, with each neuron being in either a quiescent or an active state. Let $X_{2l-1}$ and $X_{2l}$ denote a quiescent or active neuron $l$, respectively. We can assign the following two reactions to the $l^{th}$ neuron in the network:

$$
X_{2l-1} + \sum_{l' \neq l} \nu_{ll'} X_{2l'} \rightarrow X_{2l} + \sum_{l' \neq l} \nu_{ll'} X_{2l'}
$$

$$
X_{2l} \rightarrow X_{2l-1},
$$

(17)

where $\nu_{ij}$ measures the synaptic weight between neurons $i$ and $j$. Note that the first reaction models transition of the $l^{th}$ neuron from the quiescent to the active state, whereas, the second reaction models transition of the neuron from the active to the quiescent state. As a consequence, we obtain a reaction network with $N = 2L$ species and $M = 2L$ reactions.

We can describe this system by a $2L \times 1$ state vector $\mathbf{x}$ with binary-valued 0/1 elements $x_{2l-1}$, $x_{2l}$ indicating the state of the $l^{th}$ neuron. Due to the fact that a neuron must be either quiescent or active, the state variables must satisfy the mass conservation relationships $x_{2l-1} + x_{2l} = 1$, for $l = 1, 2, \ldots, L$. It has been suggested [8] that the probability of the $l^{th}$ neuron becoming active during an infinitesimally small time interval $[t, t+dt)$, given that the neuron is quiescent at time $t$, can be taken to be $x_{2l-1}[s_l(\mathbf{x}) > 0] \tanh(s_l(\mathbf{x})) dt + o(dt)$, where $[a > 0]$ is the Iverson bracket and $s_l$ is the total synaptic input to the $l^{th}$ neuron, given by

$$
s_l(\mathbf{x}) = \sum_{l' \neq l} \nu_{ll'} x_{2l'} + h_l
$$

(18)
with $h_l$ being an external input to the neuron. The term $x_{2l-1}$ ensures that the neuron becomes active within $[t, t + dt)$ only when it is quiescent at time $t$. As a consequence, the propensity of the first reaction in Equation (17) will be given by
\[
\phi_{2l-1}(x) = x_{2l-1}[s_l(x) > 0] \tanh(s_l(x)),
\]
and therefore depends on the synaptic inputs from neurons connected to the $l^{th}$ neuron and any external input to that neuron. On the other hand, if we assume that the $l^{th}$ neuron decays from an active to a quiescent state at a constant rate $d_l$, then the propensity of the second reaction will be given by
\[
\phi_{2l}(x) = d_l x_{2l},
\]
where the term $x_{2l}$ ensures that the neuron becomes inactive within $[t, t + dt)$ only when it is active at time $t$.

1.2.7 Social networks

Several models for dealing with dynamic processes on social networks are currently available, with many of them fitting nicely into the Markovian reaction framework discussed in this book. As an example, we focus on opinion spreading in social networks, a process of political, marketing, and general sociological interest, by utilizing concepts from the recently proposed heterogeneous voter model [9].

Let us consider a population of $L$ individuals whose state (opinion) takes binary 0/1 values. Possible opinions abound: democrat or republican, for or against an ideal, like or dislike of a product, etc. Let $X_{2l-1}$ and $X_{2l}$ denote the $l^{th}$ individual with opinion 0 and 1, respectively. We can model interactions between the $N = 2L$ individuals by employing the following $M = 2L$ reactions:
\[
\begin{align*}
X_{2l-1} + \sum_{l' \neq l} \nu_{l'l} X_{2l'} & \to X_{2l} + \sum_{l' \neq l} \nu_{l'l} X_{2l'} \\
X_{2l} + \sum_{l' \neq l} \nu_{l'l} X_{2l'-1} & \to X_{2l-1} + \sum_{l' \neq l} \nu_{l'l} X_{2l'-1}, \quad \text{for } l = 1, 2, \ldots, L,
\end{align*}
\]
(21)
with each pair corresponding to the $l^{th}$ individual, where $\nu_{ij}$ is a number between 0 and 1 that measures the weight of social influence of the $i^{th}$ individual on the $j^{th}$ individual. Clearly, the $2l - 1$ reaction models change
in the \( l \)th individual’s opinion from 0 to 1, whereas, the 2\( l \) reaction models opinion change of the same individual from 1 to 0.

We will now assume that, at time \( t \), the \( l \)th individual can be instantaneously contacted by all people in his social network, as determined by the nonzero coefficients \( \nu_{l'} \). Moreover, we will assume that a randomly selected \( l' \) individual may influence the opinion of the \( l \)th individual with probability \( \nu_{l,l'} \). Finally, we will assume that, if the opinion of the \( l \)th individual has been influenced at time \( t \) by the people in his social network, then he will change his opinion within the subsequent infinitesimal time interval \( [t, t + dt] \), with a probability that is proportional to \( dt \), with proportionality factor \( \kappa_{2l-1} \), if the opinion changes from 0 to 1, or \( \kappa_{2l} \), if the opinion changes from 1 to 0. In this case, the reaction system described by Equation (21) is Markovian with propensity functions given by

\[
\phi_{2l-1}(x) = \kappa_{2l-1}x_{2l-1}\sum_{l' \neq l} \nu_{l'l}x_{2l'} \\
\phi_{2l}(x) = \kappa_{2l}x_{2l}\sum_{l' \neq l} \nu_{l'l}x_{2l'-1} , \quad \text{for } l = 1, 2, \ldots, L, \tag{22}
\]

where \( x_{2l-1} \) and \( x_{2l} \) are binary 0/1 variables indicating the opinion of the \( l \)th individual. Due to the fact that the opinion of the \( l \)th individual must be either 0 or 1, the state variables must satisfy the mass conservation relationships \( x_{2l-1} + x_{2l} = 1 \), for \( l = 1, 2, \ldots, L \). The term \( x_{2l-1} \) associated with \( \phi_{2l-1}(x) \) ensures that the \( l \)th individual will not change opinion from 0 to 1 if his opinion is already 1 (in which case, \( x_{2l} = 1 \), which implies that \( x_{2l-1} = 0 \) and, therefore, the \( 2l-1 \) reaction will not occur). On the other hand, the terms \( x_{2l'} \) inside the sum ensure that only individuals with opinion 1 can cause the \( l \)th individual to change opinion from 0 to 1. Similar remarks apply for the propensity function \( \phi_{2l}(x) \). Note finally that dependence of the specific probability rate constants on \( l \) accounts for the heterogeneity of individual stubbornness and bias. Moreover, if \( \kappa_{2l-1} \neq \kappa_{2l} \), then the \( l \)th individual will have an inherent bias towards one opinion over the other.

### 1.2.8 Multi-agent networks

The study of multi-agent networks focuses on systems in which many intelligent agents, such as autonomous vehicles that observe and act upon their environment, interact with each other to achieve a certain goal. To illustrate
the fact that multi-agent systems can also be modeled as Markovian processes on reaction networks, we consider here a system comprised of $L$ autonomous unmanned vehicles (AUV’s) that can move over a two-dimensional bounded rectangular space in a discrete fashion [10]. For simplicity, we assume that, at each step, an AUV located at a discrete point $(i,j)$ in space can move towards one of four possible directions, namely east to point $(i+1,j)$, west to point $(i-1,j)$, north to point $(i,j+1)$, or south to point $(i,j-1)$. We want to develop a mathematical approach that can be used to describe vehicular motion so that the AUV’s reach a spatial configuration $\mathbf{x}$ at steady-state with desired probability $q(\mathbf{x})$. In particular, $q(\mathbf{x})$ could assign high probability over configurations that maximize a given design objective and low or zero probability over the remaining configurations.

In the following, we employ two species $X_{2l-1}$ and $X_{2l}$ whose populations $x_{2l-1}$ and $x_{2l}$ denote the position of the $l$th AUV on the two-dimensional rectangular grid. For example, if the $l$th vehicle is located at point $(i,j)$ on the grid, then $x_{2l-1} = i$ and $x_{2l} = j$. We can now characterize the motion of all AUV’s in the multi-agent network under consideration by $N = 2L$ species interacting through the following $M = 4L$ reactions:

$$\begin{align*}
X_{2l-1} + X_{2l} + \sum_{l' \neq l} (X_{2l'-1} + X_{2l'}) &\rightsquigarrow 2X_{2l-1} + X_{2l} + \sum_{l' \neq l} (X_{2l'-1} + X_{2l'}) \\
X_{2l-1} + X_{2l} + \sum_{l' \neq l} (X_{2l'-1} + X_{2l'}) &\rightsquigarrow X_{2l-1} + 2X_{2l} + \sum_{l' \neq l} (X_{2l'-1} + X_{2l'}). 
\end{align*}$$

(23)

The first two reactions model one-step motion of the $l$th AUV towards east/west, whereas, the other two reactions model one-step motion towards north/south.

Let us now define the potential $U(\mathbf{x})$ of the reaction system being in configuration $\mathbf{x}$ at steady-state by

$$U(\mathbf{x}) := \begin{cases} 
- \ln \frac{q(\mathbf{x})}{q(\mathbf{x}_0)}, & \text{for } \mathbf{x} \in \mathcal{D} \\
\infty, & \text{otherwise},
\end{cases}$$

(24)

where $\mathcal{D}$ is a set that contains all permissible vehicle configurations (e.g., $\mathbf{x}$ should not allow two vehicles to occupy the same grid position or positions occupied by obstacles, thus avoiding collisions, or assign vehicles to grid positions outside the bounded rectangular region). Moreover, $\mathbf{x}_0 \in \mathcal{D}$ is an appropriately chosen reference configuration of zero potential. Given that
$X(t) = x$, we will assume that, during the infinitesimally small time interval $[t, t + dt)$, the $l$th AUV can move one step towards east if two events take place: (a) during $[t, t + dt)$, the AUV will initiate motion with probability that is proportional to $dt$, with proportionality factor $\kappa_l$, and (b) given that the AUV initiates motion during $[t, t + dt)$, it moves with probability $\exp\{-U(x + s_{4l-3})\}$, where $s_m$ denotes the $m$th column of the net stoichiometry matrix of the reaction network given by Equation (23). As a consequence, the AUV will be moving east with higher probability if the motion produces a larger reduction in system energy. Note that parameter $\kappa_l$ controls the speed of the $l$th vehicle, with higher values of $\kappa_l$ resulting in faster motion.

By making similar assumptions for vehicle motion towards the other three directions, the dynamics of the reaction network given by Equation (23) will be Markovian with propensity functions

$$\phi_m(x) = \kappa_l e^{-U(x+s_m)}, \quad \text{for } m = 4l - 3, 4l - 2, 4l - 1, 4l, \ l = 1, 2, \ldots, L.$$  

Note that $s_{4l-3}, s_{4l-2}, s_{4l-1}, \text{ and } s_{4l}$ equal $e_{2l-1}, -e_{2l-1}, e_{2l}, \text{ and } -e_{2l},$ respectively, where $e_m$ is the $m$th column of the $2L \times 2L$ identity matrix. It turns out that the resulting master equation governing the population process $x$ has a unique stationary distribution $p(x; \infty)$, given by the Gibbs distribution

$$p(x; \infty) = \frac{1}{W} e^{-U(x)},$$

where

$$W = \sum_x e^{-U(x)}$$

is the associated partition function. As a consequence of Eqs. (24), (26), and (27), we have that $p(x; \infty) = q(x)$. Therefore, the AUV’s will asymptotically position themselves in the two-dimensional space at locations $x$ with probability $q(x)$, as expected.

2 Exact Simulation of the Master Equation

In general, it is not possible to analytically solve the ME given by Equation 7, unless we are dealing with very simple cases. Often, we do not require knowledge of the entire probability distribution $p(x, t)$. Instead, we are interested in the dynamics of various statistics of the Markov process
\{X(t), t \geq 0\}$, such as the mean $\mu(t) = \mathbb{E}[X(t)]$ and standard deviation $\sigma(t) = \sqrt{\text{Var}[X(t)]}$. Thus, if we could generate a large ensemble of sample trajectories $\{x_i(t), i = 1, 2, \ldots, I\}$ of the Markov process $\{X(t), t \geq 0\}$, we could approximate our desired quantities using the sample statistics of this ensemble; for example, we can set $\mu(t) \simeq \sum_{i=1}^{I} x_i(t)/I$. Such an approach is commonly referred to as Monte Carlo (MC) estimation algorithm. Note that we can even estimate the probability $p(x, t)$ from this ensemble using a statistic known as empirical distribution, which asymptotically converges to $p(x, t)$ according to the Glivenko-Cantelli theorem. We can do this by counting, at time $t$, the number $I(x, t)$ of samples in $\{x_i(t), i = 1, 2, \ldots, I\}$ that have value $x$ and by setting $p(x, t) \simeq I(x, t)/I$. Unfortunately, to sufficiently estimate $p(x, t)$, we need a very large number of samples $I$. Thus, even accurate computation of the probability $p(x, t)$ is most often impractical. Therefore, we usually resort to calculating only approximate values for the mean and standard deviation dynamics $\{\mu(t), t \geq 0\}$ and $\{\sigma(t), t \geq 0\}$. In the following, we discuss how to produce sample trajectories of Markov processes $\{X(t), t \geq 0\}$ and $\{Z(t), t \geq 0\}$, governed by the ME’s given by Equation 7 and Equation 9.

The algorithm under consideration is known as the Gillespie algorithm (see [11]). This algorithm is based on asking the following two questions:

(a) Given that we are at time $t$, when will a processing unit fire?

(b) Which is the next processing unit that fires?

To answer these questions, we must specify the probability density $\pi_t(\tau, m)$ that, given that we are at time $t$, the next processing unit fires at time $t + \tau$ and that this unit is the $m$th processing unit. Of course, $\pi_t(\tau, m) = 0$, for $\tau < 0$. It can be shown (see [11]) that, for $\tau \geq 0$,

\[
\pi_t(\tau, m) = \alpha_m(z(t)) \exp \left( -\tau \sum_{j \in M} \alpha_j(z(t)) \right)
\]

\[
= \frac{\alpha_m(z(t))}{\sum_{j \in M} \alpha_j(z(t))} \left[ \sum_{j \in M} \alpha_j(z(t)) \right] \exp \left( -\tau \sum_{j \in M} \alpha_j(z(t)) \right)
\]

\[
= \pi_t(m) \pi_t(\tau),
\]
where
\[ \pi_t(m) := \frac{\alpha_m(z(t))}{\sum_{j \in M} \alpha_j(z(t))}, \] (28)
and
\[ \pi_t(\tau) := \begin{cases} 
    \left[ \sum_{j \in M} \alpha_j(z(t)) \right] \exp \left( -\tau \sum_{j \in M} \alpha_j(z(t)) \right), & \text{for } \tau \geq 0 \\
    0, & \text{for } \tau < 0.
\] (29)

As a consequence, the “time to next firing” and the “index of the next firing processing unit” are statistically independent random variables with probability density functions given by \( \pi_t(\tau) \) and \( \pi_t(m) \), respectively. As a matter of fact, \( \pi_t(m) \) is a simple discrete distribution, whereas, \( \pi_t(\tau) \) is an exponential distribution with rate parameter \( \lambda(t) = \sum_{j \in M} \alpha_j(z(t)) \).

Using these results, we can construct the following simulation algorithm (see also [11]):

1. Initialize by setting \( z(0) = 0 \) and \( t = 0 \).
2. Calculate the propensity functions \( \alpha_j(z(t)) \), for \( j \in M \).
3. Choose the time \( \tau \) to next firing by drawing a sample from the exponential distribution \( \pi_t(\tau) \), given by Equation 29.
4. Choose the index \( m \) of the next firing processing unit by drawing a sample from the discrete distribution \( \pi_t(m) \), given by Equation 28.
5. Update the degrees of advancement at time \( t + \tau \) by setting \( z(t + \tau) = z(t) + e_m \), where \( e_m \) is the \( m \)th column of the \( M \times M \) identity matrix, and set \( z(t') = z(t) \), for every \( t \leq t' < t + \tau \).
6. If \( t + \tau < t_{\text{max}} \), advance the time ahead by setting \( t + \tau \to t \) and return to Step 2. Otherwise, set \( x(t) = x_0 + Sz(t) \), for \( 0 \leq t \leq t_{\text{max}} \) and STOP.

Although this is an easy to understand and implement method for generating trajectories \( \{z(t), t \geq 0\} \) and \( \{x(t), t \geq 0\} \) from the corresponding ME’s, this is not the most efficient method to do so. Many alternative methods have been proposed in the literature to speed-up this algorithm (e.g., see [12]).
References


