

Comparison of finite difference based methods to obtain sensitivities of stochastic chemical kinetic models

Rishi Srivastava, David F. Anderson, and James B. Rawlings

Citation: *J. Chem. Phys.* **138**, 074110 (2013); doi: 10.1063/1.4790650

View online: <http://dx.doi.org/10.1063/1.4790650>

View Table of Contents: <http://jcp.aip.org/resource/1/JCPSA6/v138/i7>

Published by the [American Institute of Physics](#).

Additional information on *J. Chem. Phys.*

Journal Homepage: <http://jcp.aip.org/>

Journal Information: http://jcp.aip.org/about/about_the_journal

Top downloads: http://jcp.aip.org/features/most_downloaded

Information for Authors: <http://jcp.aip.org/authors>

ADVERTISEMENT

Instruments for advanced science

Gas Analysis



- dynamic measurement of reaction gas streams
- catalysis and thermal analysis
- molecular beam studies
- dissolved species probes
- fermentation, environmental and ecological studies

Surface Science



- UHV TPD
- SIMS
- end point detection in ion beam etch
- elemental imaging - surface mapping

Plasma Diagnostics



- plasma source characterization
- etch and deposition process
- reaction kinetic studies
- analysis of neutral and radical species

Vacuum Analysis



- partial pressure measurement and control of process gases
- reactive sputter process control
- vacuum diagnostics
- vacuum coating process monitoring

contact Hiden Analytical for further details

HIDEN
ANALYTICAL

info@hideninc.com
www.HidenAnalytical.com

CLICK to view our product catalogue



Comparison of finite difference based methods to obtain sensitivities of stochastic chemical kinetic models

Rishi Srivastava,^{1,a)} David F. Anderson,^{2,b)} and James B. Rawlings^{1,c)}

¹Department of Chemical and Biological Engineering, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA

²Department of Mathematics, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA

(Received 24 September 2012; accepted 21 January 2013; published online 20 February 2013)

Sensitivity analysis is a powerful tool in determining parameters to which the system output is most responsive, in assessing robustness of the system to extreme circumstances or unusual environmental conditions, in identifying rate limiting pathways as a candidate for drug delivery, and in parameter estimation for calculating the Hessian of the objective function. Anderson [SIAM J. Numer. Anal. **50**, 2237 (2012)] shows the advantages of the newly developed coupled finite difference (CFD) estimator over the common reaction path (CRP) [M. Rathinam, P. W. Sheppard, and M. Khammash, J. Chem. Phys. **132**, 034103 (2010)] estimator. In this paper, we demonstrate the superiority of the CFD estimator over the common random number (CRN) estimator in a number of scenarios not considered previously in the literature, including the sensitivity of a negative log likelihood function for parameter estimation, the sensitivity of being in a rare state, and a sensitivity with fast fluctuating species. In all examples considered, the superiority of CFD over CRN is demonstrated. We also provide an example in which the CRN method is superior to the CRP method, something not previously observed in the literature. These examples, along with Anderson's results, lead to the conclusion that CFD is currently the best estimator in the class of finite difference estimators of stochastic chemical kinetic models. © 2013 American Institute of Physics. [<http://dx.doi.org/10.1063/1.4790650>]

I. INTRODUCTION

Recent years have seen an increasing popularity of stochastic chemical kinetic models due to their role in describing and explaining critical biological phenomena.³⁻⁹ One useful tool for understanding these models is the chemical master equation, which describes the evolution of the probability density of the system. The solution of the master equation is computationally tractable only for simple systems. Rather, approximation techniques such as finite state projection,¹⁰ that operates on a reduced state space, or the stochastic simulation algorithm (SSA),^{11,12} that generates exact sample paths, are employed to reconstruct a system's probability distribution and statistics (usually the mean and variance). Applying these techniques to solve models of biological processes leads to significant improvements in our understanding of intrinsic noise and its effect on cellular behavior.

These stochastic chemical kinetic models depend on parameters whose values are often unknown and can change due to changes in the environment. Sensitivities quantify the dependence of the system's output to changes in the model parameters. Sensitivity analysis is useful in determining parameters to which the system output is most responsive, in assessing robustness of the system to extreme circumstances or unusual environmental conditions, and in identifying rate limiting pathways as a candidate for drug delivery. However, one

of the most important applications of sensitivities is in parameter estimation. Sensitivities provide a way to approximate the Hessian of the objective function through the Gauss-Newton approximation (Ref. 13, p. 535).

A popular unbiased method of sensitivity estimation is the likelihood ratio gradient method.¹⁴⁻¹⁶ The unbiasedness of the likelihood ratio gradient method comes at the cost of a high variance of the estimator if there are several reaction events in the estimation of the output of interest. The convergence rate, which is a measure of the rate at which the mean squared error of the estimator converges to zero, of this estimator is $O(N^{-1/2})$, in which N is the number of estimator simulations. Komorowski *et al.*¹⁷ use a linear noise approximation of stochastic chemical kinetic models for sensitivity analysis. However, use of linear noise approximation limits their analysis to only stochastic differential equation models that incorporate Brownian motions. Gunawan *et al.*¹⁸ compare the sensitivity of the mean with the sensitivity of the entire distribution. They explain why the sensitivity of the mean can be inadequate in determining the sensitivity of stochastic chemical kinetic models.

Despite being easier to implement and intuitive to understand, finite difference based methods produce biased sensitivity estimates. However, implemented with consideration of the trade-off between the statistical error of the estimator and its bias, finite difference based methods can have a convergence rate close to the best possible convergence rate of $O(N^{-1/2})$.¹⁹ McGill *et al.*²⁰ compare the applicability of likelihood ratio gradient and finite difference based methods. They discuss situations where one method performs better than the

^{a)}Electronic mail: rishisrivastava1984@gmail.com.

^{b)}Electronic mail: anderson@math.wisc.edu.

^{c)}Electronic mail: rawlings@engr.wisc.edu.

other. Drew *et al.*²¹ demonstrate usefulness of sensitivity analysis on Monte Carlo simulations of copper electrodeposition.

Several different estimators using finite difference have been proposed.^{1,2,19} Anderson¹ proposed a new estimator, coupled finite difference (CFD), using a single Markov chain for the nominal and perturbed processes. The CFD estimator incorporates a tight coupling between the nominal and perturbed processes, thereby producing a significant reduction in estimator variance.¹

In this paper, we show the superiority of CFD over CRN in the estimation of sensitivities. We do not discuss the independent random number² estimator, also known as the Crude Monte Carlo¹ estimator, because either estimator, CRN or CFD, usually has several orders of magnitude smaller variance than this estimator. We calculate sensitivity estimates of five different quantities of interest. In example one, the quantity of interest is the expected value of a species. Example two looks at the likelihood of experimental data. Example three looks at the probability of a rare state. Example four looks at the expected value of a fast fluctuating species. Example five looks at the expected value of a gene product in a model of a genetic toggle switch. In this example, the CRN method is shown to be superior to the CRP method, something not previously observed in the literature.

This paper is arranged as follows. Section II defines the estimators that are used in the subsequent examples. Section III shows the results we obtain from the five examples. Finally, Sec. IV discusses the conclusions of this paper and summarizes the contributions.

II. THE ESTIMATORS

Common random number (CRN; Refs. 2 and 19): A single simulation of the CRN estimator uses two coupled SSA simulations: the first coupled SSA simulation uses the rate parameter k and randomness ω and the second one uses the perturbed rate parameter $k + \epsilon$ and the same randomness ω . By the same randomness ω , we mean that both first and second coupled simulations use the same seed of the pseudo-random number generator in an implementation of Gillespie's direct method.¹¹

Coupled Finite difference (CFD; Ref. 1): A single simulation of the CFD estimator simulates a Markov chain with an enlarged state space. The marginal processes of this Markov chain yield the realizations of the coupled processes with rate parameters k and $k + \epsilon$. The new Markov chain is constructed in such a way that there is a tight coupling between the marginal processes, yielding a low variance for the estimator. See Anderson¹ for the complete description.

Common Reaction Path (CRP; Ref. 2): A single simulation of the CRP estimator uses two SSA simulations coupled through random time change representation. Thus, it is CRN for the next reaction method.

Finite difference approximations of the sensitivity can be obtained from any of the above methods by sample averaging appropriate differences of the realizations.

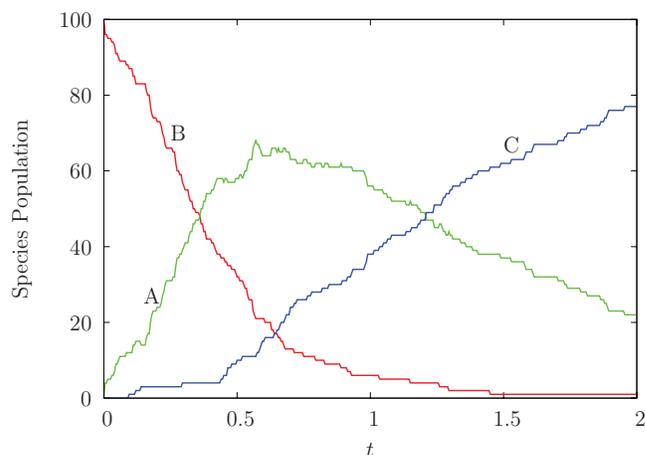


FIG. 1. A typical simulation of the network involving reactions (R1) and (R2).

III. EXAMPLES

A. Sensitivity of an expected value of a population of a species

Consider the following simple reaction network consisting of two reactions



Figure 1 shows a typical SSA simulation of the network involving reactions (R1) and (R2).

We wish to estimate the sensitivity of the expected value of B with respect to the rate constant k_1 ,

$$s(t; k_1) = \frac{d\mathbb{E}B(t; k_1)}{dk_1}, \quad (1)$$

where $B(t; k_1)$ represents the number of B molecules at time t with a choice of rate constant of k_1 . The forward finite difference approximation to Eq. (1) is

$$s(t; k_1) \approx \frac{\mathbb{E}B(t; k_1 + \epsilon) - \mathbb{E}B(t; k_1)}{\epsilon}, \quad (2)$$

which has a bias of $O(\epsilon)$. That is,

$$s(t; k_1) = \frac{\mathbb{E}B(t; k_1 + \epsilon) - \mathbb{E}B(t; k_1)}{\epsilon} + O(\epsilon).$$

Centered differences produce a bias of $O(\epsilon^2)$. Throughout the paper we use the forward finite difference to approximate the sensitivity. We denote an estimator of the right hand side of (2) using either CRN or CFD as \hat{s}_{est} in which $\text{est} \in \{\text{CRN}, \text{CFD}\}$. Let $B_i^{\text{est}}(t; k_1)$ and $B_i^{\text{est}}(t; k_1 + \epsilon)$ denote the population of B obtained through the i th simulation of estimator est . Then the estimator \hat{s}_{est} for $s(t; k_1)$ of (1) is defined as

$$\hat{s}_{\text{est}} = \frac{1}{N} \sum_{i=1}^N \frac{B_i^{\text{est}}(t; k_1 + \epsilon) - B_i^{\text{est}}(t; k_1)}{\epsilon}. \quad (3)$$

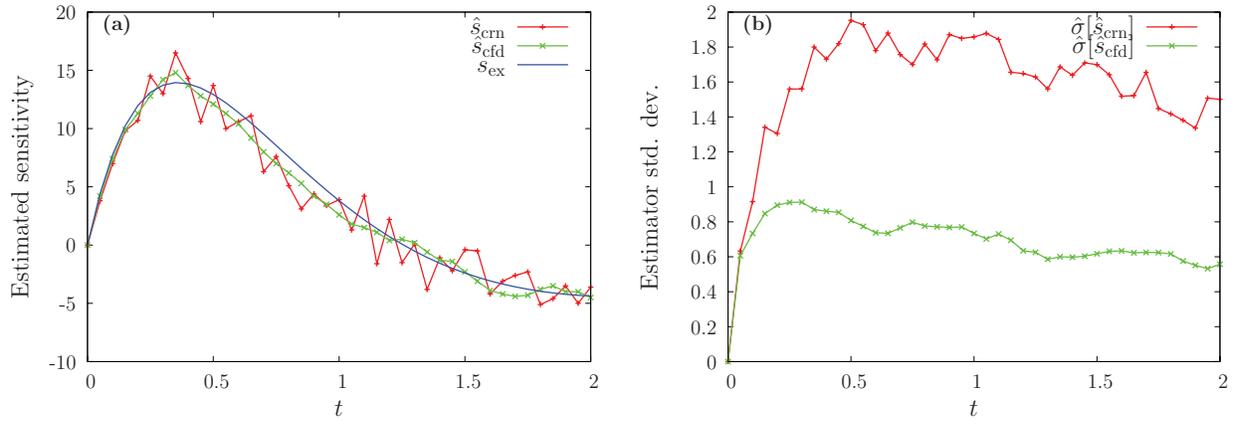


FIG. 2. Comparison of CRN and CFD estimators for the model (R1) and (R2): (a) Estimated and analytical sensitivities. (b) Sample standard deviation of the two estimators.

The sample standard deviations ($\hat{\sigma}[\hat{s}_{\text{est}}]$) of the estimator of Eq. (3) is given by

$$\hat{\sigma}[\hat{s}_{\text{est}}] = \left(\frac{1}{N(N-1)\epsilon^2} \sum_{i=1}^N [\{B_i^{\text{est}}(t; k_1 + \epsilon) - B_i^{\text{est}}(t; k_1)\} - \Delta \hat{B}^{\text{est}}]^2 \right)^{1/2} \quad (4)$$

in which

$$\Delta \hat{B}^{\text{est}} = \frac{1}{N} \sum_{i=1}^N B_i^{\text{est}}(t; k_1 + \epsilon) - B_i^{\text{est}}(t; k_1).$$

Because the model is linear, the exact expected value of B and the exact sensitivity of the expected value of B can be calculated and are given by

$$\mathbb{E}B(t; k_1) = n_{B_0}e^{-k_2t} + n_{A_0} \frac{k_1}{k_2 - k_1} (e^{-k_1t} - e^{-k_2t}), \quad (5)$$

$$s_{\text{ex}}(t; k_1) = \frac{d\mathbb{E}B}{dk_1} = \frac{n_{A_0}}{(k_1 - k_2)^2} [k_2e^{-k_1t} + k_1(k_1 - k_2)te^{-k_1t} - k_2e^{-k_2t}], \quad (6)$$

in which n_{A_0} , n_{B_0} are the initial values of A and B , respectively.

Figure 2 compares the performance of the CRN and CFD estimators, and Table I lists the parameters used to generate Figure 2. Figure 2(a) shows a comparison of the sensitivity estimates obtained from the CRN and CFD estimators. We define root mean squared error of the estimator as

$$e_{\text{est}} = \left[\frac{1}{n_d} \sum_{i=1}^{n_d} (\hat{s}_{\text{est}}(t_i, k_1) - s_{\text{ex}}(t_i, k_1))^2 \right]^{1/2} \quad (7)$$

TABLE I. Parameter values for Sec. III A.

Parameter	n_{A_0}	n_{B_0}	n_{C_0}	k_1	k_2	ϵ	N
Value	100	0	0	2.	1.	0.1	100

in which $n_d = 41$ is the total number of time points at which we calculate the sensitivity \hat{s}_{est} , $0 \leq t_i \leq 2.0$, and $t_{i+1} - t_i = 0.05$. Root mean squared errors calculated from the data of Figure 2(a) give $\frac{e_{\text{crn}}}{e_{\text{cfd}}} = 4$. A value greater than one for this ratio demonstrates that on average across all the time points considered the CFD estimator tracks the exact sensitivity better than the CRN estimator. Figure 2(b) quantifies the efficiency of the two estimators, by comparing their standard deviations. We can see that starting from $t = 0.3$, the CFD estimator has half the standard deviation of the CRN estimator. Lower standard deviation of the CFD estimator compared to the CRN estimator points to its higher efficiency.

B. Sensitivity of negative log likelihood function

Consider reactions (R1) and (R2) again. Experimental data $\mathbf{y} = (B_{t_1}, B_{t_2}, \dots, B_{t_n})^T$ from a single experiment, shown in Figure 3, are given as a time series of B .

We assume that the rate constant k_2 is known. It can be shown²² that an estimate of the likelihood of the experimental

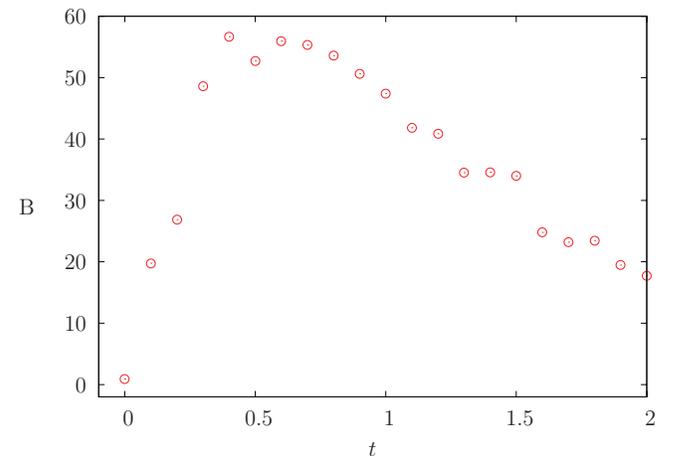


FIG. 3. Experimental data for Sec. III B generated from the model with parameter choices given in Table II.

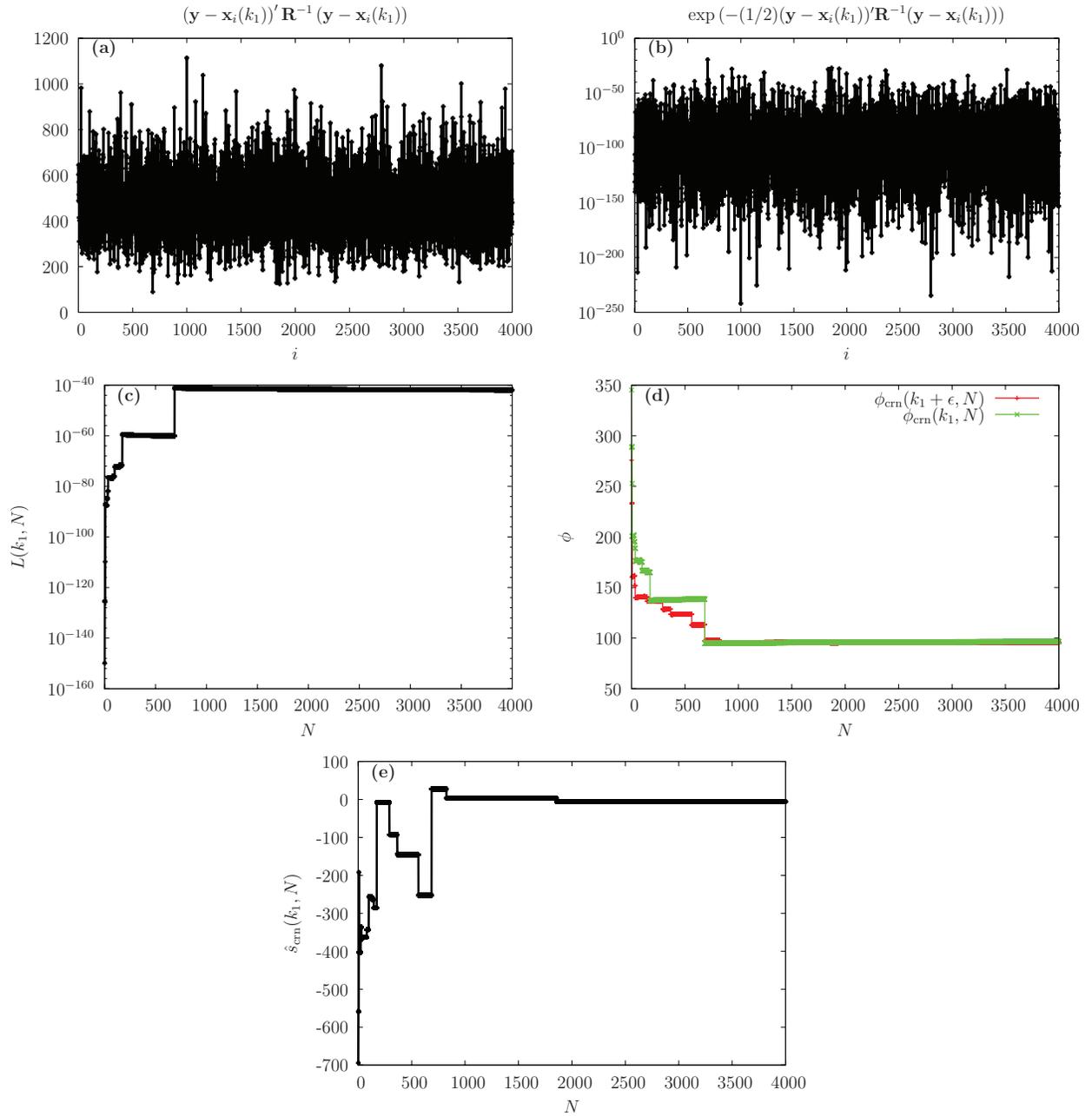


FIG. 4. Convergence of sensitivity estimate for Sec. III B from the CRN estimator: (a) Plot of the quadratic form as a function of simulation number for the unperturbed process. (b) Plot of exponential of quadratic form as a function of simulation number for the unperturbed process. (c) Likelihood as a function of total number of simulations for the unperturbed process. (d) Negative log likelihood as a function of total number of simulations. (e) Estimated sensitivity from the CRN estimator as a function of total number of simulations.

data \mathbf{y} under certain reasonable assumptions is given by

$$L(k_1, N) = \frac{1}{N(2\pi)^{n_d/2} |\mathbf{R}|^{1/2}} \sum_{i=1}^N e^{-(1/2)(\mathbf{y} - \mathbf{x}_i(k_1))^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{x}_i(k_1))} \quad (8)$$

in which \mathbf{R} is a known positive definite matrix, n_d is the number of elements in the experimental data vector \mathbf{y} , and $\mathbf{x}_i(k_1) = (B_{t_1}, B_{t_2}, \dots, B_{t_n})^T$ is the time series of the population of B , obtained by the i th SSA simulation using rate constant value k_1 for reaction (R1). Note that as the number of samples N goes to infinity, the likelihood estimate $L(k_1, N)$ from (8) approaches the true likelihood of the experimental

data. The estimate of the negative log likelihood is defined as

$$\phi(k_1, N) = -\log L(k_1, N) = -\log \left[\frac{1}{N(2\pi)^{n_d/2} |\mathbf{R}|^{1/2}} \times \sum_{i=1}^N e^{-(1/2)(\mathbf{y} - \mathbf{x}_i(k_1))^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{x}_i(k_1))} \right]. \quad (9)$$

To find the parameters that describe the experimental data, we need to minimize the estimate of the negative log likelihood function given in Eq. (9). Sensitivities can be used to obtain gradients required in any gradient based optimization

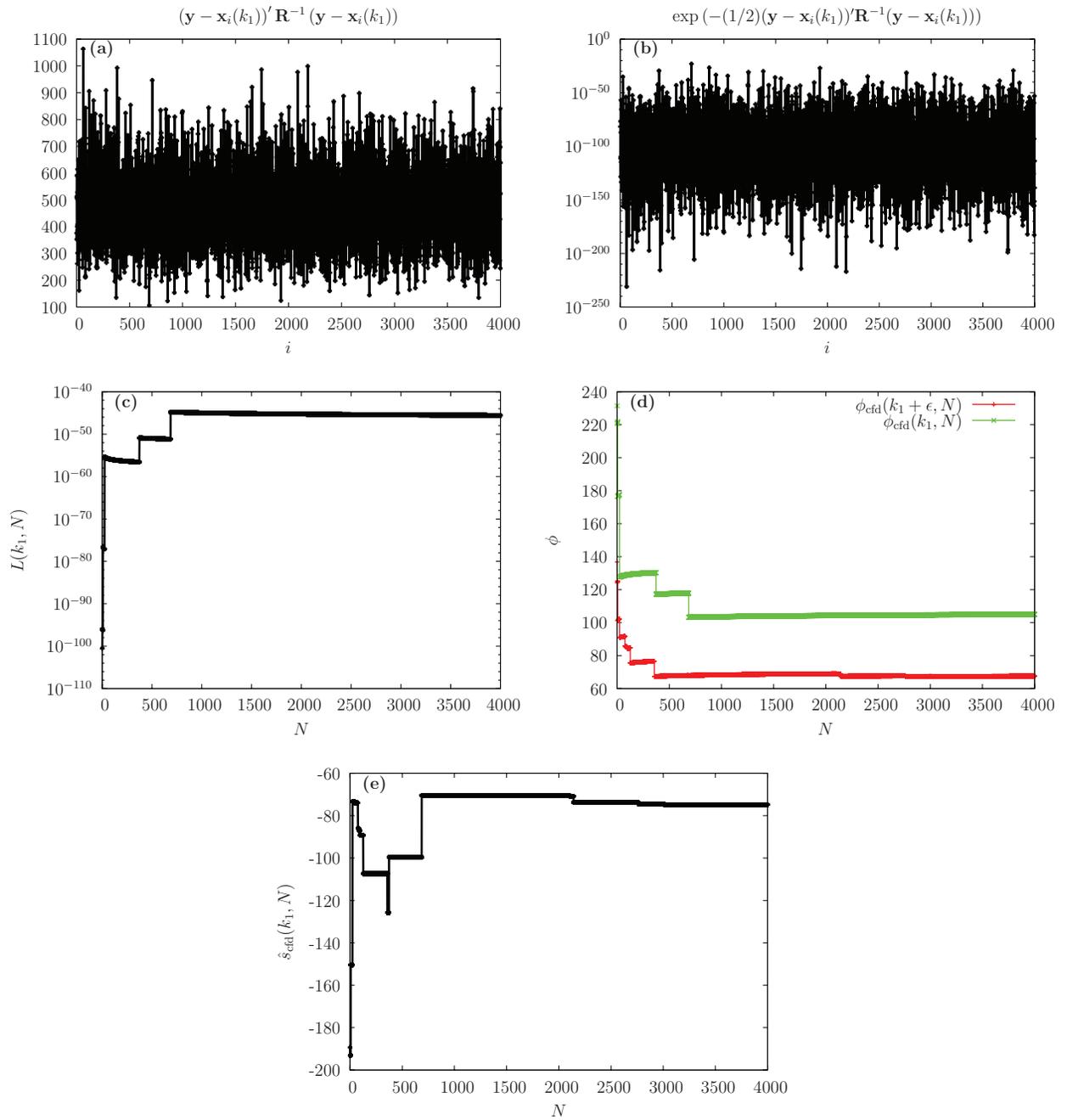


FIG. 5. Convergence of sensitivity estimate for Sec. III B from the CFD estimator: (a) Plot of quadratic form as a function of simulation number for the unperturbed process. (b) Plot of exponential of quadratic form as a function of simulation number for the unperturbed process. (c) Likelihood as a function of total number of simulations for the unperturbed process. (d) Negative log likelihood as a function of total number of simulations. (e) Estimated sensitivity from the CFD estimator as a function of total number of simulations.

algorithm. Here, we are interested in the sensitivity

$$s(k_1, N) = \frac{d\phi(k_1, N)}{dk_1} \quad (10)$$

of the estimated negative log likelihood function and the convergence of this sensitivity with the number of samples, N . The forward finite difference approximation of (10) is given by

$$s(k_1, N) \approx \frac{\phi(k_1 + \epsilon, N) - \phi(k_1, N)}{\epsilon}. \quad (11)$$

We write estimator $\text{est} \in \{\text{CRN}, \text{CFD}\}$ of the sensitivity $s(k_1, N)$ of Eq. (10) as

$$\hat{s}_{\text{est}}(k_1, N) = \frac{\phi_{\text{est}}(k_1 + \epsilon, N) - \phi_{\text{est}}(k_1, N)}{\epsilon} \quad (12)$$

in which $\phi_{\text{est}}(k_1, N)$ is the estimate of the negative log likelihood obtained from Eq. (9) using the estimator $\text{est} \in \{\text{CRN}, \text{CFD}\}$.

Figure 4 shows the steps in obtaining the sensitivity of the negative log likelihood function using the CRN estimator. Table II contains the parameters used in this example. Figure 4(a) shows the variation of the quadratic form

TABLE II. Parameter value for Sec. III B.

Parameter	n_{A_0}	n_{B_0}	n_{C_0}	k_1	k_2	ϵ	N	n_d	\mathbf{R}
Value	100	0	0	1.	1.	0.1	4000	21	$I_{n_d \times n_d}$

$(\mathbf{y} - \mathbf{x}_i(k_1))' \mathbf{R}^{-1} (\mathbf{y} - \mathbf{x}_i(k_1))$ as a function of individual SSA simulation number i . Figure 4(a) reveals the wide variation in the value of the quadratic form for different individual SSA simulations. Figure 4(b) is a plot of $e^{-(1/2)(\mathbf{y} - \mathbf{x}_i(k_1))' \mathbf{R}^{-1} (\mathbf{y} - \mathbf{x}_i(k_1))}$ as a function of the individual SSA simulation number i . The wide variation in $(\mathbf{y} - \mathbf{x}_i(k_1))' \mathbf{R}^{-1} (\mathbf{y} - \mathbf{x}_i(k_1))$ of Figure 4(a) leads to even wider variation in the exponential, as depicted in Figure 4(b). Figure 4(c) depicts convergence of $L(k_1, N)$ with N . $L(k_1, N)$ in Figure 4(c) is given by Eq. (8). Large variation in $e^{-(1/2)(\mathbf{y} - \mathbf{x}_i(k_1))' \mathbf{R}^{-1} (\mathbf{y} - \mathbf{x}_i(k_1))}$ as depicted in Figure 4(b) explains the sharp jumps in $L(k_1, N)$ which occur whenever the last SSA simulation $i = n$ dominates all the previous $1 \leq i \leq n - 1$ simulations. Figure 4(d) shows the convergence of the nominal and perturbed negative log likelihoods: $\phi_{\text{cm}}(k_1, N)$ and $\phi_{\text{cm}}(k_1 + \epsilon, N)$. The sharp jumps in $\phi_{\text{cm}}(k_1, N)$ occur at the same n values as the jumps in $L(k_1, N)$ in Figure 4(c). Finally, Figure 4(e) shows the convergence of the sensitivity of the negative log likelihood using the CRN estimator, $\hat{s}_{\text{cm}}(k_1, N)$, as a function of N .

Next, we change focus from the CRN estimator to the CFD estimator. In Figures 5(a)–5(e), we plot the analogous results to Figures 4(a)–4(e). Finally, Figure 6 compares convergence of CRN estimator, $\hat{s}_{\text{cm}}(k_1, N)$, and CFD estimator, $\hat{s}_{\text{cfd}}(k_1, N)$. We see that the convergence of the CFD estimator to the estimated true value is faster than the CRN estimator. An estimated true value is obtained by performing 50 000 simulations of the CFD estimator. The quicker convergence property makes the CFD estimator a natural choice for the estimator of sensitivity of negative log likelihood. One point to note is that, as $N \rightarrow \infty$, the final converged value for both CFD and CRN estimators is going to be the same, because both the estimators have the same bias; the CRN estimator shown in Figure 6 simply has not converged by $N = 4000$.

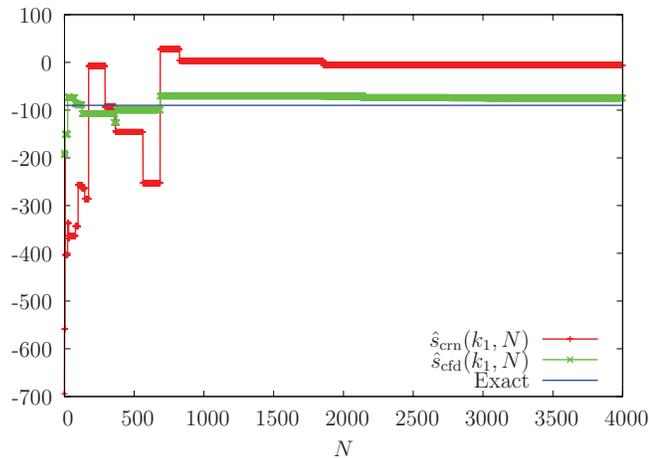


FIG. 6. Convergence of the estimated sensitivities of (11) from CRN and CFD estimators. Compared to the CRN estimator the CFD estimator shows quicker convergence to the estimated true value.

Regulatory Protein

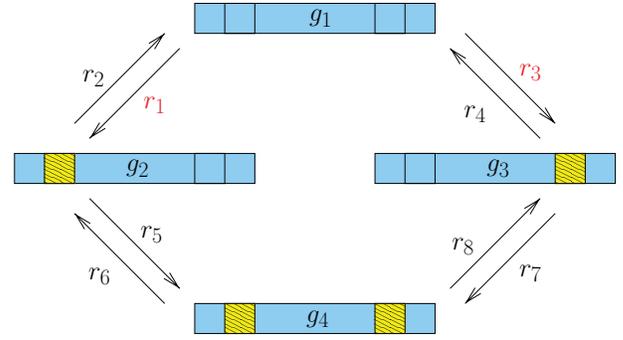


FIG. 7. Schematic diagram of the pap regulatory network. There are four possible states of the pap operon depending on the LRP-DNA binding.

C. Sensitivity of a rare state probability

Consider the pap operon regulation^{10,23} which has a biologically important rare state. In Srivastava *et al.*,²³ we obtained a significantly better estimate of the rare state probability using stochastic quasi-steady-state perturbation analysis (sQSPA), compared to both the full model and importance sampling based methods.^{24,25} Figure 7 shows the schematic of the pap operon regulation.

State g_1 is the rare state. The master equation for the system is

$$\frac{dP_1}{dt} = -(r_1 + r_3)P_1 + r_2P_2 + r_4P_3, \quad (13)$$

$$\frac{dP_2}{dt} = -(r_2 + r_5)P_2 + r_1P_1 + r_6P_4, \quad (14)$$

$$\frac{dP_3}{dt} = -(r_4 + r_7)P_3 + r_3P_1 + r_8P_4, \quad (15)$$

$$\frac{dP_4}{dt} = -(r_6 + r_8)P_4 + r_5P_2 + r_7P_3, \quad (16)$$

in which $P_i(t; r_2)$: $i = 1, 2, 3, 4$ is the probability of state g_i and r_j : $j = 1, 2, \dots, 8$ are the rates of transition defined in Table III. Define

$$S_i(t; r_2) = \frac{\partial P_i}{\partial r_2} \quad i = 1, 2, 3, 4. \quad (17)$$

The governing equations for the sensitivities S_i : $i=1,2,3,4$ are obtained by differentiating (13)–(16) with respect to r_2 ,

$$\frac{dS_1}{dt} = -(r_1 + r_3)S_1 + P_2 + r_2S_2 + r_4S_3, \quad (18)$$

$$\frac{dS_2}{dt} = -P_2 - (r_2 + r_5)S_2 + r_1S_1 + r_6S_4, \quad (19)$$

$$\frac{dS_3}{dt} = -(r_4 + r_7)S_3 + r_3S_1 + r_8S_4, \quad (20)$$

$$\frac{dS_4}{dt} = -(r_6 + r_8)S_4 + r_5S_2 + r_7S_3. \quad (21)$$

TABLE III. Reaction stoichiometry and reaction rates for Sec. III C.

Number	Reaction stoichiometry	Reaction rate (r_i)
1	$g_1 \rightarrow g_2$	100.0
2	$g_2 \rightarrow g_1$	0.625
3	$g_1 \rightarrow g_3$	100.0
4	$g_3 \rightarrow g_1$	1.033
5	$g_2 \rightarrow g_4$	0.99
6	$g_4 \rightarrow g_2$	1.033
7	$g_3 \rightarrow g_4$	0.99
8	$g_4 \rightarrow g_3$	0.625

In Srivastava *et al.*,²³ we showed that the sQSPA model reduction leads to the reaction network shown in Figure 8.

In the sQSPA model reduction, we write probabilities P_i : $i = 1, 2, 3, 4$ in a power series expression given by

$$P_i = W_{i0} + \epsilon_{sq} W_{i1} + \epsilon_{sq}^2 W_{i2} + O(\epsilon_{sq}^3),$$

in which $\epsilon_{sq} = \frac{1}{r_1+r_3}$. By comparing $O(\epsilon_{sq}^0)$ terms, the sQSPA model reduction finds the expression for W_{i0} , $i = 1, 2, 3, 4$. The master equation for the sQSPA reduced model is

$$W_{10} = 0, \quad (22)$$

$$\frac{dW_{20}}{dt} = -[\tilde{r}_1 + r_5]W_{20} + \tilde{r}_2 W_{30} + r_6 W_{40}, \quad (23)$$

$$\frac{dW_{30}}{dt} = -[\tilde{r}_2 + r_7]W_{30} + \tilde{r}_1 W_{20} + r_8 W_{40}, \quad (24)$$

$$\frac{dW_{40}}{dt} = -[r_6 + r_8]W_{40} + r_5 W_{20} + r_7 W_{30}, \quad (25)$$

in which W_{ij} are the j th-order probabilities of state g_i : $i = 1, 2, 3, 4$, $\tilde{r}_1 = r_2 r_3 / (r_1 + r_3)$, and $\tilde{r}_2 = r_1 r_4 / (r_1 + r_3)$. By comparing $O(\epsilon_{sq}^1)$ terms, one crucial equation²³ that comes out is the approximation of the probability (P_{sq_1}) of the rare state in terms of the probabilities of the other states satisfies

$$P_{sq_1} = \epsilon_{sq}(r_2 W_{20} + r_4 W_{30}) \quad (26)$$

in which $\epsilon_{sq} = 1/(r_1 + r_3)$.

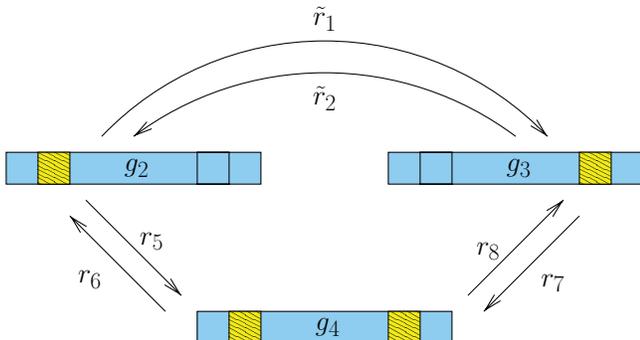


FIG. 8. Reduced system of the pap regulatory network of Sec. III C in the slow time scale regime.

We are interested in the sensitivity of probability of the rare state g_1 , with respect to r_2 ,

$$s(t; r_2) = S_1 = \frac{\partial P_1(t; r_2)}{\partial r_2}. \quad (27)$$

To estimate $s(t; r_2)$ of Eq. (27), we use three different estimators: CRN, CFD, and sQSPA with common random numbers (SRN). The CRN estimator is given as

$$\hat{s}_{cm}(t; r_2) = \frac{1}{N\epsilon} \sum_{i=1}^N [\mathbb{1}(pa_i^{cm}(t, r_2 + \epsilon) = g_1) - \mathbb{1}(pa_i^{cm}(t, r_2) = g_1)] \quad (28)$$

in which $pa_i^{cm}(t; r_2)$ is the state of the pap operon at time t with rate parameter r_2 obtained through the i th CRN simulation, and N is the number of CRN simulations. A point to note is that the i th CRN simulation uses one SSA simulation with rate parameter r_2 , i.e., $pa_i^{cm}(t; r_2)$, and one SSA simulation with rate parameter $r_2 + \epsilon$, i.e., $pa_i^{cm}(t; r_2 + \epsilon)$. The indicator random variable $\mathbb{1}(A)$ evaluates to 1 whenever the event A happens and 0 otherwise. In an analogous fashion, we have the CFD estimator for Eq. (27) as

$$\hat{s}_{cfd}(t; r_2) = \frac{1}{N\epsilon} \sum_{i=1}^N [\mathbb{1}(pa_i^{cfd}(t, r_2 + \epsilon) = g_1) - \mathbb{1}(pa_i^{cfd}(t, r_2) = g_1)]. \quad (29)$$

The SRN estimator estimates the sensitivity of P_{sq_1} of Eq. (26) with respect to r_2 . The SRN estimator is given by

$$\hat{s}_{sm} = \frac{\epsilon_{sq}}{\epsilon} [(r_2 + \epsilon)\hat{W}_{20}(t; r_2 + \epsilon) + r_4\hat{W}_{30}(t; r_2 + \epsilon) - r_2\hat{W}_{20}(t; r_2) - r_4\hat{W}_{30}(t; r_2)] \quad (30)$$

in which $\hat{W}_{20}(t; r_2)$ and $\hat{W}_{20}(t; r_2 + \epsilon)$ are obtained by simulating the reduced system shown in Figure 8 and governed by the master equation (22)–(25) using common random numbers and SSA simulations. Figure 9 shows a comparison of the CRN, CFD, and SRN estimators. The ratio of root mean squared errors of CRN and CFD estimators is $\frac{\epsilon_{cm}}{\epsilon_{cfd}} = 4.75$. As the number of reaction events is small in the pap operon, we also apply the likelihood gradient method.¹⁵ We performed 500 simulations for each of the four estimators. The likelihood method performs better than both CRN and CFD for this example, but it performs worse than SRN. The SRN estimator tracks the true sensitivity closely except for a small initial time. This example reveals the distinct advantage of analytical insight and model reduction, e.g., the sQSPA analysis, over the several proposed estimators that do not use the reduced model.

In this example, the number of reactions fired within the time interval of consideration was small, which is precisely when the likelihood method can produce a low variance estimator. In fact, in this particular application the likelihood method even outperformed the CFD method. For models with even a moderate number of reaction events, the likelihood method will not outperform CFD. We also note that in this example, the SRN method outperforms all other methods. However, SRN requires the ability to perform an analytical model

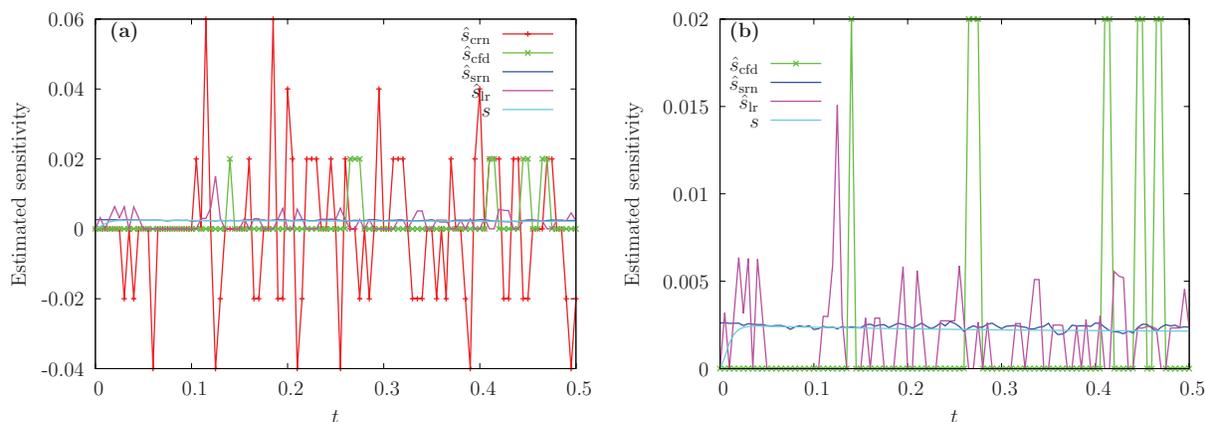
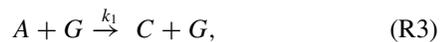


FIG. 9. Estimated sensitivity from the CRN, CFD, and SRN estimators for Sec. III C. The SRN estimator tracks the true sensitivity closely except for a small initial time. The CFD estimator performs better than the CRN estimator. On the right, we see that the likelihood method performs better than both the CRN and CFD methods but it performs worse than SRN.

reduction using sQSPA, which was possible in this example, thought not in general.

D. Sensitivity of a fast fluctuating species

In a stochastic simulation of the infection cycle of vesicular stomatitis virus (VSV), there is a fast fluctuation in a protein at low copy number along with a rapid increase in the population of the viral genome. Such a system is expensive to simulate because the frequency of the fluctuation increases as the simulation progresses leading to small time steps in the SSA simulation.²⁶ To illustrate the phenomenon, consider the following simple 3-species, 3-reaction system



with $k_2 \sim k_1 \gg k_3$. This reaction system describes the interaction of three species in a simplified VSV replication process – two forms of viral polymerase, A and C, and viral genome G. The two forms of the polymerase arise because VSV has two different complexes that serve as viral transcriptase and

replicase.²⁷ The viral transcriptase form A is a complex of constituent VSV proteins L and P. The replicase form C is a complex of L, N, and P proteins. The species A is involved in the transcription reaction (R3) to produce messenger RNA. The transcription reaction leads to the conversion of transcriptase A into replicase C. We further assume that produced mRNA from reaction (R3) is short lived and hence we do not include it in the model. Species C and G are involved in replication reaction (R4) to produce an additional viral genome G. The replication reaction (R4) leads to the conversion of replicase C into transcriptase A. Finally, there is a second-order degradation reaction (R5) of the viral genome. The model (R3)–(R5) is insufficient to predict the full viral infection cycle, but it is instructive in understanding the simulation challenges of the full infection cycle model used by Hensel *et al.*²⁶ The reaction rate constants k_1, k_2, k_3 denote macroscopic reaction rate constants with units $\mu\text{m}^3/(\text{mol s})$. We express microscopic reaction rates in terms of macroscopic rate constants (k_1, k_2, k_3) and the system size Ω ,

$$r_1 = \frac{1}{\Omega} k_1 a g, \quad r_2 = \frac{1}{\Omega} k_2 c g, \quad r_3 = \frac{1}{\Omega} k_3 g(g-1)$$

in which a, c, g represent the number of molecules of A, C, and G, respectively, and the system size appears because the reactions are second order. For the purposes of this

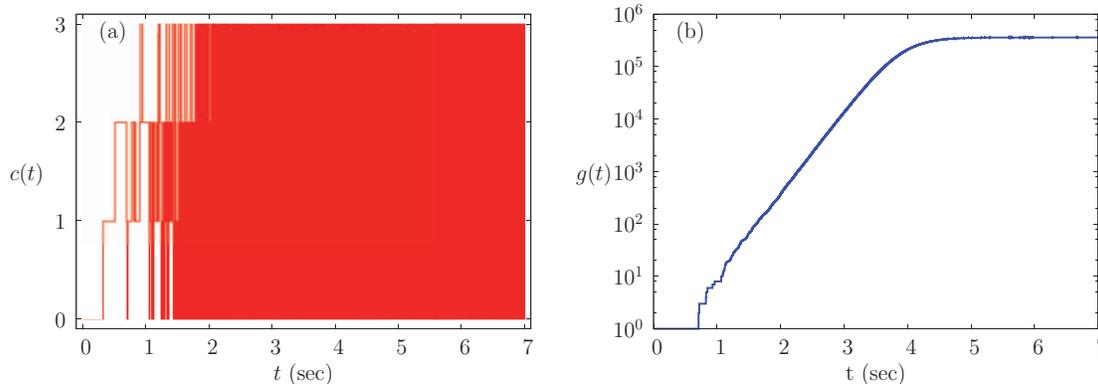


FIG. 10. A typical SSA simulation of the network of (R3)–(R5). (a) Counts of species C vs. time. (b) Counts of species G vs. time.

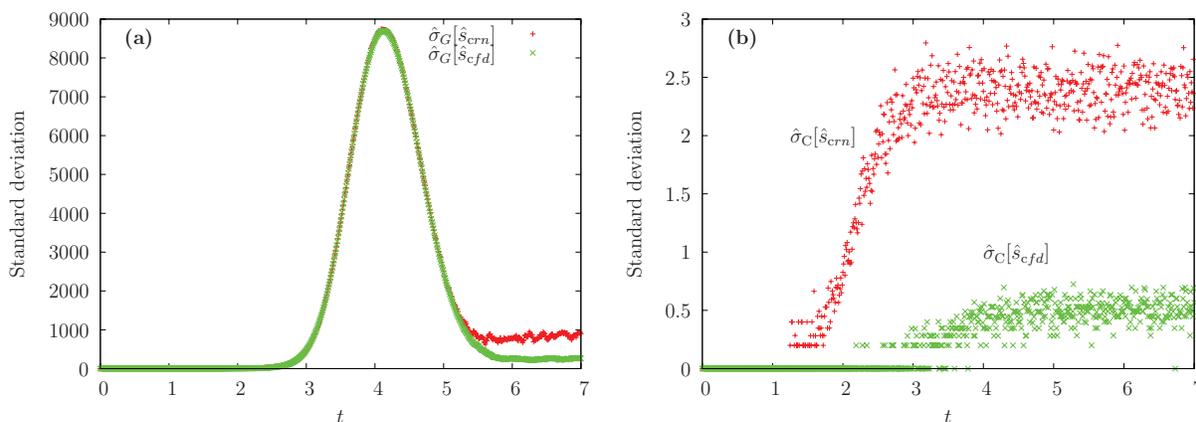


FIG. 11. Comparison of standard deviations of CFD and CRN estimators for Sec. III D: (a) Species G and (b) Species C.

example, we take $\Omega = 10^5 \mu\text{m}^3$. A stochastic simulation with the parameter values given in Table IV is shown in Figure 10. Species G increases continuously and this increase forces species C to fluctuate with increasing frequency as shown in Figure 10.

We want to investigate the sensitivity

$$s(t; k_3) = \frac{d\mathbb{E}X(t, k_3)}{dk_3}. \quad (31)$$

In which $X \in \{C, G\}$. An estimator of the sensitivities of interest (31) is

$$\hat{s}_{\text{est}} = \frac{1}{N} \sum_{i=1}^N \frac{X_i^{\text{est}}(t; k_3 + \epsilon) - X_i^{\text{est}}(t; k_3)}{\epsilon} \quad (32)$$

in which $\text{est} \in \{\text{CRN}, \text{CFD}\}$. Figure 11 shows a comparison of the standard deviations of the two estimators, CRN and CFD, obtained using Eq. (4). The parameters used to generate Figure 11 are shown in Table IV. Figure 11(a) shows that for the abundant species G, CFD, and CRN estimators have similar standard deviations, which demonstrates that both CRN and CFD are capable of obtaining good sensitivity estimates for the abundant species G. Figure 11(b) shows that for fast fluctuating species C, the CFD estimator has less than one third the standard deviation of the CRN estimator, which demonstrates that the CFD estimator captures the sensitivity of the fast fluctuating species C better than the CRN estimator. This example again demonstrates the superiority of the CFD estimator over the CRN estimator.

E. Sensitivity of genetic toggle switch

We conclude with a model of a genetic toggle switch^{1,2,28}

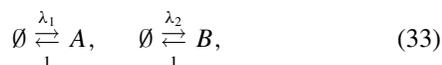


TABLE IV. Parameter values for Sec. III D.

Parameter	n_{A_0}	n_{C_0}	n_{G_0}	k_1	k_2	k_3	ϵ	N
Value	3	0	1	2×10^5	3×10^5	1	0.1	100

where the respective propensities are

$$\lambda_1(t) = \frac{b}{1 + X_B(t)^\beta} \quad \text{and} \quad \lambda_2(t) = \frac{a}{1 + X_A(t)^\alpha},$$

and where $X_A(t)$ and $X_B(t)$ will denote the number of gene products from the two interacting genes. As in Refs. 1, 2, and 28, we take parameter values of $b = 50$, $\beta = 2.5$, $a = 16$, and $\alpha = 1$. We consider the derivative of the expectation of X_A with respect to α at the value one, with $[X_A(0), X_B(0)] = [0, 0]$ as our choice of initial condition. We will consider the behavior of three finite difference methods on this example: CFD, CRN, and CRP.² For each of the three finite difference methods employed, we use a perturbation of $\epsilon = 1/50$.

Figure 12 shows a comparison of the standard deviations of the three estimators. We again see the superiority of the CFD estimator over both the CRN and CRP estimators.

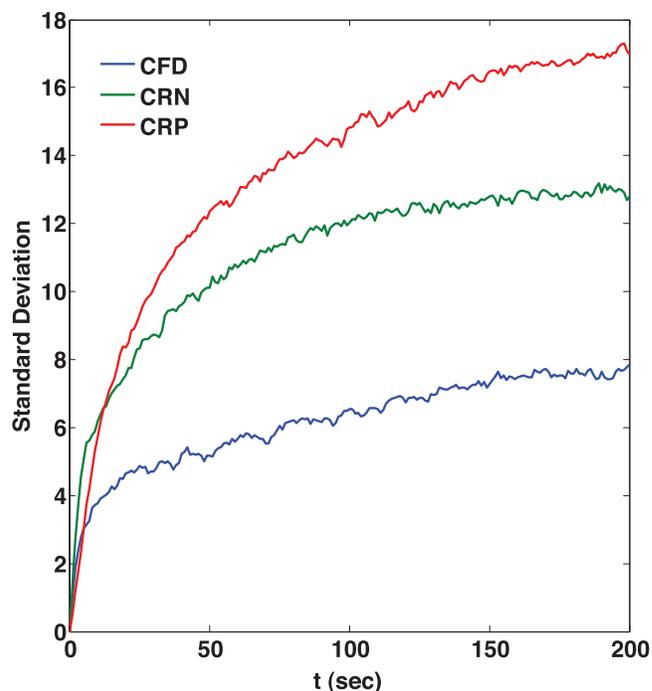


FIG. 12. Comparison of standard deviations of CFD, CRN, and CRP estimators on the model (33).

Further, starting at time $t = 5s$, the standard deviation of the CRN estimator is less than that of the CRP estimator. This result demonstrates that it is not always the case that CRP is superior to CRN, something not previously observed in the literature.

IV. CONCLUSIONS

In this paper, we compared the performance of several finite difference sensitivity estimators on a number of examples not previously considered in the literature. In all the examples and sensitivities of interest, we found that the newly developed CFD estimator performs significantly better than the CRN estimator, which is currently the most commonly used method. Further, in Sec. III E we provided a case in which the CRN estimator is superior to the CRP estimator.

In previous work, it had been shown that the CFD estimator performs better than the CRP estimator.¹ The comparisons made in this paper, along with Anderson's previous results, lead to the conclusion that CFD is currently the best available estimator in the class of finite difference estimators of stochastic chemical kinetic models.

Estimating sensitivities of stochastic chemical kinetic models accurately and efficiently remains an important problem. With variance reduction ideas incorporated in the CFD estimator through a tight coupling of the nominal and perturbed systems, we believe this estimator has significant advantages over previous estimators without such a coupling. The CFD estimator has significant potential for application in parameter estimation where it can provide accurate estimates of the Hessian of the objective function. Recently, Wolf and Anderson²⁸ have also proposed a CFD2 estimator to directly estimate the Hessian of the objective function. Evaluating the effectiveness of CFD and CFD2 estimators in parameter estimation of stochastic chemical kinetic models is a topic of ongoing research.

- ¹D. F. Anderson, *SIAM J. Numer. Anal.* **50**, 2237 (2012).
- ²M. Rathinam, P. W. Sheppard, and M. Khammash, *J. Chem. Phys.* **132**, 034103 (2010).
- ³H. H. McAdams and A. Arkin, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 814 (1997).
- ⁴A. Arkin, J. Ross, and H. McAdams, *Genetics* **149**, 1633 (1998).
- ⁵M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. S. Swain, *Science* **297**, 1183 (2002).
- ⁶E. M. Ozbudak, M. Thattai, I. Kurtser, A. D. Grossman, and A. van Oudenaarden, *Nat. Genet.* **31**, 69 (2002).
- ⁷W. J. Blake, M. A. Kaern, C. R. Cantor, and J. J. Collins, *Nature (London)* **422**, 633 (2003).
- ⁸J. M. Raser and E. K. O'Shea, *Science* **304**, 1811 (2004).
- ⁹S. Hooshangi, S. Thiberge, and R. Weiss, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 3581 (2005).
- ¹⁰B. Munsky and M. Khammash, *J. Chem. Phys.* **124**, 044104 (2006).
- ¹¹D. T. Gillespie, *J. Phys. Chem.* **81**, 2340 (1977).
- ¹²D. T. Gillespie, *Physica A* **188**, 404 (1992).
- ¹³J. B. Rawlings and J. G. Ekerdt, *Chemical Reactor Analysis and Design Fundamentals* (Nob Hill, Madison, WI, 2004), p. 640.
- ¹⁴P. W. Glynn, *Commun. ACM* **33**, 75 (1990).
- ¹⁵M. K. Nakayama, A. Goyal, and P. W. Glynn, *Oper. Res.* **42**, 137 (1994).
- ¹⁶S. Piyasunov and A. P. Arkin, *J. Comput. Phys.* **221**, 724 (2007).
- ¹⁷M. Komorowski, M. J. Costa, D. A. Rand, and M. P. H. Stumpf, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 8645 (2011).
- ¹⁸R. Gunawan, Y. Cao, L. Petzold, and F. J. Doyle, *Biophys. J.* **88**, 2530 (2005).
- ¹⁹P. W. Glynn, in *Proceedings of the 21st Conference on Winter Simulation* (ACM, 1989), pp. 90–105.
- ²⁰J. McGill, B. Ogunnaike, and D. Vlachos, *J. Comput. Phys.* **231**, 7170 (2012).
- ²¹T. Drews, R. Braatz, and R. Alkire, *J. Electrochem. Soc.* **150**, C807 (2003).
- ²²This will be a part of an upcoming publication.
- ²³R. Srivastava, E. L. Haseltine, E. A. Mastny, and J. B. Rawlings, *J. Chem. Phys.* **134**, 154109 (2011).
- ²⁴H. Kuwahara and I. Mura, *J. Chem. Phys.* **129**, 165101 (2008).
- ²⁵M. K. Roh, D. T. Gillespie, and L. R. Petzold, *J. Chem. Phys.* **133**, 174106 (2010).
- ²⁶S. Hensel, J. B. Rawlings, and J. Yin, *Bull. Math. Biol.* **71**, 1671 (2009).
- ²⁷J. K. Rose and M. A. Whitt, in *Fields Virology*, 4th ed., edited by D. M. Knipe and P. M. Howley (Lippincott, Philadelphia, 2001), Vol. 1, pp. 1221–1244.
- ²⁸E. S. Wolf and D. F. Anderson, *J. Chem. Phys.* **137**, 224112 (2012).