Available online at www.sciencedirect.com
ScienceDirect
Journal of Computational Physics 221 (2007) 724–738
JOURNAL OF COMPUTATIONAL PHYSICS
www.elsevier.com/locate/jcp

Efficient stochastic sensitivity analysis of discrete event systems

Sergey Plyasunov [a], Adam P. Arkin [b,*]

[a] *Department of Physics, University of California, Berkeley, Physical Biosciences Division, E.O. Lawrence Berkeley National Laboratory, 1 Cyclotron Road, MS 19-0175, Berkeley, CA 94720, United States*

[b] *Howard Hughes Medical Institute, Department of Bioengineering, University of California, Berkeley, Physical Biosciences Division, E.O. Lawrence Berkeley National Laboratory, 1 Cyclotron Road, MS 977-0257, Berkeley, CA 94720, United States*

Received 4 May 2005; received in revised form 6 June 2006; accepted 28 June 2006
Available online 6 September 2006

Abstract

Sensitivity analysis quantifies the dependence of a system's behavior on the parameters that could possibly affect the dynamics. Calculation of sensitivities of stochastic chemical systems using Kinetic Monte Carlo and finite-difference-based methods is not only computationally intensive, but direct calculation of sensitivities by finite-difference-based methods of parameter perturbations converges very poorly. In this paper we develop an approach to this issue using a method based on the Girsanov measure transformation for jump processes to smooth the estimate of the sensitivity coefficients and make this estimation more accurate. We demonstrate the method with simple examples and discuss its appropriate use.

Keywords: Parameter sensitivity analysis; Stochastic dynamics of reaction networks; Girsanov transformation

1. Introduction

The growing emphasis on experiment design issues and the development of increasingly sophisticated simulation techniques for the study of complex problems in industrial engineering, operations research, economics [1,2] and recently biological sciences [3–9] have challenged existing computational methodologies and have required researchers to develop methods for the estimation of sensitivity of models to variation or uncertainty in model parameters: if a small change in a parameter results in relatively large changes in the outcomes, the outcomes are said to be sensitive to that parameter. This may mean that this type of parameter has to be determined very accurately. Parameters to which a model's behavior exhibits a small sensitivity do not need rigorous measurements and are likely not good control points of the system behavior, while parameters with high sensitivity must be measured more exactly to achieve an accurate model and may be the points of control in model dynamics [10,11]. Sensitivity analysis plays a key role in implementation *response surface techniques* which use low-degree polynomials to model system responses as functions of active parameters [12,13].

* Corresponding author.
E-mail addresses: teleserg@uclink.berkeley.edu (S. Plyasunov), aparkin@lbl.gov (A.P. Arkin).

0021-9991/$ - see front matter
doi:10.1016/j.jcp.2006.06.047

In this paper we develop an unbiased estimator for sensitivity analysis of stochastic systems. The method is based on a general likelihood approach based on a Girsanov measure transformation without assumption on the shape of the probability distribution of the system state. We present a simple extension of a popular algorithm [14] for simulation of the chemical master equation that allows the efficient numerical calculation of these sensitivities concurrently with the computation of sample paths for the system under study. We also provide guidelines as to when this method obtains better accuracy and computational efficiency than other methods.

In Section 3 we describe the specific class of discrete event models relevant to the stochastic dynamics of chemical reaction networks [15,14]. After that, in Section 4, we provide the construction of the family of the jump-processes, indexed by the perturbation parameter. We demonstrate the method using simple examples and provide a straightforward algorithm to implement the method. Appendices A and B outline mathematical details of the approach.

2. Definitions and background

We suppose that the behavior of the stochastic system depends on a set of parameters and can be expressed as a family of Markov processes $\{\mathbf{X}(t), t \geqslant 0\}$ with values in some multidimensional space. Specific examples of dynamics governing $\mathbf{X}(t)$ will be outlined later on.

We assume that the complexity of the system is such that it is necessary to use simulation to estimate the behavior of the system for each set of parameter values. Given a set of observation points on the time interval $[0, T]$, $0 < t_1 \leqslant t_2 \leqslant \cdots \leqslant t_M = T$, a set of model parameters $\mathbf{k}$ and an initial condition $\mathbf{X}(0) = \mathbf{x}$, we consider the quantity whose sensitivity to parameters we wish to estimate:

$$u(\mathbf{x}, \mathbf{k}) = \mathbb{E}[F(\mathbf{X}(t_1), \ldots, \mathbf{X}(t_M))|\mathbf{X}(0) = \mathbf{x}, \mathbf{k}] \tag{1}$$

where the random variable F is the function of the underlying $\mathbf{X}(t)$ process at different observation points. Since the system is multidimensional, *a priori* analytical solutions are usually out of reach and Eq. (1) is estimated via a Monte Carlo approach. The evaluation of $u(\mathbf{x}, \mathbf{k})$ via Monte Carlo is a standard procedure and has been widely studied for many types of systems, see for example the following monographs [15,16].

To explore the behavior of a system under variation of values of model parameters and to estimate the relative importance of the initial values, sensitivity analysis applies small changes to the nominal values of model parameters or initial values and investigates change in the values of the functionals of the type (1). In particular, sensitivity analysis quantifies the local dependence of system behavior on parameters and we will be interested in the *sensitivity coefficients*, i.e. derivatives of $u(\mathbf{x}, \mathbf{k})$ with respect to the initial conditions or model parameters [17].

A very popular and straightforward approach to estimation of sensitivity coefficients is to compute, by Monte Carlo simulation, the finite difference approximation for the differentials. To simplify the discussion, let us assume that one is interested in computing the sensitivity coefficient with respect to the vector of parameters $\mathbf{k}$. To calculate the finite difference one has to compute an estimator for $u(\mathbf{x}, \mathbf{k})$ and an estimator for $u(\mathbf{x}, \mathbf{k}(1 + \boldsymbol{e}))$, where $\boldsymbol{e} = (0, \ldots, \epsilon, \ldots, 0)$ and we use shortcut notation $\mathbf{k}(1 + \boldsymbol{e})$ for the vector with components $(\ldots, k_{i-1}, k_i(1 + \epsilon), k_{i+1}, \ldots)$ i.e. perturbation is applied in the direction of only one component of the vector $\mathbf{k}$. For small enough ϵ, the estimator for the sensitivity coefficient based on the traditional finite forward difference scheme is:

$$\widehat{\partial u(\mathbf{x}, \mathbf{k})/\partial \epsilon} = \frac{u(\mathbf{x}, \mathbf{k}(1 + \mathbf{e})) - u(\mathbf{x}, \mathbf{k})}{\epsilon} \tag{2a}$$

The problem is that the meaning of the phrase "small enough ϵ" is not completely clear. First of all, that the estimator based on the finite difference scheme (2a) has a glaring weakness of being biased with the bias proportional to $O(\epsilon)$. In addition to that it possess a certain variance which decreases with the increase of the numbers of samples N used in the Monte Carlo simulation. It was shown by Glynn in [18], see also [19], that if the simulations of the two estimators for $u(\mathbf{x}, \mathbf{k})$ and $u(\mathbf{x}, \mathbf{k}(1 + \boldsymbol{e}))$ are drawn independently, then the best possible convergence rate, to the exact value, is of the order $N^{-1/4}$ with the optimal choice of ϵ which in this case is $\epsilon \propto N^{-1/5}$. If one uses the symmetric difference scheme

$$\partial u(\widehat{\mathbf{x},\mathbf{k}})/\partial\epsilon = \frac{u(\mathbf{x},\mathbf{k}(1+\mathbf{e}/2)) - u(\mathbf{x},\mathbf{k}(1-\mathbf{e}/2))}{\epsilon} \tag{2b}$$

then the overall rate of convergence is $N^{-1/3}$ [19]. Unfortunately, these optimal rates of convergence cannot be achieved in certain cases. Severe problems arise in the case when the functional of interest does not have smooth derivatives. This is the case, for example, when $F(\mathbf{X})$ is a discrete type functional and does not have a smooth derivative, e.g. $F(\mathbf{X}) = 1$ if and only if X belongs to some subset A of a multidimensional state space and 0, otherwise.

To overcome this limitation, one might use an explicit form of the density function $p(\mathbf{X}_1,\ldots,\mathbf{X}_M,\mathbf{k})$ of the underlying variable $\mathbf{X}$ and express the derivative with respect to parameters $\mathbf{k}$ by averaging over $\nabla_{\mathbf{k}}\log p(\cdot,\mathbf{k})$. This leads to the so called *likelihood method* introduced by Glynn [20], see also Boyle, Broadie and Glasserman [2,21]. For example, if function $F = F(\mathbf{X}_1,\ldots,\mathbf{X}_M)$ depends upon the values of the process $\mathbf{X}(\cdot)$ at points $0 \leqslant t_1,\ldots,t_M \leqslant T$, then the derivative with respect to the model parameters can be calculated in the following way:

$$\begin{aligned}\nabla_{\mathbf{k}} u(\mathbf{x},\mathbf{k}) &= \nabla_{\mathbf{k}} \int \mathrm{d}\mathbf{X}_1 \ldots \int \mathrm{d}\mathbf{X}_M F(\mathbf{X}_1,\ldots,\mathbf{X}_M) p(\mathbf{X}_1,\ldots,\mathbf{X}_M;\mathbf{k}) \\ &= \mathbb{E}[F(\mathbf{X}_1,\ldots,\mathbf{X}_M)\nabla_{\mathbf{k}}\log p(\mathbf{X}_1,\ldots,\mathbf{X}_M;\mathbf{k})]\end{aligned} \tag{3}$$

i.e. it essentially puts dependence upon the parameter into the underlying probability density. This method has been applied by Gunawan et al. In [22] to sensitivity analysis of some discrete event systems of biological and chemical interest. This method can bring down the order of the error to $N^{-1/2}$ but at the cost of requiring a specific form of the multi-point distribution $p(\mathbf{X}_1,\ldots,\mathbf{X}_M;\mathbf{k})$. This form cannot be found exactly in many realistic situations.

We will seek a solution which is similar in spirit to the method mentioned above, i.e. we shall search for an explicit form for a set of *weight functions*, $W(\cdot)$, which permits us to express the sensitivity coefficients as weighted, probably path dependent, averages of the functionals:

$$\nabla u(\mathbf{x},\mathbf{k}) = \mathbb{E}[F(\mathbf{X}(t_1),\ldots,\mathbf{X}(t_M)) \times W(\mathbf{X}(0 \leqslant t \leqslant T))] \tag{4}$$

however, we will not assume any specific form of $p(\cdot,\mathbf{k})$ as in the likelihood method. We will call this estimator *weighted scheme estimator*. The weigh function $W(\cdot)$ appeared in the equation above is generally called the Malliavin weight, as an expression similar to that of (4) can be obtained from the integration by parts formula of Malliavin calculus (see monographs [23–25] for authoritative reference).

3. Sensitivity analysis of a specific class of systems

Without loss of generality, here we focus on a specific class of discrete events systems: the subset of models which describe the stochastic dynamics of chemical reaction networks. Here state variables are numbers of molecules of different molecular species X_i, $i = 1,\ldots,S$ involved in different chemical reactions in the system. To simplify the discussion, we consider the case of only one reaction. As we shall see later this does not inhibit the generality of the results presented in this paper and can be easily extended to the case of the multiple reaction channels indexed by $r = 1,\ldots,R$:

$$v_{1r}^{-}\mathsf{X}_1 + v_{2r}^{-}\mathsf{X}_2 + \cdots + v_{Sr}^{-}\mathsf{X}_S \xrightarrow{a_r} v_{1r}^{+}\mathsf{X}_1 + v_{2r}^{+}\mathsf{X}_2 + \cdots + v_{Sr}^{+}\mathsf{X}_S \tag{5}$$

Integer numbers $v_{ir}^{\pm}$ represent the number of molecules of the species S_i produced (+) and consumed (−) in reaction (5). If we denote by $X_i \in \mathbb{Z}_+$ the number of molecules of the type X_i present in the system before the reaction event in channel r then, after the reaction event, the number of molecules of this type in the system is:

$$X_i \to X_i + (v_i^{+} - v_{ir}^{-}) \tag{6}$$

Then, using mass balance, the dynamics of the number species $X_i(t)$ can be represented as follows:

$$X_i(t) = X_i(0) + \sum_{r=1}^{R} \Delta v_{ir} N_r(t), \tag{7}$$

where $\Delta v_{ir} = v_{ir}^{+} - v_{ir}^{-}$ and $N_r(t)$ is the random variable representing the number of reaction events which took place in the reaction channel up to time t (reaction extent [26]) with the initial condition $N_r(0) = 0$. Events take place with state dependent transition rates also called propensity functions [26] $a_r(\mathbf{X}) = k_r h_r(\mathbf{X}) \geqslant 0$, where $h_r(\cdot)$ is function of the current state of the system [15,27], usually of a polynomial form. The transition rate no longer describes the amount of chemical species being produced or consumed per unit interval of time but rather the likelihood of a certain reaction to occur per unit time. As mentioned above, Kinetic Monte Carlo techniques [14] can be used to sample the number of events $N_r(t)$ and state $\mathbf{X}(t)$. In Section 4, we consider the importance sampling technique which will lead to the construction of the stochastic weights (4).

4. Perturbation analysis

In this section we discuss the sensitivity of functionals of type (1), defined on trajectories of the system (5), with respect to the small perturbations in values of kinetic rates k_r. The problem can be approached by looking at the perturbed process and gradually taking the strength of the perturbation to 0.

To be specific, we consider the family of solutions $\{\mathbf{X}^{\epsilon}(t)\}_{t\geqslant 0}$ parameterized by ϵ, $|\epsilon| < 1$. The parameter ϵ "tunes" the reaction constant k_r in the following way:

$$k_r^{\epsilon} = k_r(1+\epsilon), \tag{8}$$

This perturbation changes the intensity of the counting process $N_r(t)$ in channel r. Here we outline the main result. It can be shown that the stochastic weights $W(\cdot)$ can be represented as derivatives of a certain martingale process $Z^{\epsilon}(\mathbf{X}(\cdot))$, $\partial Z^{\epsilon}/\partial\epsilon$. More specifically, it turns out that the stochastic weights $W(\cdot)$ can be defined by means of the counting process $N_r(t)$ of the unperturbed system:

$$\frac{\partial}{\partial\epsilon}u^{\epsilon}(\mathbf{x}) = \mathbb{E}\left[F(\mathbf{X}(\cdot))\frac{\partial Z^{\epsilon}}{\partial\epsilon}\bigg|_{\epsilon=0}\,\bigg|\,\mathbf{X}(0)=\mathbf{x}\right], \tag{9a}$$

$$W \triangleq \frac{\partial Z^{\epsilon}}{\partial\epsilon}\bigg|_{\epsilon=0} = N_r(t) - \int_0^t \mathrm{d}s a_r(\mathbf{X}(s_-)) = M_r(t) \tag{9b}$$

Specific details of this procedure are presented in Appendix A. A Monte Carlo estimator can be readily obtained from expressions (4) and (9) by replacing the probability measure with its weighted empirical counterpart:

$$\widehat{\partial u^{\epsilon}/\partial\epsilon} = \frac{1}{N}\sum_{j=1}^{N} F^{(j)}W^{(j)} \tag{10}$$

where $F^{(j)}$, $j = 1,\ldots,N$ are the values of the function $F(\cdot)$ taken at specific time points of the trajectory $\mathbf{X}^{(j)}(0 \leqslant t \leqslant T)$, and $W^{(j)}$ are weights computed along the same set of trajectories according to (9).

It is important to note that the estimator provided by the above equations is unbiased and this property does not depend on the specific function $F(\cdot)$. This is an important difference from the finite-difference based estimators, Eqs. (2a) and (2b). We note, however, that variance of this estimator depends on the functional $F(\cdot)$. More details of this will be given in Appendix A and B. Before embarking on the numerical calculation we note here, that the derivatives $\partial Z^{\epsilon}/\partial\epsilon, \partial^2 Z^{\epsilon}/\partial\epsilon^2$, etc. *do not depend* on the functional $F(\cdot)$ itself and can be calculated along with the traditional Kinetic Monte Carlo (KMC) simulation, such as those proposed in [15,28,14]. The algorithm following Gillespie [14] is as follows:

(1) *Initialize:* $\mathbf{X} = \mathbf{X}(0)$, $t = 0$.
(2) *Select the index* $r \in \{1,\ldots,R\}$ *of the kinetic rate to be investigated with respect to the perturbation*;
(3) *Generate N indexed by* $j = 1\ldots N$; *set* $M_r^{(j)} = 0$
While $(t \leqslant T)$
  (a) At each state $\mathbf{X}^{(j)}(t)$, compute $a_0(\mathbf{X}^{(j)}(t)) = \sum_{r'=1}^{R} a_{r'}(\mathbf{X}^{(j)}(t))$
  (b) $\tau = -\frac{1}{a_0}\ln(1-u), u \in \mathscr{U}([0,1))$-random variable, uniformly distributed on $[0,1)$
  (c) Index r' of the next reaction event is given by transition probability vector $\{q_r(\mathbf{X}^{(j)}(t))\}$:

$$\{q_r(\mathbf{X}^{(j)}(t))\} = \left(\frac{a_1(\mathbf{X}^{(j)}(t))}{a_0(\mathbf{X}^{(j)}(t))},\ldots,\frac{a_R(\mathbf{X}^{(j)}(t))}{a_0(\mathbf{X}^{(j)}(t))}\right)$$

(d) Update the state: $X_i^{(j)}(t+\tau) = X_i^{(j)}(t) + \Delta v_{ir'}$, $t \leftarrow t+\tau$
(e) If $r = r'$ then $\Delta M_r^{(j)} = 1 - a_r(\mathbf{X}^{(j)}(t))\tau$, otherwise $\Delta M_r^{(j)} = -a_r(\mathbf{X}^{(j)}(t))\tau$ (see (9)).

(4) *Compute the weighted average:*

$$\sum_{j=1}^{N} F(\mathbf{X}^{(j)}), \quad \sum_{j=1}^{N} F(\mathbf{X}^{(j)}) M_r^{(j)}$$

We will not pursue the extensive investigation of the second order sensitivity coefficients $\partial^2 u^\epsilon/\partial\epsilon^2$ but they follow straightforwardly from the second derivative of Z^ϵ:

$$\frac{\partial^2 u}{\partial \epsilon^2} = \mathbb{E}\left[F(\mathbf{X}(t_1),\ldots,\mathbf{X}(t_m))\frac{\partial^2 Z^\epsilon}{\partial\epsilon^2}\bigg|_{\epsilon=0}\right], \tag{11a}$$

$$\frac{\partial^2 Z^\epsilon}{\partial\epsilon^2}\bigg|_{\epsilon=0} = N_r(t)(N_r(t)-1) + \left(\int_0^t a_r(\mathbf{X}(s_-))\,\mathrm{d}s\right)^2 - 2N_r(t)\int_0^t a_r(\mathbf{X}(s_-)\,\mathrm{d}s \tag{11b}$$

In the next section, we discuss a simple numerical example, investigate the accuracy and convergence of the weighting method, and compare it with the forward difference scheme given by (2a) for two different types of functionals $F(\mathbf{X}(\cdot))$.

5. Numerical experiment

In our first example, we consider a one specie reaction scheme:

$$(\ldots) \xrightarrow{k_2} \mathsf{X} \xrightarrow{k_1} \emptyset \tag{12}$$

which represents competing decay of X with the rate $k_1 X$ and production with the rate k_2. The probabilistic description has one SDE:

$$\mathrm{d}X(t) = \mathrm{d}N_1(t) - \mathrm{d}N_2(t), \tag{13}$$

where $N_{1,2}(t)$ are the numbers of reaction events which took place in each channel up to time t.

This simple example may be solved analytically for the time-dependent probability density $p(X,t)$ for the system to be at the state with X molecules at the moment t [29]:

$$p(X,t) = \frac{(\lambda(t))^X}{X!}\exp(-\lambda(t)), \tag{14a}$$

$$\lambda(t) = X(0)\mathrm{e}^{-k_1 t} + \frac{k_2}{k_1}(1-\mathrm{e}^{-k_1 t}), \quad 0 \leqslant t \leqslant T \tag{14b}$$

We consider independent perturbations of the kinetic parameters of the above model:

$$k_1(\epsilon) = k_1(1+\epsilon), \quad k_2(\epsilon) = k_2(1+\epsilon)$$

and two types of functionals $F(\cdot)$ will be considered:

$$F(X(\cdot)) = X(T), \tag{15a}$$

$$F(X(\cdot)) = 1_{\{a\leqslant X\leqslant b\}}, \tag{15b}$$

In Figs. 1a, 1b and 3a we present the comparison of the sensitivity coefficients computed via the direct difference method with $\epsilon = 10^{-2}$ and the weighted scheme. Figs. 1a and 1b show that the method based on introduction of the stochastic weights provides estimates of the same mean but smaller variance for the estimation of the smooth functional F (15a). In Fig. 1a, the estimate of the sensitivity from the direct method at $N = 10^4$ samples is $k_1\frac{\partial u}{\partial k_1} \approx -11.1896$ (variance 61.3147) and that from the weighted scheme is $k_1\frac{\partial u}{\partial k_1} \approx -11.2015$ (variance 2.4161). The weighted scheme converges faster and with less variance to the sensitivity estimate. The line corresponding to the exact solution ((14): $k_1\frac{\partial u}{\partial k_1} = -11.2820$) is also plotted. In Fig. 1a the estimate of the sensitivity from the direct method at $N = 10^4$ samples is $k_1\frac{\partial u}{\partial k_1} \approx 9.8386$ (variance 62.2378) and that from the

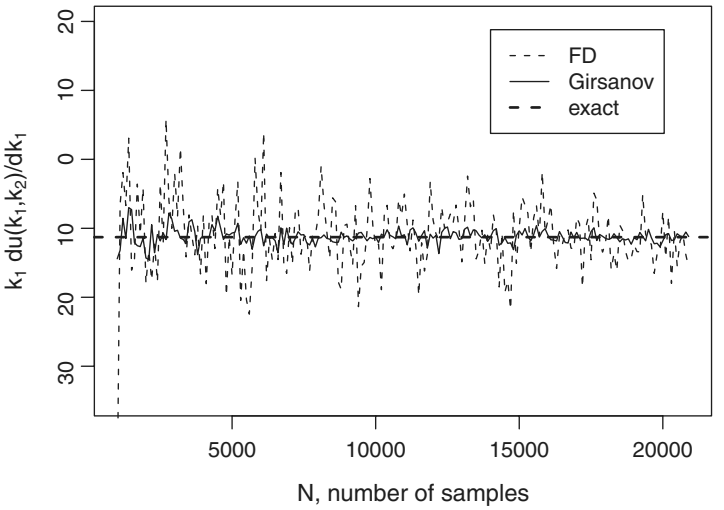

Fig. 1a. Comparison of computation of the sensitivity $k_1 \frac{\partial u}{\partial k_1}$ via the finite difference scheme (labeled "FD") (2a) and the weighted scheme based on the Girsanov transformation with the exact result for the functional $F(X(\cdot)) = X(T)$. The estimate of the sensitivity from the direct method at $N = 10^4$ samples is $k_1 \frac{\partial u}{\partial k_1} \approx -11.1896$ (variance 61.3147) and that from the weighted scheme is $k_1 \frac{\partial u}{\partial k_1} \approx -11.2015$ (variance 2.4161). The weighted scheme converges faster and with less variance to the sensitivity estimate. The line corresponding to the exact solution (14): $k_1 \frac{\partial u}{\partial k_1} = -11.2820$) is also plotted.

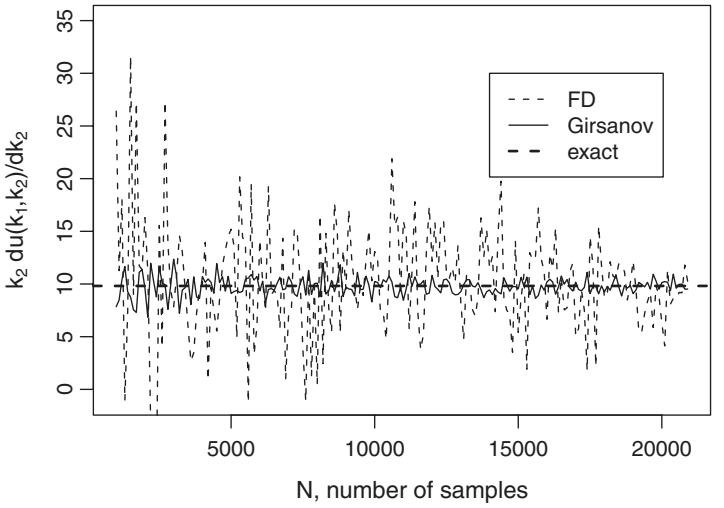

Fig. 1b. Comparison of computation of the sensitivity $k_2 \frac{\partial u}{\partial k_2}$ via the direct finite difference scheme (labeled "FD") (2a) and the weighted scheme based on the Girsanov transformation with the exact result for the functional $F(X(\cdot)) = X(T)$. The estimate of the sensitivity from the direct method at $N = 10^4$ samples is $k_1 \frac{\partial u}{\partial k_1} \approx 9.8386$ (variance 62.2378) and that from the weighted scheme is $k_1 \frac{\partial u}{\partial k_1} \approx 9.7932$ (variance 1.5815). The weighted scheme converges faster and with less variance to the sensitivity estimate. The line corresponding to the exact solution ((14): $k_2 \frac{\partial u}{\partial k_2} = -11.2820$) is also plotted.

weighted scheme is $k_1 \frac{\partial u}{\partial k_1} \approx 9.7932$ (variance 1.5815). The weighted scheme converges faster and with less variance to the sensitivity estimate. The line corresponding to the exact solution (14): $k_2 \frac{\partial u}{\partial k_2} = -11.2820$) is also plotted. The weighted scheme can also significantly outperform the direct difference method when $F(\cdot)$ is a non-smooth functional (15b) providing an estimator with smaller variance. The estimate of the sensitivity from the direct method at $N = 10^4$ samples (Fig. 3a) is $k_2 \frac{\partial u}{\partial k_2} \approx 0.3315$ (variance 0.3047) and that from the weighted scheme is $k_2 \frac{\partial u}{\partial k_2} \approx 0.3653$ (variance 0.0008). The weighted scheme converges faster and with less variance to the sensitivity estimate.

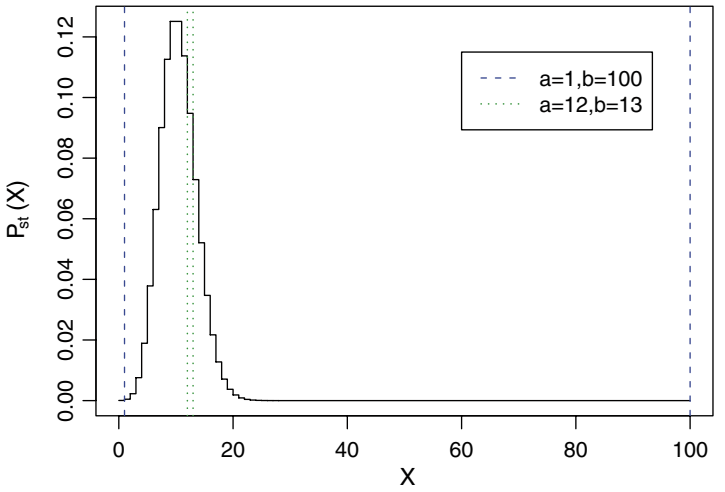

Fig. 2. Equilibrium density (Eq. (14)) profile corresponding to reaction scheme $\emptyset \overset{k_1}{\leftarrow} X \overset{k_2}{\leftarrow} (\ldots)$ and different boundaries a, b specifying the functional $F(X(\cdot)) = 1_{\{a \leqslant X(T) \leqslant b\}}$.

Examination and comparison of Figs. 3a and 3b show that the functional $F(\cdot)$ is a constant, Fig. 2, over a significant part of the state space (e.g. for parameters $a = 1$, $b = 100$ in Eq. (15b)). The estimate of the sensitivity from the direct method at $N = 10^4$ samples, Fig. 3b, is $k_2 \frac{\partial u}{\partial k_2} \approx 0.00012$ (variance 5.7×10^{-5}) and that from the weighted scheme is $k_2 \frac{\partial u}{\partial k_2} \approx -0.0103$ (variance 0.0039). The direct scheme converges faster and with less variance to the sensitivity estimate. In this situation, introduction of stochastic weights may increase the variability of the estimator as demonstrated in Fig. 3b and may be not an advantageous strategy. See Appendix B for details on this issue.

To illustrate the approach on a more complex system and functionals, we examined a simplified model of activity and transcriptional regulation of P_R promoter in bacteriophage λ [30].

The P_R promoter is an integral component of genetic circuit controlling the lysogeny vs lysis decision in the λ virus infection life-cycle in the bacterium *E. coli* [31]. The promoter P_R controls the expression of the protein *Cro*, while the activity of the promoter itself is controlled, in lytic phase, by the interaction of Cro-dimers (Cro_2) and RNA polymerase (RNAp) with the promoter control regions [30], Fig. 4. For illustration we choose a highly simplified scheme for regulation of the P_R promoter by Cro_2 as follows:

$$O_R + RNA_P \overset{\star}{\rightarrow} RNA_P - O_R, \tag{16a}$$

$$Cro_2 + O_R \overset{\star}{\rightleftarrows} Cro - O_R 1, \tag{16b}$$

$$Cro_2 + O_R \overset{\star}{\rightleftarrows} Cro - O_R 2, \tag{16c}$$

$$Cro_2 + Cro - O_R 1 \overset{\star}{\rightleftarrows} Cro - O_R, \tag{16d}$$

$$Cro_2 + Cro - O_R 2 \overset{\star}{\rightleftarrows} Cro - O_R, \tag{16e}$$

$$RNA_P - O_R \overset{a_1}{\rightarrow} N Cro + RNA_P + O_R \tag{16f}$$

We should also stress that in (16) we have employed a rather simplified model of protein-DNA interactions and translation/transcription. The preceding reactions assume a single molecule of DNA, O_R containing the promoter and operator sites which define P_R. Cro_2 can bind to one or two sites on O_R: $O_R 1$ and $O_R 2$. In the above model, for simplicity, we assert an ordered binding reaction for two site occupancy.

We assume, for example, that during RNAp binding (dissociation) events, both operator sites $O_R 1$ and $O_R 2$ become occupied (free) simultaneously. Assuming rapid equilibrium of reactions marked by a star symbol in Eqs. (16a–e), complicated mechanisms of transcription and translation are collapsed into a single reaction

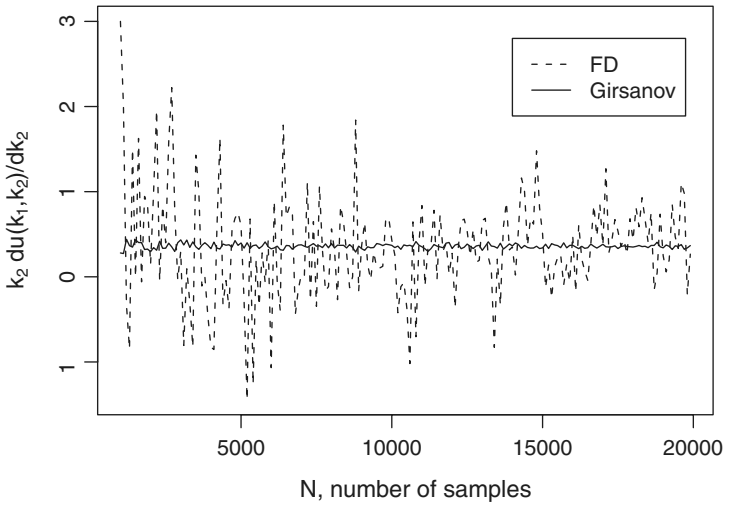

Fig. 3a. Comparison of computation of the sensitivity $k_2\frac{\partial u}{\partial k_2}$ via the finite difference scheme (labeled "FD") (2a) and the weighted scheme based on the Girsanov transformation for the non-smooth functional $F(X(\cdot)) = 1_{\{a\leqslant X(T)\leqslant b\}}$, $a = 12$, $b = 13$. The estimate of the sensitivity from the direct method at $N = 10^4$ samples is $k_2\frac{\partial u}{\partial k_2} \approx 0.3315$ (variance 0.3047) and that from the weighted scheme is $k_2\frac{\partial u}{\partial k_2} \approx 0.3653$ (variance 0.0008). The weighted scheme converges faster and with less variance to the sensitivity estimate.

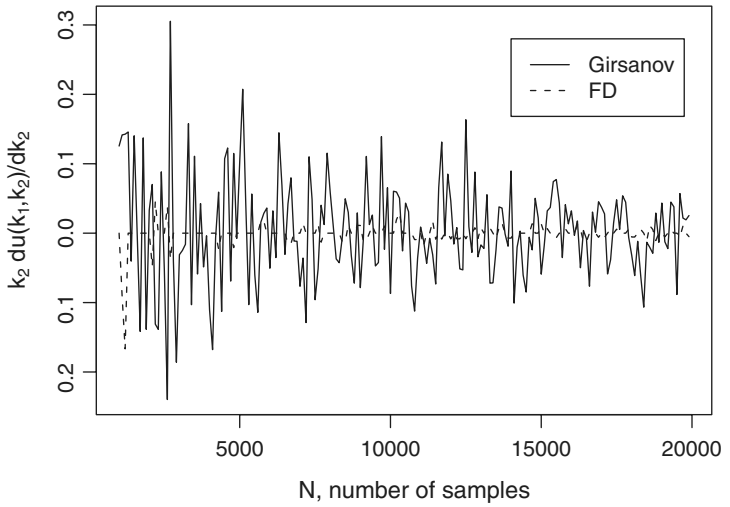

Fig. 3b. Comparison of computation of the sensitivity $k_2\frac{\partial u}{\partial k_2}$ via the finite difference scheme (labeled "FD") (2a) and the weighted scheme based on the Girsanov transformation for the non-smooth functional $F(X(\cdot)) = 1_{\{a\leqslant X(T)\leqslant b\}}$, $a = 1$, $b = 100$. The estimate of the sensitivity from the direct method at $N = 10^4$ samples is $k_2\frac{\partial u}{\partial k_2} \approx 0.00012$ (variance 5.7×10^{-5}) and that from the weighted scheme is $k_2\frac{\partial u}{\partial k_2} \approx -0.0103$ (variance 0.0039). The finite difference (FD) scheme converges faster and with less variance to the sensitivity estimate.

(16f): for example, RNAp dissociation events also simultaneously produce N Cro protein monomers. The following thermodynamical model for the activity of the P_R promoter (16) can be used [32]:

$$a_1(X_1, X_2, X_3) = k_1 N \frac{\frac{X_3}{V} e^{-\Delta G_{RNAp}/RT}}{Z}, \tag{17}$$

$$Z = 1 + \frac{X_2}{V} e^{-\frac{\Delta G_{CroO_R1}}{RT}} + \frac{X_2}{V} e^{-\frac{\Delta G_{CroO_R2}}{RT}} + \frac{X_2^2}{V} e^{-\frac{\Delta G_{CroO_R12}}{RT}} + \frac{X_3}{V} e^{-\Delta G_{RNAp}/RT}, \tag{18}$$

where the variables $X_{1,2,3}$ denote numbers of Cro monomers, Cro dimers and RNAp, respectively, present in the system at the given moment of time and the rest of the parameters are identified in Table 1. The activity

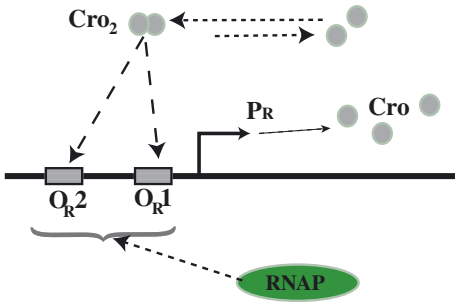

Fig. 4. Schematic representation of the processes underlying the activity of P_R promoter.

Table 1
Parameters for thermodynamical model Eq. (18) of P_R activity. Number of RNAp molecules is assumed to be constant and equal 30 [39,31]

Parameter	Value	Annotation
k_1	$0.12\ s^{-1}$	RNAp–DNA open complex formation rate [38]
k_2	$0.0025\ s^{-1}$	Cro decay rate
k_3	$0.00037\ s^{-1}$	Dimerization rate
k_4	$0.10\ s^{-1}$	Monomerization rate
N	5	Number of Cro-monomers per RNA transcript
V	$10^9\ M^{-1}$	*E. coli* volume
RT	0.617 kcal/(mol K)	Temperature
ΔG_{RNAp}	$-20.0RT$	Gibbs energy of RNAp–P_R interaction [39]
ΔG_{CroO_R1}	$-17.5RT$	Gibbs energy of Cro_2–O_R1 interaction [39]
ΔG_{CroO_R2}	$-17.5RT$	Gibbs energy of Cro_2–O_R2 interaction [39]
ΔG_{CroO_R12}	$\Delta G_{CroO_R1} + \Delta G_{CroO_R2}$	

defined above is essentially proportional to the probability that the RNAp is bound to the promoter. Parameters of the model obtained from the existing literature are summarized in Table 1.

In addition to the production of the Cro monomers there are processes of decay, dimerization and monomerization reactions:

$$\mathrm{Cro} \xrightarrow{a_2} \ldots, \tag{19a}$$

$$2\mathrm{Cro} \xrightarrow{a_3} \mathrm{Cro}_2, \tag{19b}$$

$$\mathrm{Cro}_2 \xrightarrow{a_4} 2\mathrm{Cro}, \tag{19c}$$

taking place with rates:

$$a_2(X_1, X_2, X_3) = k_2 X_1, \tag{20a}$$

$$a_3(X_1, X_2, X_3) = \frac{k_3}{2} X_1(X_1 - 1), \tag{20b}$$

$$a_4(X_1, X_2, X_3) = k_4 X_2. \tag{20c}$$

Fig. 5 presents an instance of the stochastic trajectories of Cro and Cro_2. In this example, we will be interested in estimation of the statistical properties of the random time when the number of Cro dimers will cross a certain level $Cro_2 = b$ for a first time:

$$F(\mathbf{X}(\cdot)) = \tau_b = \inf\{t > 0 : \ X_2(t) = b\} \tag{21}$$

Clearly, τ_b depends on the whole trajectory prior to the moment when $X_2 = b$ and an analytical expression for distribution of the quantity τ_b is hard to obtain given the quite complex nature of the model. Monte Carlo simulation provides us with the empirical distribution of this quantity as shown in Fig. 6.

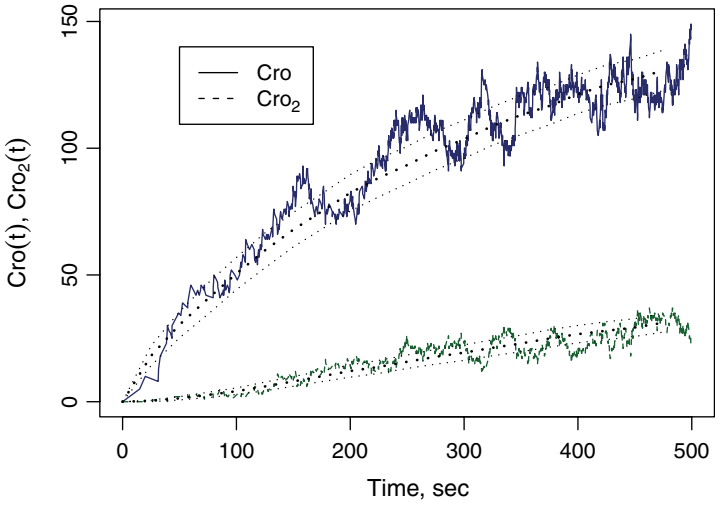

Fig. 5. Stochastic trajectories of the system corresponding to initial conditions $\mathrm{Cro}(0) = 0$, $\mathrm{Cro}_2(0) = 0$ and $\mathrm{RNAP} = 30$. Mean and variance of trajectories are also shown.

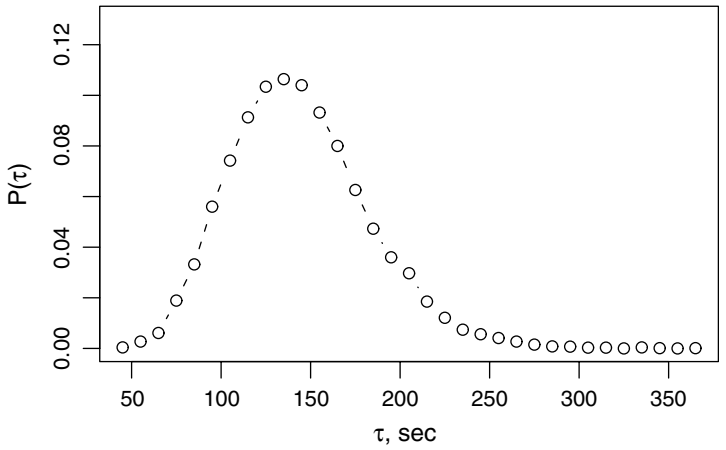

Fig. 6. Empirical distribution of the hitting time τ_b given the parameters of the model (Table 1) for $b = 10$.

Assume now that some parameters of the model are uncertain, which is usually the case in a current modeling study of transcriptional control. Specifically, we investigate the dependence of the mean value of the time τ_b,

$$u = \mathbb{E}(\tau_b | X_1(0) = 0, X_2(0) = 0)$$

with respect to perturbation in the parameters of protein–DNA interactions shown in Table 1:

$$\frac{\Delta G_{\mathrm{CroO_R1}}}{RT} \to \frac{\Delta G_{\ \mathrm{CroO_R1}}(1+\epsilon)}{RT}, \tag{22a}$$

$$\frac{\Delta G_{\mathrm{RNAp}}}{RT} \to \frac{\Delta G_{\ \mathrm{RNAp}}(1+\epsilon)}{RT}. \tag{22b}$$

The estimation of the derivatives $\frac{\partial u}{\partial \epsilon}$ for the case of (22) requires use of a more general perturbation scheme:

$$a_1^{\epsilon}(\mathbf{X}) \approx a_1(\mathbf{X})\left(1 + \epsilon \frac{\partial \ln a_1^{\epsilon}(\mathbf{X})}{\partial \epsilon}\right) = a_1(\mathbf{X})(1 + \epsilon g_1(\mathbf{X})), \quad \mathbf{X} = (X_1, X_2, X_3). \tag{23}$$

This scheme can be viewed as a Taylor expansion of the transition rate $a_1^{\epsilon}(\mathbf{X})$ in parameter ϵ. In the case of Eq. (22) we perturb only the reaction rate in the first reaction channel ($r = 1$, production of Cro), and functions $g_1(\mathbf{X})$ are:

$$g_1^{\mathrm{RNAp}}(\mathbf{X}) = \frac{\Delta G_{\mathrm{RNAp}}}{RT}\frac{\frac{X_3}{V}\mathrm{e}^{-\frac{\Delta G_{\mathrm{RNAp}}}{RT}}}{Z} - \frac{\Delta G_{\mathrm{RNAp}}}{RT}, \tag{24a}$$

$$g_1^{\mathrm{CroO_R1}}(\mathbf{X}) = -\mathrm{e}^{-\frac{\Delta G_{\mathrm{CroO_R1}}}{RT}}\frac{\Delta G_{\mathrm{CroO_R1}}}{RT}\frac{\frac{X_2}{V} + (\frac{X_2}{V})^2\mathrm{e}^{-\frac{\Delta G_{\mathrm{CroO_R1}}}{RT}}}{Z} \tag{24b}$$

For obvious reasons, it is hard to obtain the exact values of sensitivities $\frac{\partial \mathbb{E}(\tau_b)}{\partial \epsilon}$ but, intuitively, we can say that any $\epsilon > 0$ leads to an increase of Cro production rate a_1 (decrease in $\mathbb{E}(\tau_b)$) for the perturbation in ΔG_{RNAp} and to a decrease in a_1 (increase in $\mathbb{E}(\tau_b)$) for the perturbation in $\Delta G_{\mathrm{CroO_R1}}$, i.e. $\frac{\partial \mathbb{E}(\tau_b)}{\partial \epsilon} < 0$ for Eq. (24a), and $\frac{\partial \mathbb{E}(\tau_b)}{\partial \epsilon} > 0$ for Eq. (24b).

Estimation of derivatives of the mean hitting time $\mathbb{E}(\tau_b)$ for $b = 10$ calculated using methods based on the symmetric difference scheme and our new path-weighted method are presented in Figs. 7a and 7b. Estimation of sensitivities based on the weighted scheme based on $N = 10^4$ samples provides us with the estimate $\frac{\partial \mathbb{E}(\tau_b)}{\partial \epsilon} = -193.09 \pm 11.01$ s for the sensitivity with respect to variation in ΔG_{RNAp}, and $\frac{\partial \mathbb{E}(\tau_b)}{\partial \epsilon} = 11.45 \pm 0.90$ s for the sensitivity with respect to variation in ΔG_{Cro}. The direct estimator based on the finite-difference scheme does not produce a reliable estimate and its variance is several orders of magnitude higher than the variance of the path-weighted method.

In the case of smooth functionals, the application of the stochastic weights method only provides a computational speed-up in comparison to the finite difference method since there only half as many operations as required. However, for the discontinuous $F(\cdot)$, the finite difference approximation produces large errors since the contribution to be averaged is either zero or one. In this case, the weight method smooths the expectation of the functional and also provides an additional two-fold increase in computational speed since the sampling of trajectories is performed only once. We have also demonstrated that for certain classes of problems the direct estimator does not provide reliable results while the method based on path-sample weighting scheme demonstrates good convergence.

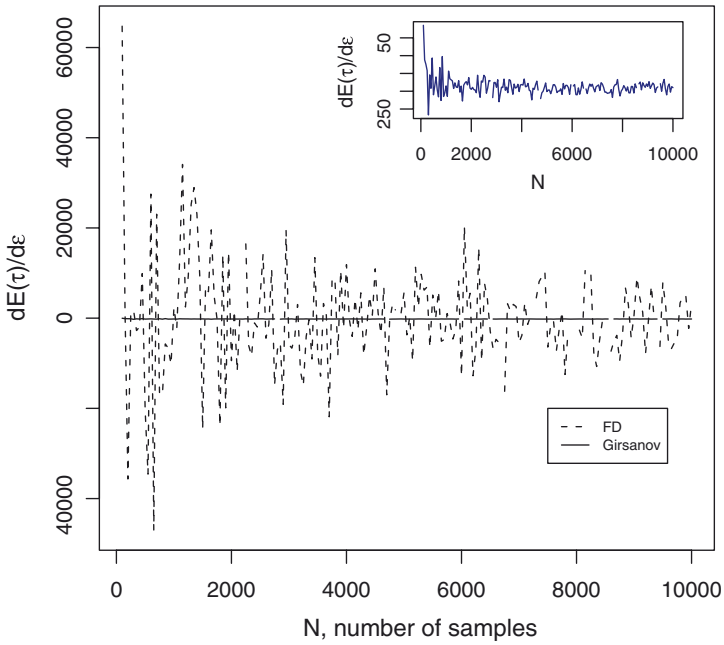

Fig. 7a. Comparison of the symmetric finite difference estimator (FD) Eq. (2b) with $\epsilon = 10^{-4}$ and the weighted scheme based on the Girsanov transformation (also shown in insert plot with different scale) for the estimation of derivative $\frac{\partial u}{\partial \epsilon} = \frac{\partial \mathbb{E}(\tau_b)}{\partial \epsilon}$, $\frac{\Delta G_{\mathrm{RNAp}}}{RT} \to \frac{\Delta G_{\mathrm{RNAp}}(1+\epsilon)}{RT}$. The weighted scheme provides us with the estimate $\frac{\partial \mathbb{E}(\tau_b)}{\partial \epsilon} = -193.09 \pm 11.01$s.

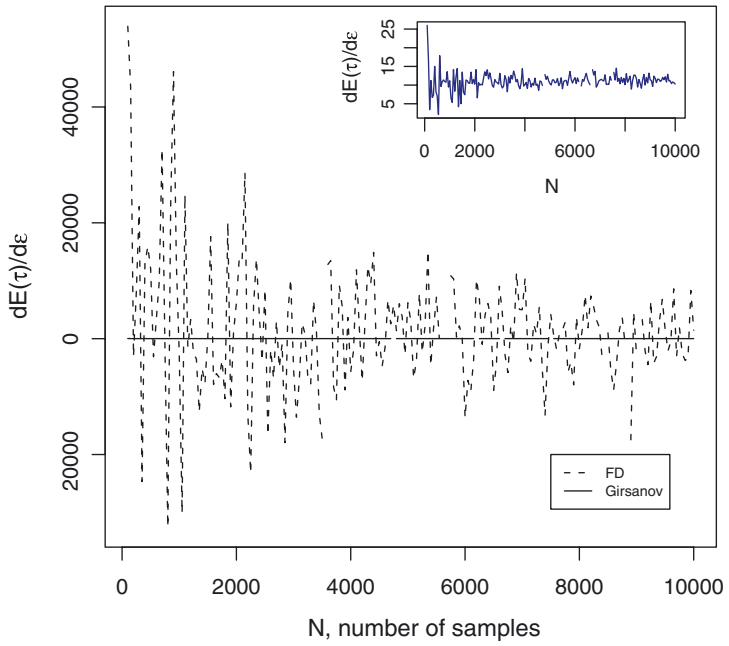

Fig. 7b. Comparison of the symmetric finite difference estimator (FD) Eq. (2b) with $\epsilon = 10^{-4}$ and the weighted scheme based on the Girsanov transformation (also shown in insert plot with different scale) for the estimation of derivative $\frac{\partial u}{\partial \epsilon} = \frac{\partial \mathbb{E}(\tau_b)}{\partial \epsilon}$, $\frac{\Delta G_{\mathrm{CroO_R}1}}{RT} \to \frac{\Delta G_{\mathrm{CroO_R}1}(1+\epsilon)}{RT}$. The weighted scheme provides us with the estimate $\frac{\partial \mathbb{E}(\tau_b)}{\partial \epsilon} = 11.45 \pm 0.90$s.

6. Conclusions

Given the uncertainty in the parameters of the models of many system and given the desire to determine the points of a system susceptible to failure or external control, sensitivity analysis provides a powerful addition to the model analysis arsenal. For stochastic systems, the computational expense of such sensitivity analysis has been prohibitive and few models, let alone stochastic models, explore the dependence of model predictions on variation of the underlying parameters. Here we develop an efficient and accurate algorithm for estimating these sensitivities. This can be accomplished efficiently by derivation of an appropriate expression for the sensitivity derivatives through introduction of stochastic weights that are derived through use of the Girsanov measure transformation. For chemical stochastic systems, the calculation involves adding just two more steps to the inner loop of a traditional stochastic simulation algorithm [14]. Using the appropriate expression for the sensitivity coefficients we have demonstrated that our new method converges more quickly and provides more robust estimate of the sensitivity coefficient for a systems with non-smooth features.

Acknowledgements

We would like to acknowledge DARPA grant # BAA-01-26-0126517 for support during the course of this research. APA would also like to acknowledge the support of the Howard Hughes Medical Institute. SP would like to acknowledge for the discussions with Dr. S. Andrews, T. Ham, A. Knizhnik and two anonymous reviewers for many helpful suggestions.

Appendix A

The usual approach to analyzing stochastic chemical systems is based on the Chemical Master Equation (CME) description (see for example [26,15,14]). We shall take an approach utilizing a path sampling representation [33]. It can be shown that this approach is equivalent to CME. In particular, the dynamics of the components of the state vector $\mathbf{X}$ can be expressed via counting processes $N_r(t)$:

$$X_i(t) = X_i(0) + \sum_{r=1}^{R} \Delta v_{ir} N_r(t). \tag{25}$$

The increment of the reaction counting process $N_r(t)$ at some generic moment of time t depends only on the state of the system $\mathbf{X}(t_-)$ just before this moment of transition $t_- = \lim_{\delta \to +0}(t-\delta)$ and is independent of the whole history of the process $\mathbf{X}(t)$ up to time t, i.e. independent of the previous history of the process:

$$\mathbb{E}(N_r(t+\Delta t) - N_r(t) | \mathbf{X}(t_-)) = a_r(\mathbf{X}(t_-))\Delta t + \mathrm{O}(\Delta t^2). \tag{26}$$

As outlined above, we consider the family of solutions $\{\mathbf{X}^\epsilon(t)\}_{t \geqslant 0}$ parameterized by ϵ, $|\epsilon| < 1$. Keeping only low-order expansion with respect to the parameter ϵ one can find that perturbation "tunes" the reaction rates $a_r(\mathbf{X})$ by changing their intensity in the following way:

$$a_r^\epsilon(\mathbf{X}) = a_r(\mathbf{X})(1 + \epsilon g_r(\mathbf{X})) = a_r(\mathbf{X}) Y_r^\epsilon(\mathbf{X}), \tag{27}$$

where we assume that function $Y_r^\epsilon(\mathbf{X}) = 1 + \epsilon g_r(\mathbf{X})$ is nonnegative. This condition can hold true for any bounded function $g(\cdot)$ if the parameter ϵ is chosen to be small enough in absolute value.

As noted before, the dependence of the intensity $a_r(\cdot)$ of the driving process on parameters of the model, ϵ, and the state $\mathbf{X}^\epsilon(t)$ makes the direct sensitivity analysis very difficult. In particular, it is not evident how to compare the influence of two different values of ϵ since different values may generate completely different trajectories. It appears useful to search for an equivalent problem setting where the driving process is *parameter independent*. The Girsanov's measure transformation gives such opportunity [33,34,24]. In probability theory, Girsanov's theorem determines how stochastic processes change under changes in measure [35,33,24]. The theorem is especially important since it tells how to convert from the physical measure which describes the probability that an underlying process $\mathbf{X}(t)$, used in the calculation of $u(\mathbf{k}, \mathbf{x}) = \mathbb{E}[F(\mathbf{X}(\cdot))]$, will take a particular value or values to a different probability measure P^ϵ, corresponding to the perturbed parameters of the model. This is a useful tool for evaluating the value of sensitivity derivatives since it allows us to construct family of the probability measures for the perturbation (27). We first construct the martingale process $M_r(t) = N_r(t) - \int_0^t a_r(\mathbf{X}(s))\,\mathrm{d}s$, i.e. $\mathbb{E}(M_r(t)) = M_r(0) = 0$. If $Y_r^\epsilon > 0$ then process $Z^\epsilon(t)$ can be viewed as a solution of the stochastic differential equation

$$Z^\epsilon(t) = 1 + \int_0^t (Y_r^\epsilon(s_-) - 1) Z^\epsilon(s_-) \mathrm{d}M_r(s), \quad Z^\epsilon(0) = 1, \tag{28a}$$

and $Z^\epsilon(t)$ is also a martingale with respect to P:

$$\mathbb{E}[Z^\epsilon(t)] = 1 \tag{28b}$$

One can interpret Z^ϵ as a Radon–Nykodim derivative [35,24] which relates measures P^ϵ and P corresponding to the perturbed ($\epsilon \neq 0$) and unperturbed ($\epsilon = 0$) trajectories:

$$\frac{\mathrm{d}P^\epsilon}{\mathrm{d}P} = Z^\epsilon = \prod_{s \leqslant t} Y_r^\epsilon(s)^{\Delta N_r(s)} \exp\left(- \int_0^t (Y_r^\epsilon(s) - 1) a_r(\mathbf{X}(s))\,\mathrm{d}s \right) \tag{29}$$

and process $M_r^\epsilon(t) = N_r^\epsilon(t) - \int_0^t Y_r^\epsilon(s) a_r(\mathbf{X}^\epsilon(s))\,\mathrm{d}s$ is a martingale under measure P^ϵ. Our objective is the estimation of the following expression and its derivatives:

$$u^\epsilon(\mathbf{x}) = \mathbb{E}^\epsilon[F(\mathbf{X}^\epsilon(\cdot)) | \mathbf{X}^\epsilon(0) = \mathbf{x}] \tag{30}$$

While estimation of $u^\epsilon(\mathbf{x})$ can be performed in a straightforward manner, the following result gives us the derivative of $u^\epsilon(\mathbf{x})$ with respect to parameter ϵ around $\epsilon = 0$ in terms of the path-dependent weight $W(\cdot)$:

$$\frac{\partial}{\partial \epsilon} u^\epsilon(\mathbf{x}) = \mathbb{E}\left[F(\mathbf{X}(\cdot)) \frac{\partial Z^\epsilon}{\partial \epsilon} | \mathbf{X}(0) = \mathbf{x} \right], \tag{31a}$$

$$W \triangleq \frac{\partial Z^\epsilon}{\partial \epsilon}|_{\epsilon=0} = \sum_{s \leqslant t} g_r(\mathbf{X}(s)) \Delta N_r(s) - \int_0^t g_r(\mathbf{X}(s)) a_r(\mathbf{X}(s))\,\mathrm{d}s = \int_0^t g_r(\mathbf{X}(s)) M_r(t). \tag{31b}$$

It is not hard to see how high order derivatives given by Eq. (11) can be computed in a similar manner.

Appendix B. Asymptotic behavior of the weighted estimator

A weighted estimator is in many ways similar to the typical estimator in importance sampling framework [36,37], in particular, it is unbiased due to the statistical properties of the likelihood process $Z^{\epsilon}(t)$ independently of the properties of the function $F(\cdot)$. Variance of the weighted estimator merits a separate discussion. Consider, for example, function $F(\cdot)$ which depends only on the final value of the process $\mathbf{X}(T)$ at the moment $t = T$: $F = F(\mathbf{X}(T))$. The Lindeberg central limit theorem for i.i.d. random variables can be invoked as follows:

$$\frac{1}{N}\sum_{j=1}^{N}(F^{(j)}W^{(j)} - \mathbb{E}[F(\mathbf{X}(T))W(T)]) \Rightarrow z\frac{\sqrt{\mathrm{Var}[F(\mathbf{X}(T))W(T)]}}{\sqrt{N}} \tag{32a}$$

where z is standard Gaussian random variable, $z \propto N(0,1)$. By the property of conditional expectation the variance of the random variable $F(\mathbf{X}(T))W(T)$ is the same as variance of the product of two independent random variables $F(\mathbf{X}(T))w(\mathbf{X}(T))$ where $w(\mathbf{X}(T))$ stands for

$$w(\mathbf{X}(T)) = \mathbb{E}[W(T)|\mathbf{X}(T)] \tag{32b}$$

i.e. the conditional expectation with respect to the terminal value. Variance in Eq. (32a) is given by:

$$\mathrm{Var}[F(\mathbf{X}(T))W(T)] = \mathrm{Var}[F(\mathbf{X}(T))]\mathrm{Var}[w(\mathbf{X}(T))] + \mathrm{Var}[F(\mathbf{X}(T))]\mathbb{E}[w(\mathbf{X}(T))]^2 + \mathrm{Var}[w(\mathbf{X}(T))]\mathbb{E}[F(\mathbf{X}(T))]^2. \tag{32c}$$

One can see that variance of the estimator given by (32c) depends on both the variance of the Monte Carlo estimator of the function F, $\mathbb{E}[F(\mathbf{X}(T))]$ and variance of the stochastic weights $\mathrm{Var}[w(\mathbf{X}(T))]$. If $\mathrm{Var}[F(\mathbf{X}(T))]$ is small then variance of the weights dominates expression (32c). This explains why traditional finite difference scheme might have an advantage in estimation of the sensitivity of smooth functions.

References

[1] A.N. Shiryaev, Essentials of Stochastic Finance: Facts, Models, Theory. Applied Series on Statistics Science and Applied Probability, vol. 3, World Scientific, 2004.
[2] P. Boyle, M. Broadie, P. Glasserman, Monte Carlo methods for security pricing, JEDC 21 (1997) 1267–1321.
[3] H.H. McAdams, A.P. Arkin, Stochastic mechanisms in gene expression, PNAS 94 (3) (1997) 814–819.
[4] E. Ozbudak, M. Thattai, I. Kurtserand, A.D. Grossman, A. van Oudenaarden, Regulation of noise in the expression of a single gene, Nature Genetics 31 (2002) 69.
[5] C.V. Rao, D.M. Wolf, A.P. Arkin, Control, exploitation and tolerance of intracellular noise, Nature 420 (6912) (2002) 231–237.
[6] M.B. Elowitz, A. Levine, E.D. Siggia, P.S. Swain, Stochastic gene expression in a single cell, Science 297 (16) (2002) 1183–1186.
[7] W.J. Blake, M. Kaern, C.R. Cantor, J.J. Collins, Noise in eukaryotic gene expression, Nature 422 (2003) 633–637.
[8] N. Rosenfeld, J.W. Young, U. Alon, P.S. Swain, M.B. Elowitz, Gene regulation at the single-cell level, Science 307 (5717) (2005) 1962–1965.
[9] L. Weinberger, J.C. Burnett, J.E. Toettcher, A.P. Arkin, D.V. Schaffer, Stochastic gene expression in a lentiviral positive-feedback loop: HIV-1 Tat fluctuations drive phenotypic diversity, Cell 122 (2005) 169–182.
[10] T.O. Drews, R.D. Braatz, R.C. Alkire, Parameter sensitivity analysis of Monte Carlo simulations of copper electrodeposition with multiple additives, J. Electrochem. Soc. 150 (11) (2003) C807.
[11] E.L. Haseltine, D.B. Patience, J.B. Rawlings, On the stochastic simulation of particulate systems, Chem. Eng. Sci. 60 (2005) 2627–2641.
[12] S. Raimondeau, P. Aghalayam, A.B. Mhadeshwar, D.G. Vlachos, Parameter optimization in molecular models: Application to surface kinetics, Ind. Eng. Chem. Res. 42 (2003) 1174.
[13] D.E. Zak, G.E. Gonye, J.S. Schwaber, F.J. Doyle III, Importance of input perturbations and stochastic gene expression in the reverse engineering of genetic regulatory networks: insights from an identifiability analysis of an in silico network, Genome Res. 13 (11) (2003) 2396–2405.
[14] D.T. Gillespie, Exact simulation of coupled chemcial reactions, J. Phys. Chem. 81 (25) (1977) 2340–2361.
[15] D.T. Gillespie, Markov Processes: An Introduction for Physical Scientists, Academic, San Diego, 1992.
[16] P.E. Kloeden, E. Platen, Numerical Solution of Stochastic Differential Equations, Springer, Berlin, 1992.
[17] A. Varma, M. Morbidelli, H. Wu, Parametric Sensitivity in Chemical Systems, Cambridge University Press, Cambridge, UK, 1999.
[18] P.W. Glynn, Optimization of stochastic systems via simulations, in: Proceedings of the 1989 Winter simulation Conference, Society for Computer Simulation, San Diego, 1989, pp. 90–105.
[19] P. L'Ecuyer, Convergence rates for steady-state derivative estimators, Ann. Operat. Res. 39 (1) (1992) 121–136.

[20] P. Glynn, Likelihood ratio derivative estimators for stochastic systems, in: Proceedings of the 1989 Winter Simulation Conference, 1989, pp. 374–380.
[21] M. Broadie, P. Glasserman, Estimating Security price Derivatives Using Simulation, Manag. Sci. 42 (2) (1996) 269–285.
[22] R. Gunawan, Y. Cao, L. Petzold, F.J. Doyle III, Sensitivity analysis of discrete stochastic systems, Biophys. J. 88 (2005) 2530–2540.
[23] D. Nualart, The Malliavin Calculus and Related Topics, Springer, Berlin, 1995.
[24] K. Bichteler, J.-B. Gravereaux, J. Jacod, Malliavin Calculus for Processes with Jumps. Stochastic Monographs, vol. 2, Gordon and Breach Science Publishers, 1987.
[25] E. Fournié, F. Lasry, J. Lebuchoux, J. Lions, P. Touzi, An application of Malliavin calculus to Monte Carlo methods in fnance, Finance Stochast. 3 (1999) 391–412.
[26] N.G. van Kampen, Stochastic Processes in Physics and Chemistry, North-Holland Publ. Co., 1992.
[27] S. Ethier, T. Kurtz, Markov Processes: Characterization and Convergance, John Wiley & Sons, 1986.
[28] A.B. Bortz, M. Kalos, J.L. Lebowitz, A new algorithm for Monte Carlo simulation of Ising spin systems, J. Comput. Phys. 17 (1975) 10–18.
[29] C.W. Gardiner, Handbook of Stochastic Methods, Springer-Verlag, 1985.
[30] M. Ptashne, A Genetic Switch: Phage λ and Higher Organisms, Blackwell Scientific Publications and Cell Press, 1992.
[31] A.P. Arkin, J. Ross, H.H. McAdams, Stochastic kinetic analysis of pathway bifurcation in phage λ-infected *Escherichia coli* cells, Genetics 149 (4) (1998) 1633–1648.
[32] M.A. Shea, G.K. Ackers, The O_R control system of bacteriophage lambda: A physical-chemical model for gene regulation, J. Mol. Biol. 181 (1985) 211–230.
[33] I.I. Gihman, A.V. Skorohod, Stochastic Differential Equations, Springer-Verlag, New York, 1972.
[34] I.I. Gihman, A.V. Skorohod, Controlled Stochastic Processes, Springer-Verlag, New York-Heidelberg, 1979.
[35] E. Wong, B. Hajek, Stochastic Processes in Engineering Systems, Springer-Verlag, New York, 1985.
[36] J.A. Bucklew, Large Deviation Techniques in Decision, Simulation and Estimation, John Wiley, New York, 1990.
[37] J.S. Liu, Monte Carlo Strategies in Scientific Computing, first ed., Springer, 2002, October.
[38] D.K. Hawley, W.R. McClure, The effect of a λ repressor mutation on the activation of transcription initiation from the λP_{RM} promoter, Cell 32 (1983) 327–333.
[39] P.J. Darling, J.M. Holt, G.K. Ackers, Coupled energetics of the λ Cro repressor self-assembly and site-specific DNA Operator Binding: Cooperative interaction of Cro dimers, JMB 302 (2000) 625–638.