Stochastic Models of Biochemical Reaction Networks: Sensitivity Analysis

Monjur Morshed[1], Dr. Brian Ingalls[1] and Dr. Silvana Ilie[2]

[1]Department of Applied Mathematics, University of Waterloo, ON, Canada, {m6morshe,bingalls}@uwaterloo.ca

[2]Department of Mathematics, Ryerson University, ON, Canada, silvana@ryerson.ca

Motivation

- In the study of Systems Biology it is necessary to simulate cellular processes and chemical reactions that comprise biochemical systems. This is achieved through a range of mathematical modelling approaches.
- Standard methods use deterministic differential equations, but because many biological processes are inherently probabilistic, stochastic models must be used to capture the random fluctuations observed in these systems.
- The presence of noise in a system can be a significant factor in determining its behavior. The Chemical Master Equation is a valuable stochastic model of biochemical kinetics.
- One critical tool in the study of biochemical systems is sensitivity analysis, which aims to quantify the dependence of a system's dynamics on model parameters. A number of approaches to sensitivity analysis of these models have been developed.
- We provide a comparison of several methodologies and introduce a strategy based on tau-leaping. We identify which approach is most efficient depending of the features of the model. This result can serve as a guide to efficient sensitivity analysis, which can serve as a foundation for the formulation, characterization, and verification of models.

Modelling Assumptions

- A well-mixed isothermal system of chemical reactions is confined to a constant volume Ω.
- The chemical species are $[S_1, S_2, ..., S_N]$, where $N \geq 1$ is the number of molecular species in the system.
- The chemical species are subject to the reaction channels $[R_1, ..., R_M]$, where $M \geq 1$ is the number of different reactions in the system. Assume that the reactions are instantaneous events.
- At time t, the system state is described by the vector $X(t) = (x_1(t), x_2(t), ...x_N(t))^T$, where $x_i(t)$ is the number of molecules of species S_i at time t. $X(t)$ is a Markov process.
- A reaction channel R_j can be characterized by:
  - The *propensity function* $a_j(x)$, defined as $a_j(x)\mathrm{d}t =$ the probability that a single reaction R_j occurs in the infinitesimal time interval $[t, t + \mathrm{d}t)$, if at time t, $X(t) = x$.
  - The *state change vector* ν_j, which describes the change in the molecular populations when reaction R_j occurs. Thus, if one reaction R_j fires over the time interval $[t, t + \mathrm{d}t]$, then the state of the system at time $t + \mathrm{d}t$ is $X(t + \mathrm{d}t) = x + \nu_j$, provided that $X(t) = x$. We denote $\nu_j = [\nu_{1j}, \nu_{2j}, ..., \nu_{Nj}]^T$, where ν_{ij} is the change in the molecular abundance of S_i when reaction R_j occurs. The matrix $v = (\nu_{ij})_{1\leq i\leq N, 1\leq j\leq M}$ is the *stoichiometric matrix*.

Mathematical Models of Chemical Kinetics

Chemical Master Equation: The behaviour of a well-mixed isothermal biochemical system is governed by the Chemical Master Equation, a stochastic discrete model. Define $P(x, t|x_0, t_0)$ to be the probability of the system to be in state x at time t, $X(t) = x$, given that initially $X(t_0) = x_0$.
The Chemical Master Equation (CME):

$$\frac{dP(x,t)}{dt} = \sum_{j=1}^{M} [a_j(x - \nu_j)P(x - \nu_j, t) - a_j(x)P(x, t)].$$

Reaction Rate Equation(RRE)

- Very near the thermodynamic limit, the dynamics of the well-stirred biochemical system may be modelled using the reaction rate equations, a continuous deterministic model.
  The reaction rate equations (RRE):

$$\frac{dX(t)}{dt} = \sum_{j=1}^{M} \nu_j a_j(X(t))$$

Gillespie's Algorithm

- With the system in state x at time t, evaluate $a_{sum}(X(t)) := \sum_{k=1}^{M} a_k(X(t))$
- Generate two independent uniform $(0, 1)$ random numbers, ξ_1 and ξ_2 and compute j and τ according to
  - $j =$ the smallest integer satisfying $\sum_{k=1}^{j} a_k(X(t)) > \xi_1 a_{sum}(X(t))$.
  - $\tau = \ln(1/\xi_2)/a_{sum}(X(t))$.
- Compute $X(t + \tau) = X(t) + \nu_j$ and update t to $t + \tau$.
- Return to step 1.

Tau-leaping Method

- The tau-leaping method:

$$X(t + \tau) = X(t) + \sum_{j=1}^{M} \nu_j P_j(a_j(X(t)), \tau)$$

  - It is valid if for all $j = \{1, \dots M\}$
    $a_j(X(s)) \simeq constant$ for $t \leq s \leq t + \tau$.
- The number of reactions R_j within $[t, t + \tau]$ are approximated using the independent Poisson random variables $\{P_j(a_j(X(t)), \tau)\}_{j=1}^{M}$, having mean and variances $a_j(X(t))\tau$, respectively.
- The tau-leaping algorithm with leap time τ can be summarized as:
  - 1. Draw samples $\{p_j\}_{j=1}^{M}$ from the distributions of independent Poisson random variables $\{P_j(a_j(X(t)), \tau)\}_{j=1}^{M}$.
  - 2. Set $X(t + \tau) = X(t) + \sum_{j=1}^{M} \nu_j P_j$ and update t to $t + \tau$.
  - 3. Return to step 1.

RTC Algorithm

1. Initial Conditions, $i = 0, T_0 = 0, S_j = 0, k_j = 1, X(T_0) = x_0$, and $I_+^j = E_1^j$ for $j = 1, ..., M$
2. begin loop
3. check break condition; break loop if met
4. find propensity function, $a_j(X(T_j))$, for each reaction
5. compute ΔT, set j^* to index of minimum found
6. set $X(T_{i+1}) = X(T_i) + v_{j^*}$
7. for each reaction, update $S_j = S_j + a_j(X(T_i))(\Delta T)$
8. set $k_{j^*} = k_{j^*} + 1$
9. set $I_+^{j^*} = I_+^{j^*} + E_{k_{j^*}}^{j^*}$
10. set $i = i + 1$
11. end loop

Summary

CME/Gillespie — discrete, stochastic

↓ $a_j(x) \simeq$ constant in $[t, t + \tau), \forall j$

Tau-leaping — discrete, stochastic

↓ $a_j(x)\tau \gg 1, \forall j$

CLE — continuous, stochastic

↓ thermodynamic limit

RRE — continuous, deterministic

Sensitivity Analysis

- Sensitivity analysis describes how properties of the system change when variations are introduced into the model parameters.
- A model output has high sensitvity to model parameter if a small change in the parameter results in a large change in the system output.
- Sensitivity analysis plays an important role in assessing the accuracy of a model, in model development and in model reduction.
- Sensitivity analysis helps to make decisions on which parts of the model are actively contributing to the system dynamics.

Monte Carlo Sensitivity Approaches

Infinitesimal Perturbation:

- Pathwise differentiation (PD)

Finite Perturbation:

- Independent samples:
  - Independent random numbers (IRN) with SSA
- Correlated sample:
  - Common random numbers (CRN) with SSA
  - Common reaction Path (CRP) with RTC
  - Coupled Leaping Sensitivity (CLS)

Common Random Numbers (CRN) Algorithm

1. begin loop over number of trajectories, N, for each i
2. generate large array of random numbers, r^*
3. choose system parameter, c_0, and execute Gillespie's algorithm for the nominal system, $X(T, c_0)$, using array of random numbers, r_j^*
4. set parameter to $c_0 + h$, and execute Gillespie's algorithm, calculating for perturbed system, $X(T, c_0 + h)$, using the same array of random numbers, r_j^*
5. find sensitivity by $s_i = [f(X(T, c_0 + h)) - f(X(T, c_0))]/h$
6. end loop over i
7. find mean and standard deviation of $\{s_i, \text{for } i = 1, ..., N\}$

Common Reaction Path (CRP) Algorithm [5]

1. loop over number of trajectories, N, for each i
2. generate large array of random numbers, r_j^* for each reaction
3. choose system parameter, c_0, and execute RTC algorithm, calculating for the nominal system, $X(T, c_0)$, using array of random numbers, r_j^*
4. set parameter to $c_0 + h$, and execute RTC algorithm, calculating for perturbed system, $X(T, c_0 + h)$, using the same array of random numbers, r_j^*
5. find sensitivity by $s_i = [f(X(T, c_0 + h)) - f(X(T, c_0))]/h$
6. end loop over i
7. find mean and standard deviation of $\{s_i, \text{for } 1, ..., N\}$

Coupled Leaping Sensitivity (CLS) Algorithm

1. loop over number of trajectories, N, for each i
2. choose system parameter, c_0, and the perturbed parameter, $c_0 + h$, and execute a leaping algorithm with tight coupling, calculating for nominal system, $X(T, c_0)$, and for perturbed system, $X(T, c_0 + h)$
3. find sensitivity by $s_i = [f(X(T, c_0 + h)) - f(X(T, c_0))]/h$
4. end loop over i
5. find mean and standard deviation of $\{s_i, \text{for } i = 1, ..., N\}$

Numerical Results

Michaelis-Menten Model:

R_i	Reaction	Propensities	Reaction rate
R_1	$S_1 + S_2 \xrightarrow{C_1} S_3$	$a_1 = C_1X_1X_2$	$C_1 = 1.661 \times 10^{-3}$
R_2	$S_3 \xrightarrow{C_2} S_1 + S_2$	$a_2 = C_2X_3$	$C_2 = 10^{-4}$
R_3	$S_3 \xrightarrow{C_3} S_4 + S_2$	$a_3 = C_3X_3$	$C_3 = 0.1000$

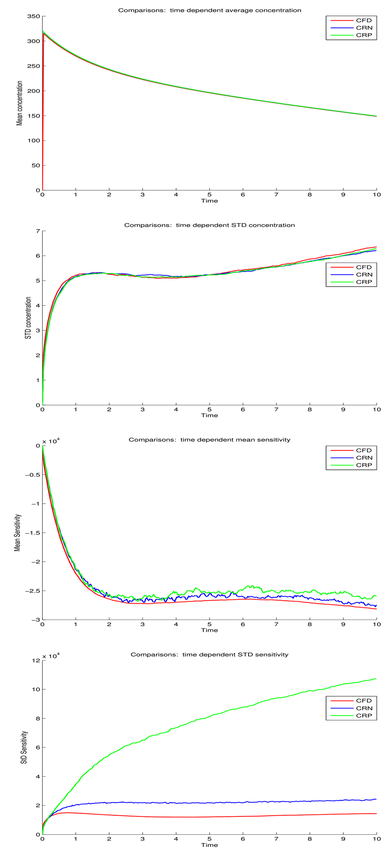

Figure 1: The Michaelis-Menten model: the comparisons for the different sensitivity methods with perturbation parameter $h = 5 \times 10^{-5}$, using $10,000$ trajectories, on the interval $[0, 10]$ (species S_1) .

Birth Death Model:

R_i	Reaction	Propensities	Reaction rate
R_1	$\oslash \xrightarrow{C_1} X$	$a_1(x) = C_1$	$C_1 = 2.5$
R_2	$X \xrightarrow{C_2} \oslash$	$a_2(x) = C_2X$	$C_2 = 0.1$

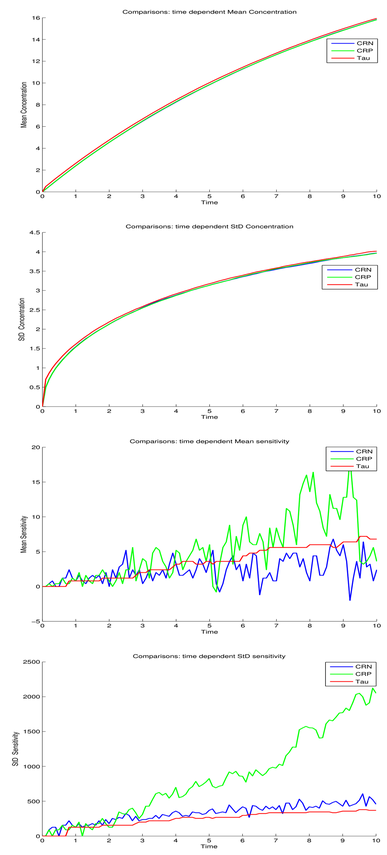

Figure 2: The Birth Death Model: the comparisons for the different sensitivity methods with perturbation parameter $h = 5 \times 10^{-5}$, using $50,000$ trajectories, on the interval $[0, 10]$ (species X)

Schlogl Model:

R_i	Reaction	Propensities	Reaction rate
R_1	$A + 2X \xrightarrow{C_1} 3X$	$a_1 = C_1AX(X - 1)/2$	$C_1 = 3 \times 10^{-7}$
R_2	$3X \xrightarrow{C_2} A + 2X$	$a_2 = C_2X(X - 1)(X - 2)/6$	$C_2 = 10^{-4}$
R_3	$B \xrightarrow{C_3} X$	$a_3 = C_3B$	$C_3 = 10^{-3}$
R_4	$X \xrightarrow{C_4} B$	$a_4 = C_4X$	$C_4 = 3.5$

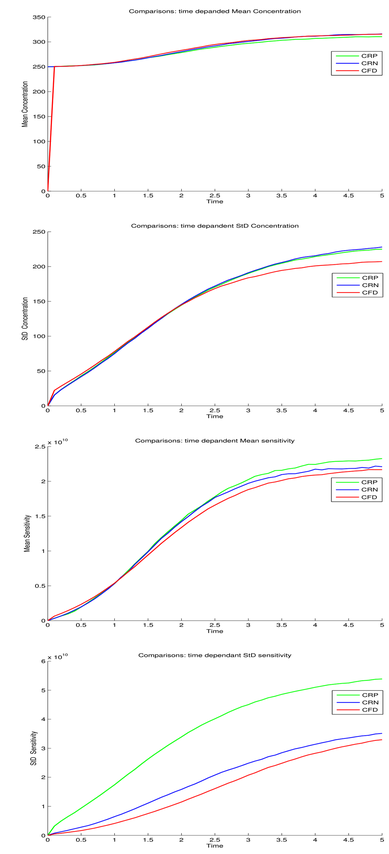

Figure 3: The Schlogl Model: the comparisons for the different sensitivity methods with perturbation parameter $h = 5 \times 10^{-9}$, using $5,000$ trajectories, on the interval $[0, 5]$ (species X).

Remark The leaping sensitivity method is 25 times faster than the CRP and 5 times faster than the CRN for the Schlogl model. It is also more accurate than CRP and CRN.

Discussion

- We discussed several sensitivity analysis strategies for the stochastic discrete model of well-stirred biochemical kinetics, the Chemical Master Equation, and introduced a strategy for approximating sensitivities based on a leaping method.
- The methods presented employ finite difference estimators to approximate sensitivities.
- The leaping based strategy uses coupling between the nominal and the perturbed processes to reduce the variance, which may lead to a reduced computational cost of the algorithm, for a similar accuracy.
- This result can serve as a guide to efficient sensitivity analysis, which can serve as a foundation for the formulation, characterization, and verification of models.

Acknowledgements

We acknowledge the financial support provided by grants from the Natural Sciences and Engineering Research Council of Canada (NSERC) and by an Ontario Graduate Scholarship (OGS).

Bibliography

1. D.T. Gillespie, *J. Chem. Phys.*, 81, 2340–2361 (1977).
2. D.T.Gillespie, *Physica A*, 188, 402–425 (1992).
3. D.T. Gillespie, *J. Chem. Phys.*, 115, 1716–1733 (2001).
4. D. F. Anderson, *SIAM J. Numer. Anal.*, 50, 2237 (2012).
5. M. Rathinam, P. W. Sheppard, and M. Khammash, *J. Chem. Phys.*, 132, 034103 (2010).