

# 系统生物学中的随机微分方程数值仿真算法

牛原玲<sup>1,\*</sup> 陈琳<sup>2</sup> 陈洛南<sup>3</sup>

(1. 中南大学数学与统计学院, 长沙, 410083;

2. 江西财经大学统计与数据科学学院, 南昌, 330013;

3. 中国科学院系统生物学重点实验室, 分子细胞科学卓越创新中心 (上海生物化学与细胞生物学研究所), 上海, 200031)

**摘要** 系统生物学中的诸多现象, 如生物化学反应过程、生态系统的演变、传染病的传播等, 都可以用随机微分方程来描述. 由于考虑了随机因素的影响, 随机微分方程模型往往能比确定性的微分方程模型更为准确地刻画变量随时间的演化规律. 但是随机微分方程的真解大多不可得到, 有的即使可以求出真解, 但解的形式极其复杂, 用起来十分不便. 因此, 在计算机上对其进行数值仿真就显得十分必要. 系统生物学中的随机微分方程模型一般呈现出高维、高度非线性、真解位于某些特定的区域等特点, 对它们的数值模拟需要做专门的研究. 本文概述求解几类常见的系统生物学模型 (生物化学反应模型、生态系统模型、传染病模型、群体遗传学模型、细胞分化模型) 的数值算法及这些数值算法各自的优缺点.

**关键词** 系统生物学 随机微分方程 数值模拟 算法

## Numerical Simulation Algorithms for Stochastic Differential Equations in Systems Biology

Niu Yuanling<sup>1,\*</sup> Chen Lin<sup>2</sup> Chen Luonan<sup>3</sup>

(1. School of mathematics and Statistics, Central South University, Changsha 410083, China;

2. School of Statistics and Data Science, Jiangxi University of Finance and Economics, Nanchang 330013, China;

3. Key Laboratory of Systems Biology, Center for Excellence in Molecular Cell Science (Shanghai Institute of Biochemistry and Cell Biology), Chinese Academy of Sciences, Shanghai 200031, China)

**Abstract** Many phenomena in systems biology, such as the biochemical reaction process, the evolution of ecosystems, the spread of infectious diseases, can be described by stochastic differential equations (SDEs). Considering the influence of randomness, stochastic differential equation models can describe the evolution of variables over time more accurately than deterministic differential equation models. However, the analytical solutions of most stochastic differential equations cannot be obtained. Even though some of them can be obtained, the forms of the solutions are usually extremely complex. One therefore requires proper numerical methods to approximate their

solutions on computers. These stochastic differential equation models in systems biology usually have the properties of high dimension, high nonlinearity, and the solutions being located in a specified region. It is difficult to simulate them numerically. This paper reviews the numerical simulation algorithms of several typical models in systems biology (biochemical reaction models, ecosystem models, infectious disease models, population genetics models, cell differentiation models), and briefly introduces their advantages and disadvantages.

**Key words** Systems biology Stochastic differential equation Numerical simulation Algorithm

**doi:** 10.3969/j.issn.1006-8074.2023.04.005

## 1 引言

随机微分方程 (stochastic differential equation) 常常用来描述一些自然社会现象中事物的发展规律, 由于考虑了随机因素的干扰, 它往往能比确定性的微分方程更为准确地刻画变量随时间的演化规律. 随机微分方程的本质是随机积分方程, 根据所用随机积分的不同, 主要分为 Itô 型和 Stratnovich 型, 两种类型的随机微分方程是可以相互转换的. 本文我们只关注如下的 Itô 型随机微分方程:

$$dX(t) = f(t, X(t))dt + g(t, X(t))dW(t), \quad (1.1)$$

其中  $t \in [0, \infty)$ ,  $X(0) = X_0$ ,  $W(t) = (W_1(t), W_2(t), \dots, W_d(t))^T$  是  $d$  维维纳过程, 又称为  $d$  维布朗运动.  $f: \mathbb{R}_+ \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  和  $g: \mathbb{R}_+ \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times d}$  分别称为漂移项系数和扩散项系数, 它们均为  $[0, \infty)$  上的 Borel 可测函数.

如果随机扰动出现在微分方程的系数中, 如

$$\frac{dx}{dt} = g(x, \eta_t), \quad (1.2)$$

其中  $\eta_t$  为一个随机过程或者随机变量, 这样的方程通常叫做带有随机输入的微分方程 (random ordinary differential equation). 本文不涉及此类微分方程的数值仿真.

作为一个重要的建模工具, 随机微分方程已经在生物、经济、控制等领域得到广泛的应用. 但是对大部分随机微分方程而言, 一般不能得到其真解. 有些随机微分方程的真解虽然可以得到, 但由于形式过于复杂, 实际应用中不太方便处理. 因此, 研究随机微分方程的数值算法就显得尤为重要. 确定性微分方程的数值算法是一个已经得到长足发展的领域<sup>[1-4]</sup>. 近几十年来, 国内外对随机微分方程数值计算的研究也取得了快速的发展, 参见 [5-20]. 常用的求解随机微分方程的数值方法有 Euler-Maruyama 方法、Milstein 方法和 Runge-Kutta 方法等. 收敛性是数值方法的一个重要评价指标. 要求数值方法在一定意义下是收敛的, 并且希望其具有较高的收敛阶以提高仿真效率. 与确定性的情形不同, 随机微分方程数值方法的收敛性有多种定义, 其中人们最为关注的分别是强收敛性和弱收敛性. 强收敛性要求数值近似的轨道充分接近真实轨道, 而弱收敛性关注的则是解过程在矩意义下的收敛性. 本文主要涉及前者. 此外, 我们还希望数值方法能够保持原系统的一些重要特征, 比如保区域、保结构等.

本文主要回顾系统生物学中的随机微分方程模型及相关的数值仿真算法. 这些随机微分方程一般具有高维、高度非线性、真解位于某个特定的区域等特点, 因此对它们的数值模拟需要做专门的研究. 鉴于篇幅有限, 我们仅从生物化学 (生化) 反应系统、生态系统、传染病模型、群体遗传学模型以及细胞分化模型的数值仿真来加以阐述.

需要说明的是,目前成熟的商业软件和开源软件中也存在一些可以用于数值求解随机微分方程的软件包:如 MATLAB 中的“SDETools”和“SDELab”及 R 中的“sde”和“yuima”.其中“SDETools”主要采用的数值格式为 Euler-Maruyama 格式和 Milstein 格式(要求随机微分方程具有对角噪声);“SDELab”则主要采用  $\theta$ -Maruyama 和 Milstein 格式(不要求随机微分方程具有对角噪声);“sde”仅能数值求解一维的随机微分方程,主要采用的数值格式为 Euler-Maruyama 格式, Milstein 格式和 Predictor-Corrector 格式;“yuima”可以数值求解多维的随机微分方程(可以是分数阶噪声驱动,也可以存在 Levy 跳),采用的数值格式为 Euler-Maruyama 格式.这些软件包中采用的数值格式均较为简单,应用于本文提到的这些复杂模型时效果往往不太理想.

## 2 生化反应系统的数值仿真

细胞中发生的大多数生化反应本质上是随机的.由于传统实验成本昂贵,生化反应系统的随机模拟近年来受到越来越多的关注.假设有一个搅拌良好的生化反应系统,包含  $N$  种分子  $S_1, S_2, \dots, S_N$ ,这些分子通过多个反应通道在定容恒温的环境下发生反应.用  $x_i(t)$  ( $i = 1, \dots, N$ ) 表示  $t$  时刻  $S_i$  的数量,则该系统的状态可用向量  $X = (x_1(t), x_2(t), \dots, x_N(t))^T$  来表示.用矩阵  $v$  表示化学计量矩阵,其元素  $v_{ij}$  表示第  $j$  个反应  $R_j$  导致分子  $S_i$  数量的变化.

### 2.1 化学主方程模型

化学主方程 (Chemical master equation, 简记为 CME) 描述了生化反应系统在不同时间点处于每种状态的概率:

$$\frac{\partial}{\partial t} P(X, t) = -\sum_{j=1}^M a_j(X) P(X, t) + \sum_{j=1}^M a_j(X - v_j) P(X - v_j, t), \quad (2.1)$$

其中  $X(t)$  是状态变量,  $a_j(X(t))$  是倾向函数,  $a_j(X(t))dt$  表示在下一个小时时间段  $[t, t + dt)$  第  $j$  个反应  $R_j$  发生的概率.

化学主方程似乎是发现随机系统时间演化规律的理想方法,它可以返回系统的概率密度函数.然而,化学主方程只有在非常简单的情况下可以求出真解.也就是说,在大多数情况下,需要借助数值方法来对其进行数值模拟.此外,由于必须在每个时间点对每个可能的状态进行评估,并且状态的数量可以呈指数增长,因此计算量非常大,这使得化学主方程模型最初只能有效地用于短期模拟小规模或者简单系统.近年来,虽然化学主方程的数值求解方法取得了长足的进步<sup>[21, 22, 23]</sup>,但当分子数和倾向函数太大以致于无法使用化学主方程计算时,必须使用新的数值模拟方法,如 Gillespie 型算法.

### 2.2 离散状态马尔可夫链模型

随机模拟算法 (stochastic simulation algorithm, 简记为 SSA) 由 Gillespie 于 1977 年提出<sup>[24]</sup>,该方法可以精确模拟化学主方程给出的完整分布中的各条路径,是一种离散状态马尔可夫链方法.由无限多个 SSA 模拟建立的概率密度函数将与化学主方程给出的系统真实分布相同.无限次模拟显然是无法达到的,不过适度多的重复次数可以获得令人满意的近似概率密度函数.尽管近年来 Gillespie 型方法取得了不小的进展<sup>[25-30]</sup>,但这些算法针对复杂的生化反应系统的计算成本仍然会非常高,尤其是在反应速率和分子数范围很广的情况下.

## 2.3 郎之万方程模型

在描述生化反应系统时, 化学主方程在零分子数时有一个自然边界条件. Fokker-Planck 方程是对化学主方程的近似<sup>[31]</sup>, 近似的过程中可能丢失了一些信息, 此时, 人们需要给 Fokker-Planck 方程加上合适的边界条件, 如 Neumann 边界条件<sup>[32]</sup>, 反射边界条件<sup>[33]</sup>, 以更好地描述生化反应过程. 利用 Feynman-Kac 公式, 某些偏微分方程的解可以用对应的随机微分方程的解来表示. 与 Fokker-Planck 方程数学上等价的随机微分方程叫做郎之万方程. 该模型是离散状态马尔可夫链方法的一种有吸引力的替代方法, 因为它可以显著提高计算速度. 然而, 随机微分方程模型的解可能会变成负数, 甚至是虚数. 显然, 负数甚至是虚数作为分子的数目是不合理的. 此时, 人们在做数值模拟时通常的做法是, 当近似值变为负值时, 将其设置为 0, 或者对维纳增量重新采样, 直到数值解为非负值. 但这些措施会给数值解带来偏差. 文献 [34] 探索了一种修改化学郎之万方程噪声项的方法, 以使变量保持非负. 他们只允许物种参与的反应步骤具有随机性. 但是, 文献 [34] 在去掉乘积噪声项后得到的改进的保持非负性的郎之万方程不能保证协方差矩阵的匹配条件. 文献 [35] 研究了漂移项和扩散项均为隐式的可保非负性的 Milstein 方法, 用于数值求解化学郎之万方程. 该格式在牛顿迭代中施加了取值非负的约束, 计算过程中需要删除因反应物耗尽而停止的反应通道  $R_j$ , 即删除化学计量矩阵  $v$  的第  $j$  列  $v_j$  和相应的倾向函数  $a_j$ , 直到反应物浓度再次升高, 再将删除的反应通道  $R_j$  恢复, 同时修改化学计量矩阵和倾向函数. 这些措施使得算法的“保非负性”得以实现. 该方案始终满足协方差矩阵匹配条件. 但是因为全隐式方法需要额外的计算, 并且计算过程中需要实时修改化学计量矩阵和倾向函数, 所以计算成本可能非常高. 文献 [36] 表明, 通过将郎之万方程的定义域扩展到复空间, 得到复数域上的郎之万方程, 该方程得到的反应物的平均浓度落在实数域. 但是此时分子数是复数, 在生物学上不合理.

当使用化学郎之万方程通过连续过程近似化学主方程描述的离散过程时, 目前尚不太清楚边界条件的行为. 文献 [37] 中, 作者通过将边界条件纳入随机微分方程, 得到了确保随机微分方程模型的生物学上合理的解决方案. 由此产生的模型称为带反射边界的随机微分方程模型. 此类方程之前曾用于模拟人类代谢过程<sup>[38]</sup>, 离子通道动力学<sup>[39]</sup> 等.

带边界条件的生化反应模型如下:

$$dX(t) = va(X(t))dt + G(X(t))dW(t) + d\xi(t), \quad (2.2)$$

$$|\xi|(t) = \int_0^t \mathbb{1}_{\{X(s) \in \partial D\}} d|\xi|(s), \quad (2.3)$$

$$\xi(t) = \int_0^t n(s) d|\xi|(s), \quad (2.4)$$

其中  $W(t)$  为维纳过程.  $\xi(t)$  有局部的有界变差,  $\xi(0) = 0$ .  $|\xi|(t)$  非负递增连续并且满足(2.3)-(2.4)式. 如果  $X(s) \in \partial D$ , 那么,  $n(s) \in N(X(s))$ , 其中  $N(y)$  表示当  $y$  在边界上时, 指向区域内部的单位向量. (2.3)式表明当且仅当  $X(t) \in \partial D$  时,  $|\xi|(t)$  会增长. 当  $X(t)$  不落在边界上时,  $d\xi(t)$  消失, 此时带反射边界的随机微分方程(2.2)退化为普通的随机微分方程

$$dX(t) = va(X(t))dt + G(X(t))dW(t). \quad (2.5)$$

由(2.2)-(2.4)建立的带反射边界的随机微分方程模型似乎是建立生化反应系统模型的理想方法, 它的解具有生物意义. 但是在大多数情况下, 此类随机微分方程并没有解析解, 因此数值近似

是获得解并研究其性质的重要途径. 这里的定义域  $D$  是凸的, 因此大多数现有的数值算法可以应用于此处提出的带反射边界的随机微分方程模型的数值模拟. 投影法是 Euler-Maruyama 方法的简单扩展, 可用于求解带反射边界的随机微分方程: 如果 Euler-Maruyama 方法获得的数值解位于  $D$  的闭包之外, 则将下一时间步的解设置为该点在闭包  $D$  上的投影, 否则该步的数值逼近等于 Euler-Maruyama 方法得到的数值解. 数值实验表明, 利用(2.2)-(2.4)确定的带反射边界的随机微分方程模型得到的数值模拟结果与 SSA 得到的数值模拟结果非常接近, 但是时间成本则大大降低.

Goldbeter-Koshland 开关是由两种具有相反作用的酶  $E_1$  和  $E_2$  促进的共价修饰生化反应系统, 它由以下 6 个反应组成:



取初始状态  $(110, 100, 30, 30, 100, 30)^T$  及参数  $k_1 = 0.05, k_2 = 0.1, k_3 = 0.1, k_4 = 0.01, k_5 = 0.1, k_6 = 0.1$ , 文献 [37] 对上述生化反应系统进行了数值模拟, 结果如图1所示. 由图1可以看到, 带反射边界的随机微分方程模型与 SSA 得到的数值模拟结果非常接近. 而 SSA 所花的 CPU 时间为 51h55min40s, 带反射边界的随机微分方程的数值模拟所花的 CPU 时间则为 3h37min47s, 可见后者的数值模拟效率大为提高.

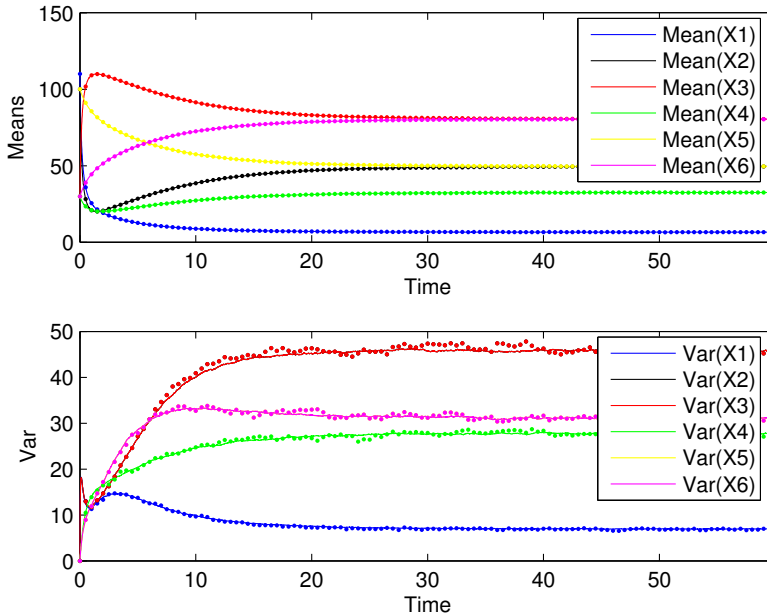


图 1 SSA 与带反射边界的随机微分方程对 Goldbeter-Koshland 开关数值模拟的均值和方差 (实线: SSA, 虚线: 带反射边界的随机微分方程)<sup>[37]</sup>

这一工作的主要贡献在于提供了一种计算高效的方法来模拟生化反应系统. 该方法不仅可以保持离散状态马尔可夫模型的随机行为, 并且可以保证真解及数值解都有生物意义, 从而使得针对大规模生化反应系统的数值仿真成为可能. 值得注意的是, 即使带反射边界的随机微分方程模型在许多情况下表现得足够好, 但它并不太适合所有反应通道的倾向函数都非常小的生化反应系统. 这种情况下需要用到 2.2 节提到的 Gillespie 型算法.

### 3 生态系统的数值仿真

生态数学中, 通过“变化率 = 输入 - 输出”这样的思路可以建立起常微分方程模型, 如在 Logistic 模型基础上得到的 Lotka-Volterra 模型. 近年来, 一些学者对 Lotka-Volterra 模型引入了随机扰动, 将其转化为随机 Lotka-Volterra 模型, 其中随机性通常是由环境的不可预测性和对生态系统的不完全认知引起的. 随机 Lotka-Volterra 模型能够更精确地描述客观现象. 随机 Lotka-Volterra 竞争模型, 随机 Lotka-Volterra 捕食模型和随机 Lotka-Volterra 合作模型是其中最为典型的代表.

#### 3.1 随机 Lotka-Volterra 竞争模型

Mao 在其专著 [40] 中研究了如下的随机 Lotka-Volterra 竞争模型:

$$dx(t) = \text{diag}(x_1(t), x_2(t), \dots, x_d(t))[(b - Ax(t))dt + \sigma dW(t)], \quad (3.1)$$

其中  $x(t) = (x_1(t), x_2(t), \dots, x_d(t))^T$  是生态系统中  $d$  个相互作用的种群在  $t$  时刻的状态,  $b = (b_1, \dots, b_d)^T \in \mathbb{R}^d$ ,  $\sigma = (\sigma_1, \dots, \sigma_d)^T \in \mathbb{R}^d$ ,  $A = (a_{ij})_{d \times d} \in \mathbb{R}^{d \times d}$  为系统的参数. 文献 [40] 指出, 当矩阵  $A$  的元素均为非负, 即  $a_{ij} \geq 0$  ( $1 \leq i, j \leq d$ ) 时, 对任意的初始值  $x(0) \in \mathbb{R}_+^d$ , 方程(3.1)有唯一的强解  $x(t)$ , 并且满足  $x(t) \in \mathbb{R}_+^d$ , 其中  $\mathbb{R}_+^d = \{(x_1, \dots, x_d)^T \in \mathbb{R}^d : x_i \geq 0, 1 \leq i \leq d\}$ . 作者还研究了带延迟的随机 Lotka-Volterra 竞争模型真解的性质.

诸如方程(3.1)这样的随机微分方程, 具有以下特点: 高度非线性, 正解, 多维. 这些特点给数值模拟带来了很大的挑战. 文献 [41] 针对方程(3.1), 在截断算法的基础上, 引入非负投影来确保算法的非负性, 给出了一种保正截断 Euler-Maruyama 方法. 作者证明了该数值算法是保正的和强收敛的. 数值算例显示, 该方法具有 1/2 阶的强收敛阶, 但作者并没有给出相关的理论证明.

文献 [42] 则在矩阵  $A$  的所有元素非负且对角线元素为正的条件下, 针对方程(3.1)提出了一种线性隐式 Milstein 方法. 该方法不仅能保正, 并且能保证解的长时间有界性. 数值实验表明, 该方法具有 1 阶的强收敛阶, 但作者理论证明方法的强收敛阶仅为 1/2.

文献 [43] 借助对数变换将方程(3.1)转换成带有加性噪声的随机微分方程, 并对转换后的方程应用显式 Euler-Maruyama 算法进行数值求解. 作者通过对数变换使得数值算法可以保持非负性, 并利用随机摄动理论证明了算法的 1/2 阶强收敛性.

#### 3.2 随机 Lotka-Volterra 捕食模型

除了 3.1 节中介绍的竞争模型, 捕食模型也是生态学中一种重要的模型. 文献 [44] 研究了这类模型解的存在唯一性. 文献 [45, 46] 分析了随机 Lotka-Volterra 捕食模型的长时间行为. 此外, 文献 [47] 研究了带时滞的随机 Lotka-Volterra 捕食模型的正解的存在性.

文献 [48] 考虑了如下的一类随机 Lotka-Volterra 捕食模型

$$\begin{cases} dX(t) = \bar{X}(t)[(-\Gamma^{(2)}Y(t) + \eta^{(2)})dt + \Sigma^{(2)}dW(t), \\ dY(t) = \bar{Y}(t)[(\Gamma^{(1)}X(t) - \eta^{(1)})dt + \Sigma^{(1)}dW(t), \end{cases} \quad (3.2)$$

其中  $X(t) = (x_1(t), \dots, x_d(t))^T$  和  $Y(t) = (y_1(t), \dots, y_d(t))^T$  分别是  $t$  时刻被捕食者和捕食者的数目,  $\eta^{(1)} = (\eta_1^{(1)}, \dots, \eta_d^{(1)})^T$  表示在没有食物的情况下捕食者的自然死亡率, 而  $\eta^{(2)} = (\eta_1^{(2)}, \dots, \eta_d^{(2)})^T$  表示没有捕食者时被捕食者的自然增长率.

作者证明了模型(3.2)具有唯一的正解, 并且具有随机辛结构, 故可以被看作是一个随机哈密顿系统. 同时, 作者提出了一种随机 Runge-Kutta 方法, 并且证明了该数值格式可以保持数值解为正和原系统的随机辛结构. 事实上, 作者是通过了对数变换来实现保正的. 此外, 作者还证明了该数值格式是 1 阶收敛的.

### 3.3 随机 Lotka-Volterra 合作模型

Lotka-Volterra 合作模型描述了自然界以及人类社会中物种或群体之间的互惠互利关系, 如蜜蜂和花. 文献 [49] 将白噪声引入到一类时滞 Lotka-Volterra 模型的内禀增长率中, 建立了持续性的充分条件, 并获得了平稳分布的存在唯一性. 文献 [50] 研究了食物有限情况下的随机 Lotka-Volterra 合作模型, 并给出了持续和灭绝的充分条件. 文献 [51] 考虑了如下带有年龄结构的随机 Lotka-Volterra 合作模型:

$$\begin{cases} d_t X(t, a) = -\frac{\partial X(t, a)}{\partial a} dt + X(t, a)(-\alpha_{11}(a)X(t, a) + \alpha_{12}(a)Y(t, a) + \gamma_1(a))dt \\ \quad + \sigma_1 X(t, a)dW(t), & (t, a) \in Q, \\ d_t Y(t, a) = -\frac{\partial Y(t, a)}{\partial a} dt + Y(t, a)(\alpha_{21}(a)X(t, a) - \alpha_{22}(a)Y(t, a) + \gamma_2(a))dt \\ \quad + \sigma_2 Y(t, a)dW(t), & (t, a) \in Q, \\ X(t, 0) = \int_0^A \gamma(t, a)X(t, a)da, & t \in [0, T], \\ Y(t, 0) = \int_0^A \beta(t, a)Y(t, a)da, & t \in [0, T], \\ X(0, a) = X_0(a), & a \in [0, A], \\ Y(0, a) = Y_0(a), & a \in [0, A], \end{cases} \quad (3.3)$$

其中  $X(t, a)$  和  $Y(t, a)$  分别是两个物种在  $t$  时刻、 $a$  年龄时的密度,  $\gamma_1(a)$  和  $\gamma_2(a)$  分别是物种的内禀增长率,  $\alpha_{11}(a)$  和  $\alpha_{22}(a)$  表示种内竞争率,  $\alpha_{12}(a)$  和  $\alpha_{21}(a)$  表示种间合作率,  $\gamma_1(a)$ ,  $\gamma_2(a)$ ,  $\alpha_{11}(a)$ ,  $\alpha_{22}(a)$ ,  $\alpha_{12}(a)$  和  $\alpha_{21}(a)$  均为正数.

与前面提到的其他模型不同, 带年龄结构的随机 Lotka-Volterra 合作模型(3.3)是一个随机偏微分方程. 事实上, 考虑到年龄结构的种群模型都是偏微分方程. 文献 [51] 证明了方程(3.3)全局解的存在唯一性, 并且分析了 Euler-Maruyama 方法用于方程(3.3)的数值求解时的收敛率. 但 [51] 并没有证明数值解的保正性.

值得注意的是, 这里谈到的 Lotka-Volterra 模型不仅可用于描述种群动力学, 还可用于物理学和经济学等方面的建模.

## 4 随机传染病模型的数值仿真

近年来,世界经济和人类健康均受到新型冠状病毒的影响,该病毒已经严重影响了人们正常的生产生活.历史上也出现过许多类似的传染病,例如黑死病,“非典”等.数理流行病模型已经成为了解疫情,掌握疫情变化的重要工具.数学家们建立了很多著名的传染病数学模型,例如 SIS, SIQS, SIR, SEIR 模型等.

事实上,在现实生活中,任何一个系统都会受到或多或少的随机干扰.正因如此,对具有随机扰动的传染病模型的研究是非常重要且是有意义的.据此我们可以得到随机版本的 SIS, SIQS, SIR, SEIR 模型等.这些模型都是具有非线性系数的随机微分方程,它们的解析解通常很难求出.因此需要使用计算机来对其进行数值模拟,从而达到对传染病模型的研究.系数不满足单边 Lipschitz 条件,数值解需要保持真解位于特定区域的性质,这是对随机传染病模型进行数值模拟的两大难点.

### 4.1 随机 SIR 模型

1927 年数学家 Kermack 与 McKendrick 在研究传染病时提出了 SIR 仓室模型<sup>[52]</sup>.此模型是最为经典的传染病数学模型,当时该模型的提出主要针对欧洲出现的黑死病.模型将人群分为 3 类,分别为易感者 (S 类)、感染者 (I 类) 和康复者 (R 类).该模型适用于治愈后康复者具有很强的免疫力,不会再次被感染的传染病,如水痘等.此外,该模型还适用于一些致死性的传染病,因为死亡的病人也可以归入“康复者 (R)”.这里的“康复”可以理解为是退出了系统.

文献 [53] 将所提出的数值算法成功运用于如下的随机 SIR 模型:

$$\begin{cases} dS(t) = [\Lambda - \beta S(t)I(t) - \mu S(t)]dt - \sigma S(t)I(t)dW(t), \\ dI(t) = [\beta S(t)I(t) - (\mu + \gamma)I(t)]dt + \sigma S(t)I(t)dW(t), \\ dR(t) = [\gamma I(t) - \mu R(t)]dt, \end{cases} \quad (4.1)$$

其中参数  $\Lambda, \beta, \mu$  和  $\gamma$  为正常数,  $\Lambda$  表示易感个体自然出生率,  $\mu$  表示每个个体类别自然死亡率,  $\beta$  为疾病接触率,  $\gamma$  表示受感染个体的恢复率.在初值假设  $S(0) + I(0) + R(0) = \Lambda/\mu$  下,模型 (4.1) 满足群体总人口不变,即在任意时刻  $t$  都有  $S(t) + I(t) + R(t) = \Lambda/\mu$ .文献 [53] 提出了一种增量驯服的 Euler 算法,并得到了模型(4.1)数值近似的强收敛性.

文献 [54] 针对模型(4.1)构造了一种维纳截断的线性隐式算法,在确保数值近似取值非负的基础上,得到了算法是强 1/2 阶收敛的,并讨论了算法的数值动力学行为.

文献 [55] 则在另一类随机扰动下研究了如下随机 SIR 模型:

$$\begin{cases} dS(t) = [\Lambda - \beta S(t)I(t) - \mu S(t) + \varepsilon I(t)]dt + \sigma S(t)dW(t), \\ dI(t) = [\beta S(t)I(t) - (\mu + \varepsilon + \gamma)I(t)]dt + \sigma I(t)dW(t), \\ dR(t) = [\gamma I(t) - \mu R(t)]dt + \sigma R(t)dW(t), \end{cases} \quad (4.2)$$

其中受感染个体以  $\varepsilon$  的速率变得易感并转移到易感类别,其他参数与模型(4.1)一致.但由于模型中康复者也受到随机因素影响,模型(4.2)的人口总数带有随机性.针对模型(4.2),作者提出了一种基于勒让德多项式的勒让德谱配置方法来求解.



## 4.2 随机 SIS 模型

在随机 SIR 模型的数值分析中, 如何得到算法的收敛阶仍然有一些困难. 但对相对简单的随机 SIS 传染病模型, 在其数值近似方面有着丰富的研究成果.

在文献 [56] 中, 作者考虑了如下的随机 SIS 模型:

$$\begin{cases} dS(t) = [\mu N - \beta S(t)I(t) + \gamma I(t) - \mu S(t)]dt - \sigma S(t)I(t)dW(t), \\ dI(t) = [\beta S(t)I(t) - (\mu + \gamma)I(t)]dt + \sigma S(t)I(t)dW(t), \end{cases} \quad (4.3)$$

其中  $S(t)$ ,  $I(t)$  表示  $t$  时刻的易感者和感染者数目,  $N$  是疾病在其中传播的人口总数,  $\mu$  是人均死亡率,  $1/\gamma$  为平均感染期,  $\beta$  是疾病传播系数. SIS 模型可以看作 SIR 模型的特殊情形, 适用于感染后不产生永久免疫力的疾病建模, 如一些性传播疾病和细菌性疾病. 对于这些疾病, 个体起初易感, 在某个阶段感染该疾病, 并在短的感染期后再次易感. 文献 [56] 证明了模型(4.3)具有唯一的全局正解  $I(t)$ , 并建立了  $I(t)$  灭绝和持续的条件.

文献 [57] 提出了一种显式的数值格式, 称为 Lamperti 平滑截断格式 (LST), 用于对随机 SIS 模型的数值模拟. 该方案将 Lamperti 型变换 (实质上为对数变换) 与显式截断方法相结合, 所得的数值近似结果保持了原始随机微分方程的解位于特定区域的性质, 并且具有 1 阶均方收敛阶. 这里使用的 Lamperti 型变换只能用于一维的随机微分方程, 变换后的随机微分方程的系数满足单边 Lipschitz 条件. 由于(4.3)式中的变量满足  $S(t) + I(t) = N$ , 所以 Lamperti 型变换在这里是可以使用的. 此时(4.3)式变为:

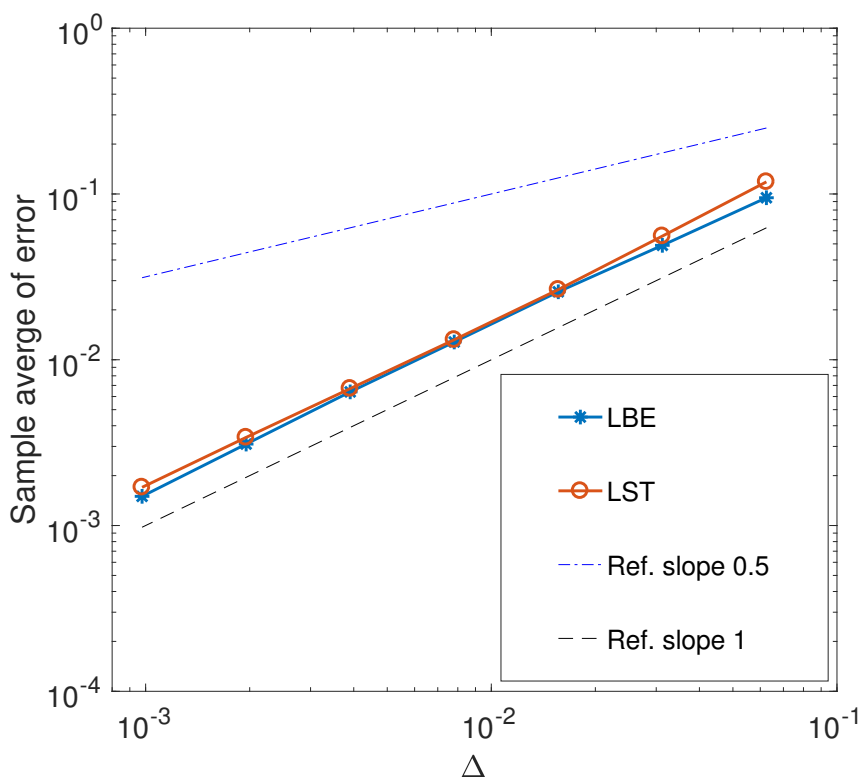
$$dI(t) = I(t)(\beta N - \mu - \gamma - \beta I(t))dt + \sigma I(t)(N - I(t))dW(t). \quad (4.4)$$

当  $\beta N - \mu - \gamma = 5$ ,  $\beta = 0.8$ ,  $\sigma = 0.1$ ,  $N = 10$  时, 作者给出了用 Lamperti 平滑截断格式求解(4.4)式的收敛阶, 如图2所示. 其中 LST 即为这里的 Lamperti 平滑截断格式. 由图 2 可知, 该方法具有 1 阶收敛阶. 虽然 LBE, 即 Lamperti 向后 Euler 格式<sup>[63]</sup>, 也具有 1 阶收敛阶, 但由于 LBE 是隐式的, 需要求解超越方程, 所以 LST 算法的运行时间明显低于 LBE 算法. 显然, LST 算法比 LBE 算法更有效.

在文献 [58] 中, 作者通过结合对数变换和 Euler-Maruyama 方法, 为随机 SIS 模型构造了一种保正的数值方法, 证明了该算法不仅保留了原始随机微分方程的值域, 而且在所有  $p > 0$  的情况下, 在有限时间间隔内具有 1 阶  $L_p$  收敛阶.

文献 [59] 考虑了随机 SIS 模型的带截断维纳过程的分裂步  $\theta$  方法, 证明了该数值格式具有 1/2 阶均方收敛阶, 数值解保正且有界. 虽然收敛阶比 [57] 提出的 LST 算法要低, 但是不需要做 Lamperti 型变换, 可以直接近似求解原方程. 事实上, 该方法通过将维纳过程截断来保正, 因此无需做变换.

此外, 文献 [60] 构造了一种带有两个随机扰动项的 SIS 模型, 证明了该模型具有唯一且有界的全局解, 分析了模型中感染人群  $I(t)$  持续、灭绝的条件. 之后文献 [61] 通过在平方根内取绝对值的方式, 针对该模型提出了一种保取值范围的数值算法, 并得到了该算法的收敛性, 但没有给出算法的收敛阶.

图2 LST 算法求解随机 SIS 模型的均方误差<sup>[57]</sup>

## 5 群体遗传学模型的数值仿真

群体遗传学中一个非常重要的模型是 Wright-Fisher 模型, 于 20 世纪 30 年代由 Wright 和 Fisher 独立引入, 它描述了种群遗传结构中的随机进化。

文献 [62] 考虑了如下的 Wright-Fisher 模型:

$$dY = (A - (A + B)Y)dt + C\sqrt{Y(1 - Y)}dW(t), \quad (5.1)$$

其中参数的具体意义见 [62]. Wright-Fisher 模型最初用于模拟有限群体内的遗传漂移, 在文献 [62] 中被用作心脏和神经细胞内离子通道动力学的模型. 该方程的解落在区间  $[0, 1]$  内, 但多数数值方法无法在数值模拟时保证此类取值范围. [62] 结合隐式平衡算法和分裂步算法, 提出了一种平衡隐式分裂步算法, 可使得数值解始终以概率 1 保持在  $[0, 1]$  内, 并证明了该方法的强收敛性. 数值实验表明, 该格式具有强  $1/2$  阶收敛性. 将该方法推广到多维情况, 数值试验表明该算法仍然具有  $1/2$  阶的强收敛性. 取  $A = 1, B = 2, C = 0.2462$ , 将平衡分裂步算法应用于 Wright-Fisher 模型(5.1)的数值实验结果如图3所示.

在文献 [63] 中, 针对这类具有非 Lipschitz 漂移或扩散系数且取值在区域  $D = (l, r)$  内的标量随机微分方程, 作者提出了一种 Lamperti-Euler 格式 (LBE). 首先借助 Lamperti 变换  $x(t) = F(y(t))$ , 使用向后 Euler 格式近似变换过程  $x(t)$ , 再用 Lamperti 逆变换进行反向变换, 得到原始随机微分

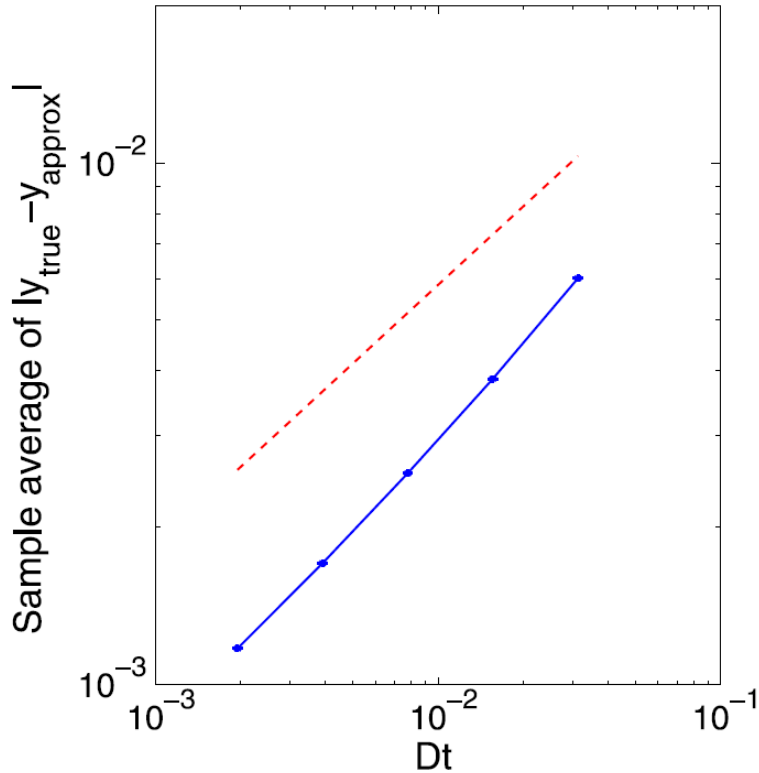


图3 平衡隐式分裂步算法求解 Wright-Fisher 模型(5.1)的收敛性. 红色虚线是斜率为 1/2 的参考线<sup>[62]</sup>

程的近似值. 该数值格式具有 1 阶收敛性, 并且可以保持原始随机微分方程的取值在  $D$  内.

这里的 LBE 格式是较早使用 Lamperti 变换实现数值解保正的格式. 随机 SIS 模型对应的 Lamperti 变换是对数变换, 而 Wright-Fisher 模型对应的则是三角变换. 需要注意的是, 变换后的隐式数值格式在每一步都需要求解超越方程, 计算成本较高. 为了克服隐式算法计算成本较大的难点, 最近文献 [64] 在 Lamperti 变换的基础上提出一种新的显式倾斜截断算法, 不仅得到了算法的强 1 阶收敛性, 还得到了数值近似的平稳分布收敛性.

## 6 细胞分化模型的数值仿真

细胞分化过程形成了形态、结构、功能各异的细胞类型, 使得生物世界丰富多彩. 对基因调控网络的建模对理解细胞的分化至关重要. 构建生物系统的表观遗传景观是研究细胞分化的重要工具.

目前主要有以下两种方法可以用来构造势能景观: 一种是基于数据的方法, 即从实验数据, 特别是单细胞 RNA 测序 (scRNA-seq) 数据中识别细胞类型、分化轨迹和细胞多能性, 如文献 [65, 66, 67] 等; 另一种方法是基于模型的方法, 即构建基因调控网络的动力学方程, 分析系统的动力学行为, 并使用数值模拟构建势能景观, 如文献 [68, 69] 等.

在文献 [68] 中, 作者用如下带生灭项的 Fokker-Planck 方程作为种群水平上的细胞分化模型:

$$\frac{\partial p(x, t)}{\partial t} = -\nabla \cdot (b(x)p(x, t)) + \varepsilon \Delta p(x, t) + R(x)p(x, t), \quad (6.1)$$

其中  $x$  表示基因表达水平,  $b(x)$  表示基因间的相互作用,  $\varepsilon$  代表噪声的强度,  $R(x)$  表示  $x$  处的细胞的出生与死亡率, 即生灭项. 同时, 作者用如下的随机微分方程作为单细胞水平的细胞分化模型:

$$\begin{cases} dX_t(\omega) = b(X_t(\omega))dt + \sqrt{2\varepsilon}dW_t(\omega), \\ X_t|_{t=0} = Y_0(\omega), \\ d\rho_t(\omega) = R(X_t(\omega))\rho_t(\omega)dt, \\ \rho_t|_{t=0} = 1, \end{cases} \quad (6.2)$$

其中  $X_t(\omega)$  为细胞  $\omega$  在  $t$  时刻的基因表达水平,  $\rho_t(\omega)$  代表细胞  $\omega$  的时变权重. 由 Itô 公式可知, (6.1), (6.2) 两个模型的等价关系可以由下式得到

$$p(x, t) = \mathbb{E}\{\rho_t(\omega)\delta(x - X_t(\omega))\}, \quad (6.3)$$

这里,  $\delta$  是狄拉克函数,  $\mathbb{E}$  表示对所有可能的轨道  $\omega$  取期望.

在构建势能景观的过程中, 针对低维模型, 作者使用有限元方法或者有限差分方法直接对种群水平上的细胞分化模型(6.1)进行数值求解直至稳态; 针对高维模型, 则对单细胞水平的细胞分化模型(6.2)进行平均场近似, 然后再根据所获得的稳态概率密度函数及作者提出的势能景观分解理论, 得到细胞类型势能景观  $U(x)$  和体现细胞干性的多潜能性势能景观  $V(x)$ :  $U(x) = -\varepsilon \log P_U(x)$ ,  $V(x) = -\varepsilon \log (P_0(x)/P_U(x))$ , 其中  $P_U(x)$  表示当  $R(x)$  给定时, 由(6.1)或(6.2)得到的稳态概率密度函数,  $P_0(x)$  表示当  $R(x) \equiv 0$  时, 由(6.1)或(6.2)得到的稳态概率密度函数.

## 7 结论

随着近年来交叉学科的蓬勃发展, 涌现了大量与系统生物学相关的随机微分方程模型. 这些模型由于来源于现实世界, 有一定的生物背景, 通常具有以下一个或者若干个特点: 解位于特定区域, 高维, 系数不满足 Lipschitz 条件等. 以上这些特点为数值算法的构造及理论分析带来了挑战. 常用的数值算法用来做这类模型的数值仿真时要么不收敛, 要么不能保持真解的某些性质, 如位于特定的区域等. Lamperti 型变换算法, 隐式高阶算法或者平衡隐式算法可能是针对这类问题构造数值格式的备选方案. 目前已有的保取值域的方法主要包括: 投影 (文献中也称截断), 相关文献有 [41] 等; 对数变换, 相关文献有 [47, 57, 58] 等; 平衡隐式方法, 相关文献有 [62] 等; 维纳增量截断, 相关文献有 [59] 等.

对于引言中提到的带有随机输入的微分方程 (random ordinary differential equation) 的数值模拟, 可参考文献 [70]. 系统生物学家通常使用这类方程对生态系统进行建模, 如文献 [71, 72] 等.

鉴于篇幅和时间有限, 本文所综述的工作难免挂一漏万, 敬请读者海涵.

## 参考文献

- [1] Hairer E, Wanner G. Solving Ordinary Differential Equations II[M]. Berlin: Springer-Verlag, 2010.

- 
- [2] 李寿佛. 刚性常微分方程及刚性泛函微分方程数值分析 [M]. 湘潭: 湘潭大学出版社, 2010.
- [3] Kuang J, Cong Y. Stability of Numerical Methods for Delay Differential Equations[M]. Beijing: Science Press, 2005.
- [4] Wu X, You X, Wang B. Structure-Preserving Algorithms for Oscillatory Differential Equations[M]. Heidelberg: Springer, 2013.
- [5] Kloeden P E, Platen E. Numerical Solution of Stochastic Differential Equations[M]. Heidelberg: Springer, 1995.
- [6] Milstein G N. Numerical Integration of Stochastic Differential Equations (Vol. 313)[M]. New York: Springer, 1995.
- [7] Wang X, Gan S. The tamed Milstein method for commutative stochastic differential equations with non-globally Lipschitz continuous coefficients[J]. Journal of Difference Equations and Applications, 2013, 19(3): 466–490.
- [8] Fan Z, Liu M, Cao W. Existence and uniqueness of the solutions and convergence of semi-implicit Euler methods for stochastic pantograph equations[J]. Journal of Mathematical Analysis and Applications, 2007, 325(2): 1142–1159.
- [9] Jiang S, Wang L, Hong J. Stochastic multi-symplectic integrator for stochastic nonlinear Schrödinger equation[J]. Communications in Computational Physics, 2013, 14(2): 393–411.
- [10] Guo Q, Liu W, Mao X. A note on the partially truncated Euler-Maruyama method[J]. Applied Numerical Mathematics, 2018, 130: 157–170.
- [11] Wang P, Liu Z. Split-step backward balanced Milstein methods for stiff stochastic systems[J]. Applied Numerical Mathematics, 2009, 59(6): 1198–1213.
- [12] Han M, Ma Q, Ding X. High-order stochastic symplectic partitioned Runge-Kutta methods for stochastic Hamiltonian systems with additive noise[J]. Applied Mathematics and Computation, 2019, 346: 575–593.
- [13] Yang H, Wu F, Kloeden P E, Mao X. The truncated Euler-Maruyama method for stochastic differential equations with Hölder diffusion coefficients[J]. Journal of Computational and Applied Mathematics, 2020, 366: 112379.
- [14] Teng L, Zhao W. High-order combined multi-step scheme for solving forward backward stochastic differential equations[J]. Journal of Scientific Computing, 2021, 87(3): 1–25.
- [15] Zhang Y, Song M, Liu M. Convergence and stability of stochastic theta method for nonlinear stochastic differential equations with piecewise continuous arguments[J]. Journal of Computational and Applied Mathematics, 2022, 403: 113849.
- [16] Zhao J, Yi Y, Xu Y. Strong convergence and stability of the split-step theta method for highly nonlinear neutral stochastic delay integro differential equation[J]. Applied Numerical Mathematics, 2020, 157: 385–404.
- [17] Tang X, Xiao A. Asymptotically optimal approximation of some stochastic integrals and its applications to the strong second-order methods[J]. Advances in Computational Mathematics, 2019, 45(2): 813–846.

- 
- [18] Li M, Huang C, Hu P, et al. Mean-square stability and convergence of a split-step theta method for stochastic Volterra integral equations[J]. *Journal of Computational and Applied Mathematics*, 2021, 382: 113077.
- [19] Hu L, Li X, Mao X. Convergence rate and stability of the truncated Euler–Maruyama method for stochastic differential equations[J]. *Journal of Computational and Applied Mathematics*, 2018, 337: 274–289.
- [20] Wang W, Chen Y. Mean-square stability of semi-implicit Euler method for nonlinear neutral stochastic delay differential equations[J]. *Applied Numerical Mathematics*, 2011, 61(5): 696–701.
- [21] Munsky B, Khammash M. The finite state projection algorithm for the solution of the chemical master equation[J]. *The Journal of Chemical Physics*, 2006, 124 (4): 044104.
- [22] MacNamara S, Bersani A M, Burrage K, et al. Stochastic chemical kinetics and the total quasi-steady-state assumption: application to the stochastic simulation algorithm and chemical master equation[J]. *The Journal of Chemical Physics*, 2008, 129 (9): 095105.
- [23] MacNamara S, Burrage K. Stochastic modeling of naive T cell homeostasis for competing clonotypes via the master equation[J]. *Multiscale Modeling and Simulation*, 2010, 8 (4): 1325–1347.
- [24] Gillespie D T. Exact stochastic simulation of coupled chemical reactions[J]. *The Journal of Physical Chemistry*, 1977, 81 (25): 2340–2361.
- [25] Yates C A, Klingbeil G. Recycling random numbers in the stochastic simulation algorithm[J]. *The Journal of Chemical Physics*, 2013, 138 (9): 094103.
- [26] Anderson D F, Higham D J, Sun Y. Complexity of multilevel Monte Carlo tau-leaping[J]. *SIAM Journal on Numerical Analysis*, 2014, 52 (6): 3106–3127.
- [27] Li T. Analysis of explicit tau-leaping schemes for simulating chemically reacting systems[J]. *Multiscale Modeling and Simulation*, 2007, 6 (2): 417–436.
- [28] Székely Jr T. *Stochastic Modelling and Simulation in Cell Biology*[D]. University of Oxford, 2014.
- [29] Niu Y, Burrage K, Zhang C. Multi-scale approach for simulating time-delay biochemical reaction systems[J]. *IET Systems Biology*, 2014, 9 (1): 31–38.
- [30] Székely T, Burrage K, Zygalakis K C, et al. Efficient simulation of stochastic chemical kinetics with the stochastic Bulirsch–Stoer extrapolation method[J]. *BMC Systems Biology*, 2014, 8 (1): 71.
- [31] van Kampen, N G. *Stochastic Processes in Physics and Chemistry*(3rd edition)[M]. London: Elsevier Publisher. 2007.
- [32] Cotter S L, Vejchodsky Tomá, Erban R. Adaptive finite element method assisted by stochastic simulation of chemical systems[J]. *SIAM Journal on Scientific Computing*, 2013, 35(1): B107–B131.
- [33] Ferm L, Loetstedt P, Sjöberg P. Conservative solution of the Fokker-Planck equation for stochastic chemical reactions[J]. *BIT Numerical Mathematics*, 2006, 46: 61–83.
- [34] Wilkie J, Wong Y M. Positivity preserving chemical Langevin equations[J]. *Chemical Physics*, 2008, 353 (1): 132–138.
- [35] Dana S, Raha S. Physically consistent simulation of mesoscale chemical kinetics: the non-negative FIS- method[J]. *Journal of Computational Physics*, 2011, 230 (24): 8813–8834.

- [36] Schnoerr D, Sanguinetti G, Grima R. The complex chemical Langevin equation[J]. The Journal of Chemical Physics, 2014, 141 (2): 024103.
- [37] Niu Y, Burrage K, Zhang C. Modelling biochemical reaction systems by stochastic differential equations with reflection[J]. Journal of Theoretical Biology, 2016, 396: 90–104.
- [38] Kawamura T, Saisho Y. Stochastic models describing human metabolic processes using SDEs with reflection[J]. Stochastic Models, 2006, 22 (2): 273–287.
- [39] Dangerfield C E, Kay D, Burrage K. Modeling ion channel dynamics through reflected stochastic differential equations[J]. Physical Review E, 2012, 85 (5): 051907.
- [40] Mao X. Stochastic Differential Equations and Applications[M]. Chichester: Horwood Publishing, 2007.
- [41] Mao X, Wei F, Wiriyaikul T. Positivity preserving truncated Euler-Maruyama method for stochastic Lotka-Volterra competition model[J]. Journal of Computational and Applied Mathematics, 2021, 394: 113566.
- [42] 戴磊. 多维噪声情形下 Lotka-Volterra 竞争模型的 Milstein 型保正算法 [D]. 中南大学, 2022.
- [43] Li Y, Cao W. A positivity preserving Lamperti transformed Euler-Maruyama method for solving the stochastic Lotka-Volterra competition model[J]. Communications in Nonlinear Science and Numerical Simulation, 2023, 112: 107206.
- [44] Chessa S, Yashima H F. The stochastic equation in predator-prey population dynamics[J]. Bollettino della Unione Matematica Italiana B, 2002, 3: 789–804.
- [45] Rudnicki R. Long-time behavior of a stochastic prey-predator model[J]. Stochastic Processes and their Applications, 2003, 108(1): 93–107.
- [46] Rudnicki R, Pichor K. Influence of stochastic perturbation on prey-predator systems[J]. Mathematical Biosciences, 2007, 206: 108–119.
- [47] Dai X, Mao Z, Li X. A stochastic prey-predator model with time-dependent delays[J]. Advances in Difference Equations, 2017, (1): 297.
- [48] Hong J, Ji L, Wang X, et al. Positivity-preserving symplectic methods for the stochastic Lotka – Volterra predator-prey model[J]. BIT Numerical Mathematics, 2022, 62(2): 493–520.
- [49] Zuo W, Jiang D, Sun X, et al. Long-time behaviors of a stochastic cooperative Lotka-Volterra system with distributed delay[J]. Physica A, 2018, 506 : 542–559.
- [50] Liu, Q. Analysis of a stochastic non-autonomous food-limited Lotka-Volterra cooperative model[J]. Applied Mathematics and Computation, 2015, 254: 1–8.
- [51] Zhang M, Zhang Q, Tian J, et al. The asymptotic stability of numerical analysis for stochastic age-dependent cooperative Lotka-Volterra system[J]. Mathematical Biosciences and Engineering, 2021, 18(2): 1425–1449.
- [52] Kermack W O, Mckendrick A G. A contribution to the mathematical theory of epidemics[J]. Proceedings of the Royal Society of London, Series A, 1927, 115(772): 700–721.
- [53] Hutzenhaler M, Jentzen A. Numerical approximations of stochastic differential equations with non-globally Lipschitz continuous coefficients[J]. Memoirs of the American Mathematical Society,

- 2015, 236(1112): 1-112.
- [54] Yang X, Yang Z, Zhang C. Numerical analysis of the Linearly implicit Euler method with truncated Wiener process for the stochastic SIR model[J]. *Mathematics and Computers in Simulation*, 2023, 208: 1-14.
- [55] Khan S U, Ali I. Numerical analysis of stochastic SIR model by Legendre spectral collocation method[J]. *Advances in Mechanical Engineering*, 2019, 11(7): 1687814019862918.
- [56] Gray A, Greenhalgh D, Hu L, et al. A stochastic differential equation SIS epidemic model[J]. *SIAM Journal on Applied Mathematics*, 2011, 71(3): 876-902.
- [57] Chen L, Gan S, Wang X. First order strong convergence of an explicit scheme for the stochastic SIS epidemic model[J]. *Journal of Computational and Applied Mathematics*, 2021, 392: 113482.
- [58] Yang H, Huang J. First order strong convergence of positivity preserving logarithmic Euler - Maruyama method for the stochastic SIS epidemic model[J]. *Applied Mathematics Letters*, 2021, 121: 107451.
- [59] Yang H, Pan Y, Liu W, et al. Numerical analysis of split-step  $\theta$  methods with truncated Wiener process for a stochastic SIS epidemic model[J]. *Journal of Computational and Applied Mathematics*, 2022: 114433.
- [60] Cai S, Cai Y, Mao X. A stochastic differential equation SIS epidemic model with two independent Brownian motions[J]. *Journal of Mathematical Analysis and Applications*, 2019, 474(2): 1536-1550.
- [61] Cai S, Hu J, Mao X. Positivity and boundedness preserving numerical scheme for the stochastic epidemic model with square-root diffusion term[J]. *Applied Numerical Mathematics*, 2022, 182: 100-116.
- [62] Dangerfield C E, Kay D, Macnamara S, et al. A boundary preserving numerical algorithm for the Wright-Fisher model with mutation[J]. *BIT Numerical Mathematics*, 2012, 52(2): 283-304.
- [63] Neuenkirch A, Szpruch L. First order strong approximations of scalar SDEs defined in a domain[J]. *Numerische Mathematik*, 2014, 128(1): 103-136.
- [64] Chen L, Gan S. Strong convergence and stationary distribution of an explicit scheme for the Wright-Fisher model[J]. *Journal of Computational and Applied Mathematics*, 2023, 424: 115017.
- [65] Tang F, Barbacioru C, Wang Y. et al. mRNA-Seq whole-transcriptome analysis of a single cell[J]. *Nature Methods*, 2009, 6: 377-382.
- [66] Shi J, Li T, Chen L, et al. Quantifying pluripotency landscape of cell differentiation from scRNA-seq data by continuous birth-death process[J]. *PLoS Computational Biology*. 2019, 15(11): e1007488.
- [67] Shi J, Teschendorff A E, Chen W, et al. Quantifying Waddington's epigenetic landscape: a comparison of single-cell potency measures[J]. *Briefings in Bioinformatics*, 2020, 21(1): 248-261.
- [68] Shi J, Kazuyuki A, Li T, et al. Energy landscape decomposition for cell differentiation with proliferation effect[J]. *National Science Review*, 2022, 9(8): nwac116.
- [69] Shi J, Kazuyuki A, Chen, L. Dynamics-based data science in biology[J]. *National Science Review*, 2021, 8(5): nwab029.
- [70] Han X, Kloeden P E. *Random Ordinary Differential Equations and Their Numerical Solutions*[M].



Singapore: Springer Nature, 2017.

- [71] Hu J, Amor D R, Barbier M, et al. Emergent phases of ecological diversity and dynamics mapped in microcosms[J]. Science, 2022, 378(6615): 85-89.
- [72] Allesina S, Tang S. Stability criteria for complex ecosystems[J]. Nature, 2012, 483(7388): 205-208.