

Nonlinear Stochastic Dynamics in Mesoscopic Biochemical Systems

Hong Qian

Department of Applied Mathematics
University of Washington, Seattle, WA 98195, U.S.A.

April 21, 2011

Contents

1	Brief Introduction	2
2	Nonlinear Stochastic Dynamics of Mesoscopic Homogeneous Biochemical Reaction Systems — An Analytical Theory	4
3	Cellular Biology in Terms of Stochastic Nonlinear Biochemical Dynamics: Emergent Properties, Isogenetic Variations and Chemical System Inheritability	35
4	The Chemical Master Equation Approach to Nonequilibrium Steady-State of Open Biochemical Systems: Linear Single-Molecule Enzyme Kinetics and Nonlinear Biochemical Reaction Networks	59
5	Computational Cellular Dynamics Based on the Chemical Master Equation: A Challenge for Understanding Complexity	88

1 Brief Introduction

This compilation contains four review articles written for four different groups of audiences. Chapter 2, “Nonlinear stochastic dynamics of mesoscopic homogeneous biochemical reaction systems — An analytical theory” was written primarily for applied mathematicians as an invited review in the IOP journal *Nonlinearity*. Chapter 3, “Cellular biology in terms of stochastic nonlinear biochemical dynamics: Emergent properties, isogenetic variations and chemical system inheritability”, was written for physicists and physical chemists. It is published in the *Journal of Statistical Physics* as a part of a special issue on physics and biology. Chapter 4, “The chemical master equation approach to nonequilibrium steady-state of open biochemical systems: Linear single-molecule enzyme kinetics and nonlinear biochemical reaction networks” coauthored with my student Lisa Bishop, was written mainly for computational and theoretical biologists. Chapter 5, “Computational cellular dynamics based on the chemical master equation: A challenge for understanding complexity” coauthored with Professor Jie Liang of University of Illinois, Chicago, tries to introduce computational and computer scientists the new development in computational biology. The latter two articles were parts of two special issues on *Quantitative Modelling in Molecular System Bioenergetics* and on *Computational Challenges From Modern Molecular Biology*, respectively.

There is no particular logical order for these four chapters. Each was written as a self-contained piece. Putting together, however, they do form a more coherent presentation of an emerging theme.

Chapter 3 provides a comprehensive mathematical theory for treating cellular biochemistry as nonlinear chemical reaction systems in a small volume, i.e., *mesoscopic*: The system is very large since it contains millions and billions of molecules, mainly water, but it is still small enough to observe significant fluctuations and stochasticity. This part tries to tie the theory of cellular dynamics with established statistical thermodynamics, nonlinear physics, and theory of phase transitions on the one hand, and the concepts and ideas from Darwin’s evolution theory on the other.

Chapter 4 gives detailed description how to develop mathematical models for specific systems based on the theory. It contains many simple examples; so it is a good place to learn the mechanics of how the modeling works. It also tries to present both nonlinear biochemical reaction systems, and the single-molecule biophysics, a hot subject in the past decade, as an integrated subject, as they should be.

All the examples in Chapter 4 are toy models. To deal with any real-

istic biochemical system, even a very simplified one (lambda phage is the canonical example), one needs to carry out mathematical analysis using computation. Interestingly the needed computational techniques have not been developed. So Chapter 5 is an introduction to computational and computer scientists this exciting opportunity. To theoretical minded computer scientists, Chapter 5 suggests that the mathematical theory of mesoscopic nonlinear chemical reaction systems, whether it is a good theory for cellular biology or not, is a great example for studying *complexity*.

This leads us back to mathematics. In Chapter 2, we argued that nonlinear, stochastic population dynamics in terms of the birth-and-death processes is a more realistic and sophisticated approach to model biological dynamics than the widely used stochastic differential equations, also known as diffusion processes by probabilists. In fact, the dynamic theory for the populations of chemical species inside a cell can be applied equally well to other biological populations: cell and virus populations in immunology, human populations in demography, and biological species in ecology, for examples. Treating this theory as a subject of applied mathematics, Chapter 2 presents the Ornstein-Uhlenbeck Gaussian processes as the *linear analysis* for the general nonlinear dynamics near a fixed point. Because of the nonlinear nature of the dynamics, this part also emphasize something one tends to forget: the Law of Large Number does imply that macroscopic system is deterministic; however, stochasticity does not really disappear, rather it manifest itself as discrete stochastic jumps on a different time scale! Transitions among discrete stochastic attractors are *rare events*, they spend most of the time in “waiting”, exhibit punctuated equilibria. It argues that these *rare events* are rather unruly from either nonlinear deterministic, or data-driven statistical perspective, the two main devices of current applied mathematics and engineering.

Stochastic dynamic modeling is indeed a different mathematical approach to scientific data: It differs from deterministic dynamics by explicitly considering uncertainties. However, it incorporates randomness at an individual, mechanistic level, and relies on mathematical deductions to provide predictions on a system’s behavior. This is very different from statistical modeling which works directly with observed data. Stochastic dynamic modeling is an analytical form of *agent based modeling* which is now quite popular in computational sociology. It provides a deeper understanding of complex, nonlinear dynamical systems.

INVITED ARTICLE

Nonlinear stochastic dynamics of mesoscopic homogeneous biochemical reaction systems—an analytical theory

Hong Qian

Department of Applied Mathematics, University of Washington, Seattle, WA 98195, USA

E-mail: hqian@u.washington.edu

Received 16 November 2010, in final form 22 March 2011

Published 21 April 2011

Online at stacks.iop.org/Non/24/R19

Recommended by J A Glazier

Abstract

The nonlinear dynamics of biochemical reactions in a small-sized system on the order of a cell are stochastic. Assuming spatial homogeneity, the populations of n molecular species follow a multi-dimensional birth-and-death process on \mathbb{Z}^n . We introduce the Delbrück–Gillespie process, a continuous-time Markov jump process, whose Kolmogorov forward equation has been known as the chemical master equation, and whose stochastic trajectories can be computed via the Gillespie algorithm. Using simple models, we illustrate that a system of nonlinear ordinary differential equations on \mathbb{R}^n emerges in the infinite system size limit. For finite system size, transitions among multiple attractors of the nonlinear dynamical system are rare events with exponentially long transit times. There is a separation of time scales between the deterministic ODEs and the stochastic Markov jumps between attractors. No diffusion process can provide a global representation that is accurate on both short and long time scales for the nonlinear, stochastic population dynamics. On the short time scale and near deterministic stable fixed points, Ornstein–Uhlenbeck Gaussian processes give linear stochastic dynamics that exhibit time-irreversible circular motion for open, driven chemical systems. Extending this individual stochastic behaviour-based nonlinear population theory of molecular species to other biological systems is discussed.

Mathematics Subject Classification: 82C31, 37N25, 92C40

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Recent studies of biochemical reaction systems in a mesoscopic volume such as a cell have firmly established the chemical master equation (CME) as the basis of an analytical theory for cellular dynamics [1–7]. A system's volume V is a natural parameter in the theory: an ideal elementary chemical reaction $A + B \rightarrow C$ has a rate constant k with the dimension of $[\text{Time}]^{-1} \times [\text{Concentration}]^{-1}$. In the CME, the probability of this reaction occurring in an infinitesimal Δt is the dimensionless $k(n_A n_B / V)(\Delta t)$, when there are n_A and n_B molecules of types A and B, respectively. With increasing system size, V , the stochastic dynamics predicted by a CME has been mathematically shown to approach the deterministic solution of the kinetic differential equations based on the law of mass action for homogeneous chemical reactions [8]. The CME, therefore, 'is not an alternative to the deterministic kinetics, it is a more complete kinetic description which is capable of modelling reactions with and without fluctuations', for systems with small and large V [9].

In this review, I would like to take this new perspective a step further. A great many nonlinear systems of ordinary differential equations (ODEs) one studies describe dynamics of populations of one type or another. Examples include molecular species in biochemical reactions, cell and virus populations in immunology, human populations in demography and biological species in ecology. At the mechanistic level, all these dynamics are concerned with *birth* and *death* of individuals whose basic unit is 1. Therefore, every such dynamic model based on a nonlinear deterministic ODE system has a corresponding stochastic counterpart based on a birth-and-death process (BDP). If a nonlinear ODE system is defined on \mathbb{R}^n , the corresponding BDP is defined on \mathbb{Z}^n . First-order ODEs correspond to Markov jump processes with continuous time [10, 11].

Ever since the work of Einstein, Smoluchowski, Langevin, and Kramers, stochastic differential equations (SDEs), also known as diffusion processes by probabilists [12], have always been considered as the stochastic counterpart of ODEs [13, 14]. However, as anyone who has developed an SDE model for an applied problem knows, the choice for the coefficient $\Gamma(x)$ in an SDE $dx(t) = b(x)dt + \Gamma(x)dW(t)$ is almost always rather arbitrary. (The only exception is the guiding principle for fluctuating equilibrium dynamics based on the *fluctuation–dissipation theory* which we shall discuss later.) A BDP, however as we shall see, provides a rather complete stochastic description for the dynamics from mechanisms based on statistics of an individual's behaviour. There is no artificial separation of the deterministic $b(x)$ and stochastic $\Gamma(x)W(t)$ as in a SDE. (See section 7.3. This is called 'intrinsic noise' in cellular biochemistry.) Even more important, as we shall discuss in section 4.4, is the 'diffusion theory's dilemma' that invalidates the diffusion-model approach to nonlinear stochastic *population* dynamics.

The BDP theory provides further insights into the theory of nonlinear population dynamics extensively studied since the 1970s. There is a fundamental concept that does not exist in the theory of deterministic nonlinear dynamics, the concept of 'rare events': something that occurs with a very small probability, but on an evolutionarily long time scale, it will occur with probability one! We have recently argued that [15] this emergent stochastic transition among different attractors, on the time scale beyond the infinity of the deterministic dynamics, is one of the origins of 'complexity' [5]. It is these dynamics that exhibit 'dynamic symmetry breaking' [16] and 'singular points' at which the dynamics are truly unpredictable [17], giving rise to complex dynamics with high information content [16].

It is safe to say that statistical inference is currently one of the key approaches to complex systems and their dynamics. Bioinformatics and statistical genomics are dominant applied mathematics in cellular molecular biology. The above nonlinear stochastic dynamic

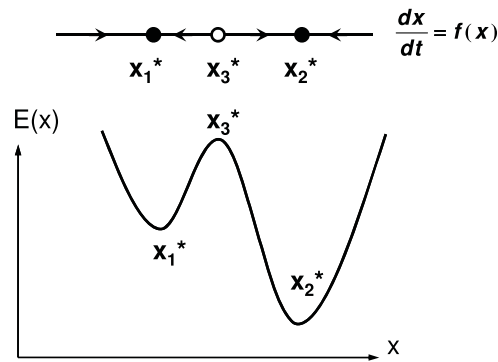


Figure 1. A pictorial introduction of the Kramers' barrier crossing problem in nonlinear, stochastic dynamics. Deterministic dynamics always go 'downhill' toward lower values of $E(x)$. Therefore, any dynamics with initial positive $x(0) < x_3^*$ will end at x_1^* , and with $x(0) > x_3^*$ will end at x_2^* . After reaching a stable fixed point, x_1^* or x_2^* , there will be no possibility of leaving. However, with stochastic elements in the dynamics, there are possibilities to go 'uphill.' With exceptional luck, continuous uphill movement leads to a transition between the two domains of attraction. Exceptional luck means the barrier crossings occur only on an extremely long time scale. x_3^* has been called the 'singular point' by James Clerk Maxwell since 'influences whose physical magnitude is too small to be taken account of by a finite being, may produce results of the greatest importance' [17].

perspective, however, clearly suggests that statistical approaches, while they can be powerful in representing data with statistical significance, cannot be useful in understanding the rare events. In fact, the very existence of multistabilities, i.e. alternative attractors, cannot be inferred from 'normal' statistical data. Mechanistic *deterministic* models can predict the existence of alternative attractors. Mechanistic *stochastic* models can further estimate the *lifetime* of an attractor. The actual time of the rare transition, however, is a random variable with exponential distribution, which is memoryless in defiance of causality.

Proving global asymptotic stability of a dynamical system, of course, has always been the ultimate goal of engineering. However, with increasing complexities, this becomes a less and less feasible task even in traditional engineering. On the other hand, one of the best understood 'rare events' is discrete chemical reactions in terms of Kramers' theory [18]. Recently in [19] we have shown that the nonlinear bifurcation theory of Thom–Zeeman's catastrophe, the phase transition theory from statistical mechanics, and Kramers' theory of barrier crossing (also known as decay of metastable states [20, 21]) are three different aspects (e.g. deterministic, steady state and kinetic) of a rare event. All these classical theories are called for in BDP dynamics.

Figure 1 shows the canonical pictorial introduction of the problem of 'barrier crossing' as a rare event. From a deterministic nonlinear dynamics standpoint, this system has three fixed points, two stable (x_1^* and x_2^* represented by the filled circles) and one unstable (x_3^* represented by the open circle). Barrier crossing requires movement against the deterministic force (shown by the arrows) which are low probability events. However, it is the cumulation of these unlikely events that leads to 'spectacular' or 'disastrous' phenomena in complex, stochastic nonlinear dynamical systems.

Our discussion of the nonlinear stochastic dynamics of biochemical reactions is organized in this paper as follows.

In section 2, using a simple example from mesoscopic chemical reaction systems, we introduce the ‘bottom-up’ approach to stochastic population dynamics based on mechanisms at the individual’s level. The example illustrates how nonlinear, bistable behaviour emerges from such a dynamical model. The analysis of stochastic dynamics gives rise to the concept of multiple time scales.

Section 3 provides a systematic treatment of the Ornstein–Uhlenbeck Gaussian process as the linear stochastic process near a fixed point of a dynamical system. In a nutshell, the stability, i.e. hyperbolicity, of a fixed point is determined by the stationary probability distribution $f^{\text{st}}(x)$, and the type of a fixed point, i.e. node *versus* focus, etc, is determined by the stationary, divergence-free circulation $j^{\text{st}}(x)$.

Section 4 presents the widely practiced ‘top-down’ approach based on SDEs and related diffusion processes. We suggest, however, that when approximating the large, but not infinite, population limit of a BDP with bistability, diffusion theory encounters a dilemma. It can provide a faithful representation for either the stationary behaviour or the fluctuating ‘downhill’ dynamics, but not both. We further illustrate the intimate relation of this problem to several other issues: the Keizer’s paradox [22], Kurtz’s convergence theorem with finite time [8], and van Kampen’s conditional diffusion equation [23, 24].

Section 5 gives a brief discussion of two types of bistability in a mesoscopic chemical reaction system: that with a macroscopic, deterministic nonlinear counterpart, and that without. It is shown that their difference can be understood from the volume dependence of the transition rates between the two attractors.

In section 6, we show how insights from studying stochastic, nonlinear chemical reaction systems can be useful to the studies of other population dynamics. We try to establish some kinetic isomorphism between chemical dynamics and ecological dynamics.

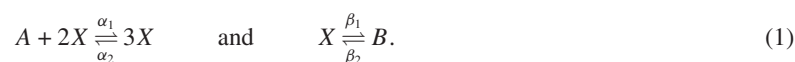
Section 7 concludes the paper with some discussions and outlooks.

In the appendices, we have given some details of the mathematical results used in the main text. Much of this material is not found in the literature.

2. Nonlinear stochastic population dynamics: the individual-based approach

In this section, we present the theoretical development of stochastic models for nonlinear chemical reaction dynamics. The approach here is ‘bottom-up’ since we use an individual’s stochastic behaviour as the starting point, considering one individual molecule at a time. As we shall see, this approach is in sharp contrast to the ‘top-down’ approach of section 4.

The approach we advance is general for any chemical and biochemical reaction system. However, we shall not present the theory in its most general form that often obscures the insights. Rather, we shall use a simple example to illustrate the theoretical approach. Let us consider the biochemical reaction system given by



This nonlinear chemical reaction system is known as the Schlögl model [25, 26]. The autocatalytic step in fact is widely observed in cellular biochemistry such as Src family kinase signalling, Rabaptin-5 mediated Rab5 GTPase activation in endocytosis, *Xenopus* oocyte maturation via a mitogen-activated protein kinase (MAPK) pathway, and self-regulating gene networks [4, 5, 27, 28].

2.1. Analysis of the deterministic dynamics

According to the law of mass action [29, 30], the nonlinear differential equation for $x(t)$, the concentration of the molecular species X in (1), is [19, 26, 31, 32]

$$\frac{dx}{dt} = \beta_2 b - \beta_1 x + \alpha_1 a x^2 - \alpha_2 x^3, \quad (2)$$

where a and b represent the concentrations of chemical species A and B , which are assumed to be sustained at constant values. Any biochemical system in living organisms has to have at least one ‘source’ and one ‘sink’ species. This feature has been called an *open driven chemical system*. A system in contact with only a single material reservoir is called *grand canonical system* according to Gibbs [33, 34]. The latter necessarily approaches a ‘dead’ chemical equilibrium.

Non-dimensionalizing equation (2) with new variables and parameters

$$z = \frac{\beta_1 x}{\beta_2 b}, \quad \tau = \beta_1 t, \quad \sigma = \frac{\alpha_2 (\beta_2 b)^2}{\beta_1^3}, \quad \gamma = \frac{a \alpha_1 \beta_1}{\alpha_2 \beta_2 b}, \quad (3)$$

we have

$$\frac{dz}{d\tau} = 1 - z + \sigma (\gamma z^2 - z^3) = f(z). \quad (4)$$

It is easy to show that for a wide range of parameter values, $f(z) = 0$ has three positive roots, corresponding to $\frac{dz}{d\tau} = f(z)$ having two stable fixed points and one unstable fixed point. See figure 2(a).

Equation (4) exhibits bistability when the parameter pair (σ, γ) is in the region bound by the curve in parametric form with

$$\sigma = \frac{z-2}{z^3} \quad \text{and} \quad \gamma = \frac{(2z-3)z}{z-2}. \quad (5)$$

Figure 2(b) shows that the region in which bistability exists has a cusp at $\sigma = \frac{1}{27}, \gamma = 9$. The nonlinear dynamics exhibits the canonical saddle-node bifurcation and Thom–Zeeman’s catastrophe [29].

2.2. The CME and stochastic models

There are four elementary reactions in system (1). In an aqueous solution, the occurrence of a reaction is a random event with exponentially distributed waiting time. The stochastic dynamics of the number of molecule X , $n(t)$, therefore, is a one-dimensional BDP. As a continuous-time Markov process, the BDP has its Kolmogorov forward equation, the CME, in the form [7, 26]

$$\frac{d}{dt} p(n, t) = p(n-1, t) \mu_{n-1} - p(n, t) (\mu_n + \lambda_n) + p(n+1, t) \lambda_{n+1}, \quad (6)$$

in which

$$\mu_n(V) = \frac{\alpha_1 a n(n-1)}{V} + \beta_2 b V \quad \text{and} \quad \lambda_n(V) = \frac{\alpha_2 n(n-1)(n-2)}{V^2} + \beta_1 n, \quad (7)$$

are the birth and death rates of the process. Note both are functions of the system’s size V .

For this simple system, it is not difficult to show heuristically that in the limit of $V \rightarrow \infty$ and $n \rightarrow \infty$, but $n/V \rightarrow x$, the stochastic dynamics following the BDP becomes the solution to the ODE

$$\frac{dx}{dt} = \mu(x) - \lambda(x), \quad (8)$$

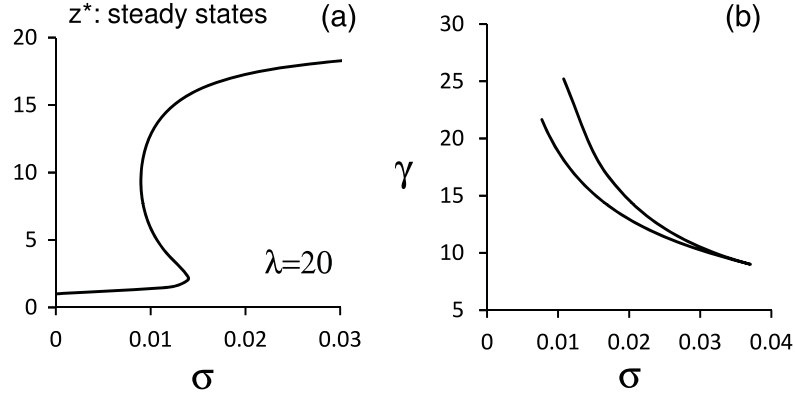


Figure 2. Nonlinear chemical reaction system (1) can exhibit bistability. (a) Fixed points of the ODE in equation (4), z^* , are obtained from $f(z) = 0$, as a function of σ with $\lambda = 20$: $\sigma = (z^* - 1)/(\gamma - z^*)/z^{*2}$. (b) The region of parameter space (σ, λ) in which the ODE is bistable has a cusp at $\sigma = \frac{1}{27}$, $\gamma = 9$. In the statistical physics theory of phase transition, the cusp is also known as a critical point [19].

with

$$\mu(x) = \lim_{n \rightarrow \infty} \frac{\mu_n(n/x)}{n/x} = \alpha_1 a x^2 + \beta_2 b \quad \text{and} \quad \lambda(x) = \lim_{n \rightarrow \infty} \frac{\lambda_n(n/x)}{n/x} = \alpha_2 x^3 + \beta_1 x.$$

Equation (8) is exactly equation (2). See [8] for a rigorous, general proof of the important limit theorem that connects the BDP following the CME and the ODE according to the law of mass action. Also see [35] for an alternative proof.

Because the problem is one-dimensional, the stationary probability distribution to equation (6) is readily obtained:

$$p^{\text{st}}(n; V) = p^{\text{st}}(0) \prod_{k=1}^n \frac{\mu_{k-1}(V)}{\lambda_k(V)}, \quad (9)$$

where $p^{\text{st}}(0)$ is determined by normalization of the probability distribution.

The distribution (9) has several important properties:

- (i) Its extrema are located at n^* where $\mu_{n^*-1} = \lambda_{n^*}$. An extreme corresponds to, therefore, a fixed point of the ODE, where $\mu(x) = \lambda(x)$.
- (ii) One can obtain an asymptotic expansion when $V, n \rightarrow \infty$ and $n/V \rightarrow x$:

$$p^{\text{st}}(n; V) \longrightarrow f^{\text{st}}(x; V) = e^{-V\phi(x)}, \quad \text{where} \quad \phi(x) = \int_0^x \ln \left[\frac{\lambda(v)}{\mu(v)} \right] dv. \quad (10)$$

Note that when $V \rightarrow \infty$, the $f^{\text{st}}(x, V)$ converges to the global minimum of $\phi(x)$.

- (iii) The function $\phi(x)$ is a Lyapunov function for the ODE (8):

$$\frac{d\phi(x(t))}{dt} = \frac{d\phi(x)}{dx} \frac{dx}{dt} = \{\mu(x) - \lambda(x)\} \ln \left(\frac{\lambda(x)}{\mu(x)} \right) \leq 0. \quad (11)$$

The rhs is equal to zero iff $\mu(x) = \lambda(x)$. While this result might not be too surprising in the case of one-dimensional dynamics, differentiable $\phi(x)$ can be obtained for many higher dimensional chemical reaction systems without detailed balance. $\phi(x)$, known as the large deviation rate function in the theory of probability, is a Lyapunov function for the ODEs from the mass-action law! [36, 37]

For the non-dimensionalized Schlögl model, one has

$$\begin{aligned} \phi(z) = \int_0^z \ln \frac{(\sigma u^2 + 1)u}{\sigma \gamma u^2 + 1} du &= z \ln \frac{(\sigma z^2 + 1)z}{\sigma \gamma z^2 + 1} + \frac{2}{\sqrt{\sigma}} \arctan(\sqrt{\sigma} z) \\ &\quad - \frac{2}{\sqrt{\sigma \gamma}} \arctan(\sqrt{\sigma \gamma} z) - z. \end{aligned} \quad (12)$$

- (iv) The two basins of attraction should be understood as two states of the chemical reaction system (1). They are the ‘emergent properties’ of the stochastic nonlinear population dynamics. For a given system, one could be in one of the states for a very long time. Still, the elementary operations at the individual level dictate the existence of the other state. This is a perspective that one cannot gain from pure statistical inference.

While the ODE predicts the existence of bistability, it cannot provide an estimation for the stability of the states. The stability can be obtained from the BDP by computing the mean first passage times (MFPTs). For the simple one-dimensional system, let n_1^* and n_2^* be the two peak positions and n_3^* the trough position of $p^{\text{st}}(n)$, then [24]

$$T_{n_1^* \rightarrow n_2^*} = \sum_{n=0}^{n_1^*} \sum_{m=n_1^*+1}^n \frac{p^{\text{st}}(n)}{\lambda_m p^{\text{st}}(m)} + \sum_{n=n_1^*+1}^{n_2^*-1} \sum_{m=n+1}^{n_2^*} \frac{p^{\text{st}}(n)}{\lambda_m p^{\text{st}}(m)}. \quad (13)$$

In the case of $V \rightarrow \infty$, equation (13) becomes (appendix B.2)

$$T_{x_1^* \rightarrow x_2^*} \approx \frac{2\pi e^{V(\phi(x_3^*) - \phi(x_1^*))}}{\lambda(x_3^*) \sqrt{-\phi''(x_1^*) \phi''(x_3^*)}}. \quad (14)$$

This time grows exponentially with system size V . Transitions between the two domains of attraction (DoA) are rare events.

- (v) A comprehensive theory emerges from analysing this simple model. There are three different time scales in the nonlinear, stochastic population dynamics: (1) the time scale of individual reactions, the α ’s and β ’s, which we call the molecular signalling time scale t_{ms} in the context of cellular biochemistry, (2) the time scale of nonlinear network dynamics t_{nd} and (3) the time scale on which the transitions between the DoA occurs, i.e. $T_{x_1^* \rightarrow x_2^*}$ and $T_{x_2^* \rightarrow x_1^*}$, which we call cellular evolution t_{ce} . In nonlinear deterministic dynamics, a long time means $t \gg t_{\text{nd}}$ but it is still $\ll t_{\text{ce}}$. On this time scale, a system settles into one attractor depending on the initial state. However, on the time scale $t \gg t_{\text{ce}}$, the system will establish a probability distribution between the two DoA.
- (vi) There is a great separation of time scales between t_{nd} and t_{ce} for a system with large populations. In this case, on the time scale $\gg t_{\text{nd}}$ but $\ll t_{\text{ce}}$, the system’s behaviour is captured by a bifurcation diagram, such as that in figure 2(a). However, on the time scale $\gg t_{\text{ce}}$, the stationary probability distribution given in equation (10) shows that the global minimum of $\phi(x)$ will have probability of almost 1, while other minimum will have only probability $\propto e^{-cV}$ where c is a positive constant. Therefore, for large systems with $t \gg t_{\text{ce}}$, the bifurcation diagram in figure 2(a) has to be modified by the Maxwell construction [19, 32]. This is the subject of phase transition theory in statistical mechanics.

3. Ornstein–Uhlenbeck processes: linear analysis of stochastic dynamics

One very useful method for analysing nonlinear dynamical systems is the local, linear analysis of fixed points. For nonlinear stochastic dynamics, the corresponding linear analysis is the theory of Gaussian–Markov processes, also known as Ornstein–Uhlenbeck processes. The

subject has been extensively studied by physicists such as Einstein, Chandrasekhar, Ornstein–Uhlenbeck–Wang, Onsager–Machlup, Lax, Keizer, Fox, and many others [13, 38, 40, 41]. Recent work has paid particular attention to the issue of time irreversibility and the breakdown of detailed balance in Gaussian processes [42, 43].

We consider linear SDE in the vector form

$$dx(t) = Bx(t) dt + \Gamma dW(t), \quad (15)$$

in which B and Γ are $n \times n$ constant matrices and x and W are n -dimensional column vectors. $W(t)$ contains n independent standard Brownian motions.

SDE (15) can be analysed by using several different methods, including the direct method

$$x(t) = e^{Bt} \left(x(0) + \int_0^t e^{-Bs} \Gamma dW(s) \right), \quad (16)$$

and the Fourier transform method

$$-i\omega \tilde{x}(\omega) = B\tilde{x}(\omega) + \Gamma \tilde{\xi}(\omega), \quad (17)$$

in which

$$x(t) = \int_{-\infty}^{\infty} \tilde{x}(\omega) e^{-i\omega t} d\omega \quad \text{and} \quad \frac{dW(t)}{dt} = \int_{-\infty}^{\infty} \tilde{\xi}(\omega) e^{-i\omega t} d\omega. \quad (18)$$

The Fourier transform of independent white noise satisfies

$$\langle \tilde{\xi}^*(\omega) \tilde{\xi}^T(\omega) \rangle = I, \quad (19)$$

the identity matrix, where $\langle \cdot \cdot \rangle$ is the ensemble average.

The stationary $x(t)$ has a multivariate Gaussian distribution

$$f^{\text{st}}(x) = \frac{1}{(2\pi)^{n/2} \det^{\frac{1}{2}}(\Xi)} \exp\left(-\frac{1}{2} x^T \Xi^{-1} x\right) \quad (20)$$

in which the symmetric matrix Ξ is the covariant matrix satisfying the Lyapunov matrix equation [44]

$$B\Xi + \Xi B^T = -A \quad \text{with} \quad A = \Gamma \Gamma^T. \quad (21)$$

3.1. Power spectrum of a stationary OU process with circulation

A stationary $x(t)$ has not only a distribution, given in equation (20), but also temporal correlation. One way to characterize the temporal dynamics is by the power spectrum. From equation (17) we have [45]

$$\tilde{x}^*(\omega) = [i\omega I - B]^{-1} \Gamma \tilde{\xi}^*(\omega) \quad \text{and} \quad \tilde{x}^T(\omega) = -\tilde{\xi}^T(\omega) \Gamma^T [i\omega I + B]^{-T}. \quad (22)$$

Thus, we have the power spectra for a multi-dimensional, stationary OU process

$$\Theta(\omega) \triangleq \langle \tilde{\xi}^*(\omega) \tilde{\xi}^T(\omega) \rangle = -[i\omega I - B]^{-1} A [i\omega I + B]^{-T}. \quad (23)$$

In other words,

$$\Theta^{-1}(\omega) = \underbrace{B^T A^{-1} B + \omega^2 A^{-1}}_{\text{symmetric}} + i\omega \underbrace{(A^{-1} B - B^T A^{-1})}_{\text{anti-symmetric}}. \quad (24)$$

In [42], it was shown that a stationary OU process is time-reversible iff $A^{-1} B = B^T A^{-1}$, and furthermore $\Xi = \frac{1}{2} B^{-1} A$. Therefore, a time-reversible OU process has

$$\Theta(\omega) = [B^2 + \omega^2 I]^{-1} A. \quad (25)$$

We see that, in this case, the power spectra $\Theta(\omega)$ is a real symmetric matrix and has a Lorentzian form with its peak located at $\omega = 0$ [46, 47].

When $A^{-1}B \neq B^T A^{-1}$, the anti-symmetric term in equation (24) indicates that the stationary $x(t)$ has a circular motion [48]. This circular motion can be best illustrated by considering the stationary Fokker–Planck equation (FPE) for the SDE (15):

$$\frac{\partial f(x)}{\partial t} = \nabla \cdot \left(\frac{1}{2} A \nabla f(x) - B x f(x) \right) = 0. \quad (26)$$

Integrating equation (26) once,

$$\frac{1}{2} A \nabla f^{\text{st}}(x) - B x f^{\text{st}}(x) = -j^{\text{st}}(x) \quad \text{with} \quad \nabla \cdot j^{\text{st}}(x) = 0. \quad (27)$$

This can be rewritten as

$$\frac{1}{2} \nabla \ln f^{\text{st}}(x) - A^{-1} B x = -A^{-1} j^{\text{st}}(x) f^{-1}(x). \quad (28)$$

Symmetry, $A^{-1}B = B^T A^{-1}$, means $A^{-1}Bx$ is a gradient force. Then by the uniqueness of the solution to the linear elliptic equation, $j^{\text{st}} = 0$ and $\ln f^{\text{st}}(x) = x^T A^{-1} B x + \text{const.}$ That is $2\Xi = B^{-1}A$.

If $A^{-1}B \neq B^T A^{-1}$, then $j^{\text{st}} \neq 0$. In fact,

$$j^{\text{st}}(x) = \left(B + \frac{1}{2} A \Xi^{-1} \right) x f^{\text{st}}(x). \quad (29)$$

The divergence-free $j^{\text{st}}(x)$ represents certain circular motion, which occurs only when a vector field is non-gradient.

It is easy to verify that the $j^{\text{st}}(x)$ and $\nabla \ln f^{\text{st}}(x)$ are orthogonal to each other [43]:

$$\begin{aligned} \nabla \ln f^{\text{st}}(x) \cdot j^{\text{st}}(x) &= x^T \Xi^{-1} \left(B + \frac{1}{2} A \Xi^{-1} \right) x f^{\text{st}}(x) \\ &= \frac{1}{2} x^T \left(\Xi^{-1} B - B^T \Xi^{-1} \right) x f^{\text{st}}(x) = 0. \end{aligned} \quad (30)$$

3.2. The Green–Kubo–Zwanzig relation

Because the SDE in (15) is linear, it is also easy to obtain

$$E^{x_0} [x(t) | x(0) = x_0] = e^{Bt} x_0, \quad (31)$$

where $E^{x_0} [\dots | x(0) = x_0]$ is the conditional ensemble average with given initial $x(0) = x_0$.

Then the stationary time correlation function matrix is, for $t \geq 0$,

$$E^{\text{st}} [E^{x_0} [x(t) | x(0) = x] x^T] = E^{\text{st}} [e^{Bt} x x^T] = e^{Bt} \Xi. \quad (32)$$

Therefore,

$$G_{xx}(t) = \langle x(\tau) x^T(\tau + t) \rangle = \begin{cases} \Xi e^{B^T t} & t \geq 0, \\ e^{-Bt} \Xi & t \leq 0. \end{cases} \quad (33)$$

We note that

$$G_{xx}(-t) = G_{xx}^T(t). \quad (34)$$

For time-reversible processes, $B\Xi = \Xi B^T$. Hence, $G_{xx}(-t) = G_{xx}^T(t) = G_{xx}(t)$ [49].

The Green–Kubo–Zwanzig formula concerns the mathematical relation between the transport coefficients and the integrals of the time-correlation function of the velocity [46, 50]. In the case of multi-dimensional Gaussian processes, the velocity is simply $v(t) = Bx(t)$ and $\langle v(\tau) v^T(\tau + t) \rangle = B G_{xx}(t) B^T$, where the correlation function matrix $G_{xx}(t) = \langle x(\tau) x^T(\tau + t) \rangle$ is the Fourier transform of $\Theta(\omega)$. Assuming all the eigenvalues of B have negative real parts,

the integral of the time-correlation function is the spectrum value of the process at $\omega = 0$. Therefore, from equation (23) we have

$$\int_{-\infty}^{\infty} \langle v(\tau) v^T(\tau + t) \rangle dt = B \left(\int_{-\infty}^{\infty} G_{xx}(t) dt \right) B^T = B \Theta(0) B^T = A. \quad (35)$$

Note that in this version, the Green–Kubo–Zwanzig formula requires no time-reversibility. It is a consequence of a linear stochastic dynamical system.

However, if one considers only $t \geq 0$, then

$$\int_0^{\infty} \langle v(\tau) v^T(\tau + t) \rangle dt = -B \Xi B^{-T} B^T = -B \Xi. \quad (36)$$

In this case, the rhs is $\frac{1}{2}A$ if and only if $B^{-1}A$ is symmetric, i.e. the Gaussian process is time reversible.

3.3. Gaussian processes in a plane

In nonlinear dynamical systems, local linear stability analysis near a fixed point in a plane has provided great insights into the nature of fixed points. In this section, we shall carry out a similar analysis for Gaussian OU processes in a plane.

We consider the two matrices

$$B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}, \quad (37)$$

where A is positive definite. Solving the Lyapunov matrix equation (21) we have the covariance matrix

$$\Xi = -\frac{1}{2(b_{11} + b_{22}) \det(B)} \times \begin{pmatrix} (b_{11}b_{22} - b_{12}b_{21} + b_{22}^2)a_{11} & -b_{21}b_{22}a_{11} + 2b_{11}b_{22}a_{12} \\ -2b_{12}b_{22}a_{12} + b_{12}^2a_{22} & -b_{11}b_{12}a_{22} \\ -b_{21}b_{22}a_{11} + 2b_{11}b_{22}a_{12} & b_{21}^2a_{11} - 2b_{11}b_{21}a_{12} \\ -b_{11}b_{12}a_{22} & +(b_{11}^2 + b_{11}b_{22} - b_{12}b_{21})a_{22} \end{pmatrix}. \quad (38)$$

Thus,

$$\Xi^{-1} = -\frac{2}{(b_{11} + b_{22})(\det(A) + \delta^2)} \times \begin{pmatrix} b_{21}^2a_{11} - 2b_{11}b_{21}a_{12} & b_{21}b_{22}a_{11} - 2b_{11}b_{22}a_{12} \\ +(b_{11}^2 + b_{11}b_{22} - b_{12}b_{21})a_{22} & +b_{11}b_{12}a_{22} \\ b_{21}b_{22}a_{11} - 2b_{11}b_{22}a_{12} & (b_{11}b_{22} - b_{12}b_{21} + b_{22}^2)a_{11} \\ +b_{11}b_{12}a_{22} & -2b_{12}b_{22}a_{12} + b_{12}^2a_{22} \end{pmatrix}, \quad (39)$$

in which

$$\delta = \frac{b_{11}a_{12} + b_{12}a_{22} - b_{21}a_{11} - b_{22}a_{12}}{b_{11} + b_{22}}.$$

Note that $-(b_{11} + b_{22}) > 0$ for a stable fixed point.

What is the relationship between the B , the linear stability matrix, and the Ξ^{-1} , the inverse of covariance matrix, that constitutes the quadratic form $\frac{1}{2}x^T \Xi^{-1}x$? We make the following observations:

(a) The determinants of B and Ξ^{-1} have same sign:

$$\det(\Xi^{-1}) = \frac{4 \det(B)}{\det(A) + \delta^2}. \quad (40)$$

- (b) The hyperbolicity of the fixed point of $\dot{x} = Bx$ is in agreement with the quadratic function $\phi(x) = \frac{1}{2}x^T \Xi x = -\ln f^{\text{st}}(x) + \text{const}$. This can be shown from

$$\dot{x}^T \cdot \nabla \phi(x) = x^T B^T \Xi^{-1} x = -\frac{1}{2}x^T \Xi^{-1} A \Xi^{-1} x \leq 0. \quad (41)$$

However, the nature of a stable fixed point, i.e. being a node or a focus, is determined by the sign of the *discriminant*

$$\text{Tr}^2(B) - 4 \det(B). \quad (42)$$

This information is not contained in the symmetric Ξ^{-1} whose discriminant is always greater than zero. To see whether the eigenvalues of B are complex, we turn to the divergence-free $j^{\text{st}}(x)$.

- (c) $j^{\text{st}}(x) = 0 \iff A^{-1}B = B^T A^{-1}$.
 (d) Since A is positive definite, we have a real, symmetric matrix $A^{\frac{1}{2}}$. If B has a pair of complex eigenvalues, then $A^{-\frac{1}{2}}BA^{\frac{1}{2}}$ has a pair of complex eigenvalues, and therefore $A^{-\frac{1}{2}}BA^{\frac{1}{2}} \neq (A^{-\frac{1}{2}}BA^{\frac{1}{2}})^T = A^{\frac{1}{2}}B^T A^{-\frac{1}{2}}$. That is, $A^{-1}B \neq B^T A^{-1}$. If the fixed point is a focus, then the Gaussian process is time-irreversible for any A .
 (e) If B has all real eigenvalues and is diagonalizable, then $\exists Q$ such that $Q^{-1}BQ$ is diagonal: $Q^{-1}BQ = Q^T B^T Q^{-T}$. One can choose $A = QQ^T$ and have $A^{-1}B = B^T A^{-1}$. If the fixed point is a node, then $\exists A$ such that the Gaussian process is time-reversible.
 (f) For an irreversible stationary process, its power spectrum can exhibit a peak at $\omega > 0$, indicating inherent frequency [51]. Spectral peaking, however, is only a sufficient condition for irreversibility, but a not necessary condition. For planar Gaussian processes, we have from equation (23)

$$\begin{aligned} \Theta(\omega) &= B^{-1}AB^{-T} + i\omega B^{-1}(B^{-1}A - AB^{-T})B^{-T} \\ &\quad - \omega^2(B^{-3}AB^{-T} - B^{-2}AB^{-2T} + B^{-1}AB^{-3T}) + O(\omega^3). \end{aligned} \quad (43)$$

Therefore, a condition for $\Theta_{11}(\omega)$ having an off-zero peak is its curvature at $\omega = 0$ being positive

$$\begin{aligned} \frac{d^2 \Theta_{11}(0)}{d\omega^2} &= ((b_{12}b_{21} - 2b_{22}^2 - 2b_{11}b_{22})b_{12}b_{21} - b_{22}^4)a_{11} \\ &\quad + (b_{11}^2 + b_{22}^2 + 2b_{21}b_{21})(2b_{12}b_{21}a_{12} - b_{12}^2a_{22}) \geq 0. \end{aligned} \quad (44)$$

The rhs can be rewritten as

$$\begin{aligned} &\underbrace{-(b_{21}, -b_{12})A \begin{pmatrix} b_{21} \\ -b_{12} \end{pmatrix}}_{\text{positive}} (b_{11}^2 + b_{22}^2 + 2b_{12}b_{21}) \\ &\quad + a_{11}(b_{11}b_{22} - b_{12}b_{21})^2 + a_{11}(b_{21}^2 - b_{22}^2)(b_{11}^2 + b_{22}^2 + 2b_{12}b_{21}). \end{aligned} \quad (45)$$

We see that the second term in (45) is positive since $a_{11} > 0$. If b_{12} and b_{21} have opposite signs, and $(b_{11}^2 + b_{22}^2) + b_{12}b_{21} < 0$, then the first term in (45) also becomes positive. But under this condition,

$$\begin{aligned} \text{Tr}^2(B) - 4 \det(B) &< -2b_{11}b_{22} + 3b_{12}b_{21} \leq 2|b_{11}b_{22}| + 3b_{12}b_{21} \\ &\leq (b_{11}^2 + b_{22}^2) + 3b_{12}b_{21} < 0. \end{aligned}$$

We shall show next, however, that it is actually possible to find an A such that for a B with a node, the corresponding OU stationary process has a power spectrum with its peak at $\omega > 0$.

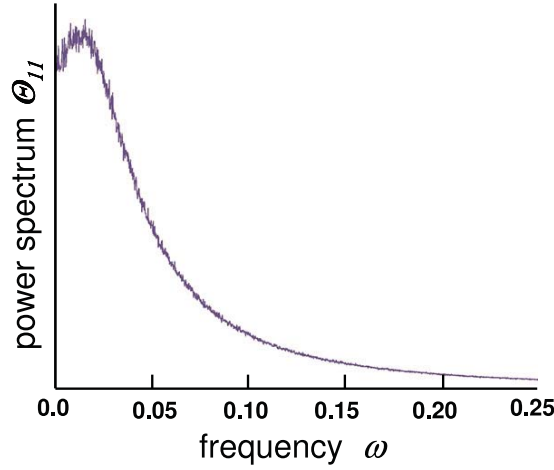


Figure 3. A planar linear dynamical system with a stable node (eigenvalues -2 and -3), when coupled to a white noise, becomes a time-irreversible OU Gaussian process exhibiting strong rotational motion. The power spectrum $\Theta_{11}(\omega)$, adopted from [45], shows an off-zero peak. (Figure provided by Dr Jia-zeng Wang.)

3.4. Noise-induced strong circular motion in a plane: an example

A Gaussian process is time irreversible iff $j^{\text{st}}(x) \neq 0$. A non-zero j^{st} indicates a certain kind of circulation in the dynamics of $x(t)$. Power spectral peaking is only a sufficient but not necessary condition for the circulation. Only when the circular motion is sufficiently ‘strong’ will its power spectrum exhibit an off-zero peak [51].

The presence of noise can induce strong circulation in an ODE system with only a node and no hint of any circular motion in the deterministic dynamics. We borrow the example given in [45], considering

$$B = \begin{pmatrix} -4 & 1 \\ -2 & -1 \end{pmatrix} \quad \text{and} \quad A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (46)$$

The two eigenvalues of B , -2 and -3 , are both real. However, the curvature of $\Theta_{11}(\omega)$ at $\omega = 0$ according to equation (44) is 10. Figure 3 shows the power spectrum.

Noise-induced circulations, or oscillations, have been extensively studied in the past in connection with the phenomena called *stochastic resonance* and *coherence resonance*. See [51–53] for studies from the perspective of irreversible stationary stochastic processes. The example in (46) and figure 3 is perhaps the most elementary version of this interesting phenomenon.

4. Diffusion theory of nonlinear population dynamics with fluctuations

The SDE,

$$dx(t) = b(x) dt + \Gamma(x) dW(t), \quad (47)$$

and the diffusion process it defines, is widely used to represent stochastic dynamics of natural and engineering systems [10, 13, 14, 24]. This is a much researched mathematical subject in both pure and applied mathematics.

Dynamics with fluctuations have been extensively studied in statistical physics, ever since the work of Einstein, Smoluchowski and Langevin, respectively, in 1905, 1906 and 1908. But it was mathematician K Itô who finally unified the mathematics of Einstein and Smoluchowski in terms of partial differential equations, and Langevin's approach in terms of SDEs (47). The description of the stochastic dynamics by diffusion processes is universal for Markov processes with continuous paths [14].

Still, one should note that the diffusion theory was conceived, in physics, as a 'correction' to deterministic dynamics. It is a phenomenological approach to stochastic fluctuations with a clear 'top-down' characteristic. When modelling with a diffusion process, it is almost always the case that one is uncertain about how to choose the Γ . It is worth noting that when modelling an equilibrium process in physics and chemistry, this problem was solved by the so called *fluctuation–dissipation theory* [45]. The Γ is not determined from the physical mechanism of the problem, but rather from requiring the stationary process to agree with known physics—Boltzmann's law, Onsager's regression hypothesis and time reversibility.

With such a rich history, it is natural for one to be interested in representing the dynamics of large, but not infinite, populations with diffusion processes [54, 55]. In particular, one may ask if the diffusion process description, which has enjoyed great successes in physics, can be the appropriate model for the CME in macroscopic volume with fluctuations? The answer turns out to be a surprising 'no' [56, 57]. The diffusion processes can fail to provide a global approximation for the nonlinear stochastic dynamics of a bistable population with birth and death [22, 26, 56, 57]. The crux of the matter turns out to rest precisely with the rare events. A rare event occurs with exponentially small probability e^{-cV} and exponentially long time e^{+cV} , where $c > 0$ and V is the system's size. The diffusive stochastic processes with continuous trajectories are not accurate enough global representations [58].

As a mathematical result, this has been known for a long time. Kurtz's 1971 limit theorem is only valid for finite time due to precisely this problem [8]. van Kampen has repeatedly emphasized that a diffusion approximation can only be obtained for master equations with *small individual jumps* [23, 24]. He actually developed a sophisticated treatment of diffusion approximations for the master equation, order-by-order, called system-size expansion [24, chapter 10]. This theory provides a satisfying approximation for the stochastic relaxation in the limit of large V . It is shown that the only mathematically valid diffusion process one can derive from a CME is a Gaussian process *conditioned* on a given deterministic solution to the corresponding ODE (see [appendix A](#)). Both excluded the rare jumps between multiple nonlinear attractors. This approach, thus, does not address how to obtain a stationary distribution with multistability.

In the present review, we shall revisit this problem from a different perspective. From Kurtz's theorem and van Kampen's system-size expansion, we know that in the limit of large system size, one can obtain a diffusion approximation near a fixed point, as we have done in section 3. This approximation, however, underestimates the time for barrier crossing (see equation (63)). On the other hand, it is also possible to obtain a different diffusion approximation which gives the correct, global stationary distribution. But this one misrepresents the downhill dynamics. We call this 'diffusion theory's dilemma' [58].

4.1. A simple example

We again use the example of a one-dimensional BDP in equation (6), which resembles a difference scheme of a diffusion equation of Fokker–Planck type:

$$\frac{\partial f(x, t)}{\partial t} = \frac{1}{2} \frac{\partial^2}{\partial x^2} (A(x) f(x, t)) - \frac{\partial}{\partial x} (b(x) f(x, t)). \quad (48)$$

If we identify $x = n/V$, and $dx = 1/V$, then the Kramers–Moyal expansion [10, 24, 40], truncated at the second-order, yields

$$A(x) = \frac{\mu(x) + \lambda(x)}{V} \quad \text{and} \quad b(x) = \mu(x) - \lambda(x). \quad (49)$$

Since the $A(x)$ term is on the order of $\frac{1}{V}$, equation (48) can also be written as

$$\frac{\partial f(x, t)}{\partial t} = \frac{\partial}{\partial x} \left(\frac{\mu(x) + \lambda(x)}{2V} \frac{\partial f}{\partial x} - (\mu(x) - \lambda(x)) f(x, t) \right), \quad (50)$$

with the divergence form for the diffusion. The difference is in the $b(x)$ term on the order of $O(1/V)$, which is negligible.

The stationary distribution to equation (50) is readily obtained as

$$\tilde{f}^{\text{st}}(x) = e^{-V\tilde{\phi}(x)}, \quad \text{where} \quad \tilde{\phi}(x) = 2 \int_0^x \frac{\lambda(v) - \mu(v)}{\lambda(v) + \mu(v)} dv. \quad (51)$$

Now comparing the $\tilde{\phi}(x)$ with the $\phi(x)$ in equation (10), we see that they are not identical. However, both have the same extrema x^*

$$\frac{d}{dx} \tilde{\phi}(x^*) = \frac{d}{dx} \phi(x^*) = 0 \quad \text{at} \quad \mu(x^*) = \lambda(x^*). \quad (52)$$

In fact, both have identical curvature near an extreme

$$\frac{d^2}{dx^2} \tilde{\phi}(x^*) = \frac{d^2}{dx^2} \phi(x^*) = \frac{1}{\mu(x^*)} \left(\frac{d\lambda(x^*)}{dx} - \frac{d\mu(x^*)}{dx} \right). \quad (53)$$

This means both have identical linear Gaussian dynamics near a fixed point. Equation (50) is a good approximation for the local dynamics.

$\tilde{\phi}(x)$ and $\phi(x)$ can have very different global behaviour [26, 56]. To illustrate this, let us consider the particular example where

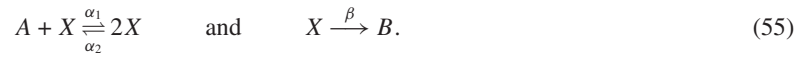
$$\mu(x) = \alpha_1 a x^2 + \beta_2 b \quad \text{and} \quad \lambda(x) = \alpha_2 x^3 + \beta_1 x \quad (54)$$

with $\alpha_1 = 6$, $\alpha_2 = 1.2$, $\beta_1 = 5.37$, $\beta_2 = 0.25$, $a = 1$ and $b = 1.4$. Figure 4 shows $\phi(x)$ as a solid blue line and $\tilde{\phi}(x)$ as a dashed orange line. The two functions are indeed very similar (figure 4(a)); however, a careful inspection shows that the $\phi(x_1^*) < \phi(x_2^*)$ but $\tilde{\phi}(x_1^*) > \tilde{\phi}(x_2^*)$ (figure 4(b)). Therefore, when $V \rightarrow \infty$, the $f^{\text{st}}(x) \rightarrow \delta(x - x_1^*)$ but $\tilde{f}^{\text{st}}(x) \rightarrow \delta(x - x_2^*)$.

4.2. Keizer's paradox

The disagreement between $\tilde{\phi}(x)$ and $\phi(x)$ in figure 4 illustrates that a naive, truncated Kramers–Moyal expansion of the BDP (6) in the form of FPE (50) yields good local approximations near every fixed point, but cannot provide a globally satisfying approximation for sufficiently long times with uniform convergence of $V \rightarrow \infty$ with respect to $\forall t$. This failure is intimately related to the rare events that connect the bistability of the corresponding ODE.

The issue can be further elucidated by an even simpler model. Keizer [40] discussed the autocatalytic reaction system with



The ODE following the law of mass action is

$$\frac{dx}{dt} = \alpha_1 a x - \alpha_2 x^2 - \beta x. \quad (56)$$

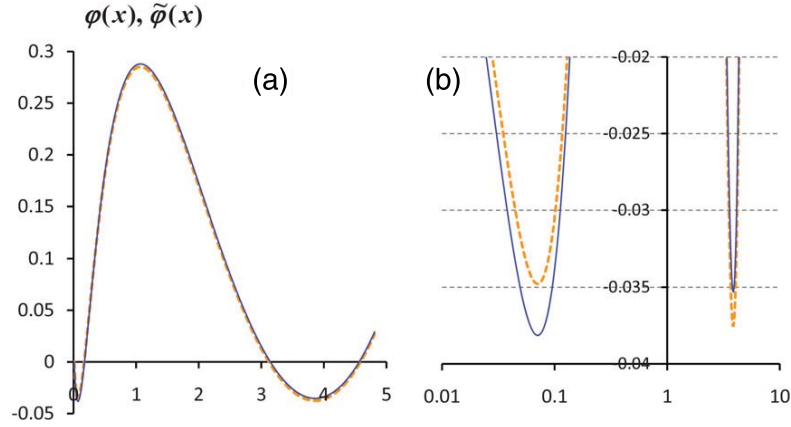


Figure 4. A comparison between the $\phi(x) = -(1/V) \ln f^{\text{st}}(x)$, in solid blue, obtained from equation (6) in the limit of large volume, given by equation (10), and the $\tilde{\phi}(x) = -(1/V) \ln \tilde{f}^{\text{st}}(x)$, in dashed orange, obtained from the diffusion approximation of the CME (6). For the Schlögl model, even though these two functions can be quite similar as shown in (a), a careful inspection in (b) shows that their global minima are different. It is at the left well for $\phi(x)$ and at the right well for $\tilde{\phi}(x)$. This difference in the infinite-volume limit becomes very significant: the probability approaches to 1 at the global minimum.

For parameters $\alpha_1 a - \beta > 0$, this is in fact the celebrated logistic equation in population dynamics, with growth rate $\alpha_1 a - \beta$ and carrying capacity $x^* = \frac{\alpha_1 a - \beta}{\alpha_2}$. The ODE has two fixed points: unstable $x = 0$ and stable x^* .

In the chemical reaction context, Keizer observed that the ODE's stable steady state is inconsistent with the stationary distribution of the CME model for the reaction system (55). The CME is again a BDP with

$$\mu_n = \alpha_1 a n \quad \text{and} \quad \lambda_n = \frac{\alpha_2 n(n-1)}{V} + \beta n. \quad (57)$$

Because $\mu_0 = 0$, the $n = 0$ is an absorbing state of the BDP, and its stationary distribution has probability 1 for $n = 0$, i.e. extinction. But the $x = 0$ is an *unstable* fixed point of the ODE!

The resolution to these seemingly paradoxical results is simple [22, 40]. As indicated in section 2.2, system (55) again has a separation of the nonlinear network dynamics time scale t_{nd} and the cellular evolution time scale t_{ce} . The fixed point of the ODE is for $t \gg t_{\text{nd}}$, but it is still $t \ll t_{\text{ce}}$. For $t \gg t_{\text{ce}}$, the system will be $n = 0$ with probability 1. However, for $t_{\text{nd}} \ll t \ll t_{\text{ce}}$, the system has a quasi-stationary distribution centred around the non-zero x^* . One can obtain this distribution as the eigenfunction associated with the largest non-zero eigenvalue of the CME. The eigenvalue, which $\propto e^{-cV}$ ($c > 0$), gives the time scale for reaching extinction.

Noting that from Kramers' theory, all barrier crossing rare events involve an exponentially slow 'climbing' to a saddle point and then a rapid 'descending' afterward (see appendix B). Keizer's paradox, therefore, is also at the root of the failure in section 4.1. The ODE predicts an infinitely long time for the climbing, but the FPE (50) predicts a time that is too short for the climbing.

4.3. The tale of two diffusion equations

Hänggi *et al* [56] proposed a different FPE that gives the correct stationary distribution

$$\frac{\partial f(x, t)}{\partial t} = \frac{1}{V} \frac{\partial^2}{\partial x^2} \left(\frac{\mu(x) - \lambda(x)}{\ln \mu(x) - \ln \lambda(x)} f \right) - \frac{\partial}{\partial x} ((\mu(x) - \lambda(x)) f). \quad (58)$$

It is not difficult to show that the stationary solution to equation (58) is the same as that in equation (10).

The ‘physics rationale’ for equation (58) is based on Onsager-type transport law as follows. First, the stationary distribution for the CME (6), in the limit of large V , is

$$p^{\text{st}}(x) = \exp \left\{ -V \int^x \ln \left(\frac{\lambda(z)}{\mu(z)} \right) dz \right\}. \quad (59)$$

The stochastic potential for the system is $\phi(x) = -(1/V) \ln p^{\text{st}}(x)$ and the thermodynamic force is $F(x) = -d\phi(x)/dx = \ln(\mu(x)/\lambda(x))$. The macroscopic ODE should be velocity \times frictional coefficient = force,

$$\frac{dx}{dt} = \mu(x) - \lambda(x) = \eta^{-1}(x) F(x). \quad (60)$$

Therefore, this yields

$$\eta^{-1}(x) = \frac{\mu(x) - \lambda(x)}{\ln \mu(x) - \ln \lambda(x)}. \quad (61)$$

And the diffusion coefficient $\propto \eta^{-1}(x)$. This relation ensures the logarithm of the stationary distribution being $\propto -\phi(x)$. This approach, therefore, amounts to enforcing the deterministic kinetics and the stationary distribution. In a one-dimensional system, these two constraints essentially determine a FPE.

The two diffusion equations in (50) and (58) have the same drift $b(x)$ given in equation (49), but different diffusion coefficients,

$$A_{\text{KM}}(x) = \mu(x) + \lambda(x) \quad \text{and} \quad A_{\text{HGTT}}(x) = \frac{2(\mu(x) - \lambda(x))}{\ln \mu(x) - \ln \lambda(x)}, \quad (62)$$

where subscripts KM and HGTT stand for Kramers–Moyal and the authors of [56], respectively.

The two diffusion coefficients are the same near x^* where $\mu(x) = \lambda(x)$, i.e. the fixed point of the $b(x)$:

$$\frac{\mu - \lambda}{\ln(\mu/\lambda)} \approx \lambda \left[1 + \frac{(\mu/\lambda) - 1}{2} - \frac{(\mu/\lambda - 1)^2}{12} + \dots \right] \approx \frac{\mu + \lambda}{2}. \quad (63)$$

However, away from the fixed point of $b(x)$, HGTT’s diffusion coefficient is always smaller than that of Kramers–Moyal’s.

The HGTT diffusion, unfortunately, is not the full solution to the problem. First, it is not clear how to generalize this approach to higher dimensional problems. More importantly, it actually poses a dilemma. As an approximation to the CME, the Kramers–Moyal’s diffusion gives the same finite time dynamics as the CME with large V , but a wrong stationary distribution. On the other hand, the HGTT diffusion gives the correct stationary distribution, but a wrong conditional diffusion equation for the finite time dynamics, as shown in [appendix A](#). The HGTT diffusion does give the correct mean time for downhill dynamics, as does the KM diffusion, since the mean time for downhill is independent of diffusion (see equation (B.3)). However, the variances and distributions of the downhill times are different from the correct KM diffusion.

Therefore, no diffusion processes, with any possible $A(x)$ and $b(x)$, will give a satisfying representation of the dynamics predicted by a CME with bistability in the limit of $V \rightarrow \infty$.

The origin of this difficulty is the tremendous separation of time scales in the ‘uphill’ and ‘downhill’ motion. For a unistable system, uphill motion will only lead to an exponentially small probability which can be safely neglected. However, for systems with multi-stability, the exponentially small probability is responsible for establishing the correct probability between the two stochastic attractors. This difficulty renders the diffusion theory not capable of reasonably representing nonlinear stochastic population fluctuations. Rather, a hybrid model that combines continuous diffusion with discrete Markov jump processes is required.

4.4. Diffusion theory’s dilemma

The problem can be even more tellingly stated as follows. According to the standard derivation of the diffusion equation from a discrete state and discrete time Markov chain with forward probability p , backward probability $q = 1 - p$, spacing δ and time step τ [59],

$$D = \frac{p+q}{2} \frac{\delta^2}{\tau} \quad \text{and} \quad b = (p-q) \frac{\delta}{\tau}. \quad (64)$$

Or

$$p = \frac{D\tau}{\delta^2} + \frac{b\tau}{2\delta} \quad \text{and} \quad q = \frac{D\tau}{\delta^2} - \frac{b\tau}{2\delta}. \quad (65)$$

We see that if both the diffusion coefficient D and the drift rate (bias) b exist, then in the limit of τ and $\delta \rightarrow 0$, $p/q \rightarrow 1$. In other words, diffusion theory assumes that in the very small spatial and temporal scale, the motion is purely random without bias. This feature, as we shall see, is inconsistent with the CME in the limit of $V \rightarrow \infty$.

If the p and q ($\neq p$) exist, then in the limit of τ and $\delta \rightarrow 0$, we have $D/b \rightarrow 0$. The diffusion is negligible. This is Kurtz’s theorem [8].

It seems to us that the stochastic trajectory of the CME in the thermodynamic limit, depends on one’s perspective. It is either a smooth, deterministic function of time, or a discontinuous stochastic function of time. There is really no diffusion process like behaviour. Thus we coined the term ‘diffusion theory’s dilemma.’

4.5. Diffusion theory’s dilemma and exponentially small asymptotics

The discussion so far has explained why the CME in general does not converge to a proper diffusion. As in the law of large numbers, the proper limit is a system of deterministic ODEs [8]. To have a proper diffusion, one has to ‘eliminate’ the ‘mean value.’ The situation is completely analogous to the sum of N identical, independently distributed random variables, $Y_N = X_1 + X_2 + \dots + X_N$. There is simply no way to capture both the mean value and the variance of Y_N with a single scaling. The scaling for the law of large numbers is N^{-1} , while for the Central Limit Theorem it is $N^{-1/2}$. This is precisely the idea behind the van Kampen’s system-size expansion, order-by-order [23, 24].

However, we still need to explain why the asymptotic form of the FPE with $1/V$ diffusion coefficient gives an erroneous stationary distribution. The insights from the present work point to this. The asymptotic order in the equation is singular. The stationary solution contains *exponentially small asymptotics* e^{-cV} ($c > 0$). This is the well-understood singularly perturbed linear two-point boundary value problem [60].

To give a better feel for the exponentially small asymptotics, let us consider the master equation with $\mu_n = \mu$ and $\lambda_n = \lambda$. Then the MFPT from n to zero, with reflecting boundary at n , is [10, 24]

$$T_n = \frac{1}{\mu - \lambda} \left(\frac{1 - (\lambda/\mu)^n}{(\lambda/\mu)^n - (\lambda/\mu)^{n+1}} \right) + \frac{n}{\lambda - \mu}. \quad (66)$$

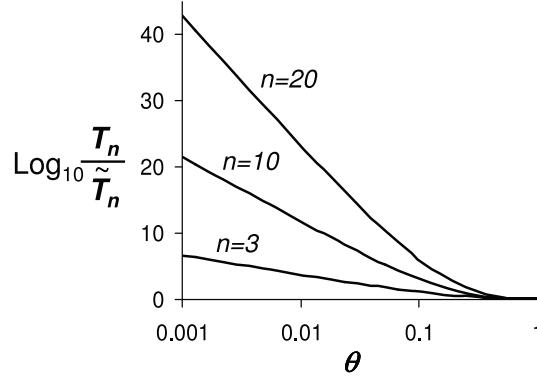


Figure 5. Comparison between the MFPTs in a BDP with constant birth and death rates, μ and λ , and corresponding Brownian motion with a constant drift $D = (\mu + \lambda)\delta^2/2$ and $V = (\mu - \lambda)\delta$. T_n is obtained from the discrete model and \tilde{T}_n is obtained from the corresponding diffusion process with $x = n\delta$. $\theta = \lambda/\mu < 1$ indicates the process to the absorbing state 0 is ‘uphill.’ With longer and longer ‘climbing’, i.e. larger n , the two times diverge exponentially.

Now if we consider the distance between n to $n + 1$ being δ , and let $\delta \rightarrow 0$ and $n \rightarrow \infty$, but $n\delta \rightarrow x$, then we have a FPE,

$$\frac{\partial f(x, t)}{\partial t} = D \frac{\partial^2 f}{\partial x^2} - V \frac{\partial f}{\partial x}, \quad (67)$$

where $D = (\mu + \lambda)\delta^2/2$ and $V = (\mu - \lambda)\delta$. The corresponding MFPT for this problem is [10, 24]

$$T_x = \frac{1}{V} \left[\frac{D}{V} \left(e^{\frac{V}{D}x} - 1 \right) - x \right]. \quad (68)$$

To compare the two results in equations (68) and (66), we rewrite equation (68) using $x = n\delta$ to obtain

$$\tilde{T}_n = \frac{(\mu + \lambda)}{2(\mu - \lambda)^2} \left(e^{\frac{2(\mu - \lambda)n}{(\mu + \lambda)}} - 1 \right) + \frac{n}{\lambda - \mu}. \quad (69)$$

Comparing \tilde{T}_n and T_n in equations (69) and (66) we have

$$\frac{T_n}{\tilde{T}_n} = \frac{\left(\frac{1 - \theta^n}{\theta^n - \theta^{n+1}} \right) - n}{\frac{(1 + \theta)}{2(1 - \theta)} \left(e^{\frac{2(1 - \theta)n}{(1 + \theta)}} - 1 \right) - n}, \quad (70)$$

where $\theta = \lambda/\mu$. We then have

$$\lim_{n \rightarrow \infty} \frac{T_n}{\tilde{T}_n} = \begin{cases} \infty & \theta < 1, \\ 1 & \theta \geq 1. \end{cases} \quad (71)$$

We note that $\theta > 1$ means the motion from positive x to zero is downhill. $\theta < 1$ means the motion from positive x to zero is uphill. The diffusion approximation for the master equation breaks down for the uphill dynamics! Figure 5 shows the ratio T_n/\tilde{T}_n as a function of θ and finite n .

We see that for the case of downhill dynamics ($\theta > 1$), both T_n and \tilde{T}_n approach $n/(\lambda - \mu)$, where $\lambda - \mu$ is the velocity. For the case of uphill dynamics ($\theta < 1$), both $\sim e^{cn}$, but with

different, positive c 's. In fact we have

$$\lim_{n \rightarrow \infty} \frac{1}{n} \ln \frac{T_n}{\tilde{T}_n} = -\ln \theta - \frac{2(1-\theta)}{1+\theta} > 0. \quad (\theta < 1). \quad (72)$$

We see the two expressions in equation (63) appear again here.

When $\theta \geq 1$, we indeed have that T_n/\tilde{T}_n converges with order $1/n$:

$$\lim_{n \rightarrow \infty} n \ln \frac{T_n}{\tilde{T}_n} = \begin{cases} 1 & \theta = 1, \\ \frac{1}{2} & \theta > 1. \end{cases} \quad (73)$$

This is at the heart of both examples. The stationary probabilities between two peaks are determined by the ratio of two exponentially long times to transition (back and forth), or two rare events with exponentially small probabilities. In other words, if an approximation can not give the 'exponent' correctly, then it is not a meaningful 'approximation"! In probability theory, this is the domain of the large deviation theory [19, 61].

4.6. The Delbrück–Gillespie process versus the Wright–Fisher model

The dynamic model for a mesoscopic, homogeneous chemical or biochemical reaction system is a stochastic process with birth and death. Any stochastic Markov process has two different mathematical representations: its stochastic trajectories and its time-dependent probability distribution following a Kolmogorov forward equation. In the context of the present review, they are the Gillespie algorithm [2] and the CME [62], respectively. While these two views of a stochastic process are mathematically equivalent, each only gives a partial understanding. For this reason, we would like to introduce the term *Delbrück–Gillespie* to refer to the stochastic process itself.

Tan [54, p 271] extensively discussed the conditions for a valid diffusion approximation (48) of finite BDs like (6). A similar analysis is presented in [appendix C](#). It showed that the necessary condition for a master equation to converge to a non-degenerate diffusion is

$$\ln \left(\frac{\mu_{n-1}(V)}{\lambda_n(V)} \right) \sim O \left(\frac{1}{V} \right) = dx. \quad (74)$$

This form yields a stationary probability density $f^{\text{st}}(x)$ which is properly supported on all x . When this condition is not met, as in the CME, one has

$$\ln \left(\frac{\mu_{n-1}(V)}{\lambda_n(V)} \right) = \left(V \ln \frac{\mu(x)}{\lambda(x)} \right) dx.$$

The V on the rhs gives rise to the form $f^{\text{st}}(x) = e^{-V\phi(x)}$, which in the limit of $V = \infty$ will have x only supported at the global minimum of $\phi(x)$.

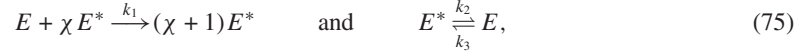
The population genetic models, on the other hand, have long enjoyed their fruitful relationship with the diffusion processes [54, 55, 63]. There is indeed a significant difference between the chemical reaction system and the genetic system. The random sampling in discrete genetic models is equivalent to a *long-range diffusion*, not just among the nearest neighbours. Hence it has a valid diffusion equation in the limit of large sample size, with both diffusion and drift terms being finite. This suggests in the infinite population size, deterministic limit, its nonlinear dynamics for the number density has the form $dx/dt = b_1(x)V + b_0(x)$, where the term $b_1(x)V$ means the individual reaction is aware of the size of the entire system. This is a *volume-dependent rate*.

This is indeed the case for the discrete population genetic drift model of Wright and Fisher [54, 55, 63]. But it does not arise from chemical kinetics in an ideal solution. The issue is as follows. In population genetic models, the conditional variance of an individual step is much greater than that in a Delbrück–Gillespie process. It is on the same order as the conditional expectation. In a chemical reaction, the former is a higher order infinitesimal.

5. Nonlinear and stochastic bistabilities

A basin of attraction around a stable fixed point in a deterministic nonlinear dynamical system corresponds to a peak in its stochastic counterpart. The converse is not necessarily true. A stochastic CME can have peaks that do not correspond to fixed points in the deterministic system of ODEs. We call the latter stochastic stability and the former nonlinear stability.

Consider the following autocatalytic reaction system [4, 5]



in which χ can be either 1 or 2. This model resembles Keizer's logistic system in equation (55) and the Schlögl system in equation (1). Assuming the system's volume is V and there are N total number of E and E^* molecules, the stationary probability distribution for the number of E^* , p_n , satisfies the steady-state CME [4, 19, 27]

$$(N - n + 1) \left(\frac{k_1(n-1) \cdots (n-\chi)}{V^\chi} + k_3 \right) p_{n-1} - \left[(N - n) \left(\frac{k_1 n \cdots (n - \chi + 1)}{V^\chi} + k_3 \right) + k_2 n \right] p_n + k_2(n+1)p_{n+1} = 0. \quad (76)$$

Solving the equation yields

$$p_n = p_0 \prod_{\ell=0}^{n-1} \frac{(N - \ell)}{(\ell + 1)} \left(\frac{k_1 \ell \cdots (\ell - \chi + 1)}{k_2 V^\chi} + \frac{k_3}{k_2} \right), \quad (77)$$

where p_0 is a normalization factor. Let $x = n/N$ be the fraction of E^* among E and E^* . The probability distribution can be written as

$$\ln p(x) = \ln p_0 + \sum_{z=0, \delta}^{x-\delta} \ln \left[\frac{(1-z)(\theta z \cdots (z - \chi\delta + \delta) + \eta)}{(z + \delta)} \right], \quad (78)$$

where $\eta = k_3/k_2$, $\delta = 1/N$, and $\theta = (k_1/k_2)(N/V)^\chi$.

When the system size tends to infinity, V and $N \rightarrow \infty$, N/V tends to a finite concentration E_t , and $\delta \rightarrow 0$, we have an integral expression of the probability distribution

$$\ln f(x) = \text{const} + N \int_0^x \ln \left[\frac{(1-z)(\theta z^\chi + \eta)}{z} \right] dz, \quad (79)$$

with continuous $x \in [0, 1]$. The distribution $f(x)$ has its extrema at the roots of the equation

$$\frac{(1-x)(\theta x^\chi + \eta)}{x} = 1. \quad (80)$$

The extrema of $f(x)$ match precisely with the macroscopic steady states from the law of mass action

$$\frac{dx}{dt} = (k_1(E_t x)^\chi + k_3)(1-x) - k_2 x = 0. \quad (81)$$

Equation (80) gives only monostability for $\chi = 1$ and the possibility of bistability for $\chi = 2$. This is the macroscopic behaviour of the chemical reaction system in (75). However, for smaller system sizes, the distribution in equation (77) can in fact have two peaks even for $\chi = 1$, if $\delta > \eta$. In this case, the peak locations of the distribution p_n are at $n_1^* = 0$ and at n_2^* , the larger root of the quadratic equation

$$\theta (n_2^*)^2 - N(\theta - 1 - \eta)n_2^* + (N - N^2\eta) = 0. \quad (82)$$

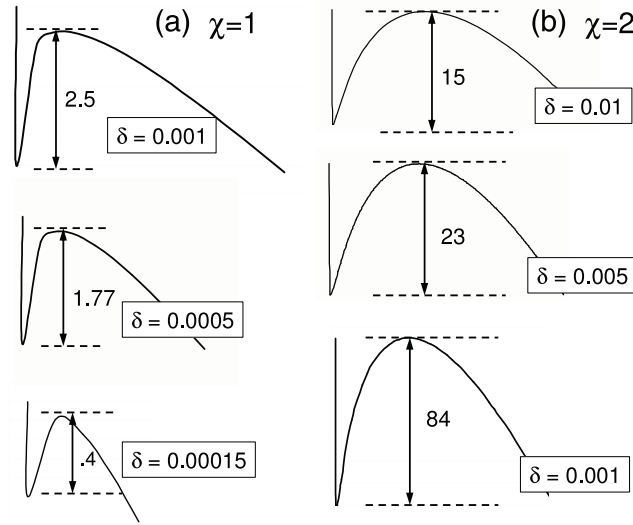


Figure 6. $-\ln p_n$ (the ordinate) as a function of n (the abscissa) and δ . (a) $\chi = 1$, $\theta = 1.5$, $\eta = 0.0001$. The peak locations, the smaller root of equation (82), are at $z = 0.002, 0.0008$, and 0.0001 for $\delta = 0.001, 0.0005, 0.00015$, respectively. (b) $\chi = 2$, $\theta = 10$, $\eta = 0.001$. The peak locations are at $z = 0.14, 0.125$, and 0.114 , respectively. With increasing system's size, i.e., $\delta \rightarrow 0$, the lifetime of the state in (a) decreases while in (b) it increases. (a) is called *stochastic bistability* and (b) is called *nonlinear bistability*.

Figure 6(a) shows the $-\ln p_n$ for three different values of δ . We see that the stability of the ‘energy well’ at $n = 0$ decreases when δ tends to zero. The well disappears when $\delta < \eta$. In contrast, for $\chi = 2$, figure 6(b) shows the stability of the energy well at $n = 0$ increases when δ tends to zero.

The distinction between nonlinear and stochastic bistabilities is related to the concept of ‘enthalpic barriers’ in the Arrhenius theory of the chemical reaction rate [64], $k = e^{-\Delta H^1/k_B T + \Delta S^1/k_B}$ in which ΔH^1 and ΔS^1 are called activation enthalpy and entropy, respectively. With decreasing temperature, i.e. decrease the thermal randomness, the rate of crossing an enthalpic barrier decreases exponentially if $\Delta H^1 > 0$ but increases if $\Delta H^1 < 0$.

6. Kinetic isomorphism and general population dynamics

While we have so far focused on biochemical reaction kinetics, the theory of nonlinear, stochastic multi-dimensional BDPs we developed in the present paper could and should be applied to many other population dynamics [65]. In this section, we shall establish a kinetic isomorphism between chemical reaction systems and general population dynamics such as predator-and-prey, competition, and cannibalism. By doing so, our understanding and development of the stochastic, nonlinear biochemical dynamics can be easily transferred to the studies of many other population systems in ecology, infection epidemics, and sociology¹.

¹ It seems to us that a distinct feature of sociological dynamics is the possibility of ‘volume-dependent rate’ discussed in section 4.6, due to the rapid information exchange and government control in modern society.

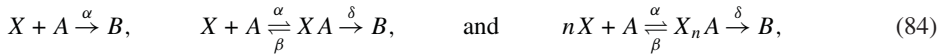
6.1. The three types of predation functional responses

In mathematical ecology [66], the predation functional response characterizes the rate of prey consumed as a function of the density of the prey population under a constant environment including the predators. There are three widely used types of functional responses. Let r be the rate and x be the prey population density, then the three types are

$$r_1(x) \propto x, \quad r_2(x) \propto \frac{x}{a+x}, \quad \text{and} \quad r_3(x) \propto \frac{x^n}{a^n + x^n} \quad (n > 1). \quad (83)$$

The most important distinction between type I and types II and III is that the latter have a saturation effect. When there is a sufficiently large population of prey, more than enough for all the predators, the rate of consumption of the prey levels off.

These three different types of functional responses in equation (83) can be precisely represented by the following three types of chemical reactions, respectively:



with the corresponding ODEs according to the law of mass action

$$\frac{dx}{dt} = -\alpha ax, \quad \frac{dx}{dt} = -\frac{\delta ax}{\left(\frac{\beta+\delta}{\alpha}\right) + x}, \quad \text{and} \quad \frac{dx}{dt} = -\frac{\delta ax^n}{\left(\frac{\beta+\delta}{\alpha}\right) + x^n}, \quad (85)$$

where a is the total concentration of molecular species A , which is kept constant in the reaction. The derivation for the expressions in equation (85) from chemical kinetic schemes in equation (84) involves singular perturbations and can be found in many enzyme kinetic textbooks [7].

It is interesting to note the correspondence between type III response and the molecular cooperativity. In ecological systems, type III response is associated with learning, that is, the natural improvement of a predator's searching and attacking efficiency as prey density increases.

6.2. Birth and death rates in multi-prey predation

Let us now consider the case of one predator population Y who has n different possible prey species X_i , $1 \leq i \leq n$. Let the consumption rate of X_i , per Y , in the absence of all the other X_j 's ($j \neq i$, $1 \leq j \leq n$) be type II functional response,

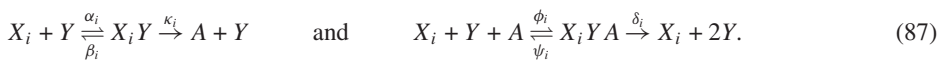
$$\frac{a_i x_i}{1 + a_i \tau_i x_i},$$

where a_i is the attack rate and τ_i is called handling time. In the presence of all the X_i , one then has ([67], section 7.2)

$$\frac{dx_i}{dt} = -x_i \left(\frac{a_i y}{1 + \sum_{j=1}^n a_j \tau_j x_j} \right) \quad \text{and} \quad \frac{dy}{dt} = y \left(\frac{\sum_{i=1}^n e_i a_i x_i}{1 + \sum_{i=1}^n a_i \tau_i x_i} \right), \quad (86)$$

where e_i is known as consumer efficiency. The first equation is the death rate of the prey population X_i caused by the predator, and the second equation is the birth rate of the predator with multiple preys.

In biochemical reaction terms, the predator Y is an autocatalytic enzyme which transforms the various X_i into Y , where



If we set the concentration of A , $a = \frac{\alpha_i}{\beta_i} \times \frac{\psi_i}{\phi_i}$, and assume the binding steps are in rapid equilibrium, then the chemical kinetic equations for the concentrations of X_i and Y are

$$\frac{dx_i}{dt} = -\frac{\kappa_i \left(\frac{\alpha_i}{\beta_i} \right) x_i y}{1 + \sum_{i=1}^n \frac{\alpha_i}{\beta_i} x_i} \quad \text{and} \quad \frac{dy}{dt} = y \left(\frac{\sum_{i=1}^n \delta_i \left(\frac{\alpha_i}{\beta_i} \right) x_i}{1 + \sum_{i=1}^n \frac{\alpha_i}{\beta_i} x_i} \right). \quad (88)$$

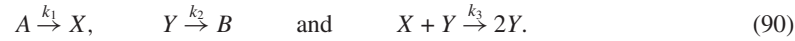
Comparing equation (88) with equation (86), we have

$$\frac{\alpha_i}{\beta_i} = a_i \tau_i, \quad \kappa_i = \frac{1}{\tau_i}, \quad \delta_i = e_i. \quad (89)$$

6.3. Four planar population systems

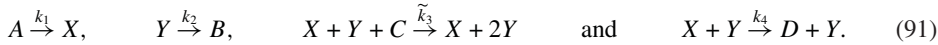
Here we consider the dynamics of two interacting populations is a planar nonlinear system and give the chemical kinetic equivalences of several well-known examples. Almost all textbooks on differential equations and in mathematical biology discuss such systems [29, 66].

Lotka–Volterra’s predator–prey model. The widely studied model for predator and prey dynamics was originally developed as a system of chemical reactions containing autocatalysis,



The corresponding ODEs from the law of mass action are given in equation (92) with $\tilde{k}_3 c = k_4 = k_3$. In ecological terms, X and Y are the prey and predator, respectively. The prey is the sole food of the predator, and it has a linear grow rate of $k_1 a$ in the absence of the predator. In the absence of the prey, the predator has a death rate of k_2 .

In an ecological context, there is no fundamental reason for the two xy terms in equation (92) to be equal. Hence, the faithful chemical reaction representation for the predator–prey model is

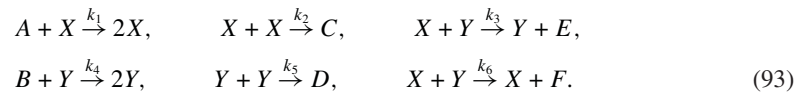


With the concentrations for chemical species X, Y, A, B, C being x, y, a, b, c , we have

$$\frac{dx}{dt} = (k_1 a) x - k_4 x y \quad \text{and} \quad \frac{dy}{dt} = -k_2 y + (\tilde{k}_3 c) x y. \quad (92)$$

See [68, 69] for a recent study of a generalization of the Lotka–Volterra system with a chemical perspective, which yields new insights to the classic problem.

Competition model. The second widely studied type of planar population dynamics involves a competition between two species. In chemical kinetic terms,

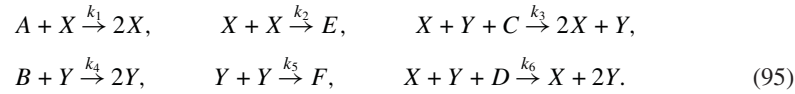


The mass-action kinetic equations for dynamical X and Y , with constant populations of A and B , are

$$\frac{dx}{dt} = (k_1 a) x - k_2 x^2 - k_3 x y \quad \text{and} \quad \frac{dy}{dt} = (k_4 b) y - k_5 y^2 - k_6 x y. \quad (94)$$

Both X and Y , in the absence of the competition, have logistic growth. Species X has a linear growth rate of $k_1 a$ and carrying capacity of $\frac{k_1 a}{k_2}$, and species Y has $k_4 b$ and $\frac{k_4 b}{k_5}$. Equation (94) is precisely the equations in section 3.5 of [29].

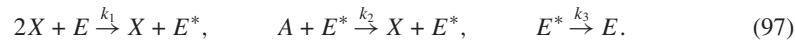
Mutualism or symbiosis. Murray [29] also presented a model for symbiosis in which species X and Y are in cooperation. In this case, the signs of the xy terms in equation (94) are positive rather than negative. The chemical reaction system that yields such dynamics is



The corresponding ODEs are

$$\frac{dx}{dt} = (k_1a)x - k_2x^2 + (k_3c)xy \quad \text{and} \quad \frac{dy}{dt} = (k_4b)y - k_5y^2 + (k_6d)xy. \quad (96)$$

Models of cannibalistic demography. A single population with cannibalism can be ‘modelled’ by the following chemical reaction system, known as energy relay [70],



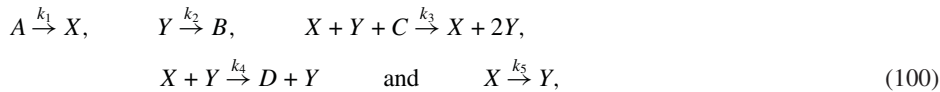
Let the concentrations for X and E^* be x and e^* , and the total E and E^* together is a constant e_t . Then, the law of mass action gives us

$$\frac{dx}{dt} = -k_1(e_t - e^*)x^2 + k_2ae^* \quad \text{and} \quad \frac{de^*}{dt} = k_1(e_t - e^*)x^2 - k_3e^*. \quad (98)$$

Treating E and E^* as an ‘enzyme’ following the Michaelis–Menten kinetics, we have

$$\frac{dx}{dt} = \frac{(k_1e_t)(k_2 - k_3)}{k_1x^2 + k_3}x^2. \quad (99)$$

A juveniles and adults two-age model with cannibalism of juveniles by adults [71] can be understood as a predator–prey system (adult as predator, juvenile as prey) with a population transfer from juveniles to adults, such that



in which X and Y are the juveniles and adults. Therefore, we have

$$\frac{dx}{dt} = (k_1a)x - k_4xy - k_5x \quad \text{and} \quad \frac{dy}{dt} = k_5x - k_2y + (\widetilde{k_3}c)xy. \quad (101)$$

7. Discussion and outlook

7.1. Nonlinear, stochastic biochemical dynamics as a new paradigm

Currently, there are mainly two mathematical approaches to biological systems and phenomena. One is based on *principles and mechanisms* and the other is based on *data*. Research on protein molecular dynamics and on the Hodgkin–Huxley equation of excitable cells belongs to the first kind, while statistical research on bioinformatics, ecology, economics, etc belong to the second kind. In between, there are *modellers* who struggle to develop mathematical principles from the data. Mathematical biologists are a large group of modellers. Accordingly, modelling of biological dynamics has been roughly divided into *deterministic* and *statistical* approaches. The study of *cellular biochemical dynamics*, as a paradigm, offers a new perspective on biological dynamics.

The *nonlinear, stochastic cellular biochemical dynamics* offers a new mathematical framework for dynamics that encompasses both deterministic and statistical aspects of

modelling. The benefits go further than this. Perhaps one of the most important insights is the emergence of *rare events* which has infinitesimal probability to occur in a regular time scale, but it will occur with probability 1 on an evolutionary time scale. Rare events can be understood by neither classical deterministic mathematics nor normal statistics. The only tool we know of is mechanistic stochastic modelling.

Cancers, ecological catastrophes, stock market crashes, and sociopolitical revolutions are all rare events. It is these rare events that are truly *unpredictable* in the classical sense, giving the appearance of *free will* [16]. John Hopfield called it *dynamic symmetry breaking*. James Clerk Maxwell has said, ‘It is manifest that the existence of unstable conditions renders impossible the prediction of future events, if our knowledge of the present state is only approximate, and not accurate. At these (unstable) points, influences whose physical magnitude is too small to be taken account of by a finite being, may produce results of the greatest importance. All great results produced by human endeavour depend on taking advantage of these singular states when they occur.’ [17]

7.2. Stochastic dynamics in terms of multi-dimensional BDPs and diffusion processes

Stochastic processes have gradually become an indispensable and powerful mathematical description of biological dynamics, from cellular biochemical to ecological systems. Yet, compared with the understanding of nonlinear deterministic dynamical systems, our current in-depth knowledge of applied stochastic processes are still rather limited. This is particularly true for stochastic processes with time-irreversibility. Time-reversible processes are appropriate models for equilibrium dynamics. Stochastic dynamics of living systems have to be time-irreversible [33, 34, 48].

The interaction between the stochastic aspect and the nonlinear aspect of dynamics creates complex behaviour. This is a subject that is yet to be fully explored. Markov processes, the stochastic counterpart of first-order ordinary differential equations, have two equally valid mathematical representations, the trajectories and the Kolmogorov forward equations. For diffusion processes that have been widely employed in physics and chemistry, these correspond to SDEs and the FPE [72]. For multi-dimensional BDPs, they correspond to the Lebowitz–Gillespie algorithm [2, 73] and master equations, respectively.

The Delbrück–Gillespie process is the stochastic counterpart of the deterministic mass-action kinetics. It is a full range analytical theory of dynamics of homogeneous chemical and biochemical reaction systems. It is more than either the wildly popular Gillespie algorithm or the CME alone. It has an emergent nonlinear differential equation system as well as the emerging stochastic jump dynamics on an evolutionary time scale [5].

Since the pioneering work of Einstein, Smoluchowski, Langevin and Kramers, the diffusion process, with its continuous but everywhere non-differentiable trajectory, has become the dominant mathematical theory for stochastic processes. The physicists’ approach to stochastic dynamics, however, is markedly ‘deterministic centric.’ The entire stochastic enterprise of statistical physics is to understand macroscopic, deterministic behaviour from the atomic nature of matters.

The stochasticity has always been considered merely as ‘fluctuations.’ In fact, physicists have long believed that there is no stochasticity in a macroscopic world. This view, of course, has been justified by the law of large number in the theory of probability. This perspective, as we have shown in the present paper, needs to be modified to embrace a macroscopic complexity with variations and stochastic jumps [16]. The law of large numbers, it turns out, requires an infinitely long time and large system. For mesoscopic systems [74], there are stochastic

dynamics beyond the deterministic limit [15, 19], and for macroscopic systems, stochasticity occurs on an evolutionary time scale.

Finally, in biological and many other non-mechanical systems, a BDP is a more fundamental approach to stochastic dynamics than the continuous diffusion. Certainly, it is more consistent with all the deterministic differential equation models based on number density, be it individuals, organisms, cells or molecules. The diffusion approach, however, is only phenomenological. It has to rely on additional information to specify the diffusion coefficient. This is why the fluctuation–dissipation relations are so essential in equilibrium statistical physics.

7.3. Intrinsic and extrinsic fluctuations: stochastic processes versus random dynamical systems (RDSs)

According to van Kampen [24], ‘internal or intrinsic fluctuations are caused by the fact that the system itself consists of discrete particles; it is inherent in the very mechanism by which the system evolves.’ External or extrinsic noises, on the other hand, often reside in systems’ parameters or environments. The distinction between these two types of randomness in a dynamic process can be best illustrated if one considers two trajectories with different initial conditions, and asks if the two dynamics utilize two different, independent sequences of realizations of random events, or the same sequence. In the theory of RDSs [75], it is the latter. A SDE can be interpreted as both. However, a Delbrück–Gillespie (DG) process does not fit the RDS perspective. The random process underlying a DG process is a time-changed Poisson process. To consider two DG processes with same realization of a Poisson process can only come from a globally synchronized clock [5]. Therefore, the fluctuations in the DG process are due to intrinsic noise. There is no separation between the deterministic nonlinear dynamics and the stochastic fluctuations.

Acknowledgments

I thank Professors R E O’Malley and W J Heuett for carefully reading the manuscript, and Professor O’Malley for numerous discussions. Collaborations with Drs Jia-zeng Wang and Da Zhou of Peking University are highly appreciated.

Appendix A. Conditional diffusion equation

Let us consider the SDE

$$dX_t = b(X_t) dt + \epsilon \sqrt{A(X_t)} dB_t, \quad X_0 = x_0. \quad (\text{A.1})$$

Let $\xi(t)$ be the deterministic solution to the ODE

$$\frac{d\xi(t)}{dt} = b(\xi(t)), \quad \xi(0) = x_0. \quad (\text{A.2})$$

Then we can consider

$$Y_t = \frac{X_t - \xi(t)}{\epsilon}, \quad Y_0 = 0. \quad (\text{A.3})$$

which satisfies the SDE

$$dY_t = \frac{b(\epsilon Y_t + \xi(t)) - b(\xi(t))}{\epsilon} dt + \sqrt{A(\epsilon Y_t + \xi(t))} dB_t. \quad (\text{A.4})$$

The probability density function for Y_t satisfies the time-inhomogeneous FPE

$$\frac{\partial}{\partial t} f_Y(y, t) = \frac{1}{2} \frac{\partial^2}{\partial y^2} (A(\epsilon y + \xi(t)) f_Y) - \frac{\partial}{\partial y} \left(\frac{b(\epsilon y + \xi(t)) - b(\xi(t))}{\epsilon} f_Y \right). \quad (\text{A.5})$$

For small ϵ , equation (A.5) can be approximated as

$$\frac{\partial}{\partial t} f_Y(y, t) = \frac{A(\xi(t))}{2} \frac{\partial^2 f_Y}{\partial y^2} - b'(\xi(t)) \frac{\partial (y f_Y)}{\partial y}. \quad (\text{A.6})$$

In particular, if the deterministic solution to (A.2) asymptotically approaches a fixed point x^* , and if we choose $X_0 = x^*$, then equation (A.6) is simplified to the FPE for an Ornstein–Uhlenbeck Gaussian process,

$$\frac{\partial}{\partial t} f_Y(y, t) = \frac{A(x^*)}{2} \frac{\partial^2 f_Y}{\partial y^2} - b'(x^*) \frac{\partial (y f_Y)}{\partial y}. \quad (\text{A.7})$$

Appendix A.1. Gaussian solution to the conditional diffusion

To solve equation (A.6), we introduce

$$\langle Y_t \rangle = \int_{-\infty}^{\infty} y f_Y(y, t) dy, \quad \langle Y_t^2 \rangle = \int_{-\infty}^{\infty} y^2 f_Y(y, t) dy, \quad (\text{A.8})$$

then we have

$$\begin{aligned} \frac{d}{dt} \langle Y_t \rangle &= b'(\xi(t)) \langle Y_t \rangle, & \langle Y_0 \rangle &= 0, \\ \frac{d}{dt} \langle Y_t^2 \rangle &= A(\xi(t)) + 2b'(\xi(t)) \langle Y_t^2 \rangle, & \langle Y_0^2 \rangle &= 0, \\ \langle Y_t \rangle &= 0, \\ \langle Y_t^2 \rangle &= \begin{cases} b^2(\xi(t)) \int_0^t \frac{a(\xi(s))}{b^2(\xi(s))} ds, & \xi'(t) \neq 0 \\ \frac{A(x^*)}{2|b'(x^*)|} (1 - e^{-2|b'(x^*)|t}), & \xi(t) \equiv x^*, b'(x^*) < 0. \end{cases} \end{aligned} \quad (\text{A.9})$$

It is easy to verify that the solution to equation (A.6) is

$$f_Y(y, t) = \frac{1}{\sqrt{2\pi \langle Y_t^2 \rangle}} \exp \left(-\frac{y^2}{2 \langle Y_t^2 \rangle} \right). \quad (\text{A.10})$$

Appendix A.2. Linear fluctuation theory according to conditional diffusion

If the $\xi(t)$ is near a stable fixed point of the $b(x)$, then $\xi(t|x_0) = x^* + (x_0 - x^*)e^{-\beta t}$ where $\beta = |b'(x^*)|$. Furthermore, $b(\xi(t)) \approx -\beta(x_0 - x^*)e^{-\beta t}$. Then equation (A.9) gives

$$\langle Y_t^2 \rangle = \frac{A(x^*)}{2\beta} (1 - e^{-2\beta t}). \quad (\text{A.11})$$

And the autocorrelation function for the stationary process is

$$\begin{aligned} \langle X_\tau X_0 \rangle^{\text{st}} &= \int_{-\infty}^{\infty} dx_0 f^{\text{st}}(x_0) x_0 \xi(\tau|x_0) \\ &= \langle X \rangle^{\text{st}} x^* (1 - e^{-\beta \tau}) + \langle X^2 \rangle^{\text{st}} e^{-\beta \tau}. \end{aligned} \quad (\text{A.12})$$

Therefore,

$$\langle \Delta X_\tau \Delta X_0 \rangle^{\text{st}} = \langle (\Delta X)^2 \rangle^{\text{st}} e^{-\beta \tau}. \quad (\text{A.13})$$

These results have been obtained many times in the theory of stochastic linear relaxation, in the work of L Onsager, M Lax, T L Hill and J Keizer. We see that the conditional diffusion shares the same principle. The stochastic dynamics is a ‘correction term’ to the deterministic behaviour.

Appendix B. The mean first passage time (MFPT)

Appendix B.1. MFPT for a 1D diffusion process

The MFPT T for a diffusion process with diffusion coefficient $A(x)/2$ and drift $b(x)$, from x_1 to x_2 , satisfies [10, 14, 24]

$$\frac{d}{dx} \frac{A(x)}{2} \frac{dT(x)}{dx} + b(x) \frac{dT(x)}{dx} = -1, \quad \frac{dT(x_1)}{dx} = 0, \quad T(x_2) = 0, \quad (\text{B.1})$$

$$T_{x_1 \rightarrow x_2} = \int_{x_1}^{x_2} e^{-\phi(x)} dx \int_x^{x_2} e^{\phi(y)} \frac{2dy}{A(y)}, \quad \phi(x) = - \int^x \frac{2b(x)}{A(x)} dx. \quad (\text{B.2})$$

If interval (x_1, x_2) contains an energy well at x_1^* , $b(x_1^*) = 0$, $b'(x_1^*) < 0$, and an energy barrier at x_3^* , $b(x_3^*) = 0$, $b'(x_3^*) > 0$, then one can use Laplace’s method to simplify equation (B.2) to

$$T_{x_1 \rightarrow x_2} = \frac{4\pi e^{\phi(x_3^*) - \phi(x_1^*)}}{A(x_3^*) \sqrt{|\phi''(x_1^*)\phi''(x_3^*)|}} + \int_{x_3^*}^{x_2} \frac{2 dx}{A(x)|\phi'(x)|}. \quad (\text{B.3})$$

The second term is for downhill relaxation. In fact, $A(x)|\phi'(x)|/2 = |b(x)|$. Thus the downhill time is essentially determined by the drift $b(x)$, independent of $A(x)$. The first term is the time for barrier crossing. The inverse of this expression is known as Kramers formulae for reaction rate. For a barrier crossing problem, the second term can be neglected.

Appendix B.2. MFPT for a 1D BDP

A same calculation can be carried out for a BDP [10]. For a one-dimensional CME with birth rate μ_n and death rate λ_n , in the limit of large V , $\mu_{xV} \rightarrow \mu(x)V$, $\lambda_{xV} \rightarrow \lambda(x)V$, and one obtains [64]

$$T_{x_1 \rightarrow x_2} = \frac{2\pi e^{V(\phi(x_3^*) - \phi(x_1^*))}}{\lambda(x_3^*) \sqrt{|\phi''(x_1^*)\phi''(x_3^*)|}} + \int_{x_3^*}^{x_2} \frac{1}{\lambda(y)|\phi'(y)|} dy, \quad (\text{B.4})$$

in which

$$\phi(x) = \int^x \ln \left[\frac{\lambda(z)}{\mu(z)} \right] dz. \quad (\text{B.5})$$

Note that at fixed point x_3^* , $\lambda(x_3^*) = \mu(x_3^*)$, which corresponds to $A(x_3^*)/2$. Hence the first terms in equations (B.3) and (B.4) are completely identical. The detailed $\phi(x)$ and $A(x)$ between x_1^* and x_3^* do not matter to the barrier crossing time.

The downhill time, however, is significantly different from that of diffusion theory. Using either $A(x) = \mu(x) + \lambda(x)$, as in the Kramers–Moyal expansion, or $A(x) = 2(\mu(x) - \lambda(x))/(\ln \mu(x) - \ln \lambda(x))$, as in Onsager’s theory, gives the same result in equation (B.7),

$$\text{diffusion : } \int \frac{dx}{\mu(x) - \lambda(x)}, \quad (\text{B.6})$$

$$\text{CME : } \int \frac{dx}{\lambda(x) (\ln \mu(x) - \ln \lambda(x))}. \quad (\text{B.7})$$

These are the different predictions based on the diffusion theory and on the CME.

Appendix C. Master equations with and without diffusion limit

Let us consider the canonical FPE for diffusion processes,

$$\frac{\partial f}{\partial t} = \frac{\epsilon}{2} \frac{\partial^2}{\partial x^2} (A(x)f) - \frac{\partial}{\partial x} (b(x)f). \quad (\text{C.1})$$

If we discretize the x in terms of a uniform interval δ , we have

$$\begin{aligned} \frac{df(x, t)}{dt} = & \left[\frac{\epsilon A(x - \delta)}{2\delta^2} + \frac{b(x - \delta)}{2\delta} \right] f(x - \delta) \\ & - \left[\left(\frac{\epsilon A(x)}{2\delta^2} + \frac{b(x)}{2\delta} \right) + \left(\frac{\epsilon A(x)}{2\delta^2} - \frac{b(x)}{2\delta} \right) \right] f(x) \\ & + \left[\frac{\epsilon A(x + \delta)}{2\delta^2} - \frac{b(x + \delta)}{2\delta} \right] f(x + \delta). \end{aligned} \quad (\text{C.2})$$

Therefore, if the birth and death rates of a master equation, $\mu_n(V)$ and $\lambda_n(V)$, are in the forms of

$$\mu_n = V \left[\frac{\epsilon A(n/V)}{2\delta} + \frac{b(n/V)}{2} \right] \quad \text{and} \quad \lambda_n = V \left[\frac{\epsilon A(n/V)}{2\delta} - \frac{b(n/V)}{2} \right], \quad (\text{C.3})$$

then we have the master equation in (6). Note that both μ_n and λ_n have to be non-negative. This is guaranteed by the δ being sufficiently small. We see that if ϵ is smaller, the δ has to be smaller as well.

The stationary distribution of the master equation is readily obtained, such that

$$p_n^{\text{st}} = p_0 \exp \left[\sum_{k=1}^n \ln \frac{\mu_{k-1}}{\lambda_k} \right] \quad (\text{C.4})$$

and

$$\frac{f^{\text{st}}(x)}{V} = p_0 \lim_{V \rightarrow \infty} \exp \left[\sum_{k=1}^{xV} \ln \frac{A\left(\frac{k-1}{V}\right) \frac{\epsilon}{\delta} + b\left(\frac{k-1}{V}\right)}{A\left(\frac{k}{V}\right) \frac{\epsilon}{\delta} - b\left(\frac{k}{V}\right)} \right], \quad (\text{C.5})$$

which yields

$$f^{\text{st}}(x) dx = A \exp \left[V \int_0^x \ln \frac{(\epsilon/\delta)A(x) + b(x)}{(\epsilon/\delta)A(x) - b(x)} dz \right]. \quad (\text{C.6})$$

Note that if we consider ϵ and $\delta \rightarrow 0$, but $\epsilon/\delta \rightarrow \nu$, then we have

$$f^{\text{st}}(x) dx = A \exp \left[-\frac{\nu}{\epsilon} \int_0^x \ln \frac{\nu A(x) - b(x)}{\nu A(x) + b(x)} dz \right]. \quad (\text{C.7})$$

The $\ln f^{\text{st}}(x)$ has its extrema at x^* with $b(x^*) = 0$. Furthermore, the curvature at x^* is $2b'(x^*)/(\epsilon A(x^*))$, which is independent of ν .

More interestingly, if $\nu = 1$, we have

$$f^{\text{st}}(x) dx = A \exp \left[-\frac{1}{\epsilon} \int_0^x \ln \frac{A(x) - b(x)}{A(x) + b(x)} dz \right]. \quad (\text{C.8})$$

This is the case of the CME in which μ_n and λ_n have the form of $\mu_n = \mu(x)V$ and $\lambda_n = \lambda(x)V$ as in equation (C.3).

If, however, $\nu = \infty$,

$$f^{\text{st}}(x) dx = A \exp \left[\frac{1}{\epsilon} \int_0^x \frac{2b(z)}{A(z)} dz \right]. \quad (\text{C.9})$$

Only equation (C.9) recovers the correct stationary distribution to the FPE in (C.1). One can not have $\nu \rightarrow 0$ because of the discussion following equation (C.3).

Equations (C.9) and (C.8) are precisely the correct and wrong stationary distributions according to [56] and Kramers–Moyal’s diffusion equations, respectively. Their difference is exactly the two expressions on the lhs and rhs of equation (63).

References

- [1] Samoilov M S and Arkin A P 2006 *Nature Biotechnol.* **24** 1235
- [2] Gillespie D T 2007 *Annu. Rev. Phys. Chem.* **58** 35
- [3] Liang J and Qian H 2010 *J. Comput. Sci. Technol.* **25** 154
- [4] Qian H and Bishop L M 2010 *Int. J. Mol. Sci.* **11** 3472
- [5] Qian H 2010 *J. Stat. Phys.* **141** 990
- [6] Wilkinson D J 2006 *Stochastic Modeling for Systems Biology* (New York: Chapman & Hall/CRC)
- [7] Beard D A and Qian H 2008 *Chemical Biophysics: Quantitative Analysis of Cellular Systems* (Cambridge: Cambridge University Press)
- [8] Kurtz T G 1971 *J. Appl. Probab.* **8** 344
- [9] Qian H, Saffarian S and Elson E L 2002 *Proc. Natl Acad. Sci. USA* **99** 10376
- [10] Gardiner C W 1983 *A Handbook of Stochastic Methods: for Physics, Chemistry and the Natural Sciences* (Berlin: Springer)
- [11] Kurtz T G 1981 *Approximation of Population Processes* (Cambridge: Cambridge University Press)
- [12] Stroock D W and Varadhan S R S 1979 *Multidimensional Diffusion Processes* (Berlin: Springer)
- [13] Wax N 1954 *Selected Papers on Noise and Stochastic Processes* (New York: Dover)
- [14] Schuss Z 2010 *Theory and Applications of Stochastic Processes: An Analytical Approach* (New York: Springer)
- [15] Qian H, Shi P-Z and Xing J 2009 *Phys. Chem. Chem. Phys.* **11** 4861
- [16] Hopfield J J 1994 *J. Theor. Biol.* **171** 53.
- [17] Prigogine I and Stengers I 1984 *Order Out of Chaos* (Boulder, CO: New Sci. Lib. Shambhala Pub.)
- [18] Hänggi P, Talkner P and Borkovec M 1990 *Rev. Mod. Phys.* **62** 251
- [19] Ge H and Qian H 2011 *J. R. Soc. Interface* **8** 107
- [20] Landauer R and Swanson J A 1961 *Phys. Rev.* **121** 1668
- [21] Langer J S 1969 *Ann. Phys.* **54** 258
- [22] Vellela M and Qian H 2007 *Bull. Math. Biol.* **69** 1727
- [23] van Kampen N G 1961 *Can. J. Phys.* **39** 551
- [24] van Kampen N G 1992 *Stochastic Processes in Physics and Chemistry* (Amsterdam: North-Holland/Elsevier Science)
- [25] Schlögl F 1972 *Z. Physik.* **253** 147
- [26] Vellela M and Qian H 2009 *J. R. Soc. Interf.* **6** 925
- [27] Bishop L M and Qian H 2010 *Biophys. J.* **98** 1
- [28] Shi P-Z and Qian H 2011 *J. Chem. Phys.* **134** 065104
- [29] Murray J D 2004 *Mathematical Biology I: An Introduction* (New York: Springer)
- [30] Epstein I R and Pojman J A 1998 *An Introduction to Nonlinear Chemical Dynamics: Oscillations, Waves, Patterns, and Chaos* (Oxford: Oxford University Press)
- [31] Qian H and Reluga T C 2005 *Phys. Rev. Lett.* **94** 028101
- [32] Ge H and Qian H 2009 *Phys. Rev. Lett.* **103** 148103
- [33] Qian H 2006 *J. Phy. Chem. B* **110** 15063 (feature article)
- [34] Qian H 2007 *Annu. Rev. Phys. Chem.* **58** 113
- [35] Keizer J 1986 *Probability, Statistical Mechanics, and Number Theory: A Volume Dedicated to Mark Kac (Advances in Mathematics Supplementary Studies vol 9)* ed G-C Rota (New York: Academic) p 1
- [36] Hu G 1986 *Zeit. Phys. B* **65** 103
- [37] Ge H and Qian H 2010 arxiv.1011.4049
- [38] Onsager L and Machlup S 1953 *Phys. Rev.* **91** 1505
- [39] Lax M 1960 *Rev. Mod. Phys.* **32** 25
- [40] Keizer J 1979 *Acc. Chem. Res.* **12** 243
- [41] Fox R F 1978 *Phys. Rep.* **48** 180
- [42] Qian H 2001 *Proc. R. Soc. A* **457** 1645
- [43] Kwon C, Ao P and Thouless D J 2005 *Proc. Natl Acad. Sci. USA* **102** 13029
- [44] Gajic Z and Qureshi M T J 2008 *Lyapunov Matrix Equation in System Stability and Control* (New York: Dover)

- [45] Jiang D-Q, Qian H and Wang J-Z 2011 in preparation
- [46] Jiang D-Q and Zhang F-X 2003 *J. Math. Phys.* **44** 4681
- [47] Chen Y, Qian M-P and Xie J-S 2007 *J. Math. Phys.* **48** 073301
- [48] Jiang D-Q, Qian M and Qian M-P 2004 *Mathematical Theory of Nonequilibrium Steady States: On the Frontier of Probability and Dynamical Systems* (New York: Springer)
- [49] Qian H and Elson E L 2004 *Proc. Natl Acad. Sci. USA* **101** 2828
- [50] Zwanzig R 1965 *Annu. Rev. Phys. Chem.* **16** 67
- [51] Qian H and Qian M 2000 *Phys. Rev. Lett.* **84** 2271
- [52] Qian M, Wang G-X and Zhang X-J 2000 *Phys. Rev. E* **62** 6469
- [53] Qian M and Zhang X-J 2001 *Phys. Rev. E* **65** 011101
- [54] Tan W Y 2000 *Stochastic Modeling of AIDS Epidemiology and HIV Pathogenesis* (Singapore: World Scientific)
- [55] Fleming W H 1975 *Suppl. Adv. Appl. Probab.* **7** 100
- [56] Hänggi P, Grabert H, Talkner P and Thomas H 1984 *Phys. Rev. A* **29** 371
- [57] Baras F, Malek-Mansour M and Pearson J E 1996 *J. Chem. Phys.* **105** 8257
- [58] Zhou D and Qian H 2011 arxiv.1101.2945
- [59] Feller W 1957 *An Introduction to Probability Theory and Its Applications* vol 1, 2nd edn (New York: Wiley)
- [60] O'Malley R E 2008 *SIAM Rev.* **50** 459
- [61] Touchette H 2009 *Phys. Rep.* **478** 1
- [62] McQuarrie D A 1967 *J. Appl. Probab.* **4** 413
- [63] Ewens W J 2004 *Mathematical Population Genetics: I. Theoretical Introduction* 2nd edn (New York: Springer)
- [64] Zhang Y, Ge H and Qian H 2010 arxiv.1011.2554
- [65] Wang Z-J, Qian M and Qian H 2011 in preparation
- [66] Kot M 2001 *Elements of Mathematical Ecology* (Cambridge: Cambridge University Press)
- [67] Dercote F and Rinaldi S 2008 *Analysis of Evolutionary Processes: The adaptive dynamics approach and its applications* (Princeton, NJ: Princeton University Press)
- [68] Li Y, Qian H and Yi Y 2008 *J. Chem. Phys.* **129** 154505
- [69] Li Y, Qian H and Yi Y 2010 *J. Dyn. Diff. Eqns* **22** 491
- [70] Qian H 2010 *J. Phys. Chem. B* **114** 16105
- [71] Cushing J M 1991 *Math. Biosci.* **107** 47
- [72] Risken H 1989 *The Fokker-Planck Equation: Methods of Solutions and Applications* 2nd edn (New York: Springer)
- [73] Bortz A B, Kalos M H and Lebowitz J L 1975 *J. Comput. Phys.* **17** 10
- [74] Laughlin R B, Pines D, Schmalian J, Stojković B P and Wolynes P G 2000 *Proc. Natl Acad. Sci. USA* **97** 32
- [75] Arnold L 1998 *Random Dynamical Systems* (New York: Springer)

Cellular Biology in Terms of Stochastic Nonlinear Biochemical Dynamics: Emergent Properties, Isogenetic Variations and Chemical System Inheritability

Hong Qian

Received: 5 October 2010 / Accepted: 1 November 2010 / Published online: 9 November 2010
© Springer Science+Business Media, LLC 2010

Abstract Based on a stochastic, nonlinear, open biochemical reaction system perspective, we present an analytical theory for cellular biochemical processes. The chemical master equation (CME) approach provides a unifying mathematical framework for cellular modeling. We apply this theory to both self-regulating gene networks and phosphorylation-dephosphorylation signaling modules with feedbacks. Two types of bistability are illustrated in mesoscopic biochemical systems: one that has a macroscopic, deterministic counterpart and another that does not. In certain cases, the latter stochastic bistability is shown to be a “ghost” of the extinction phenomenon. We argue the thermal fluctuations inherent in molecular processes do not disappear in mesoscopic cell-sized nonlinear systems; rather they manifest themselves as isogenetic variations on a different time scale. Isogenetic biochemical variations in terms of the stochastic attractors can have extremely long lifetime. Transitions among discrete stochastic attractors spend most of the time in “waiting”, exhibit punctuated equilibria. It can be naturally passed to “daughter cells” via a simple growth and division process. The CME system follows a set of nonequilibrium thermodynamic laws that include non-increasing free energy $F(t)$ with external energy drive $Q_{hk} \geq 0$, and total entropy production rate $e_p = -dF/dt + Q_{hk} \geq 0$. In the thermodynamic limit, with a system’s size being infinitely large, the nonlinear bistability in the CME exhibits many of the characteristics of macroscopic equilibrium phase transition.

Keywords Biochemistry · Cell biology · Chemical master equation · Evolution · Nonequilibrium · Nonlinear dynamics · Stochastic processes · Thermodynamics

1 Introduction

From an evolutionary biology standpoint, Kirschner and Gerhart have argued that a central task of cellular and organismal biology is to provide phenotypic variations with molecular mechanisms that connect genome to life [1]. Molecular mechanisms demystify the

H. Qian (✉)
Department of Applied Mathematics, University of Washington, Seattle, WA 98195, USA
e-mail: qian@amath.washington.edu

biological variations upon which Darwin's natural selection occurs, thus giving the "plausibility of life".

Biochemistry and molecular genetics/genomics are the two foundations of cellular molecular biology [2]. According to the Modern Synthesis School of population genetics [3] and its genomic interpretations [4], the molecular basis of biological variations is coded in the DNA sequence, which is inheritable via Mendelian genetics and Watson-Crick base-pairing. Biochemistry, on the other hand, has long been considered as a deterministic mechanics that executes the instructions in the DNA [5, 6].

Continuous theoretical investigations [7, 8] and recent experimental demonstrations of *stochastic gene expression* in single cells, however, have transformed the genomic monopoly of biological variations [9, 10]. Stochasticity has risen rapidly to prominence in cellular molecular biology [11–13]. Isogenetic biochemical variations are now widely considered as mechanisms for "novelty" in cellular processes ranging from cell differentiation to oncogenesis [14–16].

It is against this backdrop that statistical physics and physical chemistry have a defining role to play in complementing the yet descriptive cellular molecular biology with a first-principle based analytical theory. Stochastic fluctuations in atomic and molecular processes have been the rule in our fields; it is the emergent macroscopic deterministic behavior that begs for explanations. However, as we shall show, there are emergent biochemical variations from stochastic molecular systems with nonlinearity. Stochastic variations do not disappear in mesoscopic molecular systems; they simply emerge on a different, much longer "evolutionary" time scale [17, 18].

We shall also discuss statistical thermodynamics. By thermodynamics, we mean one is interested in a system's organizing properties such as entropy and energy, and their interrelations. It turns out, thermodynamics, at least the isothermal part, is a general mathematical law of any stochastic system endowed with a Markovian dynamics [19]. The concept of "entropy" in the classical thermodynamics was defined empirically via the *quasi-static process*. Therefore, there is a feeling that one can only work with this concept in systems at, or near, equilibrium. As we shall demonstrate, however, that one can introduce a mathematical concept of entropy for *any* stochastic dynamics that follows a Markov process. Therefore, the Gibbs entropy, the relative entropy, and their time derivatives (see (36)–(38)) can all be defined for a system at any give time, near or far from equilibrium, in stationarity or in a transient. When applying the "thermodynamics" to biological systems, the real question is the validity of a stochastic Markovian description of the Nature. If that is valid, so is the application of the thermodynamic and the entropy theory [20].

In the approach we take in the present work, the origin of the stochasticity is due to the "intrinsic noise" of molecular collisions [9]. It is interesting to point out that a distinction between intrinsic and extrinsic noises can be made if one considers *simultaneously* two stochastic trajectories with *different* initial conditions. The former assumes the two with different stochastic realizations, while the latter assumes the two with a same realization. The "extrinsic noise", thus, is more consistent with the *random dynamical systems* (RDS) approach which regards the origin of temporal stochasticity being in a system's parameters [21], while intrinsic noise is more consistent with the *stochastic processes* approach which considers one initial condition at a time. See [22] for an example of applying RDS models in neural biology. While a stochastic differential equation with multiplicative Brownian motion can have both interpretations, a birth-and-death process can not have a RDS interpretation.¹

¹This statement is not strictly true: In the random time-change Poisson representation of a birth-and-death process, one can have a RDS interpretation. However, this seems to pin the stochasticity entirely on the flow of a global time.

2 Stochastic Chemical Dynamics

Two papers published in 1940 have laid the theoretical foundation of stochastic chemical dynamics that connects statistical physics to cellular biology.

H.A. Kramers' paper "Brownian motion in a field of force and the diffusion model of chemical reactions", which was published in that April [23], has shown us how to compute the rate constant for a discrete, individual chemical reaction in aqueous solution, such as $X + Y \rightarrow Z$ or $A \rightarrow B$, in terms of atomic coordinates and molecular energy functions. In a nutshell, Kramers' theory connected chemistry to physics by understanding chemical reactions using the mechanics of molecular particles and their interactions. This approach, together with Smoluchowski's earlier work on diffusion-controlled chemical reactions and their later synthesis [24, 25], is now one of the main areas of theoretical chemistry [26].

Figure 1 shows the stochastic transitions between two conformations A and B of a single molecule in terms of an energy function. Kramers' theory predicted that the rate constant follows the Arrhenius' law $k_1 \propto e^{-\Delta G^\ddagger/k_B T}$. More importantly, the reaction time is spent in waiting, which is random and exponentially distributed, while the actual transition time is instantaneous. This feature is general for any "barrier crossing" process in stochastic, nonlinear systems. Stochastic trajectories of single molecule conformational transitions under room temperature were not observed experimentally until 1990s [27].

In the same year, the January of 1940, Max Delbrück published an equally ground-breaking paper, though with much less fanfare in the subsequent fifty years [28]. While Kramers' Brownian motion is in the "configuration space", Delbrück's birth-and-death process is in the "copy number space". This theory, now under the name of the *chemical master equation* (CME) and more popular Gillespie algorithm [29, 30], has recently emerged as a main workhorse in computational systems biology [31]. In a nutshell, the CME connected cell biology to chemistry by understanding cellular phenotypes and their evolutions in terms of nonlinear biochemical networks in a mesoscopic reaction volume on the order of hundred femtolitres, the size of a cell ($1 \text{ femtolitre} = 1 \mu\text{m}^3$).

Kramers' theory and the CME clearly marked two complementary domains of physical chemistry. The former computes the rate constant of a individual chemical reaction based on the molecular structures, energy functions, and the solvent environment, while the latter describes the dynamic behavior of a chemical reaction *system*, assuming that the rate constants are given for each and every reaction within.

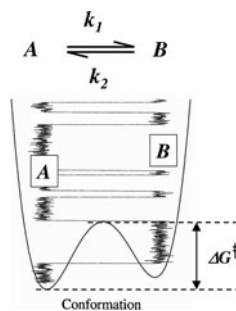
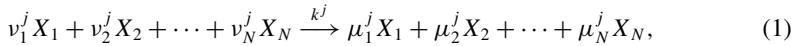


Fig. 1 Kramers 1940 theory connects the chemical reaction kinetics to the stochastic motions crossing an energy barrier. It is shown that the reaction rate constant $k_1 \propto e^{-\Delta G^\ddagger/k_B T}$. The stochastic dynamics spends most of the time in "waiting" while the actual transition is instantaneous. The waiting times are random and exponentially distributed. Barrier-crossing is a generic feature of any stochastic, nonlinear dynamical system

2.1 The Chemical Master Equation (CME)

Birth-and-death processes, to which the CME belongs, are a very special class of discrete state, continuous time, Markov processes. The discrete states are non-negative integers forming a lattice \mathbb{Z}^N . Consider a system of N chemical species X_i ($i = 1, 2, \dots, N$) with M chemical reactions, with the j th chemical reaction being represented by a set of stoichiometric coefficients ν_i^j and μ_i^j (the superscript being reaction and the subscript being species):



in which some of the integers ν 's and μ 's can be zero ($j = 1, 2, \dots, M$). If $\nu_\ell^j = \mu_\ell^j \neq 0$ for a particular ℓ , the corresponding X_ℓ is called a catalyst for the reaction j . If $\mu_\ell^j > \nu_\ell^j > 0$, then X_ℓ is an autocatalyst.

The state of the chemical reaction system at time t is characterized by the set of N integer $\mathbf{n}(t) = (n_1(t), n_2(t), \dots, n_N(t))$, i.e., a grid point on the \mathbb{Z}^N lattice, which specifies the copy number of X_i being $n_i(t)$. The dynamic of the chemical reaction system, thus, is represented by a trajectory in the copy number space \mathbb{Z}^N . The stochastic dynamics, according to Lebowitz-Gillespie's algorithm [30, 32], runs as follows: each of the M reactions by itself, say reaction j , can occur at a random time T^j , very much like the radioactive decay, which follows an exponential distribution

$$f_{T^j}(t) = \lambda^j e^{-\lambda^j t}, \quad (2)$$

with the rate

$$\lambda^j = V k^j \prod_{i=1}^N \frac{n_i(n_i - 1) \cdots (n_i - \nu_i^j + 1)}{V^{\nu_i^j}}. \quad (3)$$

The parameter V in (3) stands for the volume of the reaction system. It converts molecular copy number n_i to concentration n_i/V for species X_i ; all the rate constants in (1) are concentration based. Now with the presence of all M reactions in the system, the first one to occur is at the random time

$$T^* = \min\{T^1, T^2, \dots, T^M\}, \quad (4)$$

which follows again an exponential distribution

$$f_{T^*}(t) = \lambda^* e^{-\lambda^* t}, \quad \lambda^* = \sum_{j=1}^M \lambda^j. \quad (5)$$

The T^* determines the time for the system to move away from the current grid point. The system then moves randomly to one of the M grid points

$$(n_1, n_2, \dots, n_N) + (\mu_1^j - \nu_1^j, \mu_2^j - \nu_2^j, \dots, \mu_N^j - \nu_N^j), \quad j = 1, 2, \dots, M, \quad (6)$$

with the probability λ^j / λ^* .

The above stochastic dynamics on \mathbb{Z}^N lattice can also be described by a probability distribution $p(n_1, n_2, \dots, n_N, t)$, which satisfies the chemical master equation (CME):

$$\frac{dp(\mathbf{n}, t)}{dt} = \sum_{j=1}^M [\lambda^j(\mathbf{n} - \boldsymbol{\mu}^j + \mathbf{v}^j)p(\mathbf{n} - \boldsymbol{\mu}^j + \mathbf{v}^j) - \lambda^j(\mathbf{n})p(\mathbf{n})] \quad (7)$$

where $\mathbf{n} = (n_1, n_2, \dots, n_N)$, $\boldsymbol{\mu}^j = (\mu_1^j, \mu_2^j, \dots, \mu_N^j)$, and $\mathbf{v}^j = (v_1^j, v_2^j, \dots, v_N^j)$.

Therefore, the CME and the Gillespie algorithm are two different descriptions for the same stochastic dynamical model of chemical reaction systems in a mesoscopic volume, parallel to the diffusion (Fokker-Planck) equation and stochastic differential equation approaches to Brownian motion, developed by Einstein and Langevin respectively in 1905 and 1908.

See [33–35] for several recent reviews on the theory of the CME and [36–38] for its applications to simple nonlinear chemical reaction systems.

2.2 Nonlinear Chemical Dynamics

The parameter V , the volume of the reaction system, is a crucial parameter of the CME. For chemical reaction systems with macroscopic volume and Avogadro's number of molecules, one can introduce the concentration for species X_i , $u_i = n_i/V$. If we let both n_i and $V \rightarrow \infty$ in the CME, but keep $n_i/V \rightarrow u_i$ finite, then it can be mathematically shown that a set of deterministic, nonlinear kinetic equations arise [35, 39]:

$$\frac{du_i(t)}{dt} = \sum_{j=1}^M (\mu_i^j - v_i^j) J^j, \quad (8)$$

in which

$$J^j = k^j u_1^{v_1^j} u_2^{v_2^j} \cdots u_N^{v_N^j} = \lim_{V, \mathbf{n} \rightarrow \infty} \frac{\lambda^j(\mathbf{n})}{V}, \quad (9)$$

where $\lambda^j(\mathbf{n})$ is given in (3). We note that the system of kinetic equations (9) is nothing but the classic Law of Mass Action for the reaction scheme (1)! Therefore, the CME is not an alternative approach to biochemical kinetics. Rather, the deterministic dynamics based on the Law of Mass Action is the *skeleton* of the CME dynamics. A thorough understanding of any CME, thus, requires a full grasp of the deterministic dynamics from the corresponding nonlinear differential equation.

In fact, studying the CME, (7), together with its deterministic counterpart, (8), side-by-side leads to a series of insights into the nonlinear, stochastic chemical kinetics.

Stationary Distribution First, for most biochemical reaction systems, the CME has a unique stationary probability distribution $p_V^{ss}(\mathbf{n})$. Note that this distribution is also a function of the system's volume V . As a function of V the $p_V^{ss}(\mathbf{n})$ usually has the general form of the so called WKB (Wentzel-Kramers-Brillouin) expansion:

$$p_V^{ss}(\mathbf{n}) = \exp(-V\phi_0(\mathbf{u}) + \phi_1(\mathbf{u}) + \cdots), \quad \mathbf{u} = \frac{\mathbf{n}}{V}, \quad (10)$$

when $V \rightarrow \infty$. The function $\phi_0(\mathbf{u})$ is independent of V . It can be obtained, if it exists, from

$$\phi_0(\mathbf{u}) = \lim_{V \rightarrow \infty} -\frac{1}{V} \log p_V^{ss}(V\mathbf{u}). \quad (11)$$

One of the most important properties concerning the $\phi_0(\mathbf{u})$ is the equation [40]

$$\mathbf{b}(\mathbf{u}) \cdot \nabla \phi_0(\mathbf{u}) = -(\nabla \phi_0(\mathbf{u}))^2, \quad (12)$$

where the $\mathbf{b}(\mathbf{u}) = d\mathbf{u}/dt$ is the right-hand-side of (8). Equation (12) implies that

$$\frac{d}{dt} \phi_0(\mathbf{u}(t)) = \nabla \phi_0(\mathbf{u}) \cdot \frac{d\mathbf{u}}{dt} = \nabla \phi_0(\mathbf{u}) \cdot \mathbf{b}(\mathbf{u}) \leq 0. \quad (13)$$

In other words, the deterministic, nonlinear chemical dynamics given by (8) follow the downhill of the function $\phi_0(\mathbf{u})$. The function $\phi_0(\mathbf{u})$ can and should be considered as a *dynamic landscape*. In fact, the stable steady states of (8) are precisely the local minima of $\phi_0(\mathbf{u})$.

Maxwell Construction Many nonlinear dynamical systems have multiple, locally stable steady states. We will see one of such examples in self-regulating gene network in Sect. 3.1. A deterministic dynamical system approaches one of its stable steady states (or attractors) and then stays there forever. Which attractor it goes to depends on the initial condition of the dynamical system. This behavior gives rise to the concept of “basin of attraction”. This is the picture one obtains from studying chemical dynamics based on (8).

Is one attractor more “important” than another? This question can not be answered under strictly deterministic dynamics. However, for a dynamical system with stochasticity, different attractors can have different probabilities. In this sense, one attractor can be more “stable” than another—While jumping among different attractors, the system spends *totally* more time in a more stable attractor. This is an insight the CME offers that does not exist in the deterministic kinetics.

Now let us consider again $p_V^{ss}(\mathbf{n})$. Let us assume it has several peaks corresponding to the stable steady states of the deterministic dynamics. What is its limit when $V \rightarrow \infty$? Noting that $p_V^{ss}(\mathbf{n})$ is always normalized, and $p_V^{ss}(\mathbf{n}) \sim \exp(-V\phi(\mathbf{u}))$, we see that the entire distribution converges to the global minimum of $\phi_0(\mathbf{u})$ with probability 1. Even though a $\phi_0(\mathbf{u})$ can have many local minima, in the limit of $V \rightarrow \infty$, all their probabilities are infinitesimal but one!

All these other local minima are called *metastable*. If a stochastic dynamics start from within one of the basins of metastable states, if the V are very large, it will take an extremely long time, $\sim e^{\alpha V}$ ($\alpha > 0$) to exit. This is consistent with the result from the deterministic dynamics. The “infinite time” in the deterministic dynamics is meant to be much shorter than these extremely long exit times.

Applying the above discussions to a bistable system which undergoes saddle-node bifurcation:

$$\frac{du}{dt} = b(u, \theta) \quad (14)$$

where θ is a parameter. The steady state(s) $u^*(\theta)$ of the system is obtained from solving $b(u, \theta) = 0$. The solid line in Fig. 2, the S-shaped curve, is called bifurcation diagram. For the middle range of the θ value, the system has three steady states, the top and bottom ones are stable, while the middle one is unstable. The corresponding stochastic dynamics gives its stationary distribution and corresponding $\phi_0(u, \theta)$ which is also shown in Fig. 2. The peaks and troughs match the steady states $u^*(\theta)$. For each θ , in the limit of $V \rightarrow \infty$, the stationary probability $p_V^{ss}(u)$ converges to the global minima of $\phi_0(u, \theta)$. Therefore, the stochastic dynamics selects only one of the two minima of $\phi_0(u)$: When $\theta < \theta^*$, the lowest branch and when $\theta > \theta^*$, the uppermost branch. The dashed vertical line at θ^* is known as Maxwell construction [41].

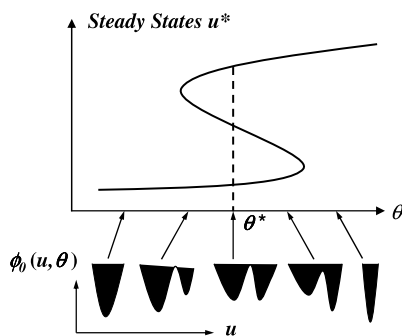


Fig. 2 The S-shaped curve $u^*(\theta)$ is known as *bifurcation diagram* for saddle-node bifurcation. It shows the steady states of a deterministic chemical dynamical system $du/dt = b(u, \theta)$ changing with parameter θ . The system is bistable for the middle ranged value of θ . The corresponding CME gives dynamic landscape $\phi_0(u, \theta)$, with its peak(s) and trough match the u^* . In the limit of $V \rightarrow \infty$, the stationary distribution of the stochastic dynamics has probability 1 located at the global minimum of $\phi_0(u)$. Hence when $\theta < \theta^*$, it is located at the lowest branch; and when $\theta > \theta^*$, it is located at the uppermost branch. There is a *discontinuity* at $\theta = \theta^*$. The *vertical dashed line* is known as Maxwell construction

Competition Between Large System Size and Long Time The stochastic CME clearly shows that there are two very different time scales in the nonlinear chemical dynamics with multistability. The two time scales are well separated by the exit times from one attractor to another. In the time scale much shorter than this, the deterministic chemical kinetics rules. However, in the time scale much longer than this, the system stochastically jumps among the multiple attractors, as a set of discrete states. The dynamics on this time scale is again stochastic. The exit time of an attractor depends exponentially on the system's volume V , $e^{\alpha V}$ ($\alpha > 0$). Hence, the larger a reaction system, the longer one has to wait to observe the stochastic jumps.

Mathematically, for a chemical reaction system with bistability, exchanging the two limits

$$\lim_{V \rightarrow \infty} \lim_{t \rightarrow \infty} p_V(\mathbf{n}, t) \neq \lim_{t \rightarrow \infty} \lim_{V \rightarrow \infty} p_V(\mathbf{n}, t). \quad (15)$$

The left-hand-side gives probability 1 at the global minimum of $\phi_0(\mathbf{u})$ independent of the initial condition; the right-hand-side goes to different local minima of $\phi_0(\mathbf{u})$ depending on the initial condition. The inequality in (15) is the origin of T. Kurtz's convergence theorem for only finite time [39], as well as the so called Keizer's paradox [36, 37].

To Kirschner and Gerhart's thesis, the most important insight from the nonlinear biochemical dynamics is that on an "evolutionary" long time scale, even a simple chemical reaction system can exhibit stochastic variations. These variations and the stochastic dynamics among them, though deeply rooted in the random fluctuations of molecular reactions as understood by Kramers, are mesoscopic, or even macroscopic, emergent properties of nonlinear interacting molecular species. Biological variations need not be solely from DNA sequences; it could come also from *isogenetic* nonlinear biochemical reaction systems.

There is a crucial conceptual issue to be resolved: From a thermodynamic standpoint, how can the above mentioned chemical variations, i.e., diversity, be maintained? Should not all chemical systems approach to equilibrium in the long-time limit?

Indeed, all the above discussed multistable systems are open chemical systems with *sustained* external chemical driving force. Therefore, they do not approach to an equilibrium but rather their respective nonequilibrium steady states [42, 43]. In fact, if one eliminates

all the chemical driving force on a system, then the CME predicts an equilibrium steady state in the long-time limit. The $p_V^{ss}(\mathbf{n})$ in this case is a simple, Gaussian-like distribution; all chemical reactions satisfy the *principle of detailed balance* [44]. The net flux is zero in each and every reaction in the system.

2.3 Nonequilibrium Steady State (NESS)

The mathematical theory of nonequilibrium steady state in stochastic dynamical systems represented by master equations and Markov processes has only been established recently [45]. The physics, however, has a long and diverse histories which can be traced back to H. Haken, T.L. Hill, J. Keizer, M.J. Klein, M. Lax, J.L. Lebowitz, G. Nicolis, I. Prigogine, J. Ross, to name a few among the many pioneers [46–53]. See [42, 43, 54] for some recent applications to chemical and biochemical systems.

Conceptually, there are four kinds of mesoscopic chemical kinetic systems: (i) stationary systems in chemical equilibrium with equilibrium fluctuations; (ii) systems with time-dependent transient relaxation to equilibrium; (iii) stationary open chemical systems which are sustained out-of-equilibrium by a sustained chemical driving force; and (iv) systems with time-dependent transient relaxation toward the (iii). Nonequilibrium steady state (NESS), or stationary nonequilibrium state, the (iii), is the most appropriate chemical dynamic model for a living cell under homeostasis [55].

Equilibrium Stochastic Dynamics and Time Reversibility It is now well understood that the stochastic fluctuations in an equilibrium, as a function of time, is *time reversible*. All statistical properties of a forward stochastic trajectory can not be distinguished from its time reversal. In fact, any sequence of events that occur will have equal probability to occur in reverse—thus nothing can be accomplished in an equilibrium dynamics. There is no energy conversion from one form to another; or transport materials from one place to another.

Detailed Balance At the chemical reaction level, all reactions are going forward and backward with equal likelihood. The net flux within each and every reaction is zero. This is the principle of detailed balance [44]. One immediately sees that any chemical kinetic scheme that assumes irreversible reactions is incompatible with an equilibrium steady state. In fact, the rate constants of a kinetic scheme for an equilibrium reaction system, or a system approaching to an equilibrium, have to satisfy the Wegscheider cycle condition [44, 56]. In the CME formulation, the Wegscheider cycle condition becomes the Kolmogorov cycle condition for reversible Markov processes [45]: each and every cycle in the \mathbb{Z}^N satisfies the detailed balance. In order to distinguish the subtle difference, we have termed the former *chemical detailed balance* and the latter *mathematical detailed balance* [37].

Gardiner [57] has shown that for chemical reaction systems with chemical detailed balance, not only its CME has mathematical detailed balance, hence its stationary distribution is solvable, the stationary distribution is also a multivariate Poissonian conditioned on the conservation laws among molecular species. It has a single peak.

Cycle Flux in NESS Any kinetic model that contains irreversible chemical reactions, therefore, implicitly assumes an chemical driving force. In fact, the force is assumed to be infinite. To have a firm thermodynamic basis, it is advised that one finds out explicitly the source(s) and sink(s) of the chemical driving force(s) applied to a biochemical system in cellular biochemical modeling. For example, in the stochastic models for motor proteins, the driving force is from the hydrolysis of $\text{ATP} \rightarrow \text{ADP} + \text{Pi}$. Their concentrations are assumed to be constant in motor protein kinetics.

With the presence of a sustained chemical driving force(s), the stochastic dynamics according to a CME approaches not to an equilibrium, but to a NESS. When a system is in a NESS, since all the probabilities are no longer changing with time, and since the system is not detailed balanced, there must be balanced *cycle fluxes*. This is a simple consequence of Kirchhoff's Law. The cycle flux and the landscape $\phi_o(\mathbf{u})$ in a chemical NESS are like the current and voltage in an electrical circuit with battery. They provide complementary information on a "living" chemical system [58, 59].

NESS on Different Time Scales The term "nonequilibrium steady state" deserves further clarification. It seems to have two different meanings in systems with multiple attractors. On the time scale shorter than the jumping times between the attractors, a NESS corresponds to a single, average chemical composition with a multivariate Gaussian-like concentration (or number) fluctuations. In this sense, a nonlinear chemical system can have multiple steady states. The long-time fate of the system depends on its initial condition. They are the attractors of deterministic dynamics.

However, on the time scale much longer than the slowest exit time between the attractors, the term NESS takes another, completely different, meaning. Here a NESS has a stationary distribution which peaks at every attractors with appropriate weights. The chemical system jumps continuously among the multiple attractors with ergodicity. In this case, a system has a unique NESS which has the stationary distribution as the solution to the CME, the $p_V^{ss}(\mathbf{u})$.

We should mention that for a chemical reaction system with individual reaction rates on the order of milli- and micro-seconds, and with a couple of thousand copies of molecules, the exit time of an attractor can easily be as long as 10^{11} years. That is an eternity! However, if the number of molecules is reduced to a few hundreds, then the time is only on the order of hours.

2.4 Nonequilibrium Phase Transition in the Bistable CME

Multiple steady states, or attractors, and bifurcations upon parameter changes are the essence of deterministic nonlinear dynamical systems. Since the 1970s, it has long been argued that cellular and physiological states of biological systems should be understood in terms of the concept of attractors in nonlinear dynamics [14, 60–63]. The CME approach to the reaction dynamics of mesoscopic biochemical systems adds a significant mathematical rigor to this still elusive idea. In particular, the interplay between nonlinearity and stochasticity seems to provide a deeper understanding of the complexity of real biological systems. The concepts such as "barrier crossing" and "nonlinear bifurcation" are unified in the CME theory.

In a recent study of the CME of an open, driven biochemical system, the phosphorylation-dephosphorylation cycle with feedback (see Sect. 3.2), it has been shown that [17] the bistable system exhibits all the characteristics of equilibrium phase transitions extensively studied in statistical physics. This includes the Maxwell construction for discontinuity in the thermodynamic ($V \rightarrow \infty$) limit, the Lee-Yang theory of a zero of the partition function being the origin of non-analyticity, and the terminal critical point in a phase diagram matching the cusp in nonlinear saddle-node bifurcation. These findings seem to suggest that isogenetic variations in biochemical systems are intimately related to phase transition.

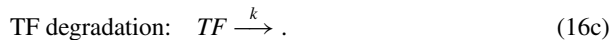
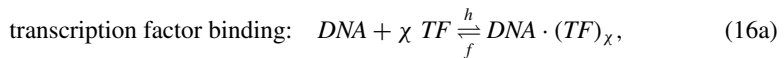
3 Self-Regulating Genes and Phosphorylation-Dephosphorylation Cycles with Feedback

We now turn our attention to two concrete biochemical reaction systems, one is the self-regulating gene network (see (16)) and the other is the phosphorylation-dephosphorylation

signaling cycle (see (23)). As we shall show, even though they are considered completely different biochemical entities, their biochemical kinetics are essentially identical. Hence, their analysis based on deterministic mass-action kinetics and stochastic CME will be carried out in parallel.

3.1 Self-regulating Gene Networks

Self-regulating gene networks have been extensively studied in recent years [64–66]. These systems can be described in terms of a biochemical kinetic scheme that consists of biosynthesis and degradation of a transcription factor (TF), as well as the TF binding to the DNA regulatory sequence of its own gene:



Hornos *et al.* [65] considered $g_1 < g_0$ with $\chi = 1$, i.e., the TF in monomeric form is a repressor for its own gene expression. Walczak *et al.* [66] studied $g_1 > g_0$ with $\chi = 2$. The gene product in this case, in dimeric form, is its own transcription enhancer.

The corresponding macroscopic kinetics of gene regulations with feedback, in terms of the deterministic Law of Mass Action, can be written in ordinary differential equations as

$$\frac{dx}{dt} = hy^{\chi}(1-x) - fx, \quad \frac{dy}{dt} = (g_0(1-x) + g_1x) - ky, \quad (17)$$

where x is the fraction of the DNA with TF bound ($\chi = 1, 2$ for monomer and dimer, respectively), y is the concentration of the TF, the gene product. Figure 3 shows that system (17) can have unique steady state as well as bistability.

To solve the steady state(s), we introduce nondimensionalization. The two equations in (17) are then simplified

$$\frac{dx}{d\tau} = \omega [\theta z^{\chi}(1-x) - x], \quad \frac{dz}{d\tau} = g + (1-g)x - z. \quad (18)$$

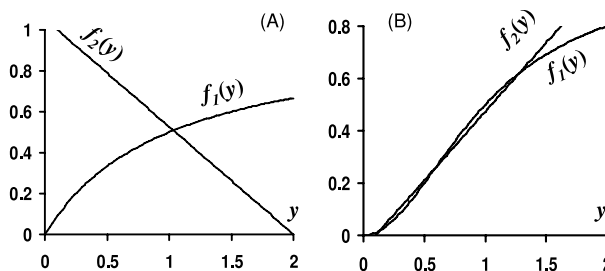


Fig. 3 The null clines (isoclines) of the system in (17) for negative feedback case (A) and positive feedback case (B), showing unique steady state and bistability, respectively. For the negative feedback case, $g_0 > g_1$ and $\chi = 1$. The two null clines are $x = f_1(y) = y/(K_d + y)$ and $x = f_2(y) = (g_0 - ky)/(g_0 - g_1)$, where $K_d = f/h$. For the positive feedback case, $g_1 > g_0$ and $\chi = 2$. The two null clines are $f_1(y) = y^2/(K_d + y^2)$ and $f_2(y) = (ky - g_0)/(g_1 - g_0)$. Parameters used in computations for (A): $K_d = 1$, $k = 0.5$, $g_0 = 1$, $g_1 = 0.05$, $\chi = 1$; and for (B) $K_d = 1$, $k = 0.5$, $g_0 = 0.05$, $g_1 = 1$, $\chi = 2$

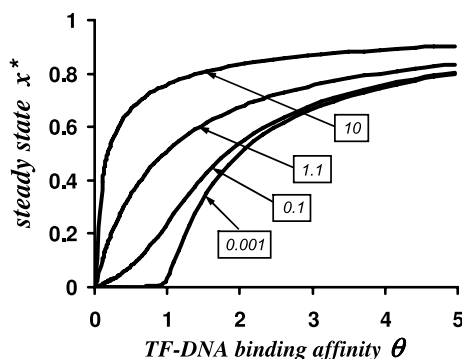


Fig. 4 Self-regulating gene network with $\chi = 1$. The steady state fraction of gene with TF bound, x^* , increases with the binding affinity parameter $\theta = (g_1/k)K_d^{-1}$ where $K_d = f/h$. The four curves are for $g = g_0/g_1 = 0.001, 0.1, 1.1, 10$ where g_0 and g_1 are the TF biosynthesis rates in the absence and presence of TF binding to DNA. $g > 1$ corresponds to the TF being a repressor, and $g < 1$ corresponds to the TF being an enhancer. For strong enhancer with very small $g = 0.001$, the TF induced gene expression is highly cooperative, exhibiting delayed onset

where

$$z = \frac{k}{g_1}y, \quad \tau = kt, \quad \theta = \frac{h}{f} \left(\frac{g_1}{k} \right)^\chi, \quad \omega = \frac{f}{k}, \quad g = \frac{g_0}{g_1}. \quad (19)$$

The three cases of $\chi = 0, 1, 2$ yield, respectively, hyperbolic gene activation, delayed onset (with possible transcritical bifurcation when $g = 0$, at $\theta = 1$), and bistability with saddle-node bifurcation.

$\chi = 0$ There is only a single steady state. The fraction of activated gene is $x^* = \theta/(1 + \theta)$ which has a hyperbolic dependence on θ . In the meantime the gene product $z^* = (g + \theta)/(1 + \theta)$.

$\chi = 1$ The quadratic equation $(1 - g)x^2 - (1 - 2g - 1/\theta)x - g = 0$ has a unique root $x^* \in (0, 1]$ for $\theta \geq 0$:

$$x^* = \frac{(1 - 2g - \frac{1}{\theta}) + \sqrt{(1 - 2g - \frac{1}{\theta})^2 + 4g(1 - g)}}{2(1 - g)}. \quad (20)$$

Figure 4 shows that x^* increases from 0 to 1 when θ increases from 0 to ∞ . When $g = 1$, there is no self-regulation and $x^* = \theta/(1 + \theta)$ which has the standard hyperbolic shape. When $g = 0$, i.e., the TF is a strong enhancer, $x^* = 0$ for $\theta \leq 1$ and $x^* = 1 - 1/\theta$ for $\theta \geq 1$. There is a transcritical bifurcation at $\theta = 1$. This type of response is called delayed onset.

$\chi = 2$ For the repressor case with negative feedback, i.e., $g > 1$ in (18), there is no bistability because the null cline for $dx/dt = 0$ is an increasing function $z = (x/(1 - x)/\theta)^{1/2}$ while the null cline for $dz/dt = 0$ is a decreasing function $z = g - (g - 1)x$. For the case of positive feedback with $g < 1$, the system (18) can have three steady states in the positive quadrant, two stable and one unstable. Figure 3 shows the qualitatively different arrangements of the null clines for the two cases.

The condition for the positive feedback case to have bistability is when (θ, g) is in the cusp region bound by the parametric curve $(\frac{1}{z(2-3z)}, \frac{z(1-2z)}{2-3z})$ as shown in Fig. 5.

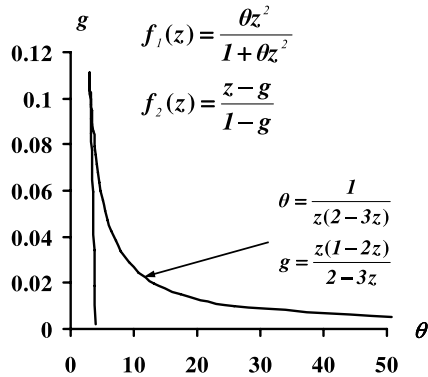


Fig. 5 The $f_1(y)$ and $f_2(y)$ in Fig. 3B has three intersection, corresponding to three steady states for the ODE system (bistability). The parameter region for the bistability typically has a “cusp”, known as *cusp catastrophe*. If we let $z = ky/g_1$, then the two null clines in Fig. 3B become the $f_1(z)$ and $f_2(z)$ with $\theta = (hg_1^2/(fk^2))$ and $g = g_0/g_1$. To obtain the cusp region, we solve simultaneously $f_1(z) = f_2(z)$ and $f_1'(z) = f_2'(z)$. This yields the (θ, g) parametrically in terms of z

Adiabatic and Non-adiabatic Limits If $\omega \gg 1$ in (18), then the FT binding to DNA is much faster than its own biosynthesis and degradation. This is known as the *adiabatic limit* [65, 66]. In this case, one can first solve the quasi-steady-state for $dx/d\tau = 0$ to obtain $x = \theta z^\chi / (1 + \theta z^\chi)$. Then the system of equations is reduced to a single one:

$$\frac{dz}{d\tau} = g + \frac{\theta(1-g)z^\chi}{1 + \theta z^\chi} - z. \quad (21)$$

On the other hand, if $\omega \ll 1$, then the FT binding to DNA is much slower than its own biosynthesis and degradation. This is known as the *non-adiabatic scenario* [65, 66]. In this case, one can solve the quasi-steady-state for $dz/d\tau = 0$ to obtain $z = g + (1 - g)x$. Then again the system in (18) is reduced:

$$\frac{dx}{d\tau} = \omega \{ \theta [g + (1 - g)x]^\chi (1 - x) - x \}. \quad (22)$$

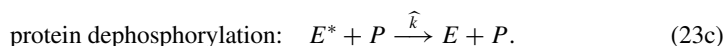
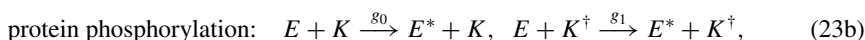
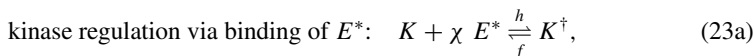
Strong Enhancer with $g = 0$ The steady state of the above kinetic system with $g = 0$ has been extensively studied in the context of phosphorylation-dephosphorylation cycles with feedback [17, 41, 67–69]. We now introduce this biochemical signaling system which is kinetically almost isomorphic to the self-regulating gene network.

We have assumed in (16) a cooperative binding of two copies of TF to the DNA in the case of $\chi = 2$. It is important to point out that the strong nonlinearity required in the bistable behavior is not from the cooperativity *per se*. Rather, it is from the fact that only the doubly occupied DNA is functional. In this case, the response function is sigmoidal even for completely independent binding: $\frac{x^2}{1+2x+x^2}$. This response function is in sharp contrast to the *fraction of binding*, $\frac{1 \cdot 2x + 2 \cdot x^2}{2 \cdot (1+2x+x^2)}$, which is hyperbolic $\frac{x}{1+x}$. See [70, 71] for recent experimental and theoretical studies on the consequences of nonlinearity from time delay and cooperative binding.

3.2 Phosphorylation-Dephosphorylation Cycles (PdPC) with Feedback

Phosphorylation-dephosphorylation cycles (PdPC) are biochemical regulatory systems in cell signaling. They consist of a substrate protein E which can be phosphorylated to become E^* , catalyzed by a protein kinase K , and protein phosphatase P . In many cases, the kinase itself can be regulated via binding to the E^* ; thus the feedback is in the form of autocatalysis. For more discussions on concrete biological examples see the Fig. 1 of [69].

The system can be described in terms of biochemical kinetic scheme



The deterministic kinetic equations for this class of models, according to the Law of Mass Action, is

$$\frac{dx}{dt} = hy^\chi(1-x) - fx, \quad \frac{dy}{dt} = (g_0(1-x) + g_1x)(y_t - y) - ky, \quad (24)$$

where x is the fraction of the kinase in the K^\dagger form, y is the concentration of phosphorylated E^* , $k = \hat{k}[P]$, and y_t is the total concentration of E and E^* . Comparing (24) to (17), we see that the two systems of equations are essentially the same except the former contains an extra term $(y_t - y)$ on the right-hand-side of dy/dt . When $k \gg g_0, g_1$, the former is reduced to the latter.

If the reaction $K + E^* \rightleftharpoons K^\dagger$ is fast, $y_t^\chi \ll f/h = K_d$, and $g_0 = 0$, then one has a quasi-steady-state for $dx/dt = 0$ and a simplified equation for $u = y/y_t$,

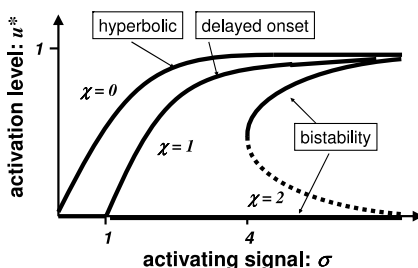
$$\frac{du}{d\tau} = \sigma u^\chi(1-u) - u, \quad (25)$$

in which u represents the fraction of phosphorylated E^* , $\tau = kt$, and $\sigma = g_1hy_t^\chi/fk$ represents the ratio of activities of a kinase to that of a phosphatase.

When $\chi = 0$, the steady state u^* is a hyperbolic function of σ : $u^* = \sigma/(1 + \sigma)$. When $\chi = 1$, $u^* = 0$ for $\sigma \leq 1$ and $1 - 1/\sigma$ for $\sigma \geq 1$, exhibiting delayed onset. When $\chi = 2$, $u_1^* = 0$ is always stable, and when $\sigma \geq 4$, a second stable steady state $u_2^* = (\sigma + \sqrt{\sigma^2 - 4\sigma})/(2\sigma)$ appears. u_1^* and u_2^* are separated by an unstable $u_3^* = (\sigma - \sqrt{\sigma^2 - 4\sigma})/(2\sigma)$. See Fig. 6.

The open, driven chemical nature of the PdPC with bistability has been studied in [68]. It was shown that if the free energy from ATP hydrolysis is below a critical value (this can be

Fig. 6 Steady state(s) of phosphorylation-dephosphorylation cycle (PdPC) with positive feedback, described by the model in (25). $\chi = 0, 1, 2$ represent no-, monomeric, and dimeric activations of the kinase, respectively. Find the equations for the three curves in the text



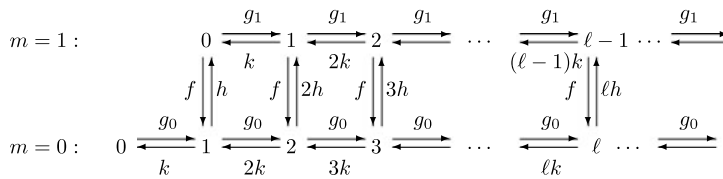


Fig. 7 $m = 0$ and $m = 1$ represent the unbound and bound state of the single copy of DNA (gene). ℓ denote the copy of free TF. Monomeric TF binds DNA with on-rate constant h and off-rate constant f . Binding reduces the copy number of free TF by 1 ($\chi = 1$). TF biosynthesis rate is g_1 and g_0 when the gene is bound and unbound, respectively. TF degradation rate is k

accomplished by either decreasing ATP concentration or increasing ADP/Pi concentrations), then the bistability disappears all together. Biochemical variations can only be maintained with an energy expenditure and free energy dissipation.

3.3 Stochastic Dynamics According to the CME

The theory in Sect. 2.2 indicates that for systems with nonlinear, deterministic bistability, the CME will have its stationary distribution with two peaks located precisely at the two fixed points; there are two stochastic attractors.

However, it comes as a surprise that the CME of a self-regulating gene network, with a monomeric repressor ($g_0 > g_1$, $\chi = 1$), also exhibits bimodal stationary distribution [64, 65]. This is not expected from (17), as illustrated in Figs. 3A and 4. In a similar vein, it has also been discovered that PdPC with feedback, even for $\chi = 1$ and $g_0 = 0$, can have bimodality [69, 72]. This phenomenon has been called *noise-induced bistability*.

It turns out, this type of stochastic bistability (to be distinguished from the nonlinear bistability) is a small copy number effect. In fact, for the self-regulating gene network in a single cell, there is only one copy of the DNA [64, 65]! In the PdPC system studied, the copy number is also small, about 30 [69]. The stochastic bistability is intimately related to the extinction phenomenon [69, 73].

Stochastic Bistability with Slow Fluctuations Stochastic bistability can be best understood in the single-molecule context with slow fluctuations [73, 74]. Consider the CME for kinetic scheme in (16) and assuming only a single copy of the DNA, then one has the detailed kinetics on a lattice, $m = 0, 1$ and $\ell = 0, 1, 2, \dots$, in Fig. 7.

If the rates h and f are sufficiently smaller than g 's and k , then one has a quasi-stationary Poisson distribution along each line in Fig. 7:

$$p(\ell|m=0) = \frac{1}{\ell!} \left(\frac{g_0}{k}\right)^\ell e^{-g_0/k}, \quad p(\ell|m=1) = \frac{1}{\ell!} \left(\frac{g_1}{k}\right)^\ell e^{-g_1/k}. \quad (26)$$

The transition rate from $m = 0$ to $m = 1$ is given as the average

$$\bar{h} = h \sum_{\ell=0}^{\infty} \ell p(\ell|m=0) = \frac{hg_0}{k}. \quad (27)$$

Then the stationary distribution is

$$p^{ss}(\ell, m) = p(\ell|m)p(m) = \begin{cases} \frac{fk}{fk+hg_0} \frac{1}{\ell!} \left(\frac{g_0}{k}\right)^\ell e^{-g_0/k}, & m=0; \\ \frac{hg_0}{fk+hg_0} \frac{1}{\ell!} \left(\frac{g_1}{k}\right)^\ell e^{-g_1/k}, & m=1. \end{cases} \quad (28)$$

Therefore, if the peak on the $m = 1$ line, g_1/k , and the peak on the $m = 0$ line, g_0/k , are well separated, then the marginal distribution $p^{ss}(\ell)$ will have two peaks. This is the stochastic bistability due to slow, non-adiabatic gene regulation.

Adiabatic Limit We now consider the case when the rates h and f are much greater than g 's and k [65, 74]. We can obtain a quasi-stationary distribution between $m = 0$ and $m = 1$ for each and every ℓ . We shall now denote the *total copy number* of TF by ℓ , so the labels along the two lines in Fig. 7 match:

$$p(m = 0|\ell) = \frac{f}{\ell h + f}, \quad p(m = 1|\ell) = \frac{\ell h}{\ell h + f}. \quad (29)$$

Then the kinetics are simplified into a 1-dimensional birth-and-death process with birth and death rates

$$b_\ell = \frac{f g_0 + \ell h g_1}{f + \ell h}, \quad d_{\ell+1} = \frac{(f + \ell h)(\ell + 1)k}{f + (\ell + 1)h}. \quad (30)$$

The marginal stationary distribution for the copy number of total TF is

$$p^{ss}(\ell) = C \prod_{i=1}^{\ell} \frac{b_{i-1}}{d_i} = C \prod_{i=1}^{\ell} \frac{[f g_0 + (i-1)h g_1][f + i h]}{i[f + (i-1)h]^2 k}, \quad (31)$$

where C is a normalization factor. We note that

$$b_\ell - d_\ell = \frac{f g_0 + \ell(h g_1 - f k + h k) - \ell^2 h k}{f + \ell h}. \quad (32)$$

The numerator of (32) is only a quadratic function of ℓ . When $\ell = 0$ it is positive and when $\ell = \pm\infty$ it is negative. Therefore, it has only a single zero for positive ℓ . The distribution $p^{ss}(\ell)$ in (31) can only have a single peak for $\ell > 0$. This result agrees with that from the deterministic kinetics. However, because of the discrete nature of ℓ , the $p^{ss}(\ell)$ can also peak at $\ell = 0$ [69]. The condition for this is $p^{ss}(0)/p^{ss}(1) = d_1/b_0 > 1$. That is,

$$\left(\frac{f}{f+h}\right) \frac{k}{g_0} > 1. \quad (33)$$

In fact, if $g_0 = 0$, then the system has an absorbing state at $\ell = m = 0$. Therefore, the stochastic bistability is the “ghost” of the extinction phenomenon.

Mathematically, we note that the $\chi = 1$ case has a quadratic nonlinearity and the $\chi = 2$ case has a cubic nonlinearity. This distinction, leading to stochastic bistability and nonlinear bistability respectively, has been discussed in the context of PdPC with feedback in [69].

Transition Rate Volume Dependence as an Indicator for Bistability Mechanism How can one determine whether the bistability in a mesoscopic system is stochastic in nature or non-linear in nature? We suggest that the volume-dependence of exit rates can be used as an indicator. By increasing volume and numbers of molecules but keeping their concentrations invariant, stochastic bistability disappears while nonlinear bistability intensifies. As a function of the system's size, the two types of bistability behave differently in a fundamental way.

In thermodynamics, one investigates the mechanism of chemical and biochemical reactions (e.g., protein folding) by measuring reaction rates as functions of temperature and

plotting the so-called Arrhenius plot with *activation enthalpy*. The widely practiced approach does not mean one is interested in a chemical reaction in high temperature or low temperature, *per se*. Rather it is understood that *temperature dependence* provides insights into the mechanism of a reaction: Is it entropic or enthalpic driven?

An analogue exists for the case of *volume dependence* of the transition rates between two stochastic attractors: The rates increase for nonlinear bistability but decrease for stochastic bistability when the volume decreases.

4 Nonequilibrium Statistical Thermodynamics

We now turn our attention to thermodynamics. Afterall, the initial motivation of statistical physics is to understand thermodynamics from a molecular perspective in terms of the theory of probability. We now have a probabilistic, stochastic description of the dynamics of open, driven biochemical reaction systems. Is there an overarching nonequilibrium thermodynamics?

The answer is “yes”. Recently, it becomes known that there is a completely statistical thermodynamics for Markovian dynamics based on a master equation [19, 75–79]. Thermodynamics, it turns out, is a general mathematical law of any Markovian dynamics. The thermodynamics of molecular systems discovered in thermal physics is simply one special example.

The Mathematical Theory of Thermodynamics Let us consider a master equation

$$\frac{dp_i}{dt} = \sum_{j=1}^N (p_j q_{ji} - p_i q_{ij}). \quad (34)$$

As we have discussed, one needs to assume that $q_{ij} \neq 0$ iff $q_{ji} \neq 0$ for any i, j in order to be able to study thermodynamics. For simplicity, we further assume the Markovian system is irreducible. Hence, it has a unique, positive stationary distribution we shall denote by $\{\pi_i\}$:

$$\sum_{j=1}^N (\pi_j q_{ji} - \pi_i q_{ij}) = 0, \quad \pi_i > 0. \quad (35)$$

Two thermodynamic quantities will be investigated [19]: The Gibbs entropy

$$S(t) = - \sum_{i=1}^N p_i(t) \ln p_i(t), \quad (36)$$

and the Gibbs free energy

$$F(t) = \sum_{i=1}^N p_i(t) \ln \left(\frac{p_i(t)}{\pi_i} \right). \quad (37)$$

Applying the chain rule, one has

$$\frac{dS(t)}{dt} = e_p - h_d, \quad \frac{dF(t)}{dt} = -f_d, \quad (38)$$

where the entropy production rate e_p , heat dissipation rate h_d , and free energy dissipation rate f_d are

$$e_p = \frac{1}{2} \sum_{i,j} (p_i q_{ij} - p_j q_{ji}) \ln \left(\frac{p_i q_{ij}}{p_j q_{ji}} \right), \quad (39)$$

$$h_d = \frac{1}{2} \sum_{i,j} (p_i q_{ij} - p_j q_{ji}) \ln \left(\frac{q_{ij}}{q_{ji}} \right), \quad (40)$$

$$f_d = \frac{1}{2} \sum_{i,j} (p_i q_{ij} - p_j q_{ji}) \ln \left(\frac{p_i \pi_j}{p_j \pi_i} \right). \quad (41)$$

One can mathematically show that

$$S(t) \geq 0, \quad F(t) \geq 0, \quad e_p(t) \geq 0, \quad \text{and} \quad f_d(t) \geq 0. \quad (42)$$

More importantly, one can define the so called *house keeping heat*, originally introduced by Oono and Paniconi [80] in a phenomenological NESS thermodynamics, to quantify the driving force applied to the system:

$$Q_{hk} = e_p - f_d = \frac{1}{2} \sum_{i,j} (p_i q_{ij} - p_j q_{ji}) \ln \left(\frac{\pi_i q_{ij}}{\pi_j q_{ji}} \right) \geq 0. \quad (43)$$

It is also non-negative.

The interpretations of these inequalities are clear: Since the h_d can be positive and negative, there is no guarantee that $dS/dt \geq 0$. As it was clear to Gibbs, for canonical systems it is not the entropy that always increases, but it is the free energy that always decrease: $dF/dt \leq 0$.

Furthermore, we have the decomposition of the entropy production rate

$$e_p = f_d + Q_{hk}, \quad \text{in which } f_d \geq 0 \text{ and } Q_{hk} \geq 0. \quad (44)$$

Now we see that the total time irreversibility, which is characterized by the entropy production rate, e_p [45], really comes from two different origins: The f_d characterizes the spontaneous relaxation (or organization) to a system's stationarity. $f_d = 0$ when a system reaches its stationary π_i . This irreversibility is Boltzmann's original thesis. However, Q_{hk} characterizes irreversibility due to sustained driving, or energy pumping, of the system out-of-equilibrium. There is a continuous dissipation even in NESS. As we have discussed in Sect. 2.3, this driving force is characterized by the breakdown of detailed balance: $\pi_i q_{ij} \neq \pi_j q_{ji}$. When there is no external driving force, $Q_{hk} = 0$. This irreversibility has long been Prigogine's thesis [49]. For systems without detailed balance, spontaneous approaching to stationary distribution π_i is a form of *self-organization* [49, 51, 63].

Systems with Detailed Balance For systems with detailed balance, which is the subject of classic statistical mechanics, we have the free energy dissipation precisely equal to entropy production rate: $f_d = e_p$. In fact, we see that in this case, $\ln \pi_i = -E_i$ is the energy of the state i ($k_B T = 1$). Then

$$F = \sum_{i=1}^N E_i p_i - S = \langle E \rangle - S. \quad (45)$$

We have recovered the fundamental equation of classical thermodynamics. Furthermore, we see that if $\pi_i = \text{constant}$, i.e., the system's equilibrium has an *equal probability a priori*, then the $S(t) = -F(t) + \text{constant}$, and $dS/dt = -dF/dt \geq 0$. This is in fact Boltzmann's statistical mechanics for isolated system with microcanonical ensemble. In this case, the Second Law states "entropy never decreases".

The Energy of a Stochastic System The foregoing theory seems to suggest that for any stochastic dynamical system, with or without detailed balance, one can define a "statistical energy" as $E_i = -\ln \pi_i$. Combined with the discussion on $\phi_0(\mathbf{u})$ from Sect. 2.2, it seems to us that Boltzmann's law, $p_i = \exp(-E_i/k_B T)$, might be understood backward and used as a way to introduce a new form of energy: The energy of stochastic systems [75].

Taking the CME as an example. If one takes $\phi_o(\mathbf{u})$ as the energy function, and takes V as $1/(k_B T)$, then one has a "partition function"

$$Z(V) = \int d\mathbf{u} e^{-V\phi_0(\mathbf{u})}. \quad (46)$$

One can in fact develop an entire system of "volumodynamics". It will be interesting to see whether this line of inquiry leads to any new insights for analyzing the CME or nonequilibrium thermodynamics [81].

5 Summary

There are implications to cellular biological systems from the stochastic, nonlinear chemical dynamics perspective. But at the onset, we shall first stress that the CME theory we have discussed assumes a spatially homogeneous chemical reaction system. This is certainly not true for a real biological cell. The CME is a highly idealized model, just as the Ising model extensively studied in statistical mechanics. While the Ising model and related interacting particle systems emphasize spatial aspect of a molecular system, the CME emphasizes compositional heterogeneity in biological systems.

The significance of the CME is its richness, depth, and sophistication. It endows a full range of dynamics from the stochastic mesoscopic scale to the deterministic macroscopic scale, and beyond. It provides insights into the nature of "thermodynamic limit".

Furthermore, it gives rigorous distinction between closed systems that approach to equilibrium and open, driven systems that exhibit spontaneous self-organization. It also allows for studying the relationship between stochastic dynamics and statistical thermodynamics. In terms of the CME, investigations into elusive ideas such as "energy cost that sustains complexity (diversity) measured by entropy" becomes possible.

Last, but not the least, the CME offers an understanding of the interaction between nonlinearity and stochasticity in dynamics. There is no doubt that these two elements are central to many biological processes.

5.1 Emergent Properties of Stochastic Nonlinear Systems

Emergent properties are central to any complex systems and processes [74, 82]. In nonlinear dynamics, emergent properties manifest themselves as the existence and locations of multiple attractors with fixed points, periodic oscillations, or chaotic motions. Simple dynamics are associated with gradient systems in which the attractors are known *a priori* and are determined *locally*; every step of the way, the system is closer to the final destiny. A non-gradient system has no such certainty.

A stochastic CME with detailed balance has its dynamics essentially following a gradient of a function. In protein folding dynamics, the energy landscape is the function. In any study of protein folding dynamics, an energy landscape is always known *a priori*; and its gradient field is the *cause* of the dynamics.

The $\phi_0(\mathbf{u})$ as a dynamic landscape, however, has a different characteristics. First, it is non-local. One requires the dynamics to move over the entire possible space many times in order to establish it. The $\phi_0(\mathbf{u})$ is a *consequence* of the dynamics. It is only known *retrospectively*.

Therefore, in this perspective, the very existence of multistability, and the average time required to move from one attractor to another, are all emergent properties. They are the results of the complex dynamics of an open biochemical system as a whole under a particular given environment condition.

Dynamics as a Sequence of Punctuated Equilibria Dynamics of nonlinear, stochastic systems with multiple attractors possess certain universal features. As we have said in the beginning, stochastic dynamics jumping among stochastic attractors spend most of the time in waiting. On an evolutionary time scale, the process exhibits a sequence of *punctuated equilibria* (see Fig. 1).

Upon a perturbation, a system's initial response is always a relaxation back to the fixed point with nearly deterministic dynamics. This occurs rapidly. The system then settles at the bottom of the attractor with Gaussian fluctuations. Such a state is often mistaken as an equilibrium. Then in a much longer time scale, the rare event of barrier-crossing leads the system to another attractor (see Fig. 8).

Protein folding kinetics is well-known to have this generic characteristic: A folded protein when immersed in a denaturing solvent, first becomes a “dry molten globule” then unfolds via a thermal activated process; a unfolded protein when immersed in a native solvent, first becomes a “wet molten globule” then folds by thermal activation [83]. These molten globules are folded protein in a denaturant and unfolded protein under a native condition, respectively.

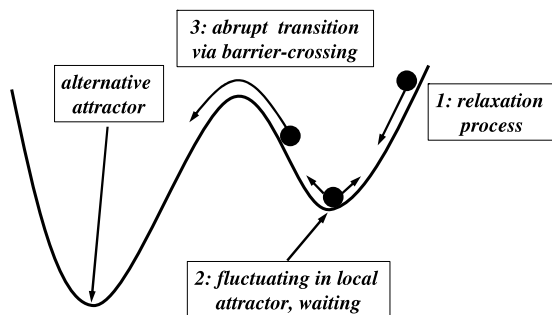


Fig. 8 A schematics showing the generic features of a nonlinear, stochastic dynamical system with multiple attractors under perturbation: Immediately after the perturbation, the system is likely residing at the slope of an attractor. Then (1) relaxation occurs and the system returns to its local steady state. At local steady state (2), the system fluctuates and spends the time in waiting until a rare event of barrier-crossing (3). The rare event usually has an exponential waiting time and the process is “memoryless”. This means it occurs without any indication

Other much more complex biological processes seem also to exhibit this feature. For example, the three Es of cancer immunoediting articulate that an immune system counteracts tumor growth with three main phases: elimination, equilibrium, and escape [84].

5.2 Isogenetic Variations of Biochemical Dynamics

Two cells with identical genomes are called *isogenetic*. With exactly same chemical environment, two isogenetic cells can have very different chemical compositions represented by different attractors of the nonlinear biochemical dynamics. Note that by the same chemical environment, we mean a sustained chemical gradient of certain nutrients and their metabolites. With the same chemical potential, however, the two cells can have different nutrient influx. Applying this idea specifically to the biochemical network responsible for gene transcriptional regulation, one can easily understand the origin of isogenetic variations in gene expression [60, 85, 86].

There are growing experimental observations on the multiple steady states of a cell population. The multiple-state nature is most convincing when a cell population is in the middle of a transition: Two peaks with comparable size rather than one can be observed. For example, in *Xenopus* oocyte maturation induced by hormone progesterone, it had been known that progesterone treatment leads to an increase in the phosphorylation of mitogen-activated protein kinase (MAPK). Ferrell and Machleder [87] have shown, however, that the level of MAPK phosphorylation in an isogenetic *Xenopus* oocyte population has a bimodal distribution. Furthermore, the relative heights of the two peaks change but their locations are invariant with the increasing progesterone. This is the hallmark of a bistable (also known as two-state, and all-or-none) system under external perturbation.

Similarly, Buckmaster *et al.* [88], using Raman spectrum as an indicator for DNA fragmentation, observed a shift in a bimodal distribution during the apoptosis of DAOY cell line (human brain tumor medulloblastoma) induced by etoposide, a topoisomerase II inhibitor.

Also in cell line U2OS, a human osteosarcoma, Xu *et al.* [89] observed a bimodal distribution in the intensity of fluorescein labeled FITC-Annexin V, a protein that preferentially binds to negatively charged phosphatidylserine (PS). Cell apoptosis involves changes on its surface with the exposure of PS. Upon irradiation, which induced DNA damage and apoptosis in U2OS cells, Xu *et al.* reported a shift in the relative heights of the two peaks. The shift is intensified with the presence of PDCD5 (programed cell death 5) protein, which is known to facilitate apoptosis. Again, the apoptotic process changes the heights of the two peaks without changing their locations.

Cancer cells are well known to be genomically very unstable and heterogeneous; it is not known to us whether these tumor cell lines are truly isogenetic. Still, assuming somatic mutations are rare, these observations strongly suggest nonlinear biochemical multistability in tumor cells.

It is interesting to note that in the 1970s, the field of protein folding had gone through a similar stage in demonstrating the two-state nature of protein folding kinetics [90]. The history of protein science can shed some light on the current development of cellular dynamics.

5.3 Inheritability of Nonlinear Chemical Attractors

DNA in terms of Watson-Crick base-pairing has been considered the only mechanism for inheritability. However, a biochemical system residing in a nonlinear attractor can also be “inherited” via cell growth and division. The CME predicts that the concentrations of the biochemical species are invariant, not their copy numbers. Therefore, if a cell has an autonomous mechanism for increasing its aqueous volume, all the copy numbers will follow

by keeping the concentrations at the steady state. This process is self-organizing and robust. Cell division also maintains the concentrations for both daughter cells. By this mechanism, two isogenetic cells in different biochemical attractors that far apart will go through “growth and division” with their respective chemical compositions inherited.

Finally, we shall also emphasize that the existence and locations of the stochastic attractors of a nonlinear biochemical system are dependent upon the environmental biochemical conditions. Therefore, “mutations” occur upon “environmental” changes in the chemical context. This possibility provides further insights into the debate on spontaneous versus adaptive mutations at the cellular biochemical systems level [91–93]. Still, whether and how such “feedback loops” in cellular evolution leading to genomic innovation is the next stage of the “plausibility of life”.

Acknowledgements I thank Lisa Bishop for carefully reading the manuscript and helpful comments, and Drs. Hao Ge and Melissa Vellela for collaborations. Discussions with Ping Ao, Jing Chen, Ken Dill, Martin Falcke, John Hopfield, Sui Huang, Annie Lemarchand, Michael Mackey, Jin Wang, Peter Wolynes and Sunney Xie are gratefully acknowledged. Part of this work was done during the KITPC (Beijing) 2010 program on “Emergent Behaviour of Biomolecular Ensembles and Networks”.

References

1. Kirschner, M.W., Gerhart, J.C.: *The Plausibility of Life: Resolving Darwin's Dilemma*. Yale Univ. Press, New Haven (2005)
2. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P.: *Molecular Biology of the Cell*, 4th edn. Garland Science, New York (2002)
3. Ewens, W.J.: *Mathematical Population Genetics: I. Theoretical Introduction*, 2nd edn. Springer, New York (2004)
4. Koonin, E.V.: Darwinian evolution in the light of genomics. *Nucleic Acids Res.* **37**, 1011–1034 (2009)
5. von Hippel, P.H.: From “simple” DNA-protein interactions to the macromolecular machines of gene expression. *Ann. Rev. Biophys. Biomol. Struct.* **36**, 79–105 (2007)
6. Alberts, B.: The cell as a collection of protein machines: Preparing the next generation of molecular biologists. *Cell* **92**, 291–294 (1998)
7. Berg, O.G.: A model for the statistical fluctuations of protein numbers in a microbial population. *J. Theor. Biol.* **71**, 587–603 (1978)
8. McAdams, H.H., Arkin, A.: Stochastic mechanisms in gene expression. *Proc. Natl. Acad. Sci. USA* **94**, 814–819 (1997)
9. Elowitz, M.B., Levine, A.J., Siggia, E.D., Swain, P.S.: Stochastic gene expression in a single cell. *Science* **297**, 1183–1186 (2002)
10. Cai, L., Friedman, N., Xie, X.: Stochastic protein expression in individual cells at the single molecule level. *Nature* **440**, 358–362 (2006)
11. Vendruscolo, M., Dobson, C.M.: Structural biology: Dynamic visions of enzymatic reactions. *Science* **313**, 1586–1587 (2006)
12. Pearson, H.: Cell biology: The cellular hullabaloo. *Nature* **453**, 150–153 (2008)
13. Shahrezaei, V., Swain, P.S.: The stochastic nature of biochemical networks. *Curr. Opin. Biotechnol.* **19**, 369–374 (2008)
14. Huang, S., Eichler, G., Bar-Yam, Y., Ingber, D.: Cell fate as a high-dimensional attractor of a complex gene regulatory network. *Phys. Rev. Lett.* **94**, 128701 (2005)
15. Noble, D.: *The Music of Life*. Oxford Univ. Press, London (2006)
16. Ao, P., Galas, D., Hood, L., Zhu, X.-M.: Cancer as robust intrinsic state of endogenous molecular-cellular network shaped by evolution. *Med. Hypotheses* **70**, 678–684 (2007)
17. Ge, H., Qian, H.: Nonequilibrium phase transition in mesoscopic biochemical systems: From stochastic to nonlinear dynamics and beyond. *J. R. Soc. Interf.* (2010), doi:[10.1098/rsif.2010.0202](https://doi.org/10.1098/rsif.2010.0202)
18. Dercole, F., Rinaldi, S.: *Analysis of Evolutionary Processes: The Adaptive Dynamics Approach and Its Applications*. Princeton Univ. Press, Princeton (2008)
19. Ge, H., Qian, H.: The physical origins of entropy production, free energy dissipation and their mathematical representations. *Phys. Rev. E* **81**, 051133 (2010)
20. Chen, J.: *The Physical Foundation of Economics—An Analytical Thermodynamic Theory*. World Scientific, Singapore (2005)

21. Arnold, L.: Random Dynamical Systems. Springer, New York (1998)
22. Lin, K.K., Shea-Brown, E., Young, L.-S.: Reliability of coupled oscillators. *J. Nonlinear Sci.* **19**, 497–545 (2009)
23. Kramers, H.A.: Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* **7**, 284–304 (1940)
24. von Smoluchowski, M.: Versuch einer mathematischen theorie der koagulationskinetik kolloider lösungen. *Z. Phys. Chem.* **92**, 129–168 (1917)
25. Shoup, D., Szabo, A.: Role of diffusion in ligand binding to macromolecules and cell-bound receptors. *Biophys. J.* **40**, 33–39 (1982)
26. Hänggi, P., Talkner, P., Borkovec, M.: Reaction-rate theory: Fifty years after Kramers. *Rev. Mod. Phys.* **62**, 251–341 (1990)
27. Lu, H.P., Xun, L., Xie, X.S.: Single-molecule enzymatic dynamics. *Science* **282**, 1877–1882 (1998)
28. Delbrück, M.: Statistical fluctuations in autocatalytic reactions. *J. Chem. Phys.* **8**, 120–124 (1940)
29. McQuarrie, D.A.: Stochastic approach to chemical kinetics. *J. Appl. Probab.* **4**, 413–478 (1967)
30. Gillespie, D.T.: Stochastic simulation of chemical kinetics. *Ann. Rev. Phys. Chem.* **58**, 35–55 (2007)
31. Wilkinson, D.J.: Stochastic Modeling for Systems Biology. Chapman & Hall/CRC, New York (2006)
32. Bortz, A.B., Kalos, M.H., Lebowitz, J.L.: A new algorithm for Monte Carlo simulation of Ising spin systems. *J. Comput. Phys.* **17**, 10–18 (1975)
33. Qian, H., Bishop, L.M.: The chemical master equation approach to nonequilibrium steady-state of open biochemical systems: Linear single-molecule enzyme kinetics and nonlinear biochemical reaction networks (Review). *Int. J. Mol. Sci.* **11**, 3472–3500 (2010)
34. Liang, J., Qian, H.: Computational cellular dynamics based on the chemical master equation: A challenge for understanding complexity (Review). *J. Comput. Sci. Technol.* **25**, 154–168 (2010)
35. Beard, D.A., Qian, H.: Chemical Biophysics: Quantitative Analysis of Cellular Systems. Cambridge Texts in Biomedical Engineering. Cambridge Univ. Press, New York (2008)
36. Vellela, M., Qian, H.: A quasistationary analysis of a stochastic chemical reaction: Keizer’s paradox. *Bull. Math. Biol.* **69**, 1727–1746 (2007)
37. Vellela, M., Qian, H.: Stochastic dynamics and nonequilibrium thermodynamics of a bistable chemical system: The Schlögl model revisited. *J. R. Soc. Interf.* **6**, 925–940 (2009)
38. Vellela, M., Qian, H.: On Poincaré-Hill cycle map of rotational random walk: Locating stochastic limit cycle in reversible Schnakenberg model. *Proc. R. Soc. A, Math. Phys. Eng. Sci.* **466**, 771–788 (2010)
39. Kurtz, T.G.: The relationship between stochastic and deterministic models for chemical reactions. *J. Chem. Phys.* **57**, 2976–2978 (1972)
40. Hu, G.: Lyapounov function and stationary probability distributions. *Z. Phys. B* **65**, 103–106 (1986)
41. Ge, H., Qian, H.: Thermodynamic limit of a nonequilibrium steady-state: Maxwell-type construction for a bistable biochemical system. *Phys. Rev. Lett.* **103**, 148103 (2009)
42. Qian, H.: Open-system nonequilibrium steady-state: Statistical thermodynamics, fluctuations and chemical oscillations (Feature Article). *J. Phys. Chem. B* **110**, 15063–15074 (2006)
43. Qian, H.: Phosphorylation energy hypothesis: Open chemical systems and their biological functions. *Ann. Rev. Phys. Chem.* **58**, 113–142 (2007)
44. Lewis, G.N.: A new principle of equilibrium. *Proc. Natl. Acad. Sci. USA* **11**, 179–83 (1925)
45. Jiang, D.-Q., Qian, M., Qian, M.-P.: Mathematical Theory of Nonequilibrium Steady States: On the Frontier of Probability and Dynamical Systems. Springer, New York (2004)
46. Klein, M.J.: Principle of detailed balance. *Phys. Rev.* **97**, 1446–1447 (1955)
47. Lebowitz, J.L.: Stationary nonequilibrium Gibbsian ensembles. *Phys. Rev.* **114**, 1192–1202 (1959)
48. Lax, M.: Fluctuations from the nonequilibrium steady state. *Rev. Mod. Phys.* **32**, 25–64 (1960)
49. Nicolis, G., Prigogine, I.: Self-Organization in Nonequilibrium Systems: From Dissipative Structures to Order Through Fluctuations. Wiley-Interscience, New York (1977)
50. Hill, T.L.: Free Energy Transduction in Biology: The Steady-State Kinetic and Thermodynamic Formalism. Academic Press, New York (1977)
51. Haken, H.: Synergetics, An Introduction: Nonequilibrium Phase-Transitions and Self-Organization in Physics, Chemistry and Biology. Springer, New York (1977)
52. Keizer, J.: Statistical Thermodynamics of Nonequilibrium Processes. Springer, New York (1987)
53. Ross, J.: Thermodynamics and Fluctuations Far From Equilibrium. Springer, New York (2008)
54. Ross, J.: From the determination of complex reaction mechanisms to systems biology. *Ann. Rev. Biochem.* **77**, 479–494 (2008)
55. Schultz, S.G.: Homeostasis, humpty dumpty, and integrative biology. *Physiology* **11**, 238–246 (1986)
56. Schuster, S., Schuster, R.: A generalization of Wegscheider’s condition. *J. Math. Chem.* **3**, 25–42 (1989)
57. Gardiner, C.W.: Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences, 2nd edn. Springer, New York (1985)

58. Wang, J., Xu, L., Wang, E.K.: Potential landscape and flux framework of non-equilibrium networks: Robustness, dissipation and coherence of biochemical oscillations. *Proc. Natl. Acad. Sci. USA* **105**, 12271–12276 (2008)
59. Qian, H., Elson, E.L.: Chemical fluxes in cellular steady states measured by fluorescence correlation spectroscopy. In: Gräslund, A., Rigler, R., Widengren, J. (eds.) *Single Molecule Spectroscopy in Chemistry, Physics and Biology*. Nobel Symposium: Springer Series in Chemical Physics, vol. 96, pp. 119–137. Springer, New York (2010)
60. Kauffman, S.A.: Gene regulation networks: A theory for their global structure and behavior. *Curr. Top. Dev. Biol.* **6**, 145–182 (1971)
61. Glass, L., Mackey, M.C.: *From Clocks to Chaos: The Rhythms of Life*. Princeton Univ. Press, Princeton (1988)
62. Amit, D.J.: *Modeling Brain Function: The World of Attractor Neural Networks*. Cambridge Univ. Press, London (1992)
63. Kauffman, S.A.: *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford Univ. Press, New York (1993)
64. Kepler, T.B., Elston, T.C.: Stochasticity in transcriptional regulation: Origins, consequences, and mathematical representations. *Biophys. J.* **81**, 3116–3136 (2001)
65. Hornos, J.E.M., Schultz, D., Innocentini, G.C., Wang, J., Walczak, A.M., Onuchic, J.N., Wolynes, P.G.: Self-regulating gene: An exact solution. *Phys. Rev. E* **72**, 051907 (2005)
66. Walczak, A.M., Onuchic, J.N., Wolynes, P.G.: Absolute rate theories of epigenetic stability. *Proc. Natl. Acad. Sci. USA* **102**, 18926–18931 (2005)
67. Ferrell, J., Xiong, W.: Bistability in cell signaling: How to make continuous processes discontinuous, and reversible processes irreversible. *Chaos* **11**, 227–236 (2001)
68. Qian, H., Reluga, T.C.: Nonequilibrium thermodynamics and nonlinear kinetics in a cellular signaling switch. *Phys. Rev. Lett.* **94**, 028101 (2005)
69. Bishop, L.M., Qian, H.: Stochastic bistability and bifurcation in a mesoscopic signaling system with autocatalytic kinase. *Biophys. J.* **98**, 1–11 (2010)
70. Choi, P.J., Cai, L., Frieda, K., Xie, X.S.: A stochastic single-molecule event triggers phenotype switching of a bacterial cell. *Science* **322**, 442–446 (2008)
71. Lepzelter, D., Feng, H., Wang, J.: Oscillation, cooperativity, and intermediates in the self-repressing gene. *Chem. Phys. Lett.* **490**, 216220 (2010)
72. Samoilov, M., Pilyasunov, S., Arkin, A.P.: Stochastic amplification and signaling in enzymatic futile cycle through noise-induced bistability with oscillation. *Proc. Natl. Acad. Sci. USA* **102**, 2310–2315 (2005)
73. Wolynes, P.G.: Gene regulation: Single-molecule chemical physics in a natural context. In: Gräslund, A., Rigler, R., Widengren, J. (eds.) *Single Molecule Spectroscopy in Chemistry, Physics and Biology*. Nobel Symposium: Springer Series in Chemical Physics, vol. 96, pp. 553–560. Springer, New York (2010)
74. Qian, H., Shi, P.-Z., Xing, J.: Stochastic bifurcation, slow fluctuations, and bistability as an origin of biochemical complexity. *Phys. Chem. Chem. Phys.* **11**, 4861–4870 (2009)
75. Qian, H.: Entropy demystified: The “thermo”-dynamics of stochastically fluctuating systems. *Methods Enzymol.* **467**, 111–134 (2009)
76. Ge, H.: Extended forms of the second law for general time-dependent stochastic processes. *Phys. Rev. E* **80**, 021137 (2009)
77. Esposito, M., Van den Broeck, C.: Three detailed fluctuation theorems. *Phys. Rev. Lett.* **104**, 090601 (2010)
78. Santillan, M., Qian, H.: Irreversible thermodynamics in multiscale stochastic dynamical systems. [arXiv:1003.3513](https://arxiv.org/abs/1003.3513) (2010)
79. Esposito, M., Van den Broeck, C.: The three faces of the Second Law: I. Master equation formulation. [arXiv:1005.1683](https://arxiv.org/abs/1005.1683) (2010)
80. Oono, Y., Paniconi, M.: Steady state thermodynamics. *Prog. Theor. Phys. Suppl.* **130**, 29–44 (1998)
81. Ao, P.: Emerging of stochastic dynamical equalities and steady state thermodynamics from Darwinian dynamics. *Commun. Theor. Phys.* **49**, 1073–1090 (2008)
82. Hopfield, J.J.: Physics, computation, and why biology looks so different? *J. Theor. Biol.* **171**, 53–60 (1994)
83. Qian, H.: From discrete protein kinetics to continuous Brownian dynamics: A new perspective. *Protein Sci.* **11**, 1–5 (2002)
84. Dunn, G.P., Old, L.J., Schreiber, R.D.: The three Es of cancer immunoediting. *Ann. Rev. Immunol.* **22**, 329–360 (2004)
85. Sasai, M., Wolynes, P.G.: Stochastic gene expression as a many-body problem. *Proc. Natl. Acad. Sci. USA* **100**, 2374–2379 (2002)

86. Slack, M.D., Martinez, E.D., Wu, L.F., Altschuler, S.J.: Characterizing heterogeneous cellular responses to perturbations. *Proc. Natl. Acad. Sci. USA* **105**, 19305–19310 (2008)
87. Ferrell, J.E., Machleder, E.M.: The biochemical basis of an all-or-none cell fate switch in *Xenopus* oocytes. *Science* **280**, 895–898 (1998)
88. Buckmaster, R., Asphahani, F., Thein, M., Xu, J., Zhang, M.-Q.: Detection of drug-induced cellular changes using confocal Raman spectroscopy on patterned single-cell biosensors. *Analyst* **134**, 1440–1446 (2009)
89. Xu, L., Chen, Y., Song, Q., Wang, Y., Ma, D.: PDCD4 interacts with Tip60 and functions as a cooperator in acetyltransferase activity and DNA damage-induced apoptosis. *Neoplasia* **11**, 345–354 (2009)
90. Lumry, R., Biltonen, R., Brandts, J.F.: Validity of the “two-state” hypothesis for conformational transitions of proteins. *Biopolymers* **4**, 917–944 (1966)
91. Foster, P.L.: Adaptive mutation: Has the unicorn landed? *Genetics* **148**, 1453–1459 (1998)
92. Hanahan, D., Weinberg, R.A.: The hallmarks of cancer. *Cell* **100**, 57–70 (2000)
93. Cady, F., Qian, H.: Open-system thermodynamic analysis of DNA polymerase fidelity. *Phys. Biol.* **6**, 036011 (2009)

Review

The Chemical Master Equation Approach to Nonequilibrium Steady-State of Open Biochemical Systems: Linear Single-Molecule Enzyme Kinetics and Nonlinear Biochemical Reaction Networks

Hong Qian * and Lisa M. Bishop *

Department of Applied Mathematics, University of Washington, Seattle, WA 98195, USA

* Authors to whom correspondence should be addressed;

E-Mails: hqian@amath.washington.edu (H.Q.); lbishop@amath.washington.edu (L.M.B.).

Received: 15 August 2010 / Accepted: 14 September 2010 / Published: 20 September 2010

Abstract: We develop the stochastic, chemical master equation as a unifying approach to the dynamics of biochemical reaction systems in a mesoscopic volume under a living environment. A living environment provides a continuous chemical energy input that sustains the reaction system in a nonequilibrium steady state with concentration fluctuations. We discuss the linear, unimolecular single-molecule enzyme kinetics, phosphorylation-dephosphorylation cycle (PdPC) with bistability, and network exhibiting oscillations. Emphasis is paid to the comparison between the stochastic dynamics and the prediction based on the traditional approach based on the Law of Mass Action. We introduce the difference between nonlinear bistability and stochastic bistability, the latter has no deterministic counterpart. For systems with nonlinear bistability, there are three different time scales: (a) individual biochemical reactions, (b) nonlinear network dynamics approaching to attractors, and (c) cellular evolution. For mesoscopic systems with size of a living cell, dynamics in (a) and (c) are stochastic while that with (b) is dominantly deterministic. Both (b) and (c) are emergent properties of a dynamic biochemical network; We suggest that the (c) is most relevant to major cellular biochemical processes such as epi-genetic regulation, apoptosis, and cancer immunoediting. The cellular evolution proceeds with transitions among the attractors of (b) in a “punctuated equilibrium” manner.

Keywords: chemical kinetics; chemical master equation; stochastic dynamics; biochemical reaction; systems biology

1. Introduction

Quantitative modelling in terms of mathematical equations is the foundation of modern physical sciences. If one deals with mechanical motions or electromagnetic issues of daily lives, he/she starts with Newton's Second Law or Maxwell's equations, respectively. For work on the subatomic and molecular level, we have the quantum mechanics of Heisenberg and Schrödinger for the small things, and Gibbs' statistical mechanics for large collections of particles. The last theory on the list, Gibbs' statistical thermodynamics, has been the foundation of molecular science [1]. Its applications to biological macromolecules have laid the groundwork for molecular biology [2,3].

However, it has long been recognized that Gibbs' theory can not be applied to a system outside chemical equilibrium. In this case, and when the deviations from an equilibrium are linear, Onsager's theory provides the unifying approach known as linear irreversible thermodynamics. However, cellular biologists have long been aware of that most living processes are not near an equilibrium, but far from it. This begs an answer to the question: What is the theory one should use in modelling a biochemical reaction system in its living environment?

1.1. The Chemical Master Equation (CME)

Both Gibbs' and Onsager's work have pointed to a new type of mathematics: random variables and stochastic processes. Gibbs' thermodynamic quantities with thermal fluctuations are random variables, and Onsager has used extensively Gaussian-Markov processes to describe the dynamics near an equilibrium [4,5]. This approach can be traced back to the earlier work of Einstein on Brownian motion and Smoluchowski on diffusion.

Quantitative modelling in chemical engineering has been based on the Law of Mass Action [6]. The applications of this theory to enzyme kinetics gives rise to the entire kinetic modelling of biochemical reactions for individual enzymes [7] and for enzymatic reaction systems [8]. However, the theory based on the Law of Mass Action considers no fluctuations so it sits in an odd position with respect to the more general theories of Gibbs and Onsager. It does have its strength though, it is a fully dynamic theory, while the Gibbs' is not, and it can be applied to system beyond Onsager's linear irreversibility.

Is there a theory which can embody all the above mentioned theories? It is clear that such a theory, even very imperfect, can provide great insights into the working of biochemical reaction systems in their living environment. While a consensus has not been reached, the recent rapid rise of applications of the Gillespie algorithm seems to suggest an interesting possibility.

It might be a surprise to some, but the Gillespie algorithm (GA) *is really an equation* that prescribes the dynamic trajectory of a stochastic process. Mathematical equations can come in many different forms: differential equations for continuous variables changing with time, stochastic differential equations (SDE) for the trajectories of continuous random variables changing with time, and the GA is simply the equation for the trajectory of discrete random variables changing with time. For a random variable changing with time one can either characterize it by its stochastic trajectories, as by the SDE and the GA, or one can characterize its probability distribution as a function of time. The corresponding equation for the SDE is the well-known Fokker-Planck equation, and the corresponding equation for the GA is called the *chemical master equation* (CME).

It has been mathematically shown that the CME approach is the mesoscopic version of the Law of Mass Action [8]. It is not a competing theory to the Law of Mass Action, rather, it extends the latter to the mesoscopic chemistry and biochemistry.

We do not expect our readers to have a background in the CME. For a quick introduction see Appendix. In particular, one should learn to draw the *chemical master equation graph*. See Chapter 11 of [8] and a more recent [9] for more on the CME. Discussions on the GA can be found in [10].

In this article, we shall follow the CME approach to study biochemical reaction networks. We are particularly interested in such systems situated in a “living environment”. It turns out, one of the precise defining characteristics of the environment is the amount of chemical energy pumped into the system – similar to the battery in a radio.

1.2. Nonequilibrium Steady State (NESS)

While the CME approach is a new methodological advance in modelling open (driven) biochemical systems, a new concept also arises from recent studies on open (driven) biochemical systems: the *nonequilibrium steady state* (NESS) as a mathematical abstraction of biochemical homeostasis. In terms of the CME approach to mesoscopic chemical and biochemical reaction systems there are three, and only three, types of dynamics [11]:

- (1) Equilibrium state with fluctuations which is well-understood according to Boltzmann’s law, and the theories of Gibbs, Einstein, and Onsager.
- (2) Time-dependent, transient processes in which systems are changing with time. In the past, this type of problems is often called “nonequilibrium problems”. As all experimentalists and computational modellers know, time-dependent kinetic experiments are very difficult to perform, and time-dependent equations are very difficult to analyze.
- (3) Nonequilibrium steady state: The system is no longer changing with time in a statistical sense, *i.e.*, all the probability distributions are stationary; nevertheless, the system is not at equilibrium. The systems fluctuate, but the fluctuations do not obey Boltzmann’s law. Such a system only exists when it is driven by a sustained chemical energy input. Complex deterministic dynamics discussed in the past, such as chemical bistability and oscillations, are all macroscopic limit of such systems.

To a first-order approximation, one can represent a biochemical cell or a subcellular network in homeostasis as a NESS. This is the theory being put forward by I. Prigogine, G. Nicolis and their Brussels group [12]. The NESS theory has recently gone through major development in terms of the fluctuation theorem in statistical physics, especially the stochastic version of J. Kurchan, J.L. Lebowitz and H. Spohn [13,14], and irreversible Markov processes in mathematics [15] (A deep mathematical result shows that the arrow of time is a sufficient and necessary condition for the positiveness of an appropriately defined entropy production rate.). The theory of NESS also has enjoyed great appreciation through works on molecular motors [16,17].

To have a better understanding of the nature of a NESS, we list three key characteristics of a system in equilibrium steady state: First, there is no flux in each and every reaction. This is known as the principle of detailed balance [18]. Second, the system is *time-reversible*. One can play a “recording

tape” of the system backward and will find it is statistically equivalent; There is no arrow of time [15]. The logical consequence of the above statements is that any process occurred in an equilibrium system will have equal probability to run backward. Hence nothing can be accomplished. There is no energy transduction, and there is no signal processing.

For more discussions on NESS and its applications to biochemical systems and modelling, the readers are referred to [19,20].

2. The Chemical Master Equation and Its Applications to Kinetics of Isolated Enzyme

Since enzyme kinetics is the workhorse of biochemical reaction networks, let us start with the CME approach to the standard Michaelis-Menten (MM) enzyme reaction scheme:



In the CME approach to chemical and biochemical kinetics, one no longer asks what are the concentrations of E , S , ES and P , but instead, what is the probability of the system having m number of S and n number of ES : $p(m, n, t) = \Pr\{N_S(t) = m, N_{ES}(t) = n\}$ where $N_S(t)$ and $N_{ES}(t)$ are the non-negative integer valued random variables, the number of S and ES . As functions of time, both are stochastic processes.

Assuming that the total number of substrate and product molecules is $m_0 = N_S(t) + N_P(t) + N_{ES}(t)$, and total number of enzyme molecules is $n_0 = N_E(t) + N_{ES}(t)$, we have the CME (See Section 4 for the details to obtain the equation.)

$$\begin{aligned} \frac{dp(m, n, t)}{dt} = & - \left(\hat{k}_1 m (n_0 - n) + \hat{k}_{-1} n + \hat{k}_2 n \right) p(m, n, t) \\ & + \hat{k}_1 (m + 1) (n_0 - n + 1) p(m + 1, n - 1, t) \\ & + \hat{k}_{-1} (n + 1) p(m - 1, n + 1, t) \\ & + \hat{k}_2 (n + 1) p(m, n + 1, t) \end{aligned} \quad (2)$$

($0 \leq m \leq m_0, 0 \leq n \leq n_0$)

There are three reactions in the kinetic scheme (1), hence there are six terms, three positive and three negative, in the CME (2). Note that the \hat{k} 's in the above equation are related, but not the same as the rate constants k 's in Equation (1). The latter is concentration based, and the former is number based:

$$\hat{k}_1 = \frac{k_1}{V}, \quad \hat{k}_{-1} = k_{-1}, \quad \hat{k}_2 = k_2 \quad (3)$$

2.1. Quasi-stationary approximation of the Michaelis-Menten enzyme kinetics

One of the most important results in deterministic enzyme kinetic theory is the quasi-steady state approximation leading to the well-known Michaelis-Menten equation for the production of the product in Equation (1):

$$\frac{d[S]}{dt} = -\frac{d[P]}{dt} = -\frac{k_1 k_2 [S] E_t}{k_{-1} + k_2 + k_1 [S]} \quad (4)$$

where E_t is the total enzyme concentration. We shall now carry out a parallel analysis for the CME (2).

As pointed out by Kepler and Elston [21], Rao and Arkin [22], and by Qian [16], the quasi-stationary approximation is best cast in terms of the *conditional probability*. If the dynamics in Equation (2) is such

that the changes in n can reach stationarity first, due to $n_0 \ll m_0$, then one can first solve the problem of steady state conditional distribution $p(n|m)$, then using this to solve the $p(m, t)$. This is done as follows.

In the first step, on a fast time scale for fixed m , the Equation (2) can be written as

$$\begin{aligned} \frac{dp(n, t|m)}{dt} = & - \left(\hat{k}_1 m(n_0 - n) + \hat{k}_{-1} n + \hat{k}_2 n \right) p(n, t|m) \\ & + \hat{k}_1 (m+1)(n_0 - n + 1) p(n-1, t|m) \\ & + (\hat{k}_{-1} + \hat{k}_2)(n+1) p(n+1, t|m) \end{aligned} \quad (0 \leq m \leq m_0, 0 \leq n \leq n_0) \quad (5)$$

Here we assumed $m-1 \approx m$. This immediate gives us the conditional stationary state distribution for n :

$$p^{ss}(n|m) = \frac{n_0!}{n!(n_0 - n)!} \frac{(\hat{k}_{-1} + \hat{k}_2)^{n_0-n} (\hat{k}_1 m)^n}{(\hat{k}_{-1} + \hat{k}_2 + \hat{k}_1 m)^{n_0}} \quad (0 \leq n \leq n_0) \quad (6)$$

which yields a mean value for n , with given m :

$$\langle n \rangle(m) = \sum_{n=0}^{n_0} n p^{ss}(n|m) = \frac{\hat{k}_1 m n_0}{\hat{k}_{-1} + \hat{k}_2 + \hat{k}_1 m} \quad (7)$$

This result agrees exactly with the deterministic model.

Now the second step, let us sum over all the n for the Equation (2). With

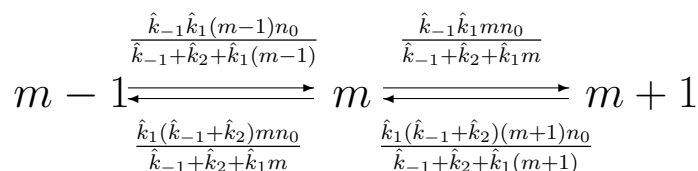
$$\sum_n p(m, n, t) = p(m, t) \quad \text{and} \quad p(m, n, t) \approx p(m, t) p^{ss}(n|m)$$

we have

$$\begin{aligned} \frac{dp(m, t)}{dt} = & - \left(\hat{k}_1 m(n_0 - \langle n \rangle(m)) + \hat{k}_{-1} \langle n \rangle(m) \right) p(m, t) \\ & + \hat{k}_1 (m+1)(n_0 - \langle n \rangle(m+1)) p(m+1, t) \\ & + \hat{k}_{-1} \langle n \rangle(m-1) p(m-1, t) \end{aligned} \quad (0 \leq m \leq m_0, 0 \leq n \leq n_0) \quad (8)$$

Equation (8) is the result of a quasi-stationary approximation. It should be compared with the deterministic equation (4). It can be graphically represented as in Figure 1.

Figure 1. Through quasi-stationary approximation, the CME in Equation (2), represented by the two-dimensional graph in Figure 6, is reduced to the one-dimensional system shown here. The corresponding master equation is shown in Equation (8).



We see that at any given time, S can increase or decrease by one. The decrease of one S is from either $ES \rightarrow E + P$ or $ES \rightarrow E + S$, with probability $\frac{\hat{k}_2}{\hat{k}_{-1}+\hat{k}_2}$ and $\frac{\hat{k}_{-1}}{\hat{k}_{-1}+\hat{k}_2}$, respectively. Hence, the number of P will increase, at any given time, with the rate of

$$\frac{\hat{k}_1 \hat{k}_2 m n_0}{\hat{k}_{-1} + \hat{k}_2 + \hat{k}_1 m} \quad (9)$$

This is exactly the CME version of Equation (4) [22]. The result shows that the waiting time between consecutive arrival of P is exponentially distributed. This surprising result can be best understood from a mathematical theorem on the superposition of N identical, independent renewal processes [23]. For a more mathematical discussion on the subject, see [24].

Even when the number of enzymes is not large, the product arrival time distribution contains no information more than the traditional Michaelis-Menten rate constant. However, this is not the case if there is truly only a single enzyme. This will be discussed below.

2.2. Single-Molecule Michaelis-Menten Enzyme Kinetics

Now in Equation (2), let us consider $n_0 = 1$. Then the equation is reduced to

$$\frac{dp(m, 0, t)}{dt} = -\hat{k}_1 m p(m, 0) + \hat{k}_{-1} p(m-1, 1) + \hat{k}_2 p(m, 1) \quad (10a)$$

$$\frac{dp(m, 1, t)}{dt} = -(\hat{k}_{-1} + \hat{k}_2) p(m, 1) + \hat{k}_1 (m+1) p(m+1, 0) \quad (10b)$$

This is the CME for a single molecule enzyme kinetics according to the MM in Equation (1). Often, one is only interested in the conformational states of the enzyme:

$$p_E(t) = \sum_{m=0}^{m_0} p(m, 0, t), \quad p_{ES}(t) = \sum_{m=0}^{m_0} p(m, 1, t)$$

Carry out the summation on the both sides of Equation (10), we have

$$\frac{dp_E(t)}{dt} = -\hat{k}_1 \langle N_S \rangle p_E(t) + (\hat{k}_{-1} + \hat{k}_2) p_{ES}(t) \quad (11a)$$

$$\frac{dp_{ES}(t)}{dt} = -(\hat{k}_{-1} + \hat{k}_2) p_{ES}(t) + \hat{k}_1 \langle N_S \rangle p_E(t) \quad (11b)$$

where

$$\langle N_S \rangle = \frac{\sum_{m=0}^{m_0} m p(m, 0, t)}{\sum_{m=0}^{m_0} p(m, 0, t)}, \quad \hat{k}_1 \langle N_S \rangle = k_1 c_S \quad (12)$$

c_S is the concentration of S . Equation (11) is the stochastic model for a single enzyme molecule dynamics.

The steady state probability for the single enzyme can be easily obtained from Equation (11):

$$p_E^{ss} = \frac{k_{-1} + k_2}{k_1 c_S + k_{-1} + k_2}, \quad p_{ES}^{ss} = \frac{k_1 c_S}{k_1 c_S + k_{-1} + k_2} \quad (13)$$

Then the steady state single enzyme turnover flux is

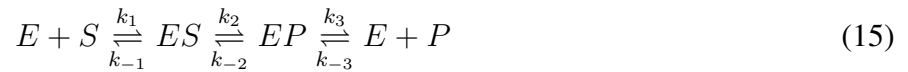
$$J^{ss} = k_2 p_{ES}^{ss} = \frac{k_1 k_2 c_S}{k_1 c_S + k_{-1} + k_2} = \frac{V_{max} c_S}{K_M + c_S} \quad (14)$$

with $k_M = \frac{k_{-1} + k_2}{k_1}$ and $V_{max} = k_2$. The last expression is precisely the Michaelis-Menten formula.

The single enzyme steady-state flux J^{ss} is exactly the inverse of the mean time duration between two product arrivals. The *time perspective* is natural for single-molecule measurements on enzyme turnover [23]. Not only one can obtain the mean time duration, one can also, according to the stochastic model in Equation (11), compute the probability density function (pdf) of the time duration between two product arrivals. The pdf is exponentially distributed if there is a single *rate-limiting step* within the enzyme cycle. In general, however, it is not exponentially distributed.

2.3. Driven Enzyme Kinetics

We now consider an enzyme kinetic scheme that is a little more complex than that in Equation (1):



With concentrations c_S and c_P for S and P being constant, as many cases in a living cell under homeostasis, we have the steady state single enzyme turnover flux

$$J^{ss} = \frac{k_1 k_2 k_3 c_S - k_{-1} k_{-2} k_{-3} c_P}{k_2 k_3 + k_{-2} k_{-1} + k_3 k_{-1} + (k_1 k_2 + k_1 k_{-2} + k_3 k_1) c_S + (k_{-1} k_{-3} + k_2 k_{-3} + k_{-3} k_{-2}) c_P} \quad (16)$$

The origin of this flux is the non-equilibrium between the chemical potentials of S and P :

$$\Delta\mu = \mu_S - \mu_P = \mu_S^o + k_B T \ln c_S - \mu_P^o - k_B T \ln c_P = k_B T \ln \frac{k_1 k_2 k_3 c_S}{k_{-1} k_{-2} k_{-3} c_P} \quad (17)$$

We see that when $\Delta\mu > 0$, $J^{ss} > 0$, when $\Delta\mu < 0$, $J^{ss} < 0$, and when $\Delta\mu = 0$, $J^{ss} = 0$. In fact, the product $\Delta\mu \times J^{ss}$ is the amount of chemical energy input to the single enzyme. If the enzyme does not do any mechanical work such as a motor protein, then all this energy becomes heat and dissipated into the aqueous solution in which the enzymatic reaction occurs.

Let us see an example of J^{ss} as function of $\Delta\mu$, for $k_i = \lambda$, $k_{-i} = 1$, $i = 1, 2, 3$, and $c_P = 1$ while changing the c_S :

$$J^{ss} = \frac{\lambda^3 c_S - 1}{(3 + 2\lambda + \lambda^2) + (1 + 2\lambda)\lambda c_S} = \frac{\lambda^2 (e^{\Delta\mu/k_B T} - 1)}{(3 + 2\lambda + \lambda^2)\lambda^2 + (1 + 2\lambda)e^{\Delta\mu/k_B T}} \quad (18)$$

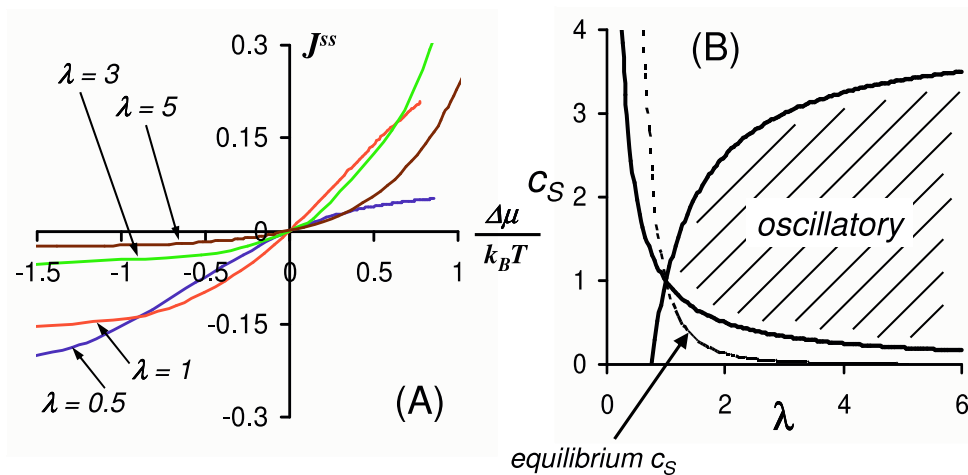
Figure 2A shows several curves. We see that their relationship is not linear over the entire range of $\Delta\mu$. Only when the $\Delta\mu$ is very small, there is a linear region $J^{ss} = \Delta\mu/k_B T (\lambda^4 + 2\lambda^3 + 3\lambda^2 + 2\lambda + 1)$. This is the region where Onsager's theory applies. In fact, the linear coefficient between $\Delta\mu$ and J^{ss} is precisely the *one-way flux* in the equilibrium. To show this, we note from Equations (16) and (17) that $J^{ss} = J^+ - J^-$ and $\Delta\mu = k_B T \ln(J^+/J^-)$. Then, when $J^{ss} \ll J^-$, we have

$$J^{ss} = J^+ - J^- = J^- (e^{\Delta\mu/k_B T} - 1) = \frac{J^-}{k_B T} \Delta\mu \quad (19)$$

Note that in equilibrium, $J^+ = J^-$. The last equation is known as *detailed balance*, which plays a central role in Onsager's theory.

Therefore, the simple enzyme kinetics is not in the region with linear irreversibility. Onsager's theory does not apply. Interestingly, we also note that the nonlinear curves in Figure 2A very much resemble the current-voltage characteristics of a semiconductant diode. It will be interesting to further explore the similarities between driven biochemical and electronic systems [25].

Figure 2. Simple enzyme kinetic system in Equation (15) exhibits nonlinear flux (J^{ss})-potential ($\Delta\mu$) relation and oscillatory behavior. The parameters used: $k_1 = k_2 = k_3 = \lambda$, $k_{-1} = k_{-2} = k_{-3} = 1$, and $c_P = 1$. **(A)** The steady state flux J^{ss} is given in Equation (18), and the chemical potential $\Delta\mu$ is given by Equation (17): $\Delta\mu/k_B T = \ln(\lambda^3 c_S)$. Each curve is obtained by fixed λ , as indicated, with varying c_S . **(B)** The region of parameter values for λ and c_S in which there are complex eigenvalues is given in Equation (20). The dashed line represents the equilibrium c_S^{eq} , which is outside the oscillatory region.



2.4. Far from Equilibrium and Enzyme Oscillations

In fact, one of the most important results in Onsager's linear theory is the *reciprocal relations* [26,27] which, based on the principle of detailed balance, dictates certain symmetry in the kinetics. One of the consequences of this symmetry is that chemical kinetics near equilibrium can not oscillate. Oscillatory kinetics are associated with the complex eigenvalues of the kinetics system. For the scheme in Equation (15), the imaginary component of its eigenvalues is $\sqrt{4\Delta - \Lambda^2}$, where

$$\Lambda = k_1 c_S + k_2 + k_3 + k_{-1} + k_{-2} + k_{-3} c_P$$

$$\Delta = k_2 k_3 + k_{-2} k_{-1} + k_3 k_{-1} + (k_1 k_2 + k_1 k_{-2} + k_3 k_1) c_S + (k_{-1} k_{-3} + k_2 k_{-3} + k_{-3} k_{-2}) c_P$$

To see an example, let us again consider $k_1 = k_2 = k_3 = \lambda$, and $k_{-1} = k_{-2} = k_{-3} = 1$, and $c_P = 1$. Then

$$\sqrt{4\Delta - \Lambda^2} = \sqrt{(3 - 4\lambda) + 2\lambda(2\lambda - 1)c_S - \lambda^2 c_S^2}$$

The oscillations exist for

$$\frac{1}{\lambda} < c_S < \frac{4\lambda - 3}{\lambda}, \text{ when } \lambda > 1 \text{ and } \frac{4\lambda - 3}{\lambda} < c_S < \frac{1}{\lambda} \text{ when } \lambda < 1 \quad (20)$$

We see from Figure 2B that in general the c_S has to be sufficiently far away from its equilibrium value in order to have the oscillation. Chemical and biochemical oscillation is a far-from-equilibrium phenomenon [28,29].

3. The CME Approach to Nonlinear Biochemical Networks in Living Environment

In a living cell, one of the most important, small biochemical regulatory networks is the phosphorylation-dephosphorylation cycle (PdPC) of an enzyme, first discovered by E.H. Fischer and E.G. Krebs in 1950s. It consists of only three players: a substrate enzyme, a kinase and a phosphatase. The phosphorylation of the substrate protein E , $ATP + E \xrightarrow{k_1} ADP + E^*$, is catalyzed by the kinase K , and its dephosphorylation $E^* \xrightarrow{k_2} E + Pi$ is catalyzed by the phosphatase P . Even though it is traditionally called *reversible chemical modification*, one should note that these two steps are different chemical reactions. In fact, a kinase should also catalyze the reaction $ADP + E^* \xrightarrow{k_{-1}} ATP + E$, as should the phosphatase for $E + Pi \xrightarrow{k_{-2}} E^*$. These latter two reactions are simply too slow, even in the presense of the respective enzymes, to be noticed, but they definitely can not be zero. The proof is that the complete of a PdPC is the hydrolysis of a single ATP to $ADP + Pi$. This reaction has an equilibrium constant of 4.9×10^5 M [30], which means

$$\frac{k_1 k_2}{k_{-1} k_{-2}} = K_{ATP} = 4.9 \times 10^5 \quad (21)$$

3.1. Phosphorylation-Dephosphorylation Cycle with Autocatalysis: A Positive Feedback Loop

Many kinase itself can exist in two different forms: an inactive state and an active state. Furthermore, the conversion from the former to the latter involves the binding of the E^* , sometime one, sometime two. Hence, we have [31–33]:



where $\chi = 1, 2$. We shall call $\chi = 1$ first-order autocatalysis and $\chi = 2$ second-order autocatalysis. Therefore, if the conversion is rapid, then the active kinase concentration is $[K^\ddagger] = K_a [K] [E^*]^\chi$. Now combining the reaction in Equation (22) with the PdPC, such a mechanism is called autocatalysis: more E^* is made, more K^\ddagger , which in turn to make more E^* . Quantitatively, the rate of phosphorylation reaction catalyzed by the active kinase is:

$$\frac{d[E^*]}{dt} = k_1 [ATP] [K^\ddagger] [E] = k_1 [ATP] K_a [K] [E^*]^\chi [E] \quad (23)$$

where $[X]$ denotes the concentration of biochemical species X . Note, however, that the same kinase K^\ddagger also catalyzed the reverse reaction of the phosphorylation. Hence, to be more realistic, we have

$$\begin{aligned} \frac{d[E^*]}{dt} &= K_a [K] [E^*]^\chi (k_1 [ATP] [E] - k_{-1} [ADP] [E^*]) \\ &= \alpha [E^*]^\chi [E] - \beta [E^*]^{\chi+1} \end{aligned} \quad (24)$$

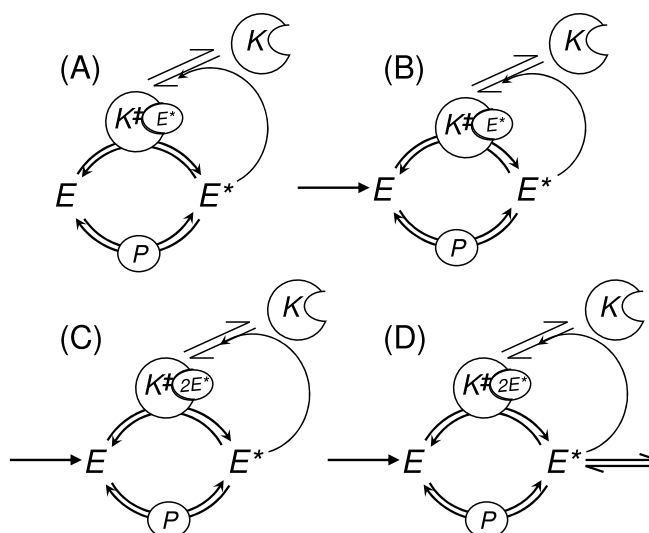
in which

$$\alpha = k_1 K_a [ATP] [K] \quad \text{and} \quad \beta = k_{-1} K_a [ADP] [K]$$

contains the concentration of ATP and ADP respectively. Equation (24) is the kinetic equation for the phosphorylation reaction catalyzed by a kinase which is activated by binding χ number of E^* .

Figure 3 shows four, with subtle differences, PdPC with such a positive feedback loop. Biochemical examples of this type of regulation are MPAK pathway [31], Src Family kinase pathway [32], and Rabaptin-5 mediated Rab5 activation in endocytosis [33]. We shall now establish the appropriate kinetic equations for each of these nonlinear biochemical networks.

Figure 3. An assorted variations of the PdPC with autocatalytic feedback. The phosphorylation of the substrate E to E^* is catalyzed by an active form of the kinase K^\ddagger , and the dephosphorylation is catalyzed by a phosphatase (P). The activation of the kinase involves the binding of K to E^* . In (A) and (B) the autocatalysis is first order: $K + E^* \rightleftharpoons K^\ddagger$; In (C) and (D), it is second order: $K + 2E^* \rightleftharpoons K^\ddagger$. The nonlinear feedback in the latter is stronger; thus they exhibit more pronounced nonlinear behavior: bistability and limit cycle oscillation.



3.2. Stochastic Bistability in PdPC with First-Order Autocatalysis

For the kinetic scheme in Figure 3A, we have $\chi = 1$ and

$$\frac{dx}{dt} = \alpha x(x_t - x) - \beta x^2 - \varepsilon x + \delta(x_t - x) \quad (25)$$

here we use x to denote the $[E^*]$, $x_t = [E] + [E^*]$. In the equation

$$\varepsilon = k_2[P] \quad \text{and} \quad \delta = k_{-2}[P][Pi]$$

represent the rates for the dephosphorylation and the rate for its reverse reaction, respectively. Both are catalyzed by the enzyme phosphatase P . For simplicity, we assume both kinase and phosphatase are operating in their linear region.

Bishop and Qian [34] have carefully studied Equation (25). While this model is very simple, the issues arise from the model are important, and have not been widely discussed. It is well-known, and as we shall discuss in Section 3.4, nonlinear open chemical and biochemical reaction systems can exhibit bistability, which plays a crucial role in cellular genetic [35] and signal regulations [20,36]. In fact, it

has been argued that bistability is one of the key origins that generate complex dynamic behavior [37]. Bistable chemical reaction systems have been extensively studied in the past [38]. In fact, it is relatively easy to theoretically construct reaction schemes that show bistability and bifurcation. Since bistability mathematically means two stable and one unstable fixed points in the positive quadrant, it is easy to show, according to the Law of Mass Action, that a tri-molecular reaction (as a reduced mechanism for multisteps of bimolecular reactions) is necessary.

In [34], however, we discovered a simpler bi-molecular chemical reaction system that possibly exhibits “bistability”. The bistability is in quotation marks since the mechanism is very different from that in traditional nonlinear reactions. The system is modelled in terms of a CME, and the bistability and (saddle-node) bifurcation are purely stochastic phenomenon. They only occur in reaction systems with small volume and small number of molecules.

3.2.1. Deterministic Kinetics of PdPC with First-Order Autocatalysis and Delayed Onset

Let y be the concentration ratio of x/x_t , the fraction of the substrate enzyme in the phosphorylated state. Also introduce nondimensional variables and parameters

$$\tau = (\alpha + \beta)x_t t, \quad a_1 = \frac{\alpha x_t - \varepsilon - \delta}{(\alpha + \beta)x_t}, \quad a_0 = \frac{\delta}{(\alpha + \beta)x_t}$$

then Equation (25) can be simplified as

$$\frac{dy}{d\tau} = -y^2 + a_1 y + a_0 \quad (26)$$

Let

$$\lambda_{1,2} = \frac{a_1 \pm \sqrt{a_1^2 + 4a_0}}{2} \quad (27)$$

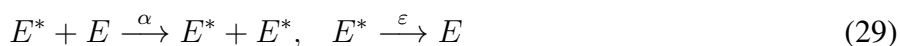
and λ_1 be the one of the two roots $\in (0, 1)$. Since $\lambda_1 \lambda_2 = -a_0 < 0$, $\lambda_2 < 0$

The solution to Equation (26) is

$$x(\tau) = \frac{\lambda_1(x_o - \lambda_2) - \lambda_2(x_o - \lambda_1)e^{(-\lambda_1 + \lambda_2)\tau}}{(x_o - \lambda_2) - (x_o - \lambda_1)e^{(-\lambda_1 + \lambda_2)\tau}} \quad (28)$$

in which $x_o = x(0)$, $x(\infty) = \lambda_1$

If both phosphorylation and dephosphorylation reactions are irreversible, as usually assumed in cell biology (When considering kinetics, but not thermodynamics, this is indeed valid for large ATP hydrolysis free energy in a living cell), then the reaction is simplified to



where α and ε are proportional the the kinase and phosphatase activity, respectively. The differential equation in Equation (25) is simplified to

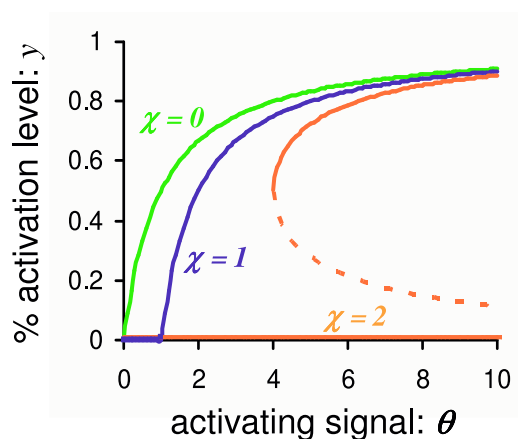
$$\frac{dy}{d\tau} = \alpha x_t y(1 - y) - \varepsilon y \quad (30)$$

Its steady state exhibits a transcritical bifurcation as a function of the activation signal, $\theta = \alpha x_t / \varepsilon$:

$$y = \begin{cases} 0 & 0 \leq \theta \leq 1 \\ 1 - \frac{1}{\theta} & \theta \geq 1 \end{cases} \quad (31)$$

Compared with the hypobolic activation curve $\frac{\theta}{1+\theta}$, Equation (31) exhibits “delayed onset” of activation [33,39]. See Figure 4. Note the curve (31) is an extreme version of a sigmoidal shape. It has a response coefficient of 9, and a Hill’s coefficient of 2. Recall that the response coefficient is defined as $\theta_{0.9}/\theta_{0.1}$, where $y(\theta_{0.9}) = 0.9$ and $y(\theta_{0.1}) = 0.1$.

Figure 4. Activation curves of PdPC with or without autocatalytic phosphorylation $E + \chi E^* \rightarrow E^* + \chi E^*$ and dephosphorylation $E^* \rightarrow E$. Curve with $\chi = 0$ is the standard hyperbolic activation without feedback: $y = \frac{\theta}{1+\theta}$. Curve with $\chi = 1$ is for the PdPC with first-order autocatalysis, following Equation (31). It exhibits an extreme version of sigmoidal shape called *delayed onset*. Curve labelled $\chi = 2$, following Equation (40), is for PdPC with second-order autocatalysis. It shows bistability when $\theta > 4$, with the dotted branch being unstable.



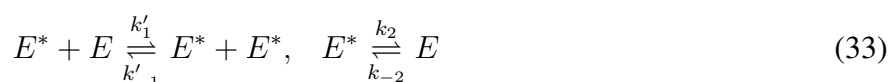
It is interesting to point out that the curve in Equation (31), the delayed onset, can be obtained from a completely different mechanism for PdPC with multiple phosphorylation sites [40]. Assuming that there is a sequential phosphorylation of sites with rate α and dephosphorylation rate β , and there are totally n sites. The active form of the substrate enzyme requires full n -sites phosphorylation. Then

$$y = \frac{\hat{\theta}^n}{1 + \hat{\theta} + \hat{\theta}^2 + \dots + \hat{\theta}^n} = \frac{\hat{\theta}^n(1 - \hat{\theta})}{1 - \hat{\theta}^{n+1}} \quad (32)$$

where $\hat{\theta} = \alpha/\beta$. One can easily show that if $n \rightarrow \infty$, the $y(\hat{\theta})$ will be precisely the one in Equation (31). Both mechanisms lead to the same mathematical expression of the activation curve.

3.2.2. Stochastic Bistability and Bifurcation without Deterministic Counterpart

Consider the first order autocatalytic system from Equation (29), adding the appropriate reverse reactions such that the system can be considered in a thermodynamic framework, yields



The stochastic model of this system was studied in depth by Bishop and Qian, [34]. The appropriate CME, where $N(t)$ is the random variable measuring the number of phosphorylated E^* molecules and N_t is the total number of kinase molecules, is

$$\begin{aligned} \frac{dp(n,t)}{dt} = & -[k_1 n(N_t - n) + k_{-1} n(n-1) + k_2 n + k_{-2}(N_t - n)] p(n,t) \\ & + [k_1(n-1)(N_t - n + 1) + k_{-2}(N_t - n + 1)] p(n-1,t) \\ & + [k_{-1}(n+1)n + k_2(n+1)] p(n+1,t) \end{aligned} \quad (34)$$

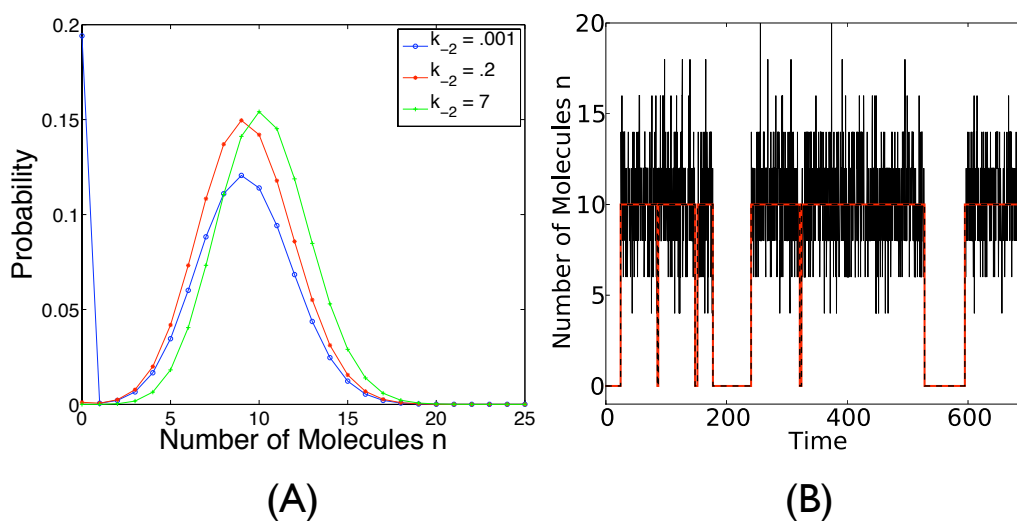
where $k_{\pm 1} = k'_{\pm 1}/V$.

Solving $\frac{dp(n,t)}{dt} = 0$ leads to the steady state distribution

$$p_{ss}(n) = C \prod_{j=0}^{n-1} \frac{(k_1 j + k_{-2})(N_t - j)}{(k_{-1} j + k_2)(j + 1)} \quad (35)$$

where C is a normalization constant. For certain parameter regimes this distribution is bimodal where the bimodality appears as a sudden second peak at zero, Figure (5). This bimodal distribution is related to traditional deterministic dynamics by considering the peaks of the probability to correspond to stable steady states, and the troughs to correspond to unstable steady states. Figure (5B) shows how this unique instance of bi-molecular bistability is related to zero being almost an absorbing state.

Figure 5. (A) The steady state distribution of the number of active kinase, N , from Equation 35. For certain parameter values the distribution is bimodal with the second peak appearing at zero. Parameter values are $k_1 = 5$, $k_{-1} = 10$, $k_2 = 10$, $N_t = 30$ and k_{-2} varied. (B) Sample trajectory of the fluctuating E^* in Equation 33 generated using the Gillespie Algorithm with parameter values, $k_1 = 5$, $k_{-1} = 10$, $k_2 = 10$, $k_{-2} = 0.001$, $N_t = 30$. For each segment of nonzero fluctuations the average was taken and plotted (*dashed line*). Data taken from and figure redrawn based on [34].



Note that this bistability is a purely stochastic phenomenon; it has no deterministic counterpart. The deterministic model of the same bi-molecular system in Equation (30) and Figure (4) has only a weak

(quadratic) nonlinearity and has no capacity bistability. Bishop and Qian showed that this stochastic bistability explains the more complex instance of the noise induced bistability first discovered in [41].

The extrema of Equation (35) can be conditioned on both the volume, V , and the energy, $\gamma = (k_1 k_2)/(k_{-1} k_{-2})$, of the system. If we consider V to be the bifurcation parameter we can find that for $0 < k'_{-1}/(k'_1 E_t - k_2 - k_{-2}) < V < k_2/(k_{-2} E_t)$ the system is bistable. Letting γ be the bifurcation parameter we find $\gamma > k_2(k_{-1} + k_2 + k_{-2})/(N_t k_{-1} k_{-2})$ with no upper bound, *i.e.*, a minimal energy input is necessary. These bounds with the parameters from Figure (5) clearly demonstrates that the stochastic bistability is dependent on having a sufficiently small volume and the presence of sufficiently large energy dissipation.

3.3. Keizer's Paradox

For the kinetic scheme in Figure 3B, we again have $\chi = 1$. However, we assume that there is a continuous biosynthesis and degradation for the E such that its concentration is sustained in the biochemical system, say at the value of a . Then, the kinetic equation for the dynamics of $[E^*]$ becomes

$$\frac{dx}{dt} = \alpha a x - \beta x^2 - \varepsilon x + \delta a \quad (36)$$

When $\delta = 0$, Equation (36) is the same equation for the generic chemical reaction scheme



J. Keizer studied this model in [42] to illustrate a very interesting observation: While the deterministic kinetics of the system has a positive steady state, the steady state of its stochastic kinetics is zero. Vellela and Qian have studied this Keizer's "paradox" [43]. It was shown that there are two very different time scales in the stochastic model: In a rather short time scale corresponding to the eigenvalues $|\lambda_2|$ and above, the system rapidly settles to a quasi-stationary distribution peaking at the deterministic positive steady state. However, in a much slower time scale corresponding to the eigenvalue $|\lambda_1|$, the above probability distribution slowly decay to zero. For very large reaction system volume V , $\lambda_1 \sim -e^{-cV}$ where c is a positive constant. Hence there is an exponentially slow decay process beyond the infinite time of the deterministic dynamics [44,45].

Keizer's paradox and its resolution is the origin of all the multi-scale dynamics in the CME system with multi-stability. It is also clear it is intimately related to the stochastic bistability in Section 3.2.2 when the k_{-2} , *i.e.*, δ in Equation (36), is very small but nonzero. k_{-2} controls the lifetime, *i.e.*, relative probability of the the zero state in Figure 5.

3.4. Schlögl's Nonlinear Bistability and PdPC with Second-Order Autocatalysis

For the kinetic scheme in Figure 3C with second-order autocatalysis, we have $\chi = 2$. We again assume that there is a continuous biosynthesis and degradation for the E such that its concentration is sustained in the biochemical system at a constant value of a . Then, the kinetic equation for the dynamics of $x = [E^*]$ becomes

$$\frac{dx}{dt} = \alpha a x^2 - \beta x^3 - \varepsilon x + \delta a \quad (38)$$

On the other hand, if we assume the rate of biosynthesis is negligible, and that both kinase and phosphatase catalyzed reactions are irreversible, then we have the kinetics



Comparing this system with that in Equation (29), the difference is in the $2E^*$ on the left-hand-side. The kinetic equation for the fraction of E^* , $y = \frac{[E^*]}{E_t}$ where E_t is the total amount of $[E] + [E^*]$. is $\frac{dy}{dt} = \alpha E_t^2 y^2 (1 - y) - \varepsilon y$. The steady state exhibits a saddle-node bifurcation at $\theta = \frac{\alpha E_t^2}{\varepsilon} = 4$:

$$y = \begin{cases} 0 & 0 \leq \theta \leq 4 \\ 0, \frac{\theta \pm \sqrt{\theta^2 - 4\theta}}{2\theta} & \theta \geq 4 \end{cases} \quad (40)$$

See the orange curve in Figure 4.

Equation (38) is precisely the same kinetic equation, according to the Law of Mass Action, for the chemical reaction system



The system (41) is known as Schlögl's model. It is the canonical example for nonlinear chemical bistability and bifurcation which has been studied for more than 30 years [46].

Qian and Reluga [47] have studied a system very similar to Equation (41) in terms of deterministic, nonlinear bifurcation theory. In particular, they established the important connection between the nonlinear bistability with nonequilibrium thermodynamics [48]. They have shown that if the concentrations of A and B are near equilibrium,

$$\left(\frac{c_B}{a}\right)^{eq} = \frac{c_B \alpha \varepsilon}{\beta \delta a}, \quad i.e., \quad \frac{\beta \delta}{\alpha \varepsilon} = \frac{k_{-1} k_{-2} [ADP][Pi]}{k_1 k_2 [ATP]} = 1 \quad (42)$$

then there would be no bistability. The last equation in (42) is precisely equivalent to free energy change of ATP hydrolysis being zero. It can be easily shown, see Section 4.1, with the equilibrium condition in Equation (42), the system has only a single, unique deterministic steady state. And also, in terms of its CME, a single peak in the equilibrium probability for the number of X . This result is much more general for all nonlinear chemical and biochemical reaction systems, not just limited to the simple reaction system in (41) [36,49,50].

Vellela and Qian [36] have recently studied the Schlögl system in great detail, with a nonequilibrium steady state perspective. In particular, it was shown that the nonlinear bistability is intimately related to nonequilibrium phase transitions in statistical physics [36,51,52]. Ge and Qian [45,51] further investigated the steady state distribution according to the CME and its relationship to the steady states according to the deterministic Law of Mass Action. They have shown that in the limit of system's volume tends infinity, *i.e.*, the so called thermodynamic limit, the CME steady state(s) differ from that of deterministic model: A *Maxwell construction* like result is obtained: According to the CME, only one of the two deterministic stable fixed point is the true global minimum, the other stable fixed point is only metastable. Hence in the thermodynamic limit, the global minimum has probability 1 while the metastable minimum has probability 0. However, the lifetime of the metastable state is infinitely long. Furthermore, using the mathematical tool of large deviation theory, [45] shows that the bistable CME system exhibits several key characteristics of nonequilibrium phase transition well-known in condensed matter physics.

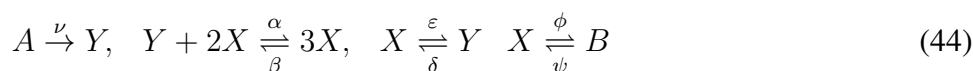
3.5. Schnakenberg's Oscillation

For the kinetic scheme in Figure 3D, we again have $\chi = 2$, and we again assume that there is a continuous biosynthesis for the E . However, we now consider the dynamics of both $[E^*]$ and $[E]$, denoted by x and y , respectively. Then, the kinetic equations becomes

$$\frac{dx}{dt} = \alpha y x^2 - \beta x^3 - \varepsilon x + \delta y - \phi x + \psi b \quad (43a)$$

$$\frac{dy}{dt} = -\alpha y x^2 + \beta x^3 + \varepsilon x - \delta y + \nu a \quad (43b)$$

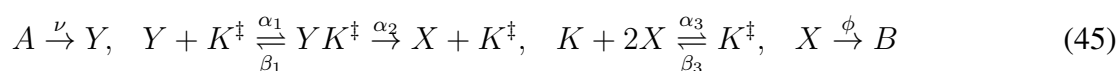
The system of Equations (43) is the same system for the kinetic scheme



with $[A] = 1$ and $[B] = 1$. If the $\beta = \varepsilon = \delta = 0$, then it becomes the celebrated Schnakenberg model which is well-known to exhibit periodic chemical oscillation. Qian *et al.* [53] first studied its stochastic dynamics in terms of the CME. Recently, Vellela and Qian [54] again have studied this system. In particular, they have introduced a novel mathematical concept of Poincaré-Hill cycle map (PHCM) to characterize the amplitude of rotational random walk. The PHCM combines the Poincaré map from nonlinear dynamic analysis [55] with the cycle kinetic analysis developed by T.L. Hill [56,57].

3.5.1. Sel'kov-Goldbeter-Lefever's Glycolytic Oscillator

So far, we have always assumed that the kinase catalysis is in its linear region, and avoided using Michaelis-Menten kinetic model for the kinase catalyzed phosphorylation. If we take the nonlinear Michaelis-Menten kinetics into account, interestingly, we discover that in this case, our model of PdPC with feedback in Figure 3D is mathematically closely related to a well-known metabolic oscillator: The Sel'kov-Goldbeter-Lefever model for glycolytic oscillation [58–60]:



In the glucolytic model, X and Y are ADP and ATP, K and K^{\ddagger} are the inactive and activated form of phosphofructokinase-1. One can find a nice nonlinear analysis of the deterministic model based on the Law of Mass Action in [60], which shows limit-cycle oscillation. As far as we know, a stochastic analysis of this model has not been carried out.

4. Conclusions

Nonlinear chemical reactions are the molecular basis of cellular biological processes and functions. Complex biochemical reactions in terms of enzymes and macromolecular complexes form “biochemical networks” in cellular control, regulation, and signaling. One of the central tasks of cellular systems biology is to quantify and integrate experimental observations into mathematical models that first reproduce and ultimately predict laboratory measurements. This review provides an introduction of the biochemical modeling paradigm in terms of the chemical master equation (CME) and explores the dynamical possibilities of various biochemical networks by considering models of homogenous,

i.e., well-mixed, reaction systems with one and two dynamic variables. From mathematical modeling perspective, these are one- and two-dimensional system, the simplest to be fully explored with sufficient depth.

The chemical master equation is a comprehensive mathematical theory that quantitatively characterizes chemical and biochemical reaction system dynamics [38,61]. Traditional chemical kinetics based on the Law of Mass Action, in terms of the concentrations of species as functions of time and differential equations, is appropriate for reaction systems in aqueous solutions [62,63]. Deterministic differential equation models have given satisfactory predictions for well mixed macroscopic chemical reaction systems. One of the most celebrated examples is the Oregonator: the mathematical theory for the Belousov-Zhabotinsky reactions [64] which display sustained oscillations in a test tube. For a recent study see [28,29].

In recent years, due to the technological advances in optical imaging, single cell analysis, and green fluorescence proteins, experimental observations of biochemical dynamics inside single living cells have become increasingly quantitative [65]. Mathematical modeling of biochemical reaction systems in a living cell requires a different approach. Chemical systems inside a cell, especially those of signaling networks involving transcription regulation, protein phosphorylation and GTPases, often involve a small number of molecules of one or more of the reactants [9,21,66,67]. Such dynamics are usually nonlinear and stochastic, exhibiting random fluctuations. Thus, the traditional method of ordinary differential equations is inappropriate. The fluctuations in the number of molecules, often called “intrinsic noise”, have been shown to have biological significance and contribute to the function of the system [41,68].

Reaction kinetics of this kind are more realistically described by stochastic models that emphasize the discrete nature of molecular reactions and the randomness of their occurrences [61]. The chemical master equation is a class of discrete-state, continuous-time Markov jump processes, known as multi-dimensional birth-death processes in probability theory [69]. Master equation is the widely used name in the physics literature [70]. In a jump process, the chemical reactions are characterized in terms of the stochastic copy numbers of the various dynamic chemical species, which differs from the traditional concentrations by a trivial volume V of the reaction system. Reactions occur at random times with exponential distribution. The expected value for the waiting period between each reaction is determined by the number of copies of each species. The differential equation models based on the Law of Mass Action should be thought of as the infinite volume limit of the Markov jump process, known as the thermodynamic limit in statistical physics. As we have seen, the volume V is critical to many phenomena which appear only in small, mesoscopic biochemical reaction systems, and thus stochastic kinetic models in theory.

The master equation approach to chemical reactions began in the 1930's with the work of M.A. Leontovich [71]. Independently, it was carried out by A.J.F. Siegert, M. Kac, M. Delbrück, A. Renyi, M. Lax and D.A. McQuarrie, among many others. Comprehensive reviews can be found in [42,61,72,73]. The chemical master equation (CME), first studied by Delbrück in 1940 [74], has become the leading mathematical theory for modeling mesoscopic nonlinear chemical reaction systems with small volume on the order of that of a living cell [8].

From a statistical mechanics point of view, each possible combination of the numbers of the chemical species defines a state of the system. The CME provides the evolution equation of the joint probability

distribution function over all system states. In open chemical systems, *i.e.*, where energy is added from an outside source, the number of system states is often infinite, leading to an infinite, coupled system of differential equations for the CME. An analytic solution to the CME for stochastic, open unimolecular reaction networks can be found in [75]. It is not possible, in general, to obtain an analytic solution for an open, non-unimolecular reaction system. However, the “steady state” solution to the master equation (also known as the stationary probability distribution) is generally unique [70] and may be algorithmically computed [76].

Continuous, diffusion approximations (also known as Fokker-Planck approximations) to the master equation were first developed by Van Kampen [77] and shown by Kurtz [78,79] to match the solution to the master equation in the thermodynamic limit for finite time. Because of the “finite time”, the stationary solution at infinite time for the Fokker-Planck equation is often not an acceptable approximation for the stationary solution of the CME. This gives rise to Keizer’s paradox. Fokker-Planck equations describe the probability distribution functions of continuous random movements known as stochastic differential equations (SDE). Approximating stochastic jump processes by diffusion processes with continuous fluctuations, however, is a delicate problem [80,81]. The delicate issue in mathematical term has to do with exchanging the limits for a large number of molecules and for a long time [82]. This limit exchange can lead to disagreements between discrete and continuous stochastic models [36].

The same issue of exchanging limits is present also between a stochastic jump process and the deterministic model. It is intimately related to the time scales for “down-hill dynamics” and “up-hill dynamics” and how their dependence upon the system size V [43]. Note that for sufficiently large V , the stochastic trajectory is close to the deterministic dynamics. However, there is no deterministic counterpart for stochastic “barrier-crossing” trajectory that moves against the deterministic vector field. A transition between stable attractors is impossible in a deterministic system, but occurs with probability 1 in stochastic dynamics, albeit with exponentially long time $\sim e^{cV}$ [44].

Kurtz carried out rigorous studies on the relation between the stochastic theory of chemical kinetics and its deterministic counterpart [83,84]. It has been shown that in the thermodynamic limit of $V \rightarrow \infty$, the CME becomes the expected deterministic ordinary differential equation (ODE) for finite time. Furthermore, solutions with given initial values to the CME approach the respective solutions to the ODE [83]. In light of this, there can still be disagreement in the steady state (*i.e.*, infinite time limit) solutions, an issue extensively revisited recently by Vellela and Qian [36,43].

Stochastic simulations of complex chemical reaction systems were carried out as early as the 1970’s [85,86]. Current software packages used for the simulation of biochemical reactions commonly make use of algorithms based on the influential work of Gillespie [8,10,87]. Microscopic particle simulations have validated the master equation as the most accurate description of a reactive process in aqueous solution [81,88]; see [89] for an up-to-date review. The CME provides the equation for the time-dependent joint probabilities of the number of molecules while the Gillespie algorithm gives the stochastic trajectories. They correspond to Fokker-Planck equation and stochastic differential equation (SDE) for diffusion processes.

In the environment of a living cell, biochemical systems are operating under a driven condition, widely called an “open system” [12,19,20,90]. There is a material and/or energy flux, from the outside, going through the system [91,92]. Such molecular systems no longer obey the traditional theory of equilibrium

thermodynamics. A closed molecular system tends to a thermal, chemical equilibrium, which is unique and in which each and every reaction has zero flux [93]. This is known as Lewis' principle of detailed balance [18]. Under equilibrium conditions, the ODE model based on the Law of Mass Action contains a unique, globally attracting equilibrium (fixed point). Accordingly, the stationary solution to the CME is a multi-Poisson distribution whose peak is located over the ODE fixed point [49].

The nonequilibrium theory for nonlinear biochemical reactions allows the possibility of multiple steady states, and nonzero steady state flux and a nonzero entropy production rate [19,20]. Recent developments in the area of fluctuation theorem [94,95] have illustrated the importance of entropy production and its relationship to the irreversible nature of a system. How is the entropy production rate related to functions of biochemical reaction networks? A correlation has been suggested between entropy production (or "dissipation cost") and the robustness of a network [96,97]. A more quantitative, if any, relationship between the entropy production rate and the dynamics of a nonequilibrium steady state is yet to be developed.

The essential difference between deterministic and stochastic models is the permanence of fixed points. According to the theory of ordinary differential equation, once the system reaches a fixed point (or an attractor), it must remain there for all time. Systems with stochasticity, however, can have trajectories being pushed away from attracting fixed points by random fluctuations. Since the noise is ever-present, it can eventually push the system out of the basin of attraction of one fixed point (attractor) and into that of another. Fixed points are no longer stationary for all time; they are only temporary, or "quasistationary" [98]. The amount of time the system spends at (or very near) a fixed point increases exponentially with the system volume. This agrees with ODE dynamics in the thermodynamic limit. However, this quasistationary behavior plays an important role at the cellular level in the "cellular evolutionary time scale" [45].

In order to systematically understand the mesoscopic cellular biochemical dynamics, this review discussed the simplest problem that is interesting: a one dimensional system with two fixed points. The systems with only one fixed point are trivial since deterministic and stochastic models are in complete agreement when there is a unique steady state [88]; the linear differential equation corresponds to a Poisson distribution in the CME [75]. The case of two fixed points, one stable and one unstable, is studied through an autocatalytic reaction first introduced by Keizer [42]. The ODE representation takes the form of a classic example in population dynamics, the logistic equation. Through this simple system, one understands the issues in the steady state predictions of the ODE and CME models [43]. This example introduces the notion of a quasistationary fixed point and a spectral analysis reveals the multiple time scales involved in the master equation formulation.

Logically, the next step is a one dimensional system with three fixed points, two stable with one unstable point between them [36,47]. For this, we use a reversible, trimolecular reaction known as Schlögl's model. This is the first case in which bistable behavior is possible, occurring through a saddle-node bifurcation. Again, the CME allows for new possibilities such as switching between the stable fixed points and a nonequilibrium phase transition in the steady state distribution function [45,51]. Because this model is fully reversible, one is also able to study thermodynamic quantities such as the chemical potential and entropy production rate and to illustrate the nonequilibrium physics [99]. The

dynamics of this system serves as a representative example for all systems with multiple stable fixed points.

Once the theory has been established for one dimensional systems with a single dynamic biochemical species, we turn our attention to planar systems with two dynamical species [54]. Here, oscillations become possible in the form of spiral nodes and limit cycles in ODE models. We explore the open question of how to define and quantify stochastic oscillations. We suggest a new method for locating oscillations in the presence of noise by extending the idea of the Poincaré return map to stochastic systems. A reversible extension of Schankenberg’s model for chemical kinetic oscillation is used to illustrate this new idea. The oscillation is represented by a rotational random walk.

In all these studies one encounters the presence of a time scale that grows exponentially with the system’s volume V . Dynamics operating on this “cellular evolution time scale” are lost in the infinite volume limit of the ODE model. To study the stability of a stochastic attractor, one must consider the chemical reactions systems in terms of the chemical master equation (CME). The ODE formulation, however, is valuable as a way to estimate the presence and location of the critical points in the landscape of the probability steady state distribution of the CME [100,101].

In summary, one of the most important insights from the CME study of biochemical reaction systems in a small, cellular volume is the realization of the *cellular evolution time scale* and the associated stochastic attractors which might indeed be the emergent cellular epigenetic states. The dynamics on this time scale is stochastic; it is completely missing from the traditional ODE theory of biochemical reaction networks. In the CME theory, deterministic fixed points become stochastic attractors [100,101]. They are the emergent properties of a complex, nonlinear biochemical network. The transitions among the emergent stochastic attractors constitute the proper cellular dynamics [102,103].

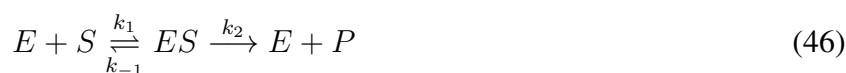
Acknowledgements

We also thank Ping Ao, Hao Ge, Kyung Kim, Jie Liang, Zhilin Qu, Michael Samoilov, Herb Sauro, Melissa Vellela, Jin Wang, and Sunney Xie for many helpful discussions. The work reported here was supported in part by NSF grant No. EF0827592.

Appendix: The Chemical Master Equation for Systems of Biochemical Networks

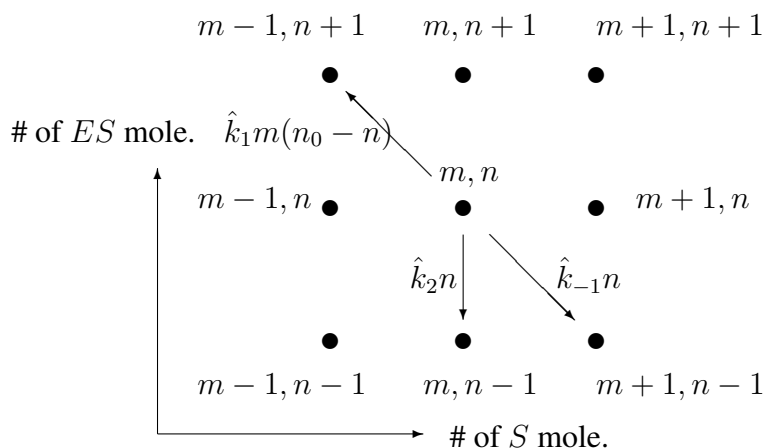
A1. Michaelis-Menten model

The canonical MM kinetic scheme is



Let m and n be the numbers of S and ES respectively. Assume the total number of enzyme is n_0 . Then, the corresponding *chemical master equation graph*, shown below, details the changes in the “state” of the system, (m, n) , due to the three reactions in the Equation (46). The graph in Figure 6 helps one to write the chemical master equation in (2). The three “leaving” terms are negative in Equation (2), and the three “into” terms are positive in Equation (2). The \hat{k} ’s in the graph is related to the k ’s in the Equation (46) according to Equation (3).

Figure 6. The chemical master equation *graph* for the stochastic Michaelis-Menten enzyme kinetics in (46). The graph only shows all the transitions associated with “leaving” the state (m, n) . What is not shown are the transitions: $(m-1, n+1) \rightarrow (m, n)$ with rate $\hat{k}_{-1}(n+1)$; $(m+1, n-1) \rightarrow (m, n)$ with rate $\hat{k}_1(m+1)(n_0 - n + 1)$; and $(m, n+1) \rightarrow (m, n)$ with rate $\hat{k}_2(n+1)$. They are all associated with “into” the state (m, n) .



A2. Keizer's Model

We are interested in the autocatalytic reaction system



in which the concentrations of A and C are hold constant. This is a modified version of an example originally discussed by J. Keizer in his book [42], where he assumed $k_{-1} = 0$. Let $n(t)$ be the stochastic number of X molecule at time t . It is then related to the concentration x by $n = xV$, where V is the volume of the system. This volume parameter V appears implicitly in the CME. For example, in the deterministic model, reaction rates k_1 and k_{-1} have units of $[\text{volume}][\text{time}]^{-1}$, and k_2 has units of $[\text{time}]^{-1}$. The reaction rates in the stochastic model are related to these rates by

$$\hat{k}_1 = k_1/V, \quad \hat{k}_{-1} = k_{-1}/V, \quad \hat{k}_2 = k_2 \quad (48)$$

These reaction rates are scaled such that the units agree in the master equation (see Equation 49 below).

Figure 7 shows how the probability of each state n is affected by the three reactions in Equation (47). The change in the probability of each state, $\frac{d}{dt}p_n(t)$ is the sum of the changes due to each of the reactions. Thus, the CME is the system of equations:

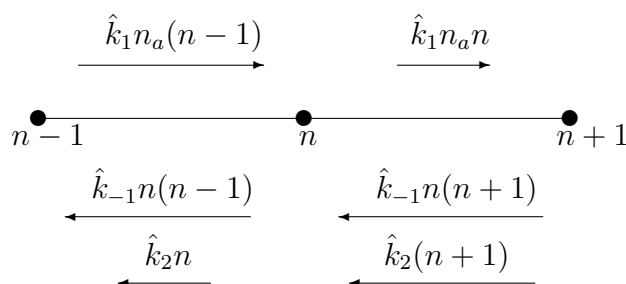
$$\frac{dp_0}{dt} = \hat{k}_2 p_1 \quad (49a)$$

$$\begin{aligned} \frac{dp_n}{dt} &= \hat{k}_1 n_a (n-1) p_{n-1} + (\hat{k}_{-1} n (n+1) + \hat{k}_2 (n+1)) p_{n+1} \\ &\quad - (\hat{k}_{-1} n (n-1) + \hat{k}_1 n_a n + \hat{k}_2 n) p_n \end{aligned} \quad (49b)$$

\vdots

where n_a represents the number of A molecules, which does not change in the model.

Figure 7. The chemical master equation *graph* shows the probability change in state P_n , where n_a is the number of A molecules.



A3. Schlögl Model

The canonical Schlögl model for chemical bistability is [46]



Following the *chemical master equation graph* in Figure 8, we have the CME for the probability of having n number of X at time t , $p_n(t) = \Pr\{n_X(t) = n\}$:

$$\frac{d}{dt}p_n(t) = \lambda_{n-1}p_{n-1} - (\lambda_n + \mu_n)p_n + \mu_{n+1}p_{n+1} \quad (n \geq 0) \quad (51)$$

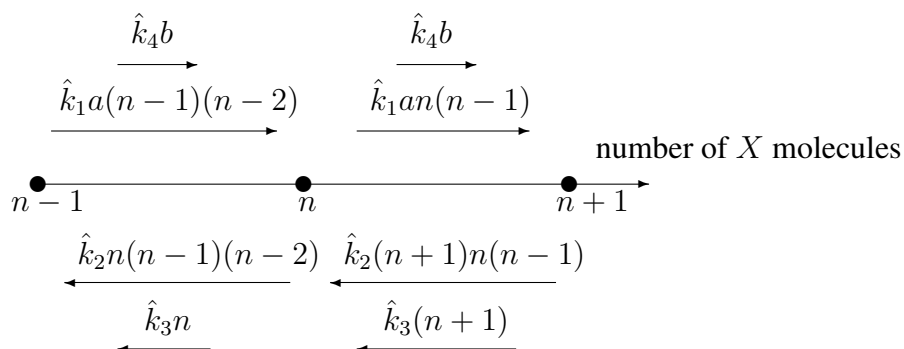
where

$$\lambda_n = \hat{k}_1 a n(n-1) + \hat{k}_4 b, \quad \mu_n = \hat{k}_2 n(n-1)(n-2) + \hat{k}_3 n \quad (52)$$

and

$$\hat{k}_1 = \frac{k_1}{V}, \quad \hat{k}_2 = \frac{k_2}{V^2}, \quad \hat{k}_3 = k_3, \quad \hat{k}_4 = k_4 V \quad (53)$$

Figure 8. The chemical master equation *graph* for the Schlögl model in Equation (50) shows the rates of probability changes into and leaving state n . a and b are concentrations of A and B . The relations between the \hat{k} 's and k 's in Equation (50) are given in Equation (53).



A3.1. Stationary Distribution: Steady State and Equilibrium

By setting the right-hand-side of Equation (51) to zero, one solves the steady state distribution

$$p_n^{ss} = C \prod_{i=0}^{n-1} \frac{\lambda_i}{\mu_{i+1}} = C \prod_{i=0}^{n-1} \frac{\hat{k}_1 a i(i-1) + \hat{k}_4 b}{\hat{k}_2(i+1)i(i-1) + \hat{k}_3(i+1)} \quad (54)$$

where C is a normalization factor.

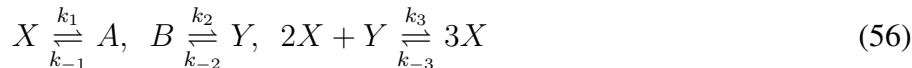
We now show a very interesting and important property of the p_n^{ss} , when the concentrations of A and B are at equilibrium: $b/a = k_1 k_3 / (k_2 k_4) = \hat{k}_1 \hat{k}_3 / (\hat{k}_2 \hat{k}_4)$. Substituting this relation into Equation (54), we have the equilibrium distribution

$$p_n^{eq} = C \prod_{i=0}^{n-1} \frac{\hat{k}_1 a}{\hat{k}_2(i+1)} = \frac{1}{n!} \left(\frac{\hat{k}_1 a}{\hat{k}_2} \right)^n e^{-\hat{k}_1 a / \hat{k}_2} \quad (55)$$

This is a Poisson distribution with the mean number of X being $\frac{\hat{k}_1 a}{\hat{k}_2} = \frac{k_1 a V}{k_2}$. That is the equilibrium concentration of X being $\frac{k_1}{k_2} a$. The Poisson distribution has only a single peak.

A4. Schnakenberg Model

We are now interested in the nonlinear chemical reaction system, the reversible Schnakenberg model, in a mesoscopic volume V :



Consider the function $p_{n,m}(t)$, the probability that there are n X molecules and m Y molecules at time t . The six reactions (three forward and three backward) in the reversible Schnakenberg model in Equation (56) define six ways to move on the lattice of possible states, *i.e.*, the (n, m) lattice, $\mathbb{Z}^+ \times \mathbb{Z}^+$ (see Figure 9). The chemical master equation (CME) is a doubly infinite set of ODEs:

$$\begin{aligned} \frac{dp_{n,m}(t)}{dt} = & \lambda_{n-1,m}^1 p_{n-1,m} + \lambda_{n,m-1}^2 p_{n,m-1} + \lambda_{n-1,m+1}^3 p_{n-1,m+1} \\ & + \mu_{n+1,m}^1 p_{n+1,m} + \mu_{n,m+1}^2 p_{n,m+1} + \mu_{n+1,m-1}^3 p_{n+1,m-1} \\ & - (\lambda_{n,m}^1 + \lambda_{n,m}^2 + \lambda_{n,m}^3 + \mu_{n,m}^1 + \mu_{n,m}^2 + \mu_{n,m}^3) p_{n,m} \end{aligned} \quad (57)$$

for $n = 0 \dots \infty, m = 0 \dots \infty$. The birth and death rates, $\lambda_{n,m}^i$ and $\mu_{n,m}^i$ respectively, for the three equations are

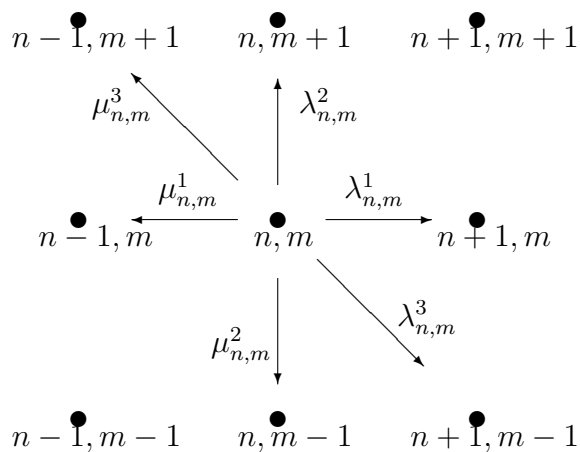
$$\lambda_{n,m}^1 = k_{-1} n_a, \quad \mu_{n,m}^1 = k_1 n \quad (58)$$

$$\lambda_{n,m}^2 = k_2 n_b, \quad \mu_{n,m}^2 = k_{-2} m \quad (59)$$

$$\lambda_{n,m}^3 = \frac{k_3}{V^2} n(n-1)m, \quad \mu_{n,m}^3 = \frac{k_{-3}}{V^2} n(n-1)(n-2) \quad (60)$$

The factor of $1/V^2$ in $\lambda_{n,m}^3$ and $\mu_{n,m}^3$ accounts for the fact that the third reaction is trimolecular, and thus k_3 and k_{-3} have units of V^2/t . Since the first and second reactions are unimolecular, the remaining rates have units of $1/t$ and do not need scaling when used in the CME.

Figure 9. The chemical master equation *graph* showing possible paths away from state (n, m) , with birth rates $\lambda_{n,m}^i$ and death rates $\mu_{n,m}^i$. Note that there will be a corresponding reverse path into state (n, m) for each of these arrows.



Because the CME is a set of linear ODEs, there will be a unique steady state to which the system tends, the probability steady state, $p^{ss}(n, m)$. Once the system reaches its steady state, the total probability flow into each point, (n, m) , will equal the total probability flow out of that point. Solving for $p^{ss}(n, m)$ involves setting each of the equations in Equation 57 equal to zero and solving a very large linear system. Cao and Liang have recently developed a method for computing the probability steady state for molecular networks in general [76]. Their method is an algorithm which enumerates the state space and solves the corresponding linear system and is optimal in both storage and time complexity.

References

1. Gibbs, J.W. *The Scientific Papers of J. Willard Gibbs*; Dover: New York, NY, USA, 1961.
2. Tanford, C. *Physical Chemistry of Macromolecules*; John Wiley & Sons: New York, NY, USA, 1961.
3. Dill, K.A.; Bromberg, S. *Molecular Driving Forces: Statistical Thermodynamics in Chemistry and Biology*; Garland Science: New York, NY, USA, 2002.
4. Onsager, L.; Machlup, S. Fluctuations and irreversible processes. *Phys. Rev.* **1953**, *91*, 1505–1512.
5. Machlup, S.; Onsager, L. Fluctuations and irreversible process. II. Systems with kinetic energy. *Phys. Rev.* **1953**, *91*, 1512–1515.
6. Lund, E.W. Guldberg and waage and the law of mass action. *J. Chem. Ed.* **1965**, *42*, 548–550.
7. Segel, I. *Enzyme Kinetics: Behavior and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems*; John Wiley & Sons: New York, NY, USA, 1993.
8. Qian, H.; Beard, D.A. Chemical biophysics: Quantitative analysis of cellular systems. *Cambridge Texts in Biomedical Engineering*; Cambridge Univ. Press: New York, NY, USA, 2008.
9. Liang, J.; Qian, H. Computational cellular dynamics based on the chemical master equation: A challenge for understanding complexity. *J. Comput. Sci. Tech.* **2010**, *25*, 154–168.

10. Gillespie, D.T. Stochastic simulation of chemical kinetics. *Annu. Rev. Phys. Chem.* **2007**, *58*, 35–55.
11. Ge, H.; Qian, H. The physical origins of entropy production, free energy dissipation and their mathematical representations. *Phys. Rev. E* **2010**, *81*, 051133.
12. Nicolis, G.; Prigogine, I. *Self-Organization in Non-Equilibrium Systems*; John Wiley & Sons: New York, NY, USA, 1977.
13. Kurchan, J. Fluctuation theorem for stochastic dynamics. *J. Phys. A: Math Gen.* **1998**, *31*, 3719–3729.
14. Lebowitz, J.L.; Spohn, H. A gallavotti-cohen-type symmetry in the large deviation functional for stochastic dynamis. *J. Stat. Phys.* **1999**, *95*, 333–365.
15. Jiang, D.Q.; Qian, M.; Qian, M.P. *Mathematical Theory of Nonequilibrium Steady States*; Lecture Notes in Mathematics. Springer: New York, NY, USA, 2004; Volume 1833.
16. Qian, H. The mathematical theory of molecular motor movement and chemomechanical energy transduction. *J. Mach. Chem.* **2000**, *27*, 219–234.
17. Qian, H. Cycle kinetics, steady-state thermodynamics and motors — a paradigm for living matter physics. *J. Phys.: Cond. Matt.* **2005**, *17*, S3783–S3794.
18. Lewis, G.N. A new principle of equilibrium. *PNAS* **1925**, *11*, 179–183.
19. Qian, H. Open-system nonequilibrium steady state: Statistical thermodynamics, fluctuations, and chemical oscillations. *J. Phys. Chem. B* **2006**, *110*, 15063–15074.
20. Qian, H. Phosphorylation energy hypothesis: Open chemical systems and their biological functions. *Annu. Rev. Phys. Chem.* **2007**, *58*, 113–142.
21. Kepler, T.B.; Elston, T.C. Stochasticity in transcriptional regulation: Origins, consequences, and mathematical representations. *Biophys. J.* **2001**, *81*, 3116–3136.
22. Rao, C.V.; Arkin, A.P. Stochastic chemical kinetics and the quasi-steady-state assumption: Application to the Gillespie algorithm. *J. Chem. Phys.* **2003**, *118*, 4999–5010.
23. Qian, H. Cooperativity and specificity in enzyme kinetics: a single-molecule time-based perspective. *Biophys. J.* **2008**, *95*, 10–17.
24. Gadgil, C.; Lee, C.H.; Othmer, H.G. A stochastic analysis of first-order reaction networks. *Bull. Math. Biol.* **2005**, *67*, 901–946.
25. Das, M.; Green, F. Landauer formula without Landauer’s assumptions. *J. Phys.: Condens. Matter* **2003**, *15*, L687.
26. Onsager, L. Reciprocal relations in irreversible processes. I. *Phys. Rev.* **1931**, *37*, 405–426.
27. Onsager, L. Reciprocal relations in irreversible processes. II. *Phys. Rev.* **1931**, *38*, 2265–2279.
28. Li, Y.; Qian, H.; Yi, Y. Oscillations and multiscale dynamics in a closed chemical reaction system: Second law of thermodynamics and temporal complexity. *J. Chem. Phys.* **2008**, *129*, 154505.
29. Li, Y.; Qian, H.; Yi, Y. Nonlinear oscillations and multiscale dynamics in a closed chemical reaction system. *J. Dyn. Diff. Eqn.* **2010**, *22*, 491–507.
30. Howard, J. *Mechanics of Motor Proteins and the Cytoskele*; Sinauer Associates: Sunderland, MA, USA, 2001.
31. Ferrell, J.E.; Xiong, W. Bistability in cell signaling: How to make continuous processes discontinuous, and reversible processes irreversible. *Chaos* **2001**, *11*, 227–236.

32. Cooper, J.A.; Qian, H. A mechanism for Src kinase-dependent signaling by noncatalytic receptors. *Biochemistry* **2008**, *47*, 5681–5688.
33. Zhu, H.; Qian, H.; Li, G. Delayed onset of positive feedback activation of Rab5 by Rabex-5 and Rabaptin-5 in endocytosis. *PLoS ONE* **2010**, *5*, e9226.
34. Bishop, L.M.; Qian, H. Stochastic bistability and bifurcation in a mesoscopic signaling system with autocatalytic kinase. *Biophys. J.* **2010**, *98*, 1–11.
35. Paulsson, J. Models of stochastic gene expression. *Phys. Life Rev.* **2005**, *2*, 157–175.
36. Vellela, M.; Qian, H. Stochastic dynamics and non-equilibrium thermodynamics of a bistable chemical system: the Schlogl model revisited. *J. R. Soc. Interface* **2009**, *6*, 925–940.
37. Qian, H.; Shi, P.Z.; Xing, J. Stochastic bifurcation, slow fluctuations, and bistability as an origin of biochemical complexity. *Phys. Chem. Chem. Phys.* **2009**, *11*, 4861–4870.
38. Epstein, I.R.; Pogman, J.A. *An Introduction to Nonlinear Chemical Dynamics: Oscillations, Waves, Patterns, and Chaos*; Oxford University Press: Oxford, UK, 1998.
39. Qian, H. Entropy demystified: The “thermo”-dynamics of stochastically fluctuating systems. *Meth. Enzym.* **2009**, *467*, 111–134.
40. Gunawardena, J. Multisite protein phosphorylation makes a good threshold but can be a poor switch. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 14617–14622.
41. Samoilov, M.; Plyasunov, S.; Arkin, A.P. Stochastic amplification and signaling in enzymatic futile cycles through noise-induced bistability with oscillations. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 2310–2315.
42. Keizer, J. *Statistical Thermodynamics of Nonequilibrium Processes*; Springer-Verlag: New York, NY, USA, 1987.
43. Vellela, M.; Qian, H. A quasistationary analysis of a stochastic chemical reaction: Keizer’s paradox. *Bull. Math. Biol.* **2007**, *69*, 1727–1746.
44. Hinch, R.; Chapman, S.J. Exponentially slow transitions on a Markov chain: The frequency of calcium sparks. *Eur. J. Appl. Math.* **2005**, *16*, 427–446.
45. Ge, H.; Qian, H. Nonequilibrium phase transition in mesoscopic biochemical systems: From stochastic to nonlinear dynamics and beyond. *J. R. Soc. Interf.* **2010**, [Epub ahead of print].
46. Schlogl, F. Chemical reaction models for non-equilibrium phase transition. *Z. Physik.* **1972**, *253*, 147–161.
47. Qian, H.; Reluga, T. Nonequilibrium thermodynamics and nonlinear kinetics in a cellular signaling switch. *Phys. Rev. Lett.* **2005**, *94*, 028101.
48. Keizer, J. Nonequilibrium thermodynamics and the stability of states far from equilibrium. *Acc. Chem. Res.* **1979**, *12*, 243–249.
49. Gardiner, C.W. *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences*, 2nd ed.; Springer: New York, NY, USA, 1985.
50. Matheson, I.; Walls, D.F.; Gardiner, C.W. Stochastic models of first-order nonequilibrium phase transitions in chemical reactions. *J. Stat. Phys.* **1974**, *12*, 21–34.
51. Ge, H.; Qian, H. Thermodynamic limit of a nonequilibrium steady-state: Maxwell-type construction for a bistable biochemical system. *Phys. Rev. Lett.* **2009**, *103*, 148103.

52. Nicolis, G.; Turner, J.W. Stochastic analysis of a nonequilibrium phase transition: Some exact results. *Phys. A* **1977**, *89*, 326–338.
53. Qian, H.; Saffarian, S.; Elson, E.L. Concentration fluctuations in a mesoscopic oscillating chemical reaction system. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 10376–10381.
54. Vellela, M.; Qian, H. On Poincare-Hill cycle map of rotational random walk: Locating stochastic limit cycle in reversible Schnakenberg model. *Proc. R. Soc. A* **2010**, *466*, 771–788.
55. Strogatz, S. *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering (Studies in Nonlinearity)*; Perseus Books Group: New York, NY, USA, 2001.
56. Hill, T.L. *Free Energy Transduction and Biochemical Cycle Kinetics*; Springer-Verlag: New York, NY, USA, 1989.
57. Hill, T.L.; Chen, Y.D. Stochastics of cycle completions (fluxes) in biochemical kinetic diagrams. *Proc. Natl. Acad. Sci. USA* **1975**, *72*, 1291–1295.
58. Sel'kov, E.E. Self-oscillations in glycolysis. *Eur. J. Biochem.* **1968**, *4*, 79–86.
59. Goldbeter, A.; Lefever, R. Dissipative structures for an allosteric model: Application to glycolytic oscillations. *Biophys. J.* **1972**, *12*, 1302–1315.
60. Keener, J.; Sneyd, J. *Mathematical Physiology*; Springer: New York, NY, USA, 1998.
61. Érdi, P.; Tóth, J. *Mathematical Models of Chemical Reactions: Theory and Applications of Deterministic and Stochastic Models*; Manchester Univ. Press: Manchester, UK, 1989.
62. Murray, J.D. *Mathematical Biology: An Introduction*, 3rd ed.; Springer: New York, NY, USA, 2002.
63. Zheng, Q.; Ross, J. Comparison of deterministic and stochastic kinetics for nonlinear systems. *J. Chem. Phys.* **1991**, *94*, 3644–3648.
64. Noyes, R.M.; Field, R.J. Oscillatory chemical reactions. *Annu. Rev. Phys. Chem.* **1974**, *25*, 95–119.
65. Xie, X.S.; Choi, P.J.; Li, G.W.; Lee, N.K.; Lia, G. Single-molecule approach to molecular biology in living bacterial cells. *Annu. Rev. Biophys.* **2008**, *37*, 417–444.
66. Smolen, P.; Baxter, D.A.; Byrne, J.H. Modeling transcriptional control in gene networks – methods, recent results, and future directions. *Bull. Math. Biol.* **2000**, *62*, 247–292.
67. Turner, T.E.; Schnell, S.; Burrage, K. Stochastic approaches for modeling *in vivo* reactions. *Comput. Biol. Chem.* **2004**, *28*, 165–178.
68. Paulsson, J.; Berg, O.G.; Ehrenberg, M. Stochastic focusing: Fluctuation-enhanced sensitivity of intracellular regulation. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 7148–7153.
69. Taylor, H.; Karlin, S. *An Introduction to Stochastic Modeling*, 3rd ed.; Academic Press: New York, NY, USA, 1998.
70. Schnakenberg, J. Network theory of microscopic and macroscopic behaviour of Master equation systems. *Rev. Mod. Phys.* **1976**, *48*, 571–585.
71. Leontovich, M.A. Basic equations of the kinetic gas theory from the point of view of the theory of random processes (in Russian). *Zh. Teoret. Eksper. Fiz.* **1935**, *5*, 211–231.
72. McQuarrie, D.A. Stochastic approach to chemical kinetics. *J. Appl. Prob.* **1967**, *4*, 413–478.
73. McQuarrie, D.A. *Stochastic Approach to Chemical Kinetics*; Methuen: London, UK, 1968.
74. Delbrück, M. Statistical fluctuations in autocatalytic reactions. *J. Chem. Phys.* **1940**, *8*, 120–124.

75. Heuett, W.J.; Qian, H. Grand canonical Markov model: A stochastic theory for open nonequilibrium biochemical networks. *J. Chem. Phys.* **2006**, *124*, 044110.
76. Cao, Y.; Liang, J. Optimal enumeration of state space of finitely buffered stochastic molecular networks and exact computation of steady state landscape probability. *BMC Syst. Biol.* **2008**, *2*, 30.
77. Van Kampen, N. *Stochastic Processes in Physics and Chemistry*; Elsevier: Amsterdam, The Netherlands, 1981.
78. Kurtz, T.G. Limit theorems and diffusion approximations for density dependent Markov chains. *Math. Progr. Stud.* **1976**, *5*, 67–78.
79. Kurtz, T.G. Strong approximation theorems for density dependent Markov chains. *Stochastic Process. Appl.* **1978**, *6*, 223–240.
80. Hanggi, P.; Grabert, H.; Talkner, P.; Thomas, H. Bistable systems: master equation versus Fokker-Planck modeling. *Phys. Rev. A* **1984**, *29*, 371–378.
81. Baras, F.; Mansour, M.; Pearson, J. Microscopic simulation of chemical bistability in homogeneous systems. *J. Chem. Phys.* **1996**, *105*, 8257.
82. Luo, J.L.; van den Broeck, C.; Nicolis, G. Stability criteria and fluctuations around nonequilibrium states. *Z. Phys. B.: Cond. Matt.* **1984**, *56*, 165–170.
83. Kurtz, T.G. Limit theorems for sequences of jump Markov processes approximating ordinary differential equations. *J. Appl. Prob.* **1971**, *8*, 344–356.
84. Kurtz, T.G. The relationship between stochastic and deterministic models for chemical reactions. *J. Chem. Phys.* **1972**, *57*, 2976–2978.
85. Sipos, T.; Tóth, J.; Érdi, J. Stochastic simulation of complex chemical reactions by digital computer, I. The model. *React. Kinet. Catal. Lett.* **1974**, *1*, 113–117.
86. Sipos, T.; Tóth, J.; Érdi, P. Stochastic simulation of complex chemical reactions by digital computer, II. Applications. *React. Kinet. Catal. Lett.* **1974**, *1*, 209–213.
87. McAdams, H.H.; Arkin, A.P. It's a noisy business! Genetic regulation at the nanomolar scale. *Trends Genet.* **1999**, *15*, 65–69.
88. Malek-Mansour, M.; van den Broeck, C.; Nicolis, G.; Turner, J.W. Asymptotic properties of Markovian master equations. *Ann. Phys. (USA)* **1981**, *131*, 283–313.
89. Baras, F.; Mansour, M. Particle simulations of chemical systems. *Adv. Chem. Phys.* **1997**, *100*, 393–474.
90. Ross, J. *Thermodynamics and Fluctuations far from Equilibrium*; Springer Series in Chemical Physics. Springer: New York, NY, USA, 2008.
91. Zia, R.K.P.; Schmittmann, B. Probability currents as principal characteristics in the statistical mechanics of non-equilibrium steady states. *J. Stat. Mech.: Theor. Exp.* **2007**, P07012.
92. Tomita, K.; Tomita, H. Irreversible circulation of fluctuation. *Prog. Theor. Phys.* **1974**, *51*, 1731–1749.
93. Wang, J.; Xu, L.; Wang, E. Potential landscape and flux framework of nonequilibrium networks: Robustness, dissipation, and coherence of biochemical oscillations. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 12271–12276.

94. Andrieux, D.; Gaspard, P. Fluctuation theorem for currents and schnakenberg network theory. *J. Stat. Phys.* **2007**, *127*, 107–131.
95. Sevick, E.M.; Prabhakar, R.; Williams, S.R.; Searles, D.J. Fluctuation theorems. *Annu. Rev. Phys. Chem.* **2008**, *59*, 603–633.
96. Han, B.; Wang, J. Quantifying robustness and dissipation cost of yeast cell cycle network: the funneled energy landscape perspective. *Biophys. J.* **2007**, *92*, 3755–3783.
97. Han, B.; Wang, J. Least dissipation cost as a design principle for robustness and function of cellular networks. *Phys. Rev. E* **2008**, *77*, 031922.
98. Nasell, I. Extinction and quasi-stationarity in the Verhulst logistic model. *J. Theor. Biol.* **2001**, *211*, 11–27.
99. Luo, J.L.; Zhao, N.; Hu, B. Effect of critical fluctuations to stochastic thermodynamic behavior of chemical reaction systems at steady state far from equilibrium. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4149–4154.
100. Ao, P.; Galas, D.; Hood, L.; Zhu, X. Cancer as robust intrinsic state of endogenous molecular-cellular network shaped by evolution. *Med. Hypotheses* **2008**, *70*, 678–84.
101. Ao, P. Global view of bionetwork dynamics: adaptive landscape. *J. Genet. Genomics* **2009**, *36*, 63–73.
102. Ge, H.; Qian, H.; Qian, M. Synchronized dynamics and nonequilibrium steady states in a yeast cell-cycle network. *Math. Biosci.* **2008**, *211*, 132–152.
103. Ge, H.; Qian, M. Boolean network approach to negative feedback loops of the p53 pathways: Synchronized dynamics and stochastic limit cycles. *J. Comput. Biol.* **2009**, *16*, 119–132.

Computational Cellular Dynamics Based on the Chemical Master Equation: A Challenge for Understanding Complexity

Jie Liang^{1,2} (梁 杰) and Hong Qian^{3,4} (钱 纭)

¹Department of Bioengineering, University of Illinois at Chicago, Chicago, IL 60607, U.S.A.

²Shanghai Center for Systems Biomedicine, Shanghai Jiao Tong University, Shanghai 200240, China

³Department of Applied Mathematics, University of Washington, Seattle, WA 98195, U.S.A.

⁴Kavli Institute for Theoretical Physics China, Chinese Academy of Sciences, Beijing 100190, China

E-mail: jliang@uic.edu; qian@amath.washington.edu

Received August 29, 2009; revised November 16, 2009.

Abstract Modern molecular biology has always been a great source of inspiration for computational science. Half a century ago, the challenge from understanding macromolecular dynamics has led the way for computations to be part of the tool set to study molecular biology. Twenty-five years ago, the demand from genome science has inspired an entire generation of computer scientists with an interest in discrete mathematics to join the field that is now called bioinformatics. In this paper, we shall lay out a new mathematical theory for dynamics of biochemical reaction systems in a small volume (i.e., mesoscopic) in terms of a stochastic, discrete-state continuous-time formulation, called the chemical master equation (CME). Similar to the wavefunction in quantum mechanics, the dynamically changing probability landscape associated with the state space provides a fundamental characterization of the biochemical reaction system. The stochastic trajectories of the dynamics are best known through the simulations using the Gillespie algorithm. In contrast to the Metropolis algorithm, this Monte Carlo sampling technique does not follow a process with detailed balance. We shall show several examples how CMEs are used to model cellular biochemical systems. We shall also illustrate the computational challenges involved: multiscale phenomena, the interplay between stochasticity and nonlinearity, and how macroscopic determinism arises from mesoscopic dynamics. We point out recent advances in computing solutions to the CME, including exact solution of the steady state landscape and stochastic differential equations that offer alternatives to the Gillespie algorithm. We argue that the CME is an ideal system from which one can learn to understand “complex behavior” and complexity theory, and from which important biological insight can be gained.

Keywords biochemical networks, cellular signaling, epigenetics, master equation, nonlinear reactions, stochastic modeling

1 Introduction

Cellular biology has two important foundations: genomics focuses on DNA sequences and their evolutionary dynamics; and biochemistry studies molecular reaction kinetics that involve both small metabolites and large macromolecules. Computational science has been an essential component of genomics. In recent years, cellular biochemistry is also increasingly relying on mathematical models for biochemical reaction networks. Two approaches have been particularly prominent: the Law of Mass Action for deterministic nonlinear chemical reactions in terms of the concentrations of chemical species, and the Chemical Master Equation (CME) for stochastic reactions in terms of the numbers of reaction species.

The Law of Mass Action and the CME are two parts of a single mathematical theory of chemical reaction systems, with the latter being fundamental. When the number of molecules in a reaction system are large, stochasticity in the CME disappears and the Law of Mass Action can be shown, mathematically, to arise as the limit^[1-2].

In this article, we shall introduce the CME approach to biochemical reaction kinetics. We use simply examples to illustrate some of the salient features of this yet to be fully developed theory. We then discuss the challenges one faces in applying this theory to computational cellular biology. There have been several recent texts which cover some of the materials we discuss. See [2-3].

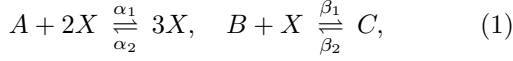
Survey

This work is supported by US NIH under Grant Nos. GM079804, GM081682, GM086145, GM068610, NSF of USA under Grant Nos. DBI-0646035 and DMS-0800257, and ‘985’ Phase II Grant of Shanghai Jiao Tong University under Grant No. T226208001.

©2010 Springer Science + Business Media, LLC & Science Press, China

2 A System of Nonlinear Reactions

To illustrate the theory of the CME and the Law of Mass Action, let us first consider a simple system of nonlinear chemical reactions first proposed by Schlögl^[4]



in which species A , B and C are at fixed concentrations a , b and c , respectively. The traditional, macroscopic kinetics of the system (1), according to the Law of Mass Action, is described by a deterministic ordinary differential equation (ODE)^[5]

$$\frac{dx}{dt} = -\alpha_2 x^3 + \alpha_1 a x^2 - \beta_1 b x + \beta_2 c, \quad (2)$$

where x represents the concentration of X . It is straightforward to show that (2) exhibits bistability (via the so called pitchfork bifurcation) when $\alpha_2 \beta_1 b / (\alpha_1 a)^2 = 1/3$ ^[4-5]: that is, the polynomial on the right-hand-side switches from having only one positive root to have three positive roots. The system also shows another bifurcations when varying another lumped parameter $\alpha_2^2 \beta_2 c / (\alpha_1 a)^3$ (this time via the so called saddle-node bifurcation).

We now turn to the CME approach to this reaction system (1). If in a small volume such as that of a cell, the number of X is sufficiently small, its concentration fluctuations become significant^[6]. The dynamics of reaction (1) then is stochastic, which should be described in terms of a master equation, also known as a birth-death process in the theory of Markov processes^[7].

The system is represented by a discrete random variable $n_X(t)$: the number of X at time t ($0 \leq n_X < \infty$). Let $P(k, t) = \Pr\{n_X(t) = k\}$, and we have

$$\begin{aligned} \frac{dP(k, t)}{dt} = & v_{k-1} P(k-1, t) + w_k P(k+1, t) \\ & - (v_k + w_{k-1}) P(k, t), \end{aligned} \quad (3)$$

where

$$v_k = \frac{\alpha_1 a k(k-1)}{V^2} + \beta_2 c,$$

and

$$w_k = \frac{\alpha_2 (k+1)k(k-1)}{V^3} + \frac{\beta_1 b (k+1)}{V}.$$

Here V is the volume of the reaction system. It is a very important parameter of the model. The basic rule is still the Law of Mass Action: the rate of one step reaction $B + X \xrightarrow{\beta_1} C$, when there are $k+1$ number of X molecules, is $\beta_1 b (k+1)/V$. This gives the above last term. Similarly, the rate of one step reaction $A + 2X \xrightarrow{\alpha_1} 3X$, when there are k number of X molecules, is $\alpha_1 a k(k-1)/V^2$.

For complex biochemical reactions, master equation like this in general cannot be solved analytically. Various algorithms exist for simulating its stochastic trajectories^[8]. For the above specific example, however, the exact stationary probability distribution to (3), i.e., after the system reaches stationarity, can be found as [9-10]:

$$P(k) = C_0 \prod_{j=0}^{k-1} \frac{v_j}{w_j}, \quad (4)$$

where C_0 is a normalization constant such that $\sum_{k=0}^{\infty} P(k) = 1$. The number of X molecules still fluctuates in the steady state. We note that for large V ,

$$\begin{aligned} \ln P(k) = & \sum_{j=0}^{k-1} \ln \frac{v_j}{w_j} + C_1 \approx \sum_{j=0}^{k-1} \ln \frac{v(k/V)}{w(k/V)} \\ & + o\left(\frac{1}{V}\right) + C_1 \approx V \int_0^{k/V} \ln \frac{v(z)}{w(z)} dz + C_1, \end{aligned}$$

in which

$$v(z) = z^2 + \sigma, \quad w(z) = z^3 + \mu z,$$

$\mu = \alpha_2 \beta_1 b / (\alpha_1 a)^2$, $\sigma = \alpha_2^2 \beta_2 c / (\alpha_1 a)^3$, and $C_1 = \ln C_0$. Therefore in terms of the concentration $x = k/V$, we have the probability distribution $f(x) = V P(Vx)$:

$$\begin{aligned} \frac{1}{2V} \ln f(x) = & \frac{1}{2V} \ln P(Vx) + C_2 \\ \approx & \frac{1}{2} \int_0^x \ln \frac{v(z)}{w(z)} dz + \hat{C} \end{aligned} \quad (5)$$

$$= \frac{1}{2} \int_0^x \ln \frac{z^2 + \sigma}{z^3 + \mu z} dz + \hat{C}. \quad (6)$$

Therefore, the stationary probability distribution of the concentration of X :

$$f(x) \approx e^{-V\phi(x)}, \quad (7)$$

where

$$\phi(x) = - \int_0^x \ln \frac{z^2 + \sigma}{z^3 + \mu z} dz, \quad (8)$$

is independent of V . It is easy to verify that $\phi(x)$ is at its extrema exactly when the ODE (2) is at its fixed points. The function $\phi(x)$ can be thought as a “landscape” for the nonlinear chemical reaction system.

Closed System, Detailed Balance and Chemical Equilibrium. A chemical equilibrium is reached in the reaction system (1) when

$$\frac{[X]^3}{[A][X]^2} = \frac{\alpha_1}{\alpha_2}, \quad \frac{[C]}{[B][X]} = \frac{\beta_1}{\beta_2}. \quad (9)$$

This leads to the equilibrium condition that

$$\left(\frac{[C]}{[A][B]} \right)^{\text{eq}} = \frac{\alpha_1 \beta_1}{\alpha_2 \beta_2}. \quad (10)$$

In term of the two model parameters μ and σ introduced above, this equilibrium (also called detailed balance) condition is expressed as

$$\frac{\sigma}{\mu} = \frac{\alpha_2^2 \beta_2 c / (\alpha_1 a)^3}{\alpha_2 \beta_1 b / (\alpha_1 a)^2} = \frac{\alpha_2 \beta_2 c}{\alpha_1 \beta_1 a b} = 1. \quad (11)$$

This equation has a very strong thermodynamic meaning: the term $\ln(\sigma/\mu) = \Delta G/(k_B T)$ is the chemical potential difference between $A + B$ and C . If one considers $A + B$ and C as two nodes in a circuit, then ΔG is the potential between them. When $\Delta G \neq 0$, there exists a nonequilibrium chemical driving force exerted on the reaction system.

Mathematically, the ODE (2) can be simplified. Let $u = \alpha_2 x / (\alpha_1 a)$ and $\tau = (\alpha_1 a)^2 t / \alpha_2$, then (2) becomes

$$\frac{du}{d\tau} = -u^3 + u^2 - \mu u + \sigma, \quad (12)$$

in which $\mu, \sigma > 0$. If $\sigma = \mu$, the right-hand-side of (12) becomes $-(u^2 + \mu)(u - 1)$. There is only one unique fixed point, i.e., $u^{\text{eq}} = 1$, the equilibrium point. This result is general. For equilibrium system, the steady state distribution obtained from the CME is always uni-modal, corresponding to the unique fixed point obtained from the Law of Mass Action ODE^[9,11].

Nonequilibrium Steady State, Gaussian Approximation, and Multiscale Dynamics. When $\sigma \neq \mu$, the chemical reaction system is not in detailed balance. In this case, there is a continuous conversion of chemical energy to heat, even in the steady state. Therefore, there is a continuous production of entropy due to the conversion of more useful chemical energy to less useful heat. The entropy production rate

$$epr = k_B T J \ln \frac{\mu}{\sigma}. \quad (13)$$

The nonequilibrium steady-state (NESS) has a net flux in the overall reaction $A + B \rightarrow C$:

$$J = u^2 - u^3 = \mu u - \sigma. \quad (14)$$

It is easy to show that the epr in (13) is always positive in the NESS. This result should be compared with “power = current \times voltage” being always positive in a stationary electrical circuit.

For certain parameter values, say $\mu = 0.25$ and $\sigma = 0.01$, the landscape function $\phi(x)$ in (8) has two minima and one maximum in-between:

$$-u^3 + u^2 - \mu u + \sigma \approx -(u - 0.05)(u - 0.32)(u - 0.63). \quad (15)$$

It is easy to see that the root of $u^3 + \mu u = u^2 + \sigma$ is

precisely the extrema of $\phi(x)$ where

$$\phi'(x) = -\ln \frac{x^2 + \sigma}{x^3 + \mu x} = 0. \quad (16)$$

Therefore, the nonlinear chemical reaction system is *bistable*. The dynamics of the system exhibits multiple time scale: the relaxation within each “well” and transitions between the two wells. The former can be accurately described by a Gaussian (linear) random process. The latter, as two-state transitions, is on a much longer time scale.

It can be shown, according to the CME, that for a closed nonlinear chemical reaction system, its stationary distribution has a unique peak, the equilibrium^[9,11]. Furthermore, the fluctuating dynamics, i.e., the stationary stochastic process in equilibrium is statistically time reversible^[12]. These theoretical results indicate that complex behavior such as chemical bistability indeed can only occur in a “living system” with dissipation, i.e., useful chemical energy is converted into heat, and the process sustains a self-organizing complex dynamical system^[13-14].

Multiscale Dynamics and the Keizer's Paradox. Every CME model contains the parameter V , the volume of the reaction system. When the number of molecules, N , and $V \rightarrow \infty$, the mathematical solution to the CME agrees with that from the Law of Mass Action which describes concentration $x = N/V$ ^[1-2]. For most biochemical models, one might also be interested in the stationary behavior of the solution to the CME. This represents all the numbers of molecules in a reaction system, which are statistically independent of time, with stationary number fluctuations due to the biochemical reactions. One naturally identifies this with the homeostasis of a cell. Mathematically, this means one is interested in the limit of $t \rightarrow \infty$. Hänggi *et al.*^[15] and Baras *et al.*^[16] correctly pointed out, however, there is a delicate computational issue of V (and N) $\rightarrow \infty$ and $t \rightarrow \infty$ and changing the order of the limits can lead to different mathematical predictions. This non-exchangability between $V \rightarrow \infty$ and $t \rightarrow \infty$ has been named Keizer's paradox. One needs to be extra careful in dealing with the steady state behavior of a CME model.

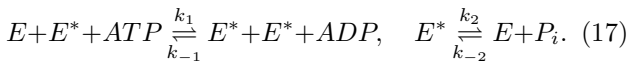
This issue has been re-examined recently^[11,17] in more details. It is shown to be intimately related to the multiple time scales of the bistability. The transition rates between the two states of a bistable system are exponentially small with increasing $V : \propto e^{-\alpha V}$ where α is a positive constant.

One naturally would like to approximate the CME in terms of a Fokker-Planck equation (second order PDE). The Fokker-Planck approximation of the CME

has been discussed in several treatises (e.g., p.116 of [18]). The approach is similar to the diffusion approximation theory for Boltzmann equation (Subsections 3.2 and 3.3 of [18]). Keizer also discussed multiple steady-states in biochemical reaction systems. However, the consequence of the multi-stability with diffusion approximation has not been fully discussed. van Kampen has repeatedly emphasized that the Fokker-Planck approximation can be obtained for master equations only with *small individual jumps*. A more sophisticated treatment of the Fokker-Planck approximation for master equation was given in terms of the Ω -expansion (Ch. 10 of [10]). This theory provides a more satisfying approximation for the stochastic relaxation in the limit of large V . However, it does not address how to obtain the stationary distribution with multistability. Computing such a stationary distribution is a major challenge.

3 Stochastic Bistability in the CME

In the previous section we stated that for sufficiently large V , the CME gives a stationary probability distribution for the numbers of all the dynamical species, which is in complete agreement with the prediction from the Law of Mass Action. A bistable system according to the Law of Mass Action, thus, corresponds to a bimodal distribution in the CME. The converse is not true, however. In recent years, there have been increasingly more examples showing that nonlinear biochemical reaction systems with macroscopic unistability can exhibit bistable behavior in a small volume. These results have important implications to cellular biochemistry. We shall give one example: the phosphorylation-dephosphorylation cycle (PdPC) with autocatalytic kinase^[19]:



If we use x to denote the fraction of the phosphorylated E^* , then according to the Law of Mass Action:

$$\begin{aligned} \frac{dx}{dt} &= \tilde{k}_1 x(1-x) - \tilde{k}_{-1} x^2 - k_2 x + k_{-2}(1-x) \\ &= -(\tilde{k}_1 + \tilde{k}_{-1})x^2 + (\tilde{k}_1 - k_2 - k_{-2})x + k_{-2}, \end{aligned} \quad (18)$$

where $\tilde{k}_1 = k_1 E_t c_T$, $\tilde{k}_{-1} = k_{-1} E_t c_D$, E_t is the total concentration of E and E^* , c_T and c_D are ATP and ADP concentrations. (18) has two steady states, only one is positive and chemically meaningful. Hence there is no bistability in macroscopic size reaction system, with any parameters.

However, if the exactly same nonlinear PdPC is in a small reaction volume such as a cell, then according

to the CME, the stationary probability distribution for the number of E^* is

$$p_{ss}(n) = C \prod_{j=0}^{n-1} \frac{(\hat{k}_1 j + k_{-2})(N_t - j)}{(\hat{k}_{-1} j + k_2)(j+1)}, \quad (19)$$

where $\hat{k}_i = \tilde{k}_i/V$, $i = \pm 1$. C is a normalization factor.

It is easy to check that the distribution in (19) has two peaks, one at $n_1^* = 0$ and the other at n_2^* :

$$n_2^* = \frac{k_2 + k_{-2} + \hat{k}_{-1} - \hat{k}_1 N_t + \left(\frac{(k_2 + k_{-2} + \hat{k}_{-1} - \hat{k}_1 N_t)^2}{-4(\hat{k}_{-1} + \hat{k}_1)(k_2 - k_{-2} N_t)} \right)^{\frac{1}{2}}}{2(\hat{k}_1 + \hat{k}_{-1})}. \quad (20)$$

It is usually not an integer. Hence it exhibits stochastic bistability in a small volume.

4 Biochemical Bistability in a Cell and Epigenetic Inheritance with a Distributive Code

Since the discovery of DNA double helix, it has been well understood that DNA replication is the molecular basis of biological inheritance. However, in addition to DNA based inheritance, epigenetic inheritance has become an increasingly important concept in cell differentiation, stem cell research, as well as bacterial persistence^[20]. Current research has been focusing on several specific molecular processes as the possible “code” for epigenetics, e.g., histone acetylation^[21] and DNA methylation^[22-23]. One of the key issues is that the code has to be sufficiently stable. This leads researchers to look for specific covalent modifications of transcriptional regulation apparatus.

However, specific covalent modifications might not be necessary in some cases. According to the theory of the CME, the stability of a state of a biochemical reaction system, i.e., the peak in the stationary distribution, is due to the biochemical reaction network^[24]. In other words, the epigenetic code could be *distributive*, namely, properties such as state stabilities are the outcome of the collective behavior of many components of a biochemical network^[23]. Therefore, detailed molecular mechanism(s) aside, the nonlinear biochemical reaction network(s) as the foundation of cellular epigenetics has to be valid.

Ptashne has recently re-emphasized the importance of heritability in the term of “epigenetics”^[25]. We shall point out that the states of bi- or multistable nonlinear biochemical reaction systems, as defined above, naturally give rise to heritability. It is important to recall that the function $\phi(x)$ above is *independent* of V , and

the variable x is the concentration. Hence, assuming there is no specific mechanism of regulating the production of molecule X , if the system's volume is increased, the concentration x will go down. However, the nonlinear dynamic nature of the network automatically regulates the system and the steady state concentration of X is regained. Thus, as long as the volume of the system is *slowly* increasing in the synthesis phase of the cell cycle, the concentrations of all the key biochemical species (i.e., transcriptional regulators) are always maintained at its steady state value. Only when the volume change occurs in short time period and the amount of change is sufficiently large, there would be a chance that the system “jumps” into another basin of attraction (Fig.1). If the basins of attraction of states are broad, then a daughter cell will still be in the same state as the parent cell without the need for any additional signal and regulation.

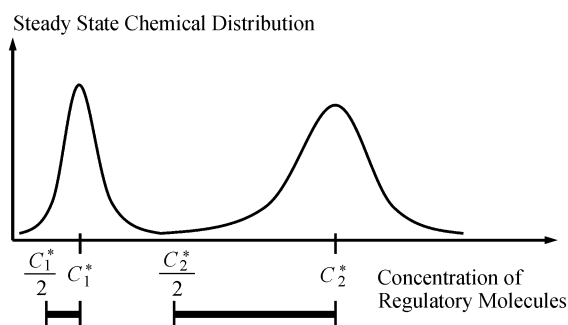


Fig.1. Schematics showing how two biochemical states of a nonlinear biochemical reaction system can be inheritable if the volume of the reaction system is increased, and then divided into two. Note that the abscissa is concentration, not number of molecules. In the figure, an increase in volume with a factor of 2, corresponding to a decrease of concentration to one half, will still maintain the system in its original attractors. Division does not change the concentration.

5 Computational Challenges from the Chemical Master Equation

In the theory of the CME, the dynamics of a biochemical reaction system, in a small volume, is represented by a multi-dimensional, integer-valued stochastic jump process in \mathbb{Z}^n . The process is a discrete-state, continuous-time Markov process. As any Markov process, it can be mathematically characterized either in terms of its ensemble of stochastic trajectories, or by its probability distribution as function of time. These correspond to the stochastic differential equation and the Fokker-Planck equation representations of a Brownian dynamics. The CME is the differential equation for the probability distribution; the stochastic trajectory is defined by the well-known Gillespie

algorithm. In analyzing a CME model, these two approach complement to each other.

One type of chemical reaction systems, the single molecules or uni-molecular reaction system, has been extensively studied in the past. It is important to note that such systems are linear chemical reaction systems. Since all the molecules in uni-molecular reaction systems are statistically independent, it can be represented by either the particle-state-tracking (PST) algorithm or particle-number-tracking (PNT) algorithm^[26]. The simulation can also be carried out approximately, but satisfactorily, by a continuous model of Langevin dynamics^[27]. There is no multistability in such systems; nor complex dynamics.

The difference between PST and PNT is as follows: one either considers the discrete states of the particles in the simulation, or considers the number of particles in a particular state. These two approaches correspond precisely to the Lagrangian and Euler descriptions of fluid particles — in terms of trajectories of particles and in terms of the density^[28]. In the current research on stochastic simulation of biochemical reaction systems, these correspond to the StochSim/MCell^[29] and the StochKit, respectively. The Langevin approximated algorithm is closely related to the linear noise approximation (LNA)^[30]. The LNA can be only valid within each “peak” region, i.e., a basin of attraction, of the CME. For nonlinear reaction systems with multistability, the Keizer’s paradox can occur which invalidates the Langevin approximation for the longer time scale dynamics.

On more general terms, there are many reasons to seek accurate solution to the CME directly, although much has been learned about the overall probabilistic landscape of many biochemical networks through stochastic simulations (Gillespie, StochSim/MCell, and StochKit) and approximated continuous models based on stochastic differential equations. First, details of the topological features and their quantification such as the existence and location of basins of attraction, craters, peaks, and saddle points of various dimensions, their widths, breadth, and depths on the probabilistic landscape, as well as their possible biological implications such as the inheritable epigenetic state arising from the properties of the network are largely unexplored. This is true even for simple reaction systems such as the 2-dimensional Schnakenberg model^[31], which is only slightly more complex than the 1-dimensional Schlögl model discussed above, as there are no general exact probabilistic solutions available yet. Second, accurate solution to the CME problems can facilitate development of approximation methods that are capable of solving large-size problems. There is a large body of studies on theoretical approaches approximating the

CME through the Fokker-Planck, and equivalently Langevin, equations. For effective design of these models and efficient computations of accurate solutions to large biochemical systems, it is essential to have some *a priori* knowledge of the ground truth. Third, perhaps most importantly, an accurate solution to the CME of simpler model systems can reveal important insights into basic principles on how biological networks function and how they respond to various environmental perturbations. A shining example of studying complex systems using manageably simple models is the study of protein folding. Models such as two- and three-dimensional lattice self-avoiding walks with only hydrophobic and polar (HP) interactions allow complete enumeration of all feasible conformations and calculation of exact thermodynamic parameters for molecules with short chain lengths. They have played important roles in elucidating the principles of protein folding^[32], including collapse and folding transitions^[33–40], influence of packing on secondary structure and void formation^[41–44], the evolution of protein function^[45–46], nascent chain folding^[47], and the effects of chirality and side chains^[44].

6 State Space of the Chemical Master Equation and Exact Calculation of Steady State Probability Landscape

The state space of the CME is that of M -dimensional vectors with non-negative integers, which represents the copy numbers of molecular species in a network; M is the number of dynamic species. These states are microscopic in nature, as they provide a detailed, chemical amount of each and every molecular species. An important advantage of treating these microscopic states of copy numbers explicitly is that both linear and nonlinear reactions (such as synthesis, degradation, bimolecular association, and polymerization) can be modeled as Markovian transitions between two microstates, one reaction at a time. Here the transition rates between states are determined by the intrinsic propensities of the reaction, and the copy numbers of molecules involved.

For any biochemical systems beyond the simplest toy problems, a challenging issue in obtaining an accurate solution to the CME is the characterization of the state space. That is, what are all the possible combinations of concentrations (or copy numbers) of the molecular species for a given set of reactions represented by a network? An accurate description of the state space is a prerequisite for computationally obtaining solutions to the CME. In principle, the size of the state space grows exponentially with the number of molecular species and the copy numbers of molecules in the

system. For example, if there are 16 molecular species in a network, and there are only a total of 33 copies of molecules in the whole system, one can estimate somewhat naively the upper bound of the state space as $(33 + 1)^{16} = 3.19 \times 10^{24}$. Note the +1 counts the zero copy as a state. This is an astronomic number that is well beyond what can be computed with current and for-seeable future computing technology.

Below we discuss the enumeration of the state space of CME and exam how to obtain exact steady state solutions to the CME for biochemical systems with small and moderate sizes.

Optimal Enumeration of State Space. Although in principle the size of the state space grows exponentially with the number of molecular species and the copy numbers of molecules in the system, all is not lost. There are two important observations about general biochemical networks. First, the Markovian transition matrix is very sparse. For any given microstate, the number of reactions that could occur in a short time interval is small, which could be bounded by the total number of possible reactions in a biochemical network. Second, as an open system, molecules are synthesized and degraded constantly. However, the number of molecules that can be synthesized is never infinite, as synthesis is constrained by the time and resources required. With these two considerations, an algorithm to enumerate the state space of CME has recently been developed^[48]. The algorithm is optimal in memory requirement, as it allows the enumeration of all states that can be reached from a given initial state, without including any irrelevant states. In addition, all possible transitions are recorded, and no infeasible transitions are attempted. The resulting transition matrix based on the enumerated state space is compact without redundant information, and is minimal in size. In addition, its computational time is also optimal^[48].

Exact Calculation of Probability Distribution of the Steady State. Once the states reachable from a given initial state are enumerated, the rates of chemical reactions connecting two of these states can be computed. For example, we can study a simplified model of protein-DNA interaction. For the process of two protein monomer (*ProteinA*) dimerize and bind to a segment of DNA (*GeneB*), we can use the simplified model below. If we denote the rate of the reaction that brings the before-state i to the after-state j as $a_{j,i}$, we have for the third order reaction:



with the following reaction rate coefficient $a_{j,i}$:

$$a_{j,i} = b \cdot n_{gB,i} \cdot n_{pA,i} \cdot (n_{pA,i} - 1)/2,$$

where b is the intrinsic reaction rate which contains

hidden systems volume V , $n_{pA,i}$ is the copy number of protein A in state i , and $n_{gB,i}$ is the copy number of unbound gene B . Here the combination number of the protein for this second order reaction is $\binom{n_{pA,i}}{2} = n_{pA,i} \cdot (n_{pA,i} - 1) / 2$. Note that in addition to the volume V , there is a factor of 2 difference between the intrinsic reaction rate here and the macroscopic rate constant discussed in Section 2.

Once the full reaction rate matrix $\mathbf{A} = \{a_{j,i}\} \in \mathbb{R}^{n \times n}$ is filled with computed rates, the chemical master equation can be written in a matrix-vector form as:

$$\dot{\mathbf{P}}(t) = \mathbf{A}\mathbf{P}(t). \quad (21)$$

Here the matrix \mathbf{A} represents the *infinitesimal generator* of a continuous time Markov process. The diagonal elements a_{ii} is set as: $a_{ii} = -\sum_{i \neq j} a_{j,i}$, and all off-diagonal elements are nonnegative. The analytical solution at time t to (21) can be written as a matrix exponential:

$$\mathbf{P}(t) = \exp(\mathbf{A}t)\mathbf{P}(0). \quad (22)$$

The matrix $e^{\mathbf{A}t}$ is the Markovian state transition probability matrix with time duration t . We can also obtain its discrete equivalent \mathbf{M} as^[40]:

$$\mathbf{M} = \mathbf{I} + \mathbf{A} \cdot \Delta t, \quad (23)$$

where \mathbf{I} is the identity matrix, Δt is a small time interval during which one reaction occurs. When the system has reached the steady state, the probability landscape over the enumerated states \mathbf{P} can be computed by solving the equation:

$$\mathbf{P} = \mathbf{M}\mathbf{P}.$$

Here \mathbf{P} can be obtained with an iterative solver such as that based on the successive over-relaxation (SOR) technique^[49]. Alternatively, since \mathbf{P} for the steady state corresponds to the eigenvector of \mathbf{M} with eigenvalue 1.0, one can obtain \mathbf{P} by using eigenvector method such as the Arnoldi method^[50], as done in [48].

By examining computationally the stochastic behavior of genetic circuits for wild type and mutant networks, and by studying the probabilities of rare events, one can gain further understanding of the regulation mechanism of genetic circuits, its system stability against perturbation (such as fluctuations in nutritional conditions), and its robustness against genetic mutations (such as those due to DNA damage)^[51].

7 Two Examples of Stochastic Biochemical Systems and Their CMEs

In this section we give two examples on how exact stationary probability landscapes of a biochemical network can be computed from its CME. The CME, of

course, gives more than just a stationary distribution, but solving the steady state is almost obligatory in any analysis of mathematical models.

Toggle Switch. In Section 3, we already discussed how bistability arises from stochasticity. Another example is the well studied genetic toggle-switch system. This is a small network consisting of two genes, A and B , each inhibits the other (Fig.2). It was the first synthetic network constructed in a wet lab from two repressible promoters arranged in a mutually inhibitory network in *Escherichia coli* by Gardner *et al.*^[52]. It is flippable between two stable states by chemical or thermal induction and exhibits an ideal switching threshold. This toggle switch forms a synthetic cellular memory unit^[52]. Although this is the simplest network with bistability that can already be identified from ODE models based on the Law of Mass Action, important questions such as switching probability between the “on” and “off” states requires a treatment of the stochasticity. Although there have been great recent progresses in deriving analytical solutions^[53-56], they are applicable under special conditions, such as fast transition between the on- and off-states, or overall

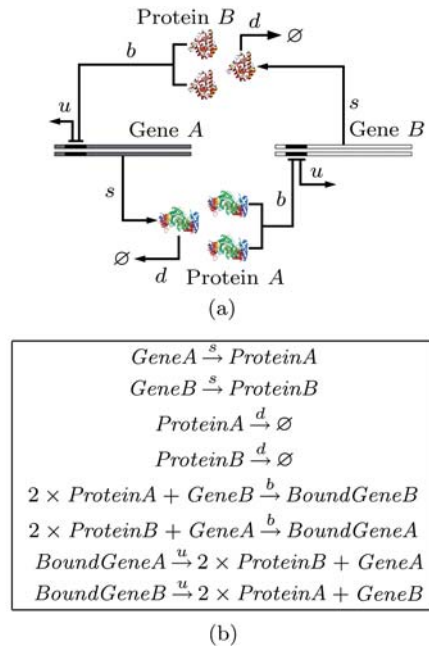


Fig.2. Model of a toggle switch. (a) The network model and the reaction rates. Single copies of gene A and gene B encode protein products. Two protein monomers can repress the transcription of the other gene. The synthesis of protein product of gene A and gene B depends on the bound or unbound state of the gene. (b) The chemical reactions of the 8 stochastic processes involved in the toggle-switch system. The reaction rates include s for protein synthesis, d for protein degradation, b for protein-gene binding, and u for protein-gene unbinding (adapted from [48]).

small noise associated with high concentrations. With the algorithm for state enumeration, the steady state landscape probability of the toggle-switch can be solved exactly for models with arbitrary parameter specifications.

Epigenetic Switch in Phage Lambda. Exact solution of the CME can also be obtained for larger systems in which biological phenomenon are modeled more realistically. An example is the epigenetic switch of phage lambda. Phage lambda is a virus that infects *E. coli* bacteria. It is the system in which gene regulation was first studied. Upon infection, phage lambda can choose two different life styles. In the lysogenic pathway, the DNA of phage lambda becomes integrated into the chromosome of the host, and can replicate for many generations along with the host. Upon adverse environmental perturbations such as UV irradiation, phage lambda switches from the lysogenic pathway to the lytic pathway, in which it uses the protein synthesis machinery of the host, and replicate to 100s of copies, which leads to the burst of the host cell. The lytic pathway offers critical evolutionary advantage for

phage lambda to survive, as it allows phage to escape from hopelessly distressed *E. coli* host cells. In phage lambda, a gene regulatory circuit controls the switching between the maintenance of the lysogenic state and the induction of the lytic state. The CME model analysis clearly demonstrates the idea of a distributive epigenetic code. As a paradigm for understanding cell development, phage lambda has been extensively studied, with the molecular components and reaction rates well characterized (see the seminal book by Ptashne^[57]). The key components of the switch of the genetic circuits and their wirings can be summarized in Fig.3. There are three operators (OR1, OR2, and OR3) and two promoters (Pr and Prm). These are used to control the transcription of CI and Cro proteins, which dimerize and bind to the operator sites with different affinity and inhibit the expression of each other^[57].

The importance of stochasticity in the genetic circuit of lambda phage is well recognized, and its effects have been studied using stochastic simulations^[58] and stochastic differential equations^[24,59]. The steady state probability landscape of the CME model based on the

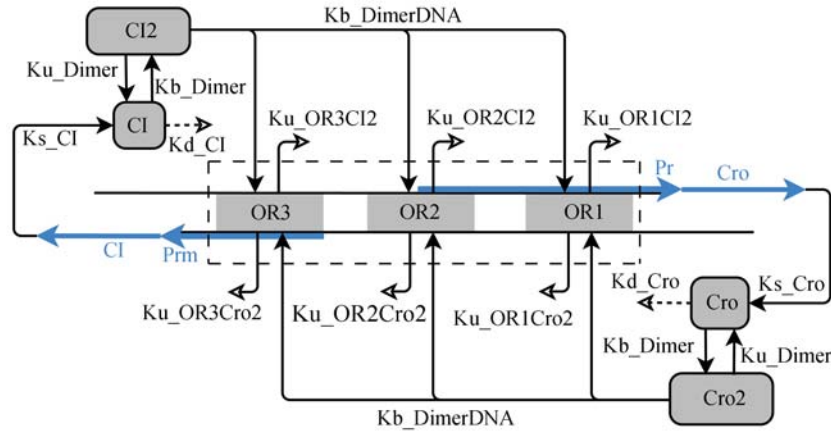


Fig.3. Phage λ switching network. Reactions including binding and unbinding, synthesis and degradation, dimerization are labeled as arrows, along with the corresponding kinetic constants (adapted from [51]).

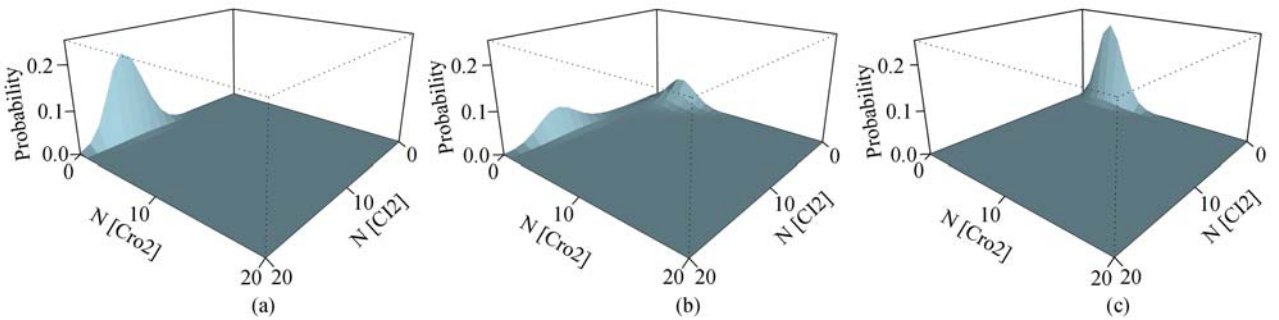


Fig.4. Lysogenic and lytic states and CI synthesis rate. (a) Lysogenic state, $K_{s_CI}=0.045/s$. (b) The switching state, $K_{s_CI}=0.0245/s$. (c) Lytic state, $K_{s_CI}=0.0077/s$. X and Y axes are copy numbers of CI and Cro dimers; and Z axis is the marginal probability (adapted from [51]).

network depicted in Fig.3 can be solved directly^[51]. Fig.4 shows the probability landscape under several physiological conditions when the system is in the lysogenic state, in transitory switching state, and in lytic state^[51]. Fig.5 shows the phase diagram of concentrations of CI and Cro at different CI synthesis rate.

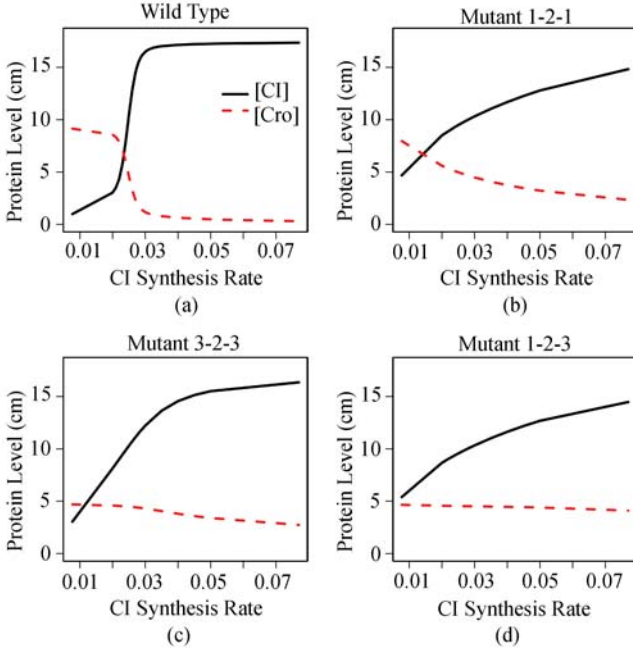


Fig.5. Relative CI and Cro dimer levels for wild type and mutant lambda phage. The lysogenic state has high CI (solid line) concentration, and the lytic state has high Cro (dashed line) concentration. Wild type lambda phage has a well-behaving switch, while mutants are all leaky (adapted from [51]).

By examining computationally the stochastic behavior of genetic circuits for wild type and mutant network, and by studying the probabilities of rare events, one can gain further understanding of the regulation mechanism of genetic circuits, its system stability against perturbation (such as fluctuations in nutritional conditions), and its robustness against genetic mutations (such as those due to DNA damage)^[51].

8 Methods for State Space Simplification

For large systems in which enumeration is no longer feasible, one approach for numerical computation is to reduce the large number of microstates to a smaller finite number^[60].

Finite State Projection. Minsky and Khammash made two key observations about projecting the high dimensional state space to a lower dimensional finite space by including only a subset of the original states. Denote two sets of indices of the states being chosen as J_1 and J_2 , and assume $J_1 \subseteq J_2$. The reduced rate

matrix obtained by selecting states in J_1 and J_2 are \mathbf{A}_{J_1} and \mathbf{A}_{J_2} , respectively. The first observation is:

$$(\mathbf{e}^{\mathbf{A}_{J_2}})_{J_1} \geq \mathbf{e}^{\mathbf{A}_{J_1}} \geq 0. \quad (24)$$

This relation implies that by increasing the size of the selected subset of states, the approximation improves monotonically. Second, if one obtains a reduced state space by selecting states contained in the index set J and if $\mathbf{1}^T \mathbf{e}^{t\mathbf{A}_J} \mathbf{P}_J(0) \geq 1 - \epsilon$ for $\epsilon > 0$ and $t \geq 0$, then:

$$\mathbf{e}^{t\mathbf{A}_J} \mathbf{P}_J(0) \leq \mathbf{P}_J(t) \leq \mathbf{e}^{t\mathbf{A}_J} \mathbf{P}_J(0) + \epsilon \mathbf{I}. \quad (25)$$

That is, starting with the initial probability of the reduced vector $\mathbf{P}_J(0)$, compute the probability vector in the reduced space $\mathbf{e}^{t\mathbf{A}_J} \mathbf{P}_J(0)$ at time t using the reduced rate matrix \mathbf{A}_J . If the inner-product of this vector for time t with $\mathbf{1}$ is no less than $1 - \epsilon$, then the error of this vector with the projected true vector $\mathbf{P}_J(t)$ from the true probability $\mathbf{P}(t)$ is no more than $\epsilon \mathbf{I}$. This inequality guarantees that the approximation obtained with reduced state space will never exceed the actual solution, and its error is bounded by ϵ ^[60].

These key observations led to the Finite State Project Algorithm, which iteratively adds new states to an initial reduced state space, until the approximation error is within a prescribed bound^[60]. Minsky and Khammash further extended the original Finite State Projection method^[61], and recommends running a few steps of stochastic simulation to obtain the initial probability vector $\mathbf{P}(0)$ that is non-sparse. However, there are no generally applicable strategies as to what states to add to a finite projection to most efficiently improve the approximation accuracy.

Krylov Subspace Method. The analytical solution to the CME can be expressed in the form of a matrix exponential $\mathbf{P}(t) = \mathbf{e}^{t\mathbf{A}} \mathbf{P}(0)$. As discussed before, the rate matrix \mathbf{A} has a very large dimension but is sparse. An alternative approach to reduced the state space is to convert the problem of exponentiating a large sparse matrix to that of exponentiating a small dense matrix in the Krylov subspace \mathcal{K}_m ^[62]:

$$\mathcal{K}_m(\mathbf{A}t, \mathbf{P}(0)) \equiv \text{Span}\{\mathbf{P}(0), \dots, (\mathbf{A}t)^{m-1} \mathbf{P}(0)\}. \quad (26)$$

The idea is that the Krylov subspace used is of a very small dimension of $m = 30-60$. Denoting $\|\cdot\|_2$ as the 2-norm of a vector or matrix, the approximation then becomes $\mathbf{P}(t) \approx \|\mathbf{P}(0)\|_2 \mathbf{V}_{m+1} \exp(t\overline{\mathbf{H}}_{m+1}) \mathbf{e}_1$, where \mathbf{e}_1 is the first unit basis vector, \mathbf{V}_{m+1} is a $(m+1) \times (m+1)$ matrix formed by the orthonormal basis of the Krylov subspace, and $\overline{\mathbf{H}}_{m+1}$ the upper Hessenberg matrix, both computed from an Arnoldi algorithm^[63]. The

error can be bounded by

$$\mathcal{O}(e^{m-t\|\mathbf{A}\|_2}(t\|\mathbf{A}\|_2/m)^m).$$

One only needs to compute explicitly $\exp(\overline{\mathbf{H}}_{m+1}t)$. This is a simpler problem as m is much smaller. A special form of the well-known Padé rational of polynomials instead of Taylor expansion is used^[64-65]:

$$e^{t\overline{\mathbf{H}}_{m+1}} \approx N_{\text{pp}}(t\overline{\mathbf{H}}_{m+1})/N_{\text{pp}}(-t\overline{\mathbf{H}}_{m+1}),$$

where $N_{\text{pp}}(t\overline{\mathbf{H}}_{m+1}) = \sum_{k=0}^p c_k(t\overline{\mathbf{H}}_{m+1})^k$ and $c_k = c_{k-1} \cdot \frac{p+1-k}{(2p+1-k)k}$. The EXPOKIT software by Sidje provides an excellent implementation of this method^[65]. This approach has been shown to be very effective in studying large dynamic system ($n = 8.0 \times 10^5$) such as protein folding^[40] and signaling transmission in macromolecular assembly of GroEL-GroES^[66].

The Krylov subspace method concurrently evaluate the matrix exponential. The overall scheme can be expressed as follows:

$$\begin{aligned} \mathbf{P}(t) &\approx \exp(\tau_K \mathbf{A}_K) \dots \exp(\tau_0 \mathbf{A}_0) \mathbf{P}(0), \\ t &= \sum_{k=0}^K \tau_k, \end{aligned} \quad (27)$$

in which the evaluation is from right to left. Here $\{\tau_i\}$ are size of time steps, and K is the total number of time steps^[62].

MacNamara *et al.* further extends the Krylov subspace method by splitting the rate matrix \mathbf{A} . Based on the reachability criteria, one can divide the states into the “fast partition” and the “slow partition”^[67]. Here the condition is that two states belong to the same subset of the fast partition if and only if one can be reached from the other via a sequence of finite fast reactions^[67]. Correspondingly, the matrix can be splitted into two:

$$\mathbf{A} = \mathbf{A}_f + \mathbf{A}_s,$$

where \mathbf{A}_f corresponds to the fast CME, and \mathbf{A}_s corresponds to the slow CME, and one has:

$$\dot{\mathbf{P}}_f(t) = \mathbf{A}_f \mathbf{P}_f(t)$$

and

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

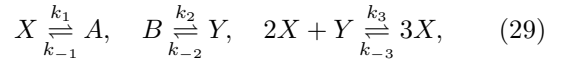
With this deliberate separation, both \mathbf{A}_f and \mathbf{A}_s maintain the important property of being infinitesimal generators of continuous time Markov processes by themselves^[67]. With more elaborated splitting scheme for aggregation of Markov processes, the Krylov subspace projection method have been shown to be computationally very efficient^[62].

Approximation by Continuous Stochastic Differential Equation. An effective approach to study biochemical networks whose chemical master equations cannot be solved directly is to approximate them with stochastic differential equations. One widely used approach is that of the Fokker-Planck-Langevin model^[10]. The Langevin equation for concentration flux consists of a *drift* term and a *diffusion* term. The drift term models the macroscopic deterministic component of the system. It reflects the time-dependent evolution of the mean concentrations of the molecular species. The diffusion term models the intrinsic stochasticity of the system. The basic form of a Langevin, stochastic differential equation is:

$$\frac{d\mathbf{X}}{dt} = \mu(\mathbf{X}) + \sigma(\mathbf{X})\mathcal{N}\left(0, \frac{1}{dt}\right). \quad (28)$$

Here \mathbf{X} is the vector of concentrations of molecular species in the reaction system, $\mu(\mathbf{X})$ the drift term, and the second term is the diffusion term. Here $\mathcal{N}(0, 1/dt)$ is a vector of one-dimensional Gaussians, with zero mean and $1/dt$ variance. The coefficient $\sigma(\mathbf{X})$ controls the amplitude of the Gaussian noise. It can be either a function of \mathbf{X} or a constant. The key issue in developing Langevin models for biochemical networks is to determine $\mu(\mathbf{X})$ and $\sigma(\mathbf{X})$. When $\sigma(\mathbf{X})$ is a vector of constants, one adjusts its values so the variance of the Gaussian noise produce the correct fluctuations in the system^[10].

One of the most important issues to keep in mind when developing Fokker-Planck-Langevin approximations for a CME is the Keizer’s paradox previously discussed. For dynamical system with a single, globally attracting attractor, however, this is not an issue. We shall use the Schnakenberg model to demonstrate how well the Langevin approach works. Originally developed for studying the limit cycle behavior in a simple chemical reaction system^[31,69], the Schnakenberg model is a simple system with two reacting components and three reversible reactions:



where X and Y are reacting species of the system, and A and B are external reactants whose concentrations (or copy numbers) are fixed constants. Each reaction has a corresponding microscopic reaction rate. The fixed copy numbers or concentrations of A and B can be adjusted, which lead to different behavior of the system. This simple system already produces complex behavior such as oscillation and has a single stable limit cycle (see [6, 70] for recent examples).

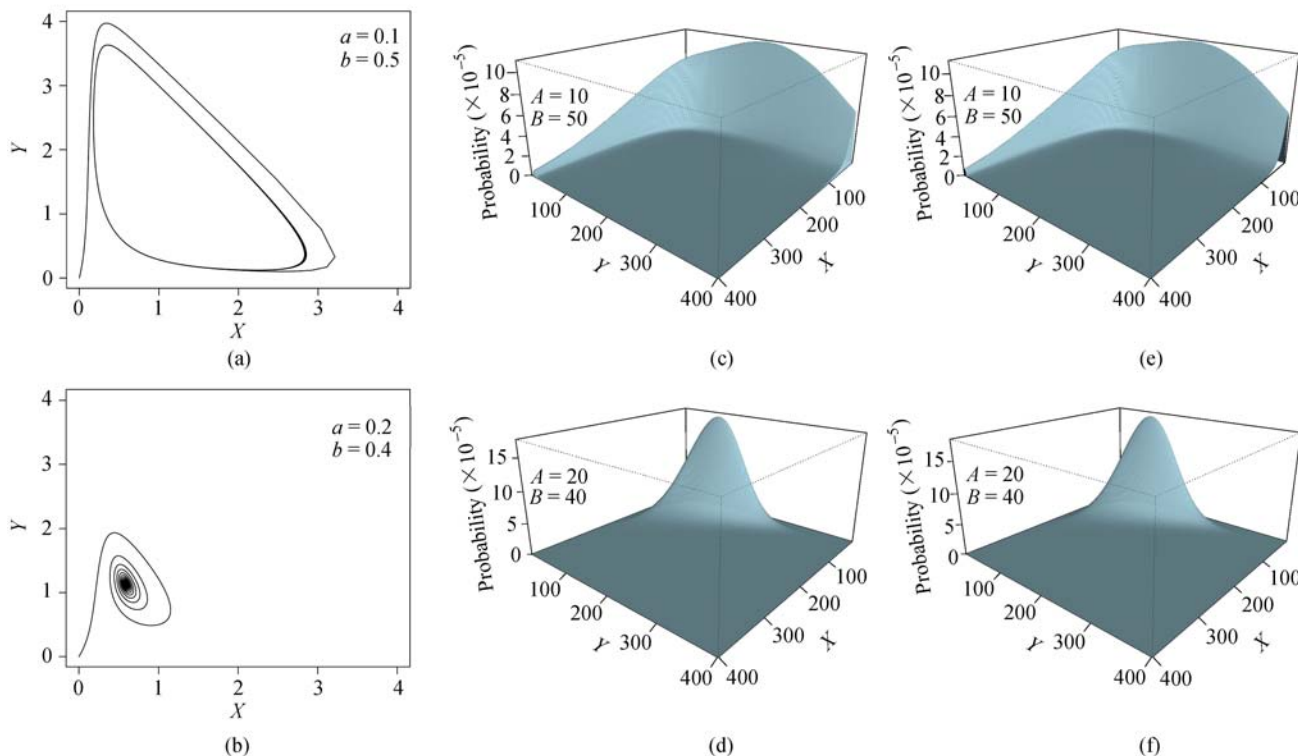


Fig.6. Calculated steady state probability distributions over different copy numbers of X and Y and the trajectories of evolving concentrations of X and Y of the Schnakenberg model. (a) and (b): Trajectories of evolving concentrations of X and Y according to the deterministic ordinary differential equation (ODE). Here (a) shows the well-known oscillating limit cycle behavior of the Schnakenberg model, and (b) shows the convergence towards a fixed point. The concentrations of A and B are set at values equivalent to the copy numbers used in stochastic models. (c) and (d): Reconstructed probability distributions over X and Y obtained from 200 000 simulations of the Langevin equation (LE). (e) and (f): Exact probability distributions over copy numbers X and Y obtained by solving the chemical master equation (CME). Two sets of copy numbers of (A, B) at $(10, 50)$ and $(20, 40)$ are used for the fixed parameters A and B (adapted from [68]).

The macroscopic concentration obtained by solving the corresponding ODE model, the approximated steady state probability distribution obtained by integrating the Langevin model, and the exact probability distributions obtained by solving the chemical master equation^[48] are shown in Fig.6. At the parameter values of $A = 10$ and $B = 50$, the well-known oscillating limit cycle behavior of the Schnakenberg model can be seen in Fig.6(a). At $A = 20$ and $B = 40$, the behavior of the system converges towards a fixed point (Fig.6(b)). The landscapes of the steady state probability distributions obtained from solving the Langevin equation (Fig.6(c)) and the chemical master equation (Fig.6(e)) all show a crater, or a basin surrounded by a mountainous ridge for the parameter set of $A = 10$ and $B = 50$ (details not shown). This corresponds well with the limit cycle behavior observed in the ODE model. At $A = 20$ and $B = 40$, the landscapes show a single peak (Figs. 6(c) and 6(d)), which again corresponds well with the fixed-point behavior observed in the ODE model.

As can be seen in Figs. 6, and 7, the model of Langevin equation approximates well the true probability landscape of the chemical master equation. This demonstrates that the diffusion term models the intrinsic stochasticity of the Schnakenberg model well.

Alternative models account for the stochasticity by replacing the diffusion term with a term for the variance-covariance between pairs of the molecular reactions^[71], or between concentrations of different molecular species, without the explicit inclusion of a random process^[72]. The magnitude of the covariance is determined by the Hessian matrix of the second-order partial derivative of the propensity functions of the reactions^[71-72]. This inclusion of the second moments to account for the stochasticity is the basis of the stochastic kinetic model^[71] and the mass fluctuation kinetic model (MFK)^[72]. These models can model reactions involving one or two molecules well^[71-72]. They are similar in spirit to the Fokker-Planck equation model of the CME as described in [73] by including a second moment term for better approximation, but are

different from that of [73] as they are macroscopic in nature and do not involve any random processes.

Yet another approach is to directly model explicitly the stochastic coupling of the macroscopic concentrations of molecular species, in addition to the Gaussian noise of the original Langevin model^[68]. The steady state probability landscape of the Schnakenberg model resulting from this approach is shown in Fig.7. Significant improvement after incorporating the coupling term can be seen in Fig.7.

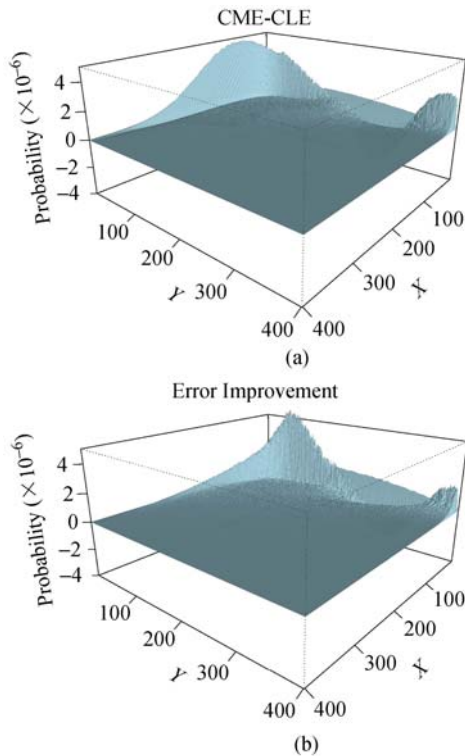


Fig.7. Comparison of errors between different steady state solutions of the Schnakenberg model. (a) Difference between the probability landscapes of the Langevin equation and that of the chemical master equation. This represents errors in the Langevin model. (b) The amount of the errors in (a) that are corrected by introducing explicitly a coupling term between X and Y (adapted from [68]).

Remark. The complex nature of the stochastic dynamics arising from biochemical networks bears much resemblance to another complex system, namely, that of protein folding. Both have very large space of microstates, and both can be modeled by transitions between micro-states using master equations (for master equation approach in protein folding studies, see [37, 39-40]). However, these two systems differ in several important aspects. First, while protein folding can be modeled as a relaxation process towards the equilibrium state, biochemical networks are intrinsically open, with synthesis

and degradation of molecules an integral part of the system, hence there are no equilibrium states. Instead, one frequently seeks to study the non-equilibrium steady state. Second, once the energy of a protein conformation is known, the relative probability of its sequence adopting this conformation in the equilibrium state can be calculated from the Boltzmann distribution, without the need of knowing all other possible conformations and their associated probabilities. In other words, the protein folding problem is *local* in the energy landscape. In contrast, it is not possible to calculate the relative probability of a specific microstate of copy numbers *a priori* without solving the entire CME, as the probability distribution of network states do not generally follow any specific analytical forms (no detailed balance and the existence of cycle fluxes). Third, transitions between microstates are clearly defined in biochemical networks by the reactions, whereas transitions between different protein conformations often technically depend on specific move sets, which are different in terms of allowable transitions between states and transition rates, although all such move-sets are developed with the goal to mimic physical movement of molecules.

9 Discussions and Outlooks

In this review, we have discussed the significance of the chemical master equation (CME) as a theoretical framework for modeling nonlinear, biochemical reaction networks inside cells, and the possible mechanism of cellular states, or attractors, as the inheritable phenotypes with a distributive epigenetic code. The validity of such a grand theory requires close comparisons between theoretical predictions with experiments. Solving a given CME, however, is a computationally challenging task at the present time. We have outlined several key difficulties, as well as some of the progresses that have been made so far.

In addition to the subject of studying algorithmic complexity, complex system, in a broad sense, is a major scientific problem of computer science and computational science^[74]. One needs not to be reminded of the complex phenomena exhibited in the natural world of her/his surroundings. How to characterize and quantify such complex behavior is of great interests for understanding our physical and biological worlds. But what is complexity and how to define complex behaviors? Through studies of the CME, one seems to be able to gain some deeper understanding of the issues involved through concrete physical and biology examples. Recently, one of us has suggested that a key to mesoscopic complexity^[75] is in the multi-stability with multiple time scale dynamics^[76]. Nonlinear biochemical reaction systems in a cell-size volume can be a prototype

for studying complexity.

Acknowledgement We thank Professors Ping Ao, Bai-Lin Hao, Shou-Dan Liang, Kim Sneppen, Jin Wang, and Peter Wolynes for helpful discussions, and Youfang Cao for reading a draft of the manuscript.

References

- [1] Kurtz T G. The relationship between stochastic and deterministic models for chemical reactions. *J. Chem. Phys.*, 1972, 57(7): 2976-2978.
- [2] Beard D A, Qian H. Chemical Biophysics: Quantitative Analysis of Cellular Systems. London: Cambridge Univ. Press, 2008.
- [3] Wilkinson D J. Stochastic Modeling for Systems Biology. New York: Chapman & Hall/CRC, 2006.
- [4] Schlögl F. Chemical reaction models for non-equilibrium phase transition. *Z. Physik.*, 1972, 253(2): 147-161.
- [5] Murray J D. Mathematical Biology: An Introduction. 3rd Ed., New York: Springer, 2002.
- [6] Qian H, Saffarian S, Elson E L. Concentration fluctuations in a mesoscopic oscillating chemical reaction system. *Proc. Natl. Acad. Sci. USA*, 2002, 99(16): 10376-10381.
- [7] Taylor H M, Karlin S K. An Introduction to Stochastic Modeling. 3rd Ed., New York: Academic Press, 1998.
- [8] Resat H, Wiley H S, Dixon D A. Probability-weighted dynamic Monte Carlo method for reaction kinetics simulations. *J. Phys. Chem. B*, 2001, 105(44): 11026-11034.
- [9] Gardiner C W. Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences. 3rd Ed., New York: Springer, 2004.
- [10] van Kampen N G. Stochastic Processes in Physics and Chemistry. 3rd Ed., Amsterdam: Elsevier Science, 2007.
- [11] Vellela M, Qian H. Stochastic dynamics and nonequilibrium thermodynamics of a bistable chemical system: The Schlögl model revisited. *J. R. Soc. Interf.*, 2009, 6(39): 925-940.
- [12] Qian H, Qian M, Tang X. Thermodynamics of the general diffusion process: Time-reversibility and entropy production. *J. Stat. Phys.*, 2002, 107(5/6): 1129-1141.
- [13] Schrödinger E. What Is Life? The Physical Aspect of the Living Cell. New York: Cambridge Univ. Press, 1944.
- [14] Nicolis G, Prigogine I. Self-Organization in Nonequilibrium Systems. New York: Wiley-Interscience, 1977.
- [15] Hänggi P, Grabert H, Talkner P, Thomas H. Bistable systems: Master equation versus Fokker-Planck modeling. *Phys. Rev. A*, 1984, 29(1): 371-378.
- [16] Baras F, Mansour M M, Pearson J E. Microscopic simulation of chemical bistability in homogeneous systems. *J. Chem. Phys.*, 1996, 105(18): 8257-8261.
- [17] Vellela M, Qian H. A quasistationary analysis of a stochastic chemical reaction: Keizer's paradox. *Bull. Math. Biol.*, 2007, 69(5): 1727-1746.
- [18] Keizer J. Statistical Thermodynamics of Nonequilibrium Processes. New York: Springer-Verlag, 1987.
- [19] Bishop L, Qian H. Stochastic bistability and bifurcation in a mesoscopic signaling system with autocatalytic kinase. *Biophys. J.*, 2010. (in the press)
- [20] Kussell E, Kishony R, Balaban N Q, Leibler S. Bacterial persistence: A model of survival in changing environments. *Genetics*, 2005, 169(4): 1804-1807.
- [21] Turner B M. Histone acetylation and an epigenetic code. *Bioessays*, 2000, 22(9): 836-845.
- [22] Jones P A, Takai D. The role of DNA methylation in mammalian epigenetics. *Science*, 2001, 293(5532): 1068-1070.
- [23] Dodd I B, Micheelsen M A, Sneppen K, Thon G. Theoretical analysis of epigenetic cell memory by nucleosome modification. *Cell*, 2007, 129(4): 813-822.
- [24] Zhu X M, Yin L, Hood L, Ao P. Robustness, stability and efficiency of phage λ genetic switch: Dynamical structure analysis. *J. Bioinf. Compt. Biol.*, 2004, 2(4): 785-817.
- [25] Ptashne M. On the use of the word "epigenetic". *Curr. Biol.*, 2007, 17(7): R233-R236.
- [26] Mino H, Rubinstein J T, White J A. Comparison of algorithms for the simulation of action potentials with stochastic sodium channels. *Ann. Biomed. Eng.*, 2002, 30(4): 578-587.
- [27] Fox R F. Stochastic versions of the Hodgkin-Huxley equations. *Biophys. J.*, 1997, 72(5): 2069-2074.
- [28] Lamb H. Hydrodynamic. New York: Dover, 1945.
- [29] Morton-Firth C J, Bray D. Predicting temporal fluctuations in an intracellular signalling pathway *J. Theoret. Biol.*, 1998, 192(1): 117-128.
- [30] Elf J, Ehrenberg M. Fast evaluation of fluctuations in biochemical networks with the linear noise approximation. *Genome Res.*, 2003, 13(11): 2475-2484.
- [31] Vellela M, Qian H. On Poincaré-Hill cycle map of rotational random walk: Locating stochastic limit cycle in reversible Schnakenberg model. *Proc. Roy. Soc. A: Math. Phys. Engr. Sci.*, 2009. (in the press)
- [32] Dill K A, Bromberg S, Yue K, Fiebig K M, Yee D P, Thomas P D, Chan H S. Principles of protein-folding — A perspective from simple exact models. *Prot. Sci.*, 1995, 4(4): 561-602.
- [33] Šali A, Shakhnovich E I, Karplus M. How does a protein fold? *Nature*, 1994, 369(6477): 248-251.
- [34] Socci N D, Onuchic J N. Folding kinetics of protein like heteropolymer. *J. Chem. Phys.*, 1994, 101: 1519-1528.
- [35] Shrivastava I, Vishveshwara S, Cieplak M, Maritan A, Banavar J R. Lattice model for rapidly folding protein-like heteropolymers. *Proc. Natl. Acad. Sci. U.S.A.*, 1995, 92(20): 9206-9209.
- [36] Klimov D K, Thirumalai D. Criterion that determines the foldability of proteins. *Phys. Rev. Lett.*, 1996, 76(21): 4070-4073.
- [37] Cieplak M, Henkel M, Karbowski J, Banavar J R. Master equation approach to protein folding and kinetic traps. *Phys. Rev. Lett.*, 1998, 80(16): 3654-3657.
- [38] Mélin R, Li H, Wingreen N, Tang C. Designability, thermodynamic stability, and dynamics in protein folding: A lattice model study. *J. Chem. Phys.*, 1999, 110(2): 1252-1262.
- [39] Ozkan S B, Bahar I, Dill K A. Transition states and the meaning of ϕ -values in protein folding kinetics. *Nature Struct. Biol.*, 2001, 8(9): 765-769.
- [40] Kachalo S, Lu H, Liang J. Protein folding dynamics via quantification of kinematic energy landscape. *Phys. Rev. Lett.*, 2006, 96(5): 058106.
- [41] Chan H S, Dill K A. Compact polymers. *Macromolecules*, 1989, 22(12): 4559-4573.
- [42] Chan H S, Dill K A. The effects of internal constraints on the configurations of chain molecules. *J. Chem. Phys.*, 1990, 92(5): 3118-3135.
- [43] Liang J, Zhang J, Chen R. Statistical geometry of packing defects of lattice chain polymer from enumeration and sequential Monte Carlo method. *J. Chem. Phys.*, 2002, 117(7): 3511-3521.
- [44] Zhang J, Chen Y, Chen R, Liang J. Importance of chirality and reduced flexibility of protein side chains: A study with square and tetrahedral lattice models. *J. Chem. Phys.*, 2004, 121(1): 592-603.
- [45] Williams P D, Pollock D D, Goldstein R A. Evolution of functionality in lattice proteins. *J. Mole. Graph. Modelling*, 2001, 19(1): 150-156.

- [46] Bloom J D, Wilke C O, Arnold F H, Adami C. Stability and the evolvability of function in a model protein. *Biophys. J.*, 2004, 86(5): 2758-2764.
- [47] Lu H M, Liang J. A model study of protein nascent chain and cotranslational folding using hydrophobic-polar residues. *Prot. Struct. Funct. Bioinf.*, 2008, 70(2): 442-449.
- [48] Cao Y, Liang J. Optimal enumeration of state space of finitely buffered stochastic molecular networks and exact computation of steady state landscape probability. *BMC Syst. Biol.*, 2008, 2: 30.
- [49] Barrett R, Berry M, Chan T F, Demmel J, Donato J, Dongarra J, Eijkhout V, Pozo R, Romine C, van der Vorst H. Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods. 2nd Ed., Philadelphia, PA: SIAM, 1994.
- [50] Lehoucq R, Sorensen D, Yang C. Arpack Users' Guide: Solution of Large Scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods. Philadelphia, PA: SIAM, 1998.
- [51] Cao Y, Lu H M, Liang J. Stochastic probability landscape model for switching efficiency, robustness, and differential threshold for induction of genetic circuit in phage λ . In *Proc. the 30th Annual Int. Conf. IEEE Eng. Med. Biol. Soc.*, Vancouver, Canada, Aug. 20-24, 2008, pp.611-614.
- [52] Gardner T S, Cantor C R, Collins J J. Construction of a genetic toggle switch in *Escherichia coli*. *Nature*, 2000, 403(6767): 339-342.
- [53] Kepler T B, Elston T C. Stochasticity in transcriptional regulation: Origins, consequences, and mathematical representations. *Biophys. J.*, 2001, 81(6): 3116-3136.
- [54] Schultz D, Onuchic J N, Wolynes P G. Understanding stochastic simulations of the smallest genetic networks. *J. Chem. Phys.*, 2007, 126(24): 245102.
- [55] Kim K Y, Wang, J. Potential energy landscape and robustness of a gene regulatory network: Toggle Switch. *PLoS Comput. Biol.*, 2007, 3(3): e60.
- [56] Wang J, Xu L, Wang E. Potential landscape and flux framework of nonequilibrium networks: Robustness, dissipation, and coherence of biochemical oscillations. *Proc. Natl. Acad. Sci. U.S.A.*, 2008, 105(34): 12271-12276.
- [57] Ptashne M. Genetic Switch: Phage Lambda Revisited. New York: Cold Spring Harbor Laboratory Press, 2004.
- [58] Arkin A, Ross J, McAdams H H. Stochastic kinetic analysis of developmental pathway bifurcation in phage λ -infected *Escherichia coli* cells. *Genetics*, 1998, 149(44): 1633-1648.
- [59] Aurell E, Brown S, Johanson J, Sneppen K. Stability puzzles in phage λ . *Phys. Rev. E*, 2002, 65(5): 051914.
- [60] Munsky B, Khammash M. The finite state projection algorithm for the solution of the chemical master equation. *J. Chem. Phys.*, 2006, 124(4): 044104.
- [61] Munsky B, Khammash M. A multiple time interval finite state projection algorithm for the solution to the chemical master equation. *J. Comput. Phys.*, 2007, 226(1): 818-835.
- [62] Macnamara S, Bersani A M, Burrage K, Sidje R B. Stochastic chemical kinetics and the total quasi-steady-state assumption: Application to the stochastic simulation algorithm and chemical master equation. *J. Chem. Phys.*, 2008, 129(9): 095105.
- [63] Datta B N. Numerical Linear Algebra and Applications. Brooks/Cole Pub. Co., 1995.
- [64] Golub G H, van Loan C F. Matrix Computations. Johns Hopkins Univ. Press, 1996.
- [65] Sidje R B. Expokit: A software package for computing matrix exponentials. *ACM Trans. Math. Softw.*, 1998, 24(1): 130-156.
- [66] Lu H M, Liang J. Perturbation-based Markovian transmission model for probing allosteric dynamics of large macromolecular assembling: A study of GroEL-GroES. *PLoS Comput. Biol.*, 2009, 5(10): e1000526.
- [67] Cao Y, Gillespie D T, Petzold L R. The slow-scale stochastic simulation algorithm. *J. Chem. Phys.*, 2005, 122(1): 14116.
- [68] Cao Y, Liang J. Nonlinear coupling for improved stochastic network model: A study of Schnakenberg model. In *Proc. the 3rd Symp. Optimiz. Syst. Biol.*, Zhangjiajie, China, Sept. 20-22, 2009, pp.379-386.
- [69] Schnakenberg J. Simple chemical reaction systems with limit cycle behaviour. *J. Theoret. Biol.*, 1979, 81(3): 389-400.
- [70] Qian H. Open-system nonequilibrium steady state: Statistical thermodynamics, fluctuations, and chemical oscillations. *J. Phys. Chem. B*, 2006, 110(31): 15063-15074.
- [71] Goutsias J. Classical versus stochastic kinetics modeling of biochemical reaction systems. *Biophys. J.*, 2007, 92(7): 2350-2365.
- [72] Uribe C A, Verghese G C. Mass fluctuation kinetics: Capturing stochastic effects in systems of chemical reactions through coupled mean-variance computations. *J. Chem. Phys.*, 2007, 126(2): 024109.
- [73] Keizer J. On the macroscopic equivalence of descriptions of fluctuations for chemical reactions. *J. Math. Phys.*, 1977, 18: 1316-1321.
- [74] Mitchell M. Complexity: A Guided Tour. London: Oxford Univ. Press, 2009.
- [75] Laughlin R B, Pines D, Schmalian J, Stojković B P, Wolynes P G. The middle way. *Proc. Natl. Acad. Sci. USA*, 2000, 97(1): 32-37.
- [76] Qian H, Shi P Z, Xing J. Stochastic bifurcation, slow fluctuations, and bistability as an origin of biochemical complexity. *Physical Chemistry Chemical Physics*, 2009, 11(24): 4861-4870.



Jie Liang is a professor in the Department of Bioengineering at the University of Illinois at Chicago, and holds a visiting position at Shanghai Jiao Tong University. He received his B.S. degree from Fudan University (1986), MCS and Ph.D. degree from the University of Illinois at Urbana-Champaign (1994). He was an NSF CISE postdoctoral research associate

(1994~1996) at the Beckman Institute and National Center for Supercomputing and its Applications (NCSA) in Urbana, Illinois. He was a visiting fellow at the Institute of Mathematics and Applications at Minneapolis, Minnesota in 1996, and an Investigator at SmithKline Beecham Pharmaceuticals in Philadelphia from 1997 to 1999. He was a recipient of the NSF CAREER award in 2003. He is a fellow of American Institute of Medical and Biological Engineering, and served as regular member of the NIH Biological Data Management and Analysis study section. His research interests are in biogeometry, biophysics, computational proteomics, stochastic molecular networks, and cellular pattern formation. His recent work can be accessed at (<http://www.uic.edu/~jliang>).



Hong Qian received his B.S. degree in astrophysics from Peking University. He worked on fluorescence correlation spectroscopy (FCS) and single-particle tracking (SPT) and obtained his Ph.D. degree in biochemistry from Washington University (St. Louis). His research interests turned to theoretical biophysical chemistry and mathematical biology

when he was a postdoctoral fellow at the University of Oregon and at the California Institute of Technology. In that period of time, he worked on protein thermodynamics, fluctuations and folding. Between 1994 and 1997, he was with the Department of Biomathematics at the UCLA School of Medicine, where he worked on the theory of motor proteins and single-molecule biophysics. This work led to his current interest in mesoscopic open chemical systems. He joined the University of Washington (Seattle) in 1997 and is now professor of applied mathematics, and an adjunct professor of bioengineering. His current research is in stochastic analysis and statistical physics of cellular systems. His recent book “Chemical Biophysics: Quantitative Analysis of Cellular Systems”, co-authored with Daniel A. Beard, has been published by the Cambridge University Press.