

Effective Finite-Difference Techniques for Estimating Sensitivities for Stochastic Biochemical Systems

Fauzia Jabeen, Silvana Ilie*

Department of Mathematics, Ryerson University, Toronto, Canada
Email: *silvana@ryerson.ca

How to cite this paper: Jabeen, F. and Ilie, S. (2022) Effective Finite-Difference Techniques for Estimating Sensitivities for Stochastic Biochemical Systems. *Applied Mathematics*, 13, 878-895.
<https://doi.org/10.4236/am.2022.1311056>

Received: October 5, 2022

Accepted: November 21, 2022

Published: November 24, 2022

Copyright © 2022 by author(s) and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Cellular environments are in essence stochastic, owing to the random character of the biochemical reaction events in a single cell. Stochastic fluctuations may substantially contribute to the dynamics of systems with small copy numbers of some biochemical species. Then, stochastic models are indispensable for properly portraying the behaviour of the system. Sensitivity analysis is one of the central tools for studying stochastic models of cellular dynamics. Here, we propose some finite-difference strategies for estimating parametric sensitivities of higher-order moments of the system state for stochastic discrete biochemical kinetic models. To reduce the variance of the sensitivity estimator, we employ various coupling techniques. The advantages of the proposed methods are illustrated in several models of biochemical systems of practical relevance.

Keywords

Stochastic Simulation Algorithm, Stochastic Biochemical Kinetics, Sensitivity Analysis

1. Introduction

In recent years, stochastic models have been successfully utilized for capturing the discrete nature of intracellular networks and the randomness of the molecular interactions, when some biochemical species are available in low amounts [1] [2] [3]. By contrast, deterministic formulations of the standard theory of chemical kinetics are unable to describe the effects of such random fluctuations and, on many occasions, can not reliably depict the average behaviour of the system [4]. For such networks, the evolution of the molecular populations can be

represented by continuous time discrete space Markov processes. Chemical Master Equation is a discrete stochastic model which adequately describes the evolution of well-stirred biochemical networks [5]. This formulation can be solved directly just for a few simple systems. In most cases, the solution of the Chemical Master Equation can solely be approximated numerically, frequently by applying Monte Carlo strategies. Among the Monte Carlo methods which compute exact realizations of the Markov process modelling the evolution of a biochemical system are the stochastic simulation algorithm of Gillespie [6] [7] and the algorithm based on the random time change representation of Kurtz [8]. Numerical methods to simulate approximate realizations of this Markov process were also proposed in the literature (see, e.g. [9] [10] [11] [12] [13] and references therein).

Biochemical reaction models depend on a number of reaction rate parameters [14]. In many instances, the values of these parameters are inadequately measured. Hence, it is important to investigate the sensitivity of the model behaviour with respect to changes in the parameter values [15]. Studying how the dynamics of the system respond to small fluctuations in the model parameters is the subject of local sensitivity analysis. Among the applications of parametric sensitivity, we mention identifiability analysis, parameter estimation, fine-tuning of the model and guiding model reduction [16] [17] [18]. The existing methods to estimate parameter sensitivities of biochemical reaction networks often rely on simulations of exact methods such as the stochastic simulation algorithm (SSA) [6] [7], the Random Time Change (RTC) algorithm [19] or the Next Reaction Method (NRM) [20] [21]. Monte Carlo techniques are utilized together with finite-difference schemes to approximate local parametric sensitivities [16] [19] [20] [22] [23]. More precisely, the sensitivity of the expected value $\partial \mathbb{E}[\mathbf{f}(\mathbf{X}(t,c))]/\partial c$ is estimated by finite differences such as $(\mathbb{E}[\mathbf{f}(\mathbf{X}(t,c+h))] - \mathbb{E}[\mathbf{f}(\mathbf{X}(t,c))])/h$. Here, $\mathbb{E}[\cdot]$ is the expected value, \mathbf{f} represents the function of interest, $\mathbf{X}(t,c)$ denotes the state of the system at time t , depending on the parameter c . Among the existing finite-difference sensitivity estimators of the stochastic biochemical kinetic models are the Common Random Number (CRN), the Common Reaction Path (CRP) [19] and the Coupled Finite Difference (CFD) [20] methods, which rely on the SSA, the RTC and the NRM algorithms, respectively, to simulate sample paths. These methods employ coupling of the nominal and perturbed trajectories to reduce the variance of the sensitivity estimator. Unfortunately, the existing applications focused exclusively on approximating the sensitivity of the expected value of various molecular species, $\partial \mathbb{E}[\mathbf{X}(t,c)]/\partial c$, and there are no studies dedicated to sensitivity estimations of higher moments, such as the variance of the state vector $\mathbf{X}(t,c)$. Nonetheless, the sensitivity of the variance of the system state may be critical for many models arising in practice, such as for systems with large levels of noise or for which the intrinsic noise drives the system behaviour. For such models, examining only the robustness of the mean trajectory with respect to changes in various parameters will not be sufficient for

studying the influence of these parameters on the system dynamics and the level of noise.

In this paper, we introduce some sensitivity estimators with respect to kinetic parameters of higher-order moments (e.g. variance) of the system state of a stochastic discrete biochemical model. These estimators make use of the Common Random Number, the Common Reaction Path and the Coupled Finite Difference schemes to generate coupled trajectories, and central finite-difference strategies to accurately approximate the sensitivities of higher-order moments. The coupling reduces the variance of the estimator, yielding a more efficient simulation for a similar accuracy of the output. We note that the analysis tools developed in this work have no counterpart for the simplified, deterministic model of well-stirred biochemical systems, unlike the sensitivity analysis methods for the expected value of the system state. The proposed methodologies are particularly useful for models with moderate to large levels of noise, for which a more comprehensive sensitivity analysis is crucial for studying the model behaviour. Additionally, we expect these tools to be valuable for a range of applications of local sensitivity analysis, such as guiding model reduction and identifiability studies, for biochemical networks whose intrinsic noise exerts a significant influence on the dynamics. The intrinsic noise consists of random fluctuations due to the biochemical reaction events, in particular associated with low molecular amounts of some reactants.

This work is organized as follows: Section 2 discusses the background of stochastic discrete models of well-stirred biochemical networks and their simulation strategies. Some existing parametric sensitivity techniques are described in Section 3. In Section 4, the proposed finite-difference sensitivity estimators of higher-order moments of the biochemical system state are presented. The features of the proposed techniques are reported in Section 5, where numerical tests are performed on three models arising in applications. Lastly, our conclusions are drawn in Section 6.

2. Background

2.1. Stochastic Modelling of Homogeneous Biochemical Systems

Let us consider a system of N biochemical species S_1, \dots, S_N , experiencing M chemical reactions, R_1, \dots, R_M . The system is contained in a constant volume Ω and is held at a constant temperature T , in a spatially homogeneous regime. Denote by the vector $\mathbf{X}(t) = (X_1(t), \dots, X_N(t))^T$ the state of the system at time t , where $X_i(t)$ is a non-negative integer indicating the number of molecules of species S_i at time t . The number of molecules for each species at the initial time $t_0 = 0$ is given, $X(t_0) = x_0$.

A reaction R_j is specified by the associated propensity function $a_j(\cdot)$ and stoichiometric vector ν_j . The propensity of the j -th reaction is such that $a_j(x)dt$ measures the probability that one reaction R_j fires during the time interval $[t, t + dt)$, assuming that $\mathbf{X}(t) = \mathbf{x}$. The state change vector ν_j de-

picts the variation in the molecular amounts of all chemical species when a reaction R_j happens. More precisely, if $\mathbf{X}(t) = \mathbf{x}$ is the system state at time t , then the firing of one reaction R_j in the time interval $[t, t + dt)$ brings the system to state $\mathbf{X}(t + dt) = \mathbf{x} + \nu_j$. The entry ν_{ij} of the state change vector $\nu_j = (\nu_{1j}, \nu_{2j}, \dots, \nu_{Nj})^T$ measures the variation in the S_j molecular count after a single reaction R_j occurs. Finally, $\nu = (\nu_{ij})_{1 \leq i \leq N, 1 \leq j \leq M}$ denotes the stoichiometric matrix of the biochemical reaction system.

The Chemical Master Equation [5] represents the evolution in time of the conditional probability $P(\mathbf{x}, t | \mathbf{x}_0, t_0)$ of the system being at time t in state $\mathbf{X}(t) = \mathbf{x}$, on the assumption that at time $t_0 = 0$ it was in state $\mathbf{X}(t_0) = \mathbf{x}_0$. The Chemical Master Equation,

$$\frac{dP(\mathbf{x}, t | \mathbf{x}_0, t_0)}{dt} = \sum_{j=1}^M \left[a_j(\mathbf{x} - \nu_j) P(\mathbf{x} - \nu_j, t | \mathbf{x}_0, t_0) - a_j(\mathbf{x}) P(\mathbf{x}, t | \mathbf{x}_0, t_0) \right], \quad (1)$$

is a stochastic discrete model of well-stirred biochemical networks. For most networks arising in applications, this model is very difficult to solve directly [24]. To overcome this difficulty, Monte Carlo simulations may be employed that generate trajectories in agreement with the solution of the Chemical Master Equation. Let us remark that the system state $\mathbf{X}(t)$, governed by the Chemical Master Equation, may be expressed using the Random Time Change representation of Kurtz [8],

$$\mathbf{X}(t) = \mathbf{X}(0) + \sum_{j=1}^M \nu_j Z_j \left(\int_0^t a_j(\mathbf{X}(\theta)) d\theta \right), \quad (2)$$

where Z_j are independent Poisson processes with unit rates, for $1 \leq j \leq M$.

2.2. Stochastic Simulations of Homogeneous Biochemical Systems

Some of the widely used exact Monte Carlo simulation strategies for the Chemical Master Equation are the Stochastic Simulation Algorithm (SSA) due to Gillespie [6] [7] and the Reaction Time Change (RTC) method [19] [25]. These strategies are utilized for reconstructing the probability distribution of the biochemical system state. Furthermore, the SSA and RTC schemes are employed for computing the coupled paths for estimating sensitivities with the Common Random Number and the Common Reaction Path algorithms, respectively. A modified Next Reaction Method, which is an exact Monte Carlo scheme, is applied by the Coupled Finite Difference strategy (see [20] for details).

Stochastic Simulation Algorithm

The SSA proposed by Gillespie [6] [7] generates an exact realization of the Markov process representing the system state. At each step, the algorithm computes the time and the index of the reaction that occurs first and updates the system accordingly. Gillespie's algorithm is outlined below.

- 1) Initialize the state $\mathbf{X} = \mathbf{x}_0$ at initial time $t = t_0$.
- 2) Calculate the propensity functions $a_j(\mathbf{X})$, for $1 \leq j \leq M$, and their sum $a_0(\mathbf{X}) = \sum_{j=1}^M a_j(\mathbf{X})$.

3) Pick two independent samples u_1 and u_2 from the uniform distribution $U(0,1)$.

4) Choose the time to the next reaction as $\tau = \frac{1}{a_0(\mathbf{X})} \ln \frac{1}{u_1}$.

5) Determine the index ℓ of the next reaction, namely the smallest integer for which $u_2 a_0(\mathbf{X}) < \sum_{j=1}^{\ell} a_j(\mathbf{X})$.

6) Update the state $\mathbf{X} \leftarrow \mathbf{X} + \nu_{\ell}$ as well as the time $t \leftarrow t + \tau$.

7) Go to step 2 or stop.

Random Time Change Algorithm

An alternative exact Monte Carlo simulation method for stochastic discrete biochemical kinetic models is the Random Time Change algorithm proposed by Rathinam *et al.* [19]. Remark that the internal time of each reaction R_j with $1 \leq j \leq M$, associated with the physical time t , can be expressed as

$$s_j(t) = \int_0^t a_j(\mathbf{X}(\theta)) d\theta.$$

According to the representation (2), the number of occurrences of R_j during $[0, t]$ is $Z_j(s_j(t))$. Here Z_j , with $1 \leq j \leq M$, are independent until rate Poisson processes. Given M streams of unit exponential random numbers, E_n^j , with $n = 1, 2, \dots$ and $1 \leq j \leq M$, execute the following steps (see [19] for a comprehensive description):

1) Initialize $n = 0$, the time $T_n = 0$ and the state $\mathbf{X}(T_n) = \mathbf{x}_0$.

For $j = 1, 2, \dots, M$, take $k_j = 1$, $I_+^j = E_1^j$ and the internal time of the R_j reaction, $\sigma_j = 0$.

2) For $j = 1, 2, \dots, M$, evaluate $a_j(\mathbf{X}(T_n))$.

3) Pick $\delta t = \min_{1 \leq j \leq M} \left\{ \frac{1}{a_j(\mathbf{X}(T_n))} (I_+^j - \sigma_j) \right\}$. Determine ℓ the index for which minimum is obtained.

4) Update the time $T_{n+1} \leftarrow T_n + \delta t$ and the state $\mathbf{X}(T_{n+1}) \leftarrow \mathbf{X}(T_n) + \nu_{\ell}$.

5) For $j = 1, 2, \dots, M$, update the internal time $\sigma_j \leftarrow \sigma_j + a_j(\mathbf{X}(T_n)) \cdot \delta t$.

6) Set $k_{\ell} \leftarrow k_{\ell} + 1$.

7) Take $I_+^{\ell} \leftarrow I_+^{\ell} + E_{k_{\ell}}^{\ell}$.

8) Set $n \leftarrow n + 1$ and go to step 2 or stop.

3. Parametric Sensitivity

Parametric sensitivity assesses the influence of the model parameters on the biochemical system dynamics. Particularly, it is an important technique when the values of some of these parameters are inaccurate or uncertain. Local sensitivity analysis investigates how the system behaviour is altered by small variations in model parameters. A large local sensitivity of the system state with respect to a parameter suggests that the effect of inaccuracies in that parameter measurement is also expected to be substantial. This calls for a more precise estimation of the parameter of interest. In such cases, we say that the model is sensitive with respect to fluctuations in that parameter. Otherwise, the biological model is ro-

bust relative to variations in the parameter under consideration. For the deterministic framework, computing the sensitivities $\partial X_i(t,c)/\partial c$ involves solving systems of coupled ordinary differential equations. By contrast, in stochastic models, estimating the parametric sensitivities $\partial \mathbb{E}[\mathbf{f}(\mathbf{X}(t,c))]/\partial c$, where \mathbf{f} is a function of interest and $\mathbb{E}[\cdot]$ is the expectation, is quite a challenging problem. Here c is one of the model parameters. It often happens that the Chemical Master Equation is a very high dimension model. Indeed, it is a system of ordinary differential equations with one equation for any possible state of the system.

In the literature, there are several approaches for estimating local sensitivities, finite-difference strategies being among the most efficient and popular ones [16] [19] [20]. The sensitivity $\partial \mathbb{E}[\mathbf{f}(\mathbf{X}(t,c))]/\partial c$ may be approximated, for example, by the forward finite-difference scheme $(\mathbb{E}[\mathbf{f}(\mathbf{X}(t,c+h))] - \mathbb{E}[\mathbf{f}(\mathbf{X}(t,c))])/h$, where h is a small perturbation of the parameter of interest c . The expected values $\mathbb{E}[\mathbf{f}(\mathbf{X}(t,c+h))]$ and $\mathbb{E}[\mathbf{f}(\mathbf{X}(t,c))]$ are approximated by Monte Carlo simulation methods. To reduce the variance of the sensitivity estimation and thus to more efficiently compute the approximation with a comparable accuracy, we couple the nominal and perturbed trajectories. This is done by employing common random numbers. The coupled perturbed and nominal trajectories may be simulated by Gillespie's direct method [6] [7], the RTC algorithm [19], the next reaction method and the Random Time Change representation [20] or the tau-leaping schemes [16] [22]. These methods were applied for determining the sensitivities of the expected value $\mathbb{E}[\mathbf{X}(t,c)]$. Some effective techniques for approximating local parametric sensitivities are the Common Random Number, the Common Reaction Path [19] and the Coupled Finite Difference [20] algorithms, which are summarized below.

Common Random Number Algorithm

The Common Random Number method [19] employs the Stochastic Simulation Algorithm to generate pairs of trajectories. The nominal and perturbed paths are coupled by considering the same array of $[0,1]$ uniform random numbers. The algorithm is presented below.

- 1) For $k = 1$ to L .
- 2) Draw a sequence of independent uniform $(0,1)$ random numbers.
- 3) For this sequence of random numbers, apply SSA to compute the k -th pair of nominal and perturbed trajectories, $\mathbf{X}^{[k]}(t,c)$ and $\mathbf{X}^{[k]}(t,c+h)$.
- 4) Use a finite-difference scheme to estimate the sensitivity of the k -th trajectory as

$$\Gamma_k(t) = \frac{\mathbf{f}(\mathbf{X}^{[k]}(t,c+h)) - \mathbf{f}(\mathbf{X}^{[k]}(t,c))}{h}.$$

- 5) End for loop.

Common Reaction Path Algorithm

The Common Reaction Path strategy [19] generates the trajectories by applying the Random Time Change Algorithm. In other words, this algorithm com-

putes two trajectories based on the next reaction method, coupled based on the random time change representation (2). In this case, the perturbed and unperturbed paths are coupled utilizing the same array of unit exponential random numbers.

- 1) For $k = 1$ to L .
- 2) Generate M sequences of unit exponential random numbers E_1^j, E_2^j, \dots for $j = 1, 2, \dots, M$.
- 3) For these sequences of random numbers, apply the RTC algorithm to compute the k -th pair of nominal and perturbed trajectories, $\mathbf{X}^{[k]}(t, c)$ and $\mathbf{X}^{[k]}(t, c + h)$.
- 4) Use a finite-difference scheme to estimate the sensitivity of the k -th trajectory as

$$\Gamma_k(t) = \frac{\mathbf{f}(\mathbf{X}^{[k]}(t, c + h)) - \mathbf{f}(\mathbf{X}^{[k]}(t, c))}{h}.$$

- 5) End for loop.

Coupled Finite Difference Algorithm

The Coupled Finite Difference technique [20] applies a strong coupling of the paths utilized for estimating the sensitivity by a finite-difference scheme. As a result of this strong coupling between the nominal and perturbed trajectories, the method reduces the estimation variance. This lowers the computational cost of obtaining the desired accuracy of the sensitivity estimation. The CFD relies on the Next Reaction Method and the Random Time Change representation [20].

- 1) Initialize time and states: $t = 0, \mathbf{X}^{[k]}(t, c + h) = \mathbf{X}^{[k]}(t, c) = x_0$.
- 2) For j and for $i = 1:3$, take $T_{j,i} = 0$ and $r_{j,i} = rand(0,1), P_{j,i} = \ln(1/r_{j,i})$.
- 3) While $t < T$.
- a) For $j = 1: M$.
- i) Evaluate $a_j(\mathbf{X}^{[k]}(t, c), c)$ and $a_j(\mathbf{X}^{[k]}(t, c + h), c + h)$.
- ii) Determine.

$$A_{j,1} = \min\{a_j(\mathbf{X}^{[k]}(t, c), c), a_j(\mathbf{X}^{[k]}(t, c + h), c + h)\},$$

$$A_{j,2} = a_j(\mathbf{X}^{[k]}(t, c + h), c + h) - A_{j,1},$$

$$A_{j,3} = a_j(\mathbf{X}^{[k]}(t, c), c) - A_{j,1}.$$

- iii) For $i = 1:3$ if $A_{j,i} > 0, \delta t_{j,i} = (P_{j,i} - T_{j,i}) / A_{j,i}$, else $\delta t_{j,i} = \infty$.
- b) Obtain $\Delta t = \min_{j,i} \{\delta t_{j,i}\}$ and find the indices for minimum, $r = \{j^*, i^*\}$.
- c) if $i^* = 1, \mathbf{X}^{[k]}(t + \Delta t, c + h) = \mathbf{X}^{[k]}(t, c + h) + \mu_{j^*}$ and $\mathbf{X}^{[k]}(t + \Delta t, c) = \mathbf{X}^{[k]}(t, c) + \mu_{j^*}$, else if $i^* = 2, \mathbf{X}^{[k]}(t + \Delta t, c + h) = \mathbf{X}^{[k]}(t, c + h) + \mu_{j^*}$ and $\mathbf{X}^{[k]}(t + \Delta t, c) = \mathbf{X}^{[k]}(t, c)$, else $\mathbf{X}^{[k]}(t + \Delta t, c) = \mathbf{X}^{[k]}(t, c) + \mu_{j^*}$ and $\mathbf{X}^{[k]}(t + \Delta t, c + h) = \mathbf{X}^{[k]}(t, c + h)$.
- d) $t = t + \Delta t$.
- e) For $j = 1: M$ and $i = 1:3$, let $T_{j,i} = T_{j,i} + \Delta t \cdot A_{j,i}$.
- f) Take $P_r = P_r + \ln(1/\eta)$, with $\eta = rand(0,1)$.

- 4) End while.
- 5) Estimate the sensitivity of the k -th trajectory by

$$\Gamma_k(t) = \frac{\mathbf{f}(\mathbf{X}^{[k]}(t, c+h)) - \mathbf{f}(\mathbf{X}^{[k]}(t, c))}{h}.$$

For the Common Random Number, the Common Reaction Path and the Coupled Finite Difference algorithms, an estimation of the sensitivity of $\mathbb{E}[\mathbf{f}(\mathbf{X}(t, c))]$ with respect to parameter c is obtain by taking the component-wise sample average of $\{\Gamma_k(t)\}_{k=1, \dots, L}$. Remark that the smaller the standard deviation of the sensitivity estimator, the more effective the estimator is for a similar accuracy. Frequently, the CFD method outperforms the CRP and the CRN schemes in terms of computational efficiency, for a comparable accuracy of the sensitivity approximation [20].

4. Sensitivity of Higher-Order Moments

In this section, we introduce some techniques for estimating local sensitivities of higher-order moments of the state of a well-stirred biochemical network. As discussed above, the dynamics of the biochemical network is governed by the Chemical Master Equation, a stochastic discrete model. Parametric sensitivities of higher-order moments of $\mathbf{X}(t, c)$ (e.g. its variance) with respect to model parameters quantify the influence of the parameters on the intrinsic noise in the system. In this work, the approximations of the sensitivities are determined by finite-difference strategies. While finite-difference sensitivity estimators were proposed in the literature [16] [19] [20], these estimators were applied exclusively to compute the local sensitivities of the expected value of the system state, $\partial \mathbb{E}[\mathbf{X}(t, c)] / \partial c$. Still, for many models, the local sensitivity of the average path may not provide sufficient information on the robustness of the system with respect to variations in a parameter. Then, a sensitivity analysis of higher-order moments of the state is required for a more thorough study of the effect of the parameters on the system behaviour. Such an analysis is particularly important for models with a significant level of noise. To estimate local sensitivities of higher-order moments, we need to approximate the expectations $\mathbb{E}[\mathbf{f}(\mathbf{X}(t, \cdot))]$ of the coupled paths corresponding to small variations around the nominal parameter value. This approximation is performed for $f(x) = x = (x_1, \dots, x_N)^T$, $\mathbf{f}(x) = (x_1^2, \dots, x_N^2)^T$, \dots , $\mathbf{f}(x) = (x_1^n, \dots, x_N^n)^T$. For example, let us consider the parametric sensitivity $\frac{\partial}{\partial c} \mathbb{V}[\mathbf{f}(\mathbf{X}(t, c))]$, where $\mathbb{V}[\mathbf{X}(t, c)]$ is the variance of the system state $\mathbf{X}(t, c)$ and c is the parameter of interest. To estimate this sensitivity by finite-difference schemes, we need to approximate the expectations $\mathbb{E}[X_i(t, \cdot)]$ and $\mathbb{E}[X_i^2(t, \cdot)]$ on coupled paths. Indeed, we observe that for each $i = 1, \dots, N$ we have

$$\frac{\partial}{\partial c} \mathbb{V}[X_i(t, c)] = \frac{\partial}{\partial c} \mathbb{E}[X_i^2(t, c)] - 2\mathbb{E}[X_i(t, c)] \frac{\partial}{\partial c} \mathbb{E}[X_i(t, c)]. \tag{3}$$

Adopting a central finite-difference strategy, we derive the following approxima-

tion of the sensitivity of the variance

$$\frac{\partial}{\partial c} \mathbb{V}[X_i(t, c)] \approx \frac{1}{2h} \left\{ \mathbb{V}[X_i(t, c+h)] - \mathbb{V}[X_i(t, c-h)] \right\}, \quad (4)$$

which may be written as

$$\begin{aligned} \frac{\partial}{\partial c} \mathbb{V}[X_i(t, c)] \approx & \frac{1}{2h} \left\{ \mathbb{E} \left[\left(X_i(t, c+h) - \mathbb{E}[X_i(t, c+h)] \right)^2 \right] \right\} \\ & - \frac{1}{2h} \left\{ \mathbb{E} \left[\left(X_i(t, c-h) - \mathbb{E}[X_i(t, c-h)] \right)^2 \right] \right\}, \end{aligned} \quad (5)$$

or as

$$\begin{aligned} \frac{\partial}{\partial c} \mathbb{V}[X_i(t, c)] \approx & \frac{1}{2h} \left\{ \mathbb{E}[X_i^2(t, c+h)] - \mathbb{E}[X_i^2(t, c-h)] \right\} \\ & - \frac{1}{2h} \left\{ \left(\mathbb{E}[X_i(t, c+h)] \right)^2 - \left(\mathbb{E}[X_i(t, c-h)] \right)^2 \right\}. \end{aligned} \quad (6)$$

Likewise, a central finite-difference estimation of the sensitivity of the skewness is

$$\frac{\partial}{\partial c} \mathbb{S}_k[X_i(t, c)] \approx \frac{1}{2h} \left\{ \mathbb{S}_k[X_i(t, c+h)] - \mathbb{S}_k[X_i(t, c-h)] \right\}, \quad (7)$$

where h is a small perturbation of the parameter c . The sensitivity of the mean, $\frac{\partial}{\partial c} \mathbb{E}[X_i(t, c)]$, may also be approximated by central finite-difference techniques.

Employing Monte Carlo methods, we compute pairs of trajectories, $\mathbf{X}^{[k]}(t, c+h)$ and $\mathbf{X}^{[k]}(t, c-h)$, for $1 \leq k \leq L$. In what follows, the coupled trajectories $\mathbf{X}^{[k]}(t, c+h)$ and $\mathbf{X}^{[k]}(t, c-h)$ are generated numerically by the CRN, CRP or CFD strategies. The coupling will reduce the variance of each sensitivity estimator. This will increase the computational efficiency of approximating the parametric sensitivity with a specified accuracy. Note that these trajectories may also be simulated by approximate methods, such as those coupling tau-leaping strategies [16] [22]. With the coupled paths computed above for the parameters $(c+h)$ and $(c-h)$, we obtain the following estimator for the sensitivity of the variance, $\frac{\partial}{\partial c} \mathbb{V}[\mathbf{X}(t, c)]$, has the following components

$$s_{i,L}^{var} = \frac{1}{2hL} \sum_{k=1}^L \left[\left(X_i^{[k]}(t, c+h) - \overline{X_{i,L}}(t, c+h) \right)^2 - \left(X_i^{[k]}(t, c-h) - \overline{X_{i,L}}(t, c-h) \right)^2 \right] \quad (8)$$

where

$$\overline{X_{i,L}}(t, c+h) = \frac{1}{L} \sum_{k=1}^L X_i^{[k]}(t, c+h), \quad \overline{X_{i,L}}(t, c-h) = \frac{1}{L} \sum_{k=1}^L X_i^{[k]}(t, c-h) \quad (9)$$

are the averages over L trajectories, for $1 \leq i \leq N$. Observe also that a Monte Carlo sensitivity estimator of the mean $\mathbb{E}[X_i^2(t)]$, with $1 \leq i \leq N$, by central finite-difference schemes is

$$\mu_{i,L}^{var} = \frac{1}{2hL} \sum_{k=1}^L \left[\left(X_i^{[k]}(t, c+h) \right)^2 - \left(X_i^{[k]}(t, c-h) \right)^2 \right].$$

5. Numerical Experiments

Below, we test the proposed sensitivity estimators of higher-order moments of the system state on three models of biochemical systems. In each case, the Monte Carlo simulations are performed on ensembles of 10^7 trajectories. Parameter sensitivities are approximated by central finite-difference methods. In our computations, the perturbation h used for the finite-difference techniques represents 5% of the value of the parameter of interest c . The numerical results indicate that sensitivity estimations adopting the CFD method are more accurate than those based on the other methods considered, the CRN and the CRP. For the first model, we applied the proposed strategies to approximate the sensitivities of the variance and the skewness, whereas for the remaining models the sensitivity of the variance is estimated. In each case, we show the evolution in time of the parametric sensitivities, which are normalized with respect to the parameter of interest. A similar analysis may be performed for other higher-order moments of the biochemical system state.

5.1. Birth-Death Model

First, we illustrate the accuracy of the sensitivity estimation methods on a simple reaction system, the birth-death model [4]. The biochemical reaction system is described in Table 1, which includes the reaction channels, the propensities and the reaction rate parameter values. The system consists of one species, S , undergoing two reactions. We integrated the system with the initial condition $X(0) = 10$, on the time interval $[0, 115]$.

For this model, we applied the CRN, the CRP and the CFD strategies to generate 10 million coupled trajectories $X^{[k]}(t, c+h)$ and $X^{[k]}(t, c-h)$. These simulations are then used to estimate, by central finite difference schemes, the sensitivities in the kinetic parameters for the first three moments of the system state. Figure 1(a), Figure 1(c) and Figure 1(e) show the mean, variance and skewness of the evolution in time of the amount of S molecules, generated with the SSA, RTC and the NRM techniques. Figure 1(b), Figure 1(d) and Figure 1(f) provide the plots of the central finite-difference approximations of the sensitivities in parameters c_1 and c_2 of the mean, variance and skewness of the molecular S counts, as functions of time. The CRP, the CRN and the CFD algorithms are employed in each case. For this model, the estimates of the sensitivities in the parameter of interest for each moment computed with the CRP, CRN and the CFD schemes agree very well and are accurate. Finally, the dependence on time of the standard deviations of the central finite-difference sensitivity estimators of the means $\mathbb{E}[X(t, c)]$ and $\mathbb{E}[X^2(t, c)]$ are displayed in Figure 2(a)

Table 1. Birth-death model.

	Reaction	Propensity	Parameter value
R_1	$\emptyset \xrightarrow{c_1} S$	$a_1 = c_1$	$c_1 = 2.5$
R_2	$S \xrightarrow{c_2} \emptyset$	$a_2 = c_2 X$	$c_2 = 0.1$

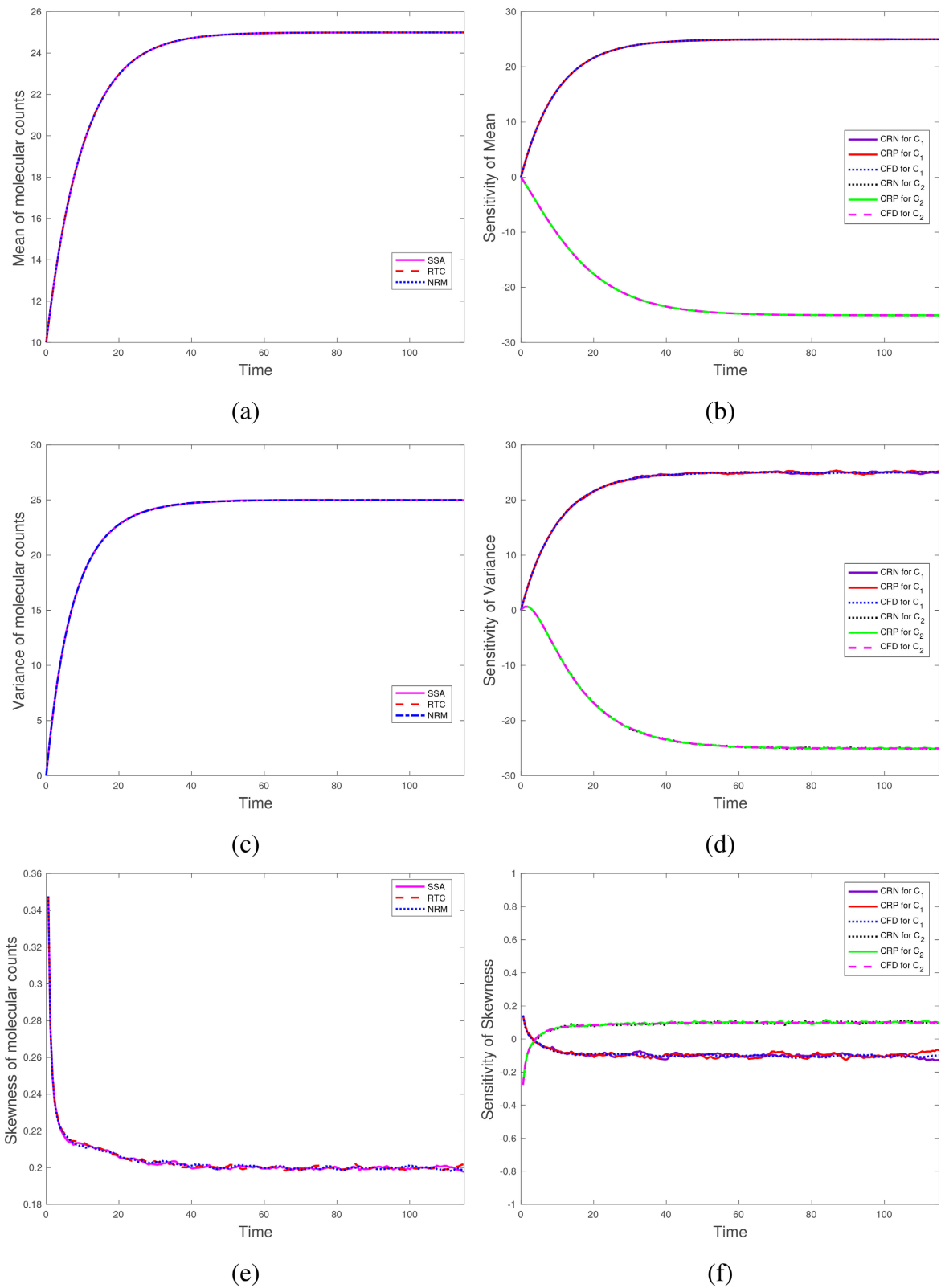


Figure 1. Birth-death model: (a) Mean, (c) variance and (e) skewness of the number of S molecules by the SSA, the RTC and the NRM schemes. (b), (d) and (f) Central finite-difference estimators of the sensitivity to all parameters of the mean, variance and skewness of the S molecular counts, respectively. The 10^7 pairs of trajectories computed on the interval $[0, 115]$ are determined with the CRN, CRP and CFD algorithms.

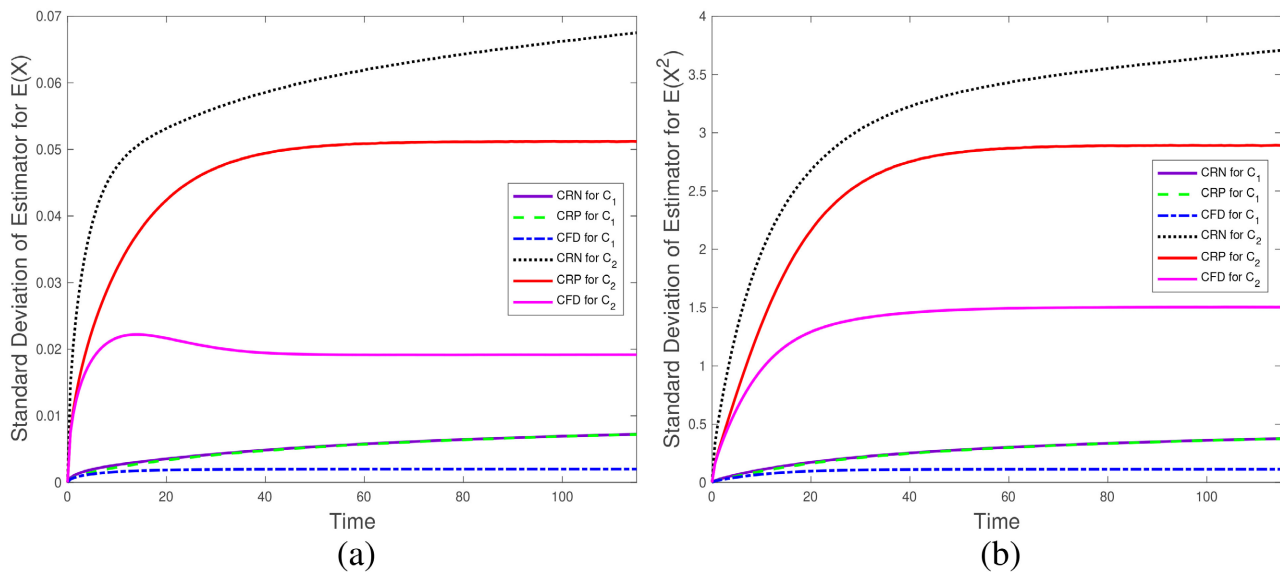


Figure 2. Birth-death model: Standard deviation of the finite-difference sensitivity estimators for (a) $\mathbb{E}[X(t, c)]$ and (b) $\mathbb{E}[X^2(t, c)]$. The CRN, CRP and CFD algorithms are applied to generate 10^7 coupled sample paths on $[0, 115]$.

and **Figure 2(b)**, respectively. The CFD sensitivity estimators have lower standard deviations than their CRN and CRP counterparts, thus are expected to be faster to simulate.

5.2. Infectious Disease Model

The infectious disease model [26] consists of two species which take part in five reactions. The species S_1 represents the infected particles, while the species S_2 becomes infected from S_1 according to the fifth reaction. The first two reaction channels show the death of the species S_1 and S_2 . Finally, the production of these species is considered in the third and fourth reactions. The reaction channels of the infectious disease model are presented in **Table 2**, along with their propensities and the rate parameter values. The model is simulated on the time interval $[0, 2]$ over 10^7 trajectories, for the initial species counts $X(0) = [40, 40]$.

The sensitivity of the variance of the molecular counts with respect to kinetic parameters is estimated by central finite-difference schemes, utilizing the coupling of the CRP, CRN and the CFD methods. The dependence on time of the mean and variance of the number of S_1 and S_2 molecules are depicted in **Figure 3(a)** and **Figure 3(b)**. In each case, the simulations are advanced in time with the SSA, the RTC and the NRM algorithms. Also, the central finite-difference sensitivity estimations of the variance with respect to the parameter c_1 , as functions of time, are shown in **Figure 3(c)** for species S_1 and S_2 . The local sensitivity approximations for the variance using the CRN, CRP and CFD algorithms are in very good agreement. Similar results were obtained for the sensitivity estimations with respect to the other kinetic parameters, but they are not included for brevity.

Table 2. Infectious disease model.

	Reaction	Propensity	Parameter value
R_1	$S_1 \xrightarrow{c_1} \emptyset$	$a_1 = c_1 X_1$	$c_1 = 2$
R_2	$S_2 \xrightarrow{c_2} \emptyset$	$a_2 = c_2 X_2$	$c_2 = 2.5$
R_3	$\emptyset \xrightarrow{c_3} S_1$	$a_3 = c_3$	$c_3 = 25$
R_4	$\emptyset \xrightarrow{c_4} S_2$	$a_4 = c_4$	$c_4 = 75$
R_5	$S_1 + S_2 \xrightarrow{c_5} 2S_1$	$a_5 = c_5 X_1 X_2$	$c_5 = 0.05$

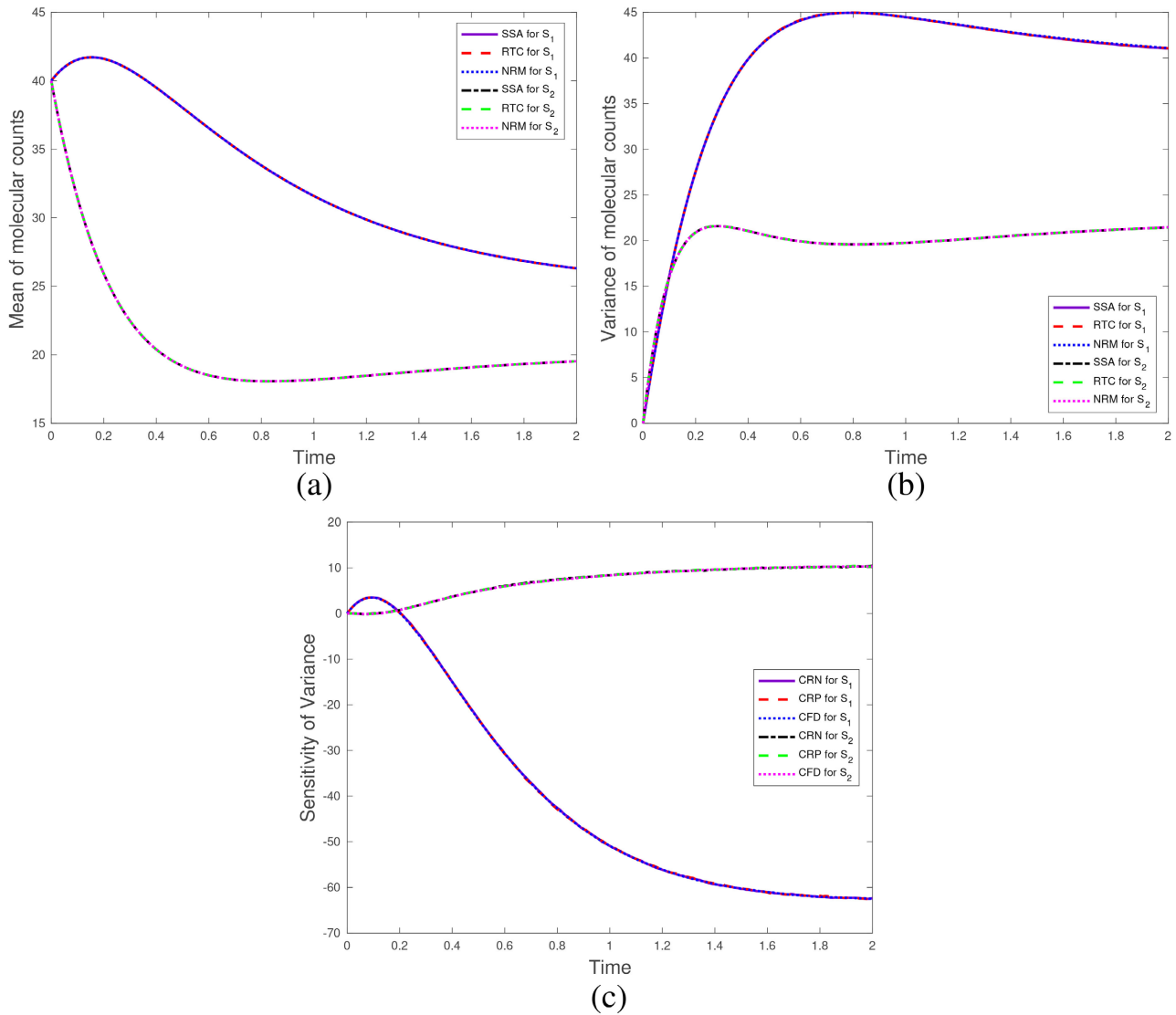


Figure 3. Infectious disease model: (a) Mean and (b) variance of the molecular counts of all species, with the SSA, the RTC and the NRM strategies. (c) Central finite-difference estimators of the sensitivity to the parameter c_1 of the variance of the S_1 and S_2 molecular amounts. As before, the CRN, CRP and CFD strategies were adopted to simulate 10^7 coupled trajectories on the time interval $[0, 2]$.

Lastly, we plotted the timecourse of the standard deviations of the central finite-difference sensitivity estimators with respect to c_1 , of the means $\mathbb{E}[X_i(t, c)]$

in **Figure 4(a)** and $\mathbb{E}[X_i^2(t,c)]$ in **Figure 4(b)**, for $i=1$ or 2. As with the previous model, the sensitivity estimators based on the CFD implementation provide lower standard deviations compared to the CRN and to the CRP ones. As a consequence, the CFD sensitivity estimators are more efficient.

5.3. Michaelis-Menten Model

The third numerical investigation considers the Michaelis-Menten model [11]. This model describes the enzyme-substrate system, which is a crucial mechanism, and consists of a set of three reactions involving four biochemical species. A substrate S_1 and an enzyme S_2 react to create an enzyme-substrate complex S_3 . This complex can unbind to yield the substrate and the enzyme. Furthermore, the complex may also create a product S_4 , a modified form of the substrate, and the enzyme.

Table 3 specifies the reactions, the propensities and the corresponding parameter values for this example. The biochemical system’s volume is $vol = 10^{-15}$ and the Avogadro’s number is $n_A = 6.023 \times 10^{23}$.

The initial state of the Michaelis-Menten model is $\mathbf{X}(0) = [301, 120, 0, 0]^T$. In our investigation, the numerical simulations are carried out on 10 million coupled trajectories with the Common Random Number, the Common Reaction Path and the Coupled Finite Difference methods, on the time-interval $[0, 25]$. As before, the perturbation h corresponds to 5% of the nominal value of the parameter.

Our findings are presented below. In **Figure 5(a)** and **Figure 5(b)**, we compare the evolution of the mean and the variance of the molecular counts of all species, computed using the Stochastic Simulation, the Random Time Change

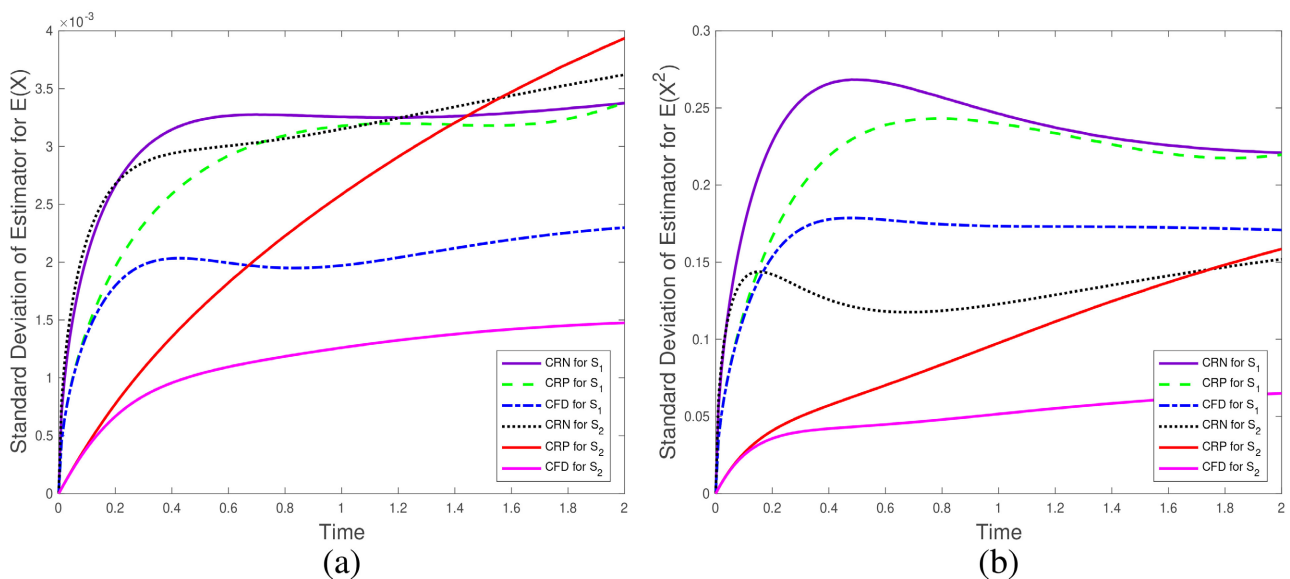


Figure 4. Infectious disease model: Standard deviation of the central finite-difference sensitivity estimators to parameter c_1 of all species for (a) $\mathbb{E}[X_i(t,c)]$ and (b) $\mathbb{E}[X_i^2(t,c)]$ with $i=1$ or 2. In this case, 10^7 pairs of trajectories were computed on the time-interval $[0, 2]$, coupled with the CRP, CRN and CFD methods.

Table 3. Michaelis-Menten model.

	Reaction	Propensity	Parameter value
R_1	$S_1 + S_2 \xrightarrow{c_1} S_3$	$a_1 = c_1 X_1 X_2$	$c_1 = \frac{10^6}{n_A \cdot vol}$
R_2	$S_3 \xrightarrow{c_2} S_1 + S_2$	$a_2 = c_2 X_3$	$c_2 = 10^{-4}$
R_3	$S_3 \xrightarrow{c_3} S_2 + S_4$	$a_3 = c_3 X_3$	$c_3 = 0.1$

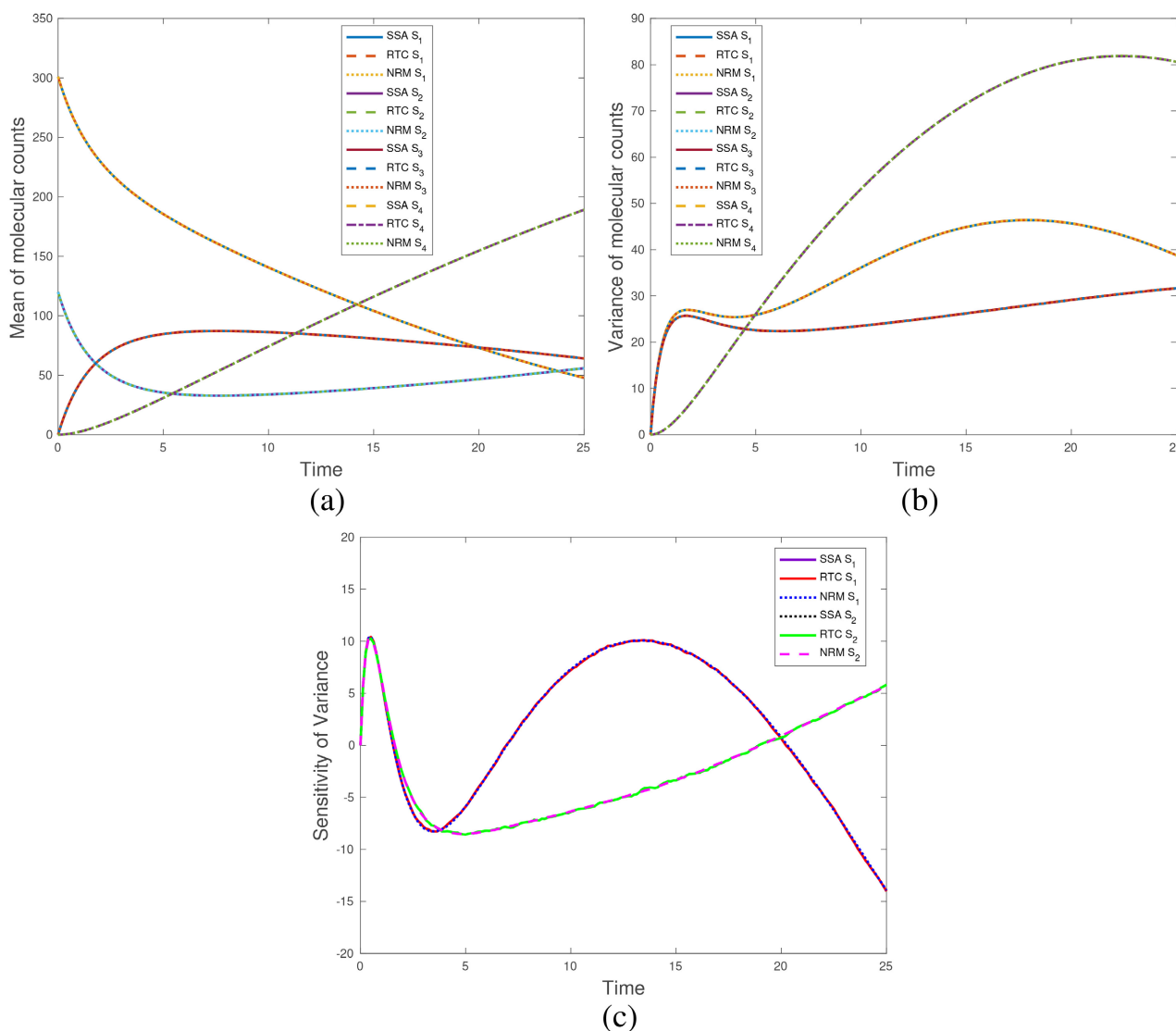


Figure 5. Michaelis-Menten model: (a) Mean and (b) variance of the molecular quantities of all species, with the SSA, the RTC and the NRM algorithms. (c) Central finite-difference sensitivity estimators of the variance of the S_1 and S_2 abundances to the parameter c_1 . The simulation of 10^7 pairs of trajectories on $[0, 25]$ utilized the coupling of the CRN, CRP and CFD strategies.

and the Next Reaction Method algorithms, for the nominal values of the parameters. The behaviour predicted by the three exact Monte Carlo strategies is almost identical. Further, we conducted a sensitivity analysis by employing the finite-difference techniques introduced above, with coupling given by the CRN,

CRP and CFD algorithms. The time dependence of the local sensitivity estimates of the variances for the S_1 and S_2 molecular counts with respect to c_1 is shown in **Figure 5(c)**. The parametric sensitivities are approximated by the central finite-difference methods outlined in Chapter 4, and are normalized with respect to the parameter c_1 . The CRP, CRN and CFD approximations have a comparable accuracy in this case. Moreover, equivalent outcomes are found for the sensitivity estimates of the second moment of the S_3 and S_4 molecular numbers (omitted for conciseness). To conclude, the numerical results reveal the excellent agreement of the proposed finite-difference sensitivity estimators of the variance, for all species.

6. Conclusions

In the present work, we present some numerical strategies for estimating parametric sensitivities for stochastic discrete models of homogeneous biochemical systems. Specifically, we introduce computational tools to approximate local sensitivities of higher-order moments, such as the variance, of the system state. The model studied here is the Chemical Master Equation, which has been successfully applied to a large class of biochemical reaction networks in a cell. This discrete and stochastic model is quite challenging to simulate and analyze for many realistic biochemical networks. The proposed methods are valuable for investigating the robustness properties of higher-order moments of a biochemical system state, while previous works focused on the sensitivity of its expected value. Such tools are particularly important for models with moderate to large levels of noise, or for which the random fluctuations due to molecular interactions direct the system behaviour. For such systems, investigating the mean trajectory is often not sufficient for an in-depth analysis of the features of the stochastic model. Parametric sensitivities of higher-order moments quantify how parameters influence the intrinsic noise and give insight into the dynamics of the biochemical system.

The methods we introduced utilize central finite-difference schemes and Monte Carlo strategies for estimating the parametric sensitivities. More precisely, we considered exact Monte Carlo techniques and coupled the pairs of trajectories through existing strategies, namely the Common Random Number, Common Reaction Path and Coupled Finite Differences schemes. The coupling reduced the variance of the sensitivity estimators, leading to a decreased computational cost. Hence, the proposed techniques provide an *accurate* and *effective* sensitivity estimation of the moments of interest, for a stochastic process governed by the Chemical Master Equation. These methods were successfully tested on different models arising in applications and shown to be accurate and in excellent agreement with each other. It is expected that these strategies are far more efficient than the sensitivity estimators of higher-order moments which are not based on coupling. While the Common Random Number, Common Reaction Path and Coupled Finite Difference algorithms were applied in this work, it is worth ob-

serving that the approach proposed above may be modified to adopt other finite-difference sensitivity methods to investigate higher-order moments, such as those based on tau-leaping schemes.

Acknowledgements

This work was partially supported by a grant from the Natural Sciences and Engineering Research Council of Canada (NSERC).

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Elowitz, M.B., *et al.* (2002) Stochastic Gene Expression in a Single Cell. *Science*, **297**, 1183-1186. <https://doi.org/10.1126/science.1070919>
- [2] Federoff, N. and Fontana, W. (2002) Small Numbers of Big Molecules. *Science*, **297**, 1129-1131. <https://doi.org/10.1126/science.1075988>
- [3] Raser, J.M. and O'Shea, E.K. (2004) Control of Stochasticity in Eukaryotic Gene Expression. *Science*, **304**, 1811-1814. <https://doi.org/10.1126/science.1098641>
- [4] Wilkinson, D.J. (2006) Stochastic Modelling for Systems Biology. Chapman & Hall/CRC, Boca Raton. <https://doi.org/10.1201/9781420010664>
- [5] Gillespie, D.T. (1992) A Rigorous Derivation of the Chemical Master Equation. *Physica A*, **188**, 402-425. [https://doi.org/10.1016/0378-4371\(92\)90283-V](https://doi.org/10.1016/0378-4371(92)90283-V)
- [6] Gillespie, D.T. (1976) A General Method for Numerically Simulating the Stochastic Time Evolution of Coupled Chemical Reactions. *Journal of Computational Physics*, **22**, 403-434. [https://doi.org/10.1016/0021-9991\(76\)90041-3](https://doi.org/10.1016/0021-9991(76)90041-3)
- [7] Gillespie, D.T. (1977) Exact Stochastic Simulation of Coupled Chemical Reactions. *Journal of Computational Physics*, **81**, 2340-2361. <https://doi.org/10.1021/j100540a008>
- [8] Kurtz, T.G. (1982) Representation and Approximation of Counting Processes. In: *Advances in Filtering and Optimal Stochastic Control*, Lecture Notes in Control and Information Sciences, Vol. 42, Springer, Berlin, 177-191. <https://doi.org/10.1007/BFb0004537>
- [9] Gillespie, D.T. (2001) Approximate Accelerated Stochastic Simulation of Chemically Reacting Systems. *Journal of Computational Physics*, **115**, 1716-1733. <https://doi.org/10.1063/1.1378322>
- [10] Padgett, J.M.A. and Ilie, S. (2016) An Adaptive Tau-Leaping Method for Stochastic Simulations of Reaction-Diffusion Systems. *AIP Advances*, **6**, Article ID: 035217. <https://doi.org/10.1063/1.4944952>
- [11] Higham, D.J. (2008) Modeling and Simulating Chemical Reactions. *SIAM Review*, **50**, 347-368. <https://doi.org/10.1137/060666457>
- [12] Li, T. (2007) Analysis of Explicit Tau-Leaping Schemes for Simulating Chemically Reacting Systems. *Multiscale Modeling and Simulation*, **6**, 417-436. <https://doi.org/10.1137/06066792X>
- [13] Simoni, G., Reali, F., Priami, C. and Marchetti, L. (2019) Stochastic Simulation Algorithms for Computational Systems Biology: Exact, Approximate, and Hybrid Me-

- thods. *Wiley Interdisciplinary Reviews. Systems Biology and Medicine*, **11**, e1459. <https://doi.org/10.1002/wsbm.1459>
- [14] Sekiguchi, T., Hamada, H. and Okamoto, M. (2019) Inference of General Mass Action-Based State Equations for Oscillatory Biochemical Reaction Systems Using k-Step Genetic Programming. *Applied Mathematics*, **10**, 627-645. <https://doi.org/10.4236/am.2019.108045>
- [15] Varma, A., Morbidelli, M. and Wu, H. (1999) Parametric Sensitivity in Chemical Systems. Cambridge University Press, Cambridge. <https://doi.org/10.1017/CBO9780511721779>
- [16] Morshed, M., Ingalls, B. and Ilie, S. (2017) An Efficient Finite-Difference Strategy for Sensitivity Analysis of Stochastic Models of Biochemical Systems. *Biosystems*, **151**, 43-52. <https://doi.org/10.1016/j.biosystems.2016.11.006>
- [17] Gholami, S. and Ilie, S. (2021) Reducing Stochastic Discrete Models of Biochemical Networks. *Applied Mathematics*, **12**, 449-469. <https://doi.org/10.4236/am.2021.125031>
- [18] Ilie, S. and Gholami, S. (2013) Simplifying Stochastic Mathematical Models of Biochemical Systems. *Applied Mathematics*, **4**, 248-256. <https://doi.org/10.4236/am.2013.41A038>
- [19] Rathinam, M., Sheppard, P.W. and Khammash, M. (2010) Efficient Computation of Parameter Sensitivities of Discrete Stochastic Chemical Reaction Networks. *The Journal of Chemical Physics*, **132**, Article ID: 034103. <https://doi.org/10.1063/1.3280166>
- [20] Anderson, D.F. (2012) An Efficient Finite Difference Method for Parameter Sensitivities of Continuous Time Markov Chains. *SIAM Journal on Numerical Analysis*, **50**, 2237-2258. <https://doi.org/10.1137/110849079>
- [21] Gibson, M.A. and Bruck, J. (2000) Exact Stochastic Simulation of Chemical Systems with Many Species and Many Channels. *Journal of Computational Physics*, **105**, 1876-1889. <https://doi.org/10.1021/jp993732q>
- [22] Morshed, M., Ingalls, B. and Ilie, S. (2018) An Effective Implicit Finite-Difference Method for Sensitivity Analysis of Stiff Stochastic Discrete Biochemical Systems. *IET Systems Biology*, **12**, 123-130. <https://doi.org/10.1049/iet-syb.2017.0048>
- [23] Thanh, V.H. (2019) A Critical Comparison of Rejection-Based Algorithms for Simulation of Large Biochemical Reaction Networks. *Bulletin of Mathematical Biology*, **81**, 3053-3073. <https://doi.org/10.1007/s11538-018-0462-y>
- [24] Sayyidmousavi, A. and Ilie, S. (2017) An Efficient Hybrid Method for Stochastic Reaction-Diffusion Biochemical Systems with Delay. *AIP Advances*, **7**, Article ID: 125305. <https://doi.org/10.1063/1.5001760>
- [25] Ethier, S.N. and Kurtz, T.G. (1986) Markov Processes: Characterization and Convergence. Wiley, New York. <https://doi.org/10.1002/9780470316658>
- [26] Jahnke, T. (2011) On Reduced Models for the Chemical Master Equation. *Multiscale Modeling and Simulation*, **9**, 1646-1676. <https://doi.org/10.1137/110821500>