Sensitivity Estimation and Inverse Problems in Spatial Stochastic Models of Chemical Kinetics

Pavol Bauer and Stefan Engblom

Abstract We consider computational stochastic modeling of diffusion-controlled reactions with applications mainly in molecular cell biology. A complication from the traditional 'well-stirred' case is that our models have a spatial dimension. Our aim here is to put forward a practical algorithm by which perturbations can be propagated through these types of simulations. This is important since the quality of experimental data calls for frequently estimating stability constants. Another use is in inverse formulations which generally relies on being able to effectively and accurately judge the effects of small perturbations. For this purpose we present our implementation of an "all events method" and give two concrete examples of its use. One case studied is the effect of stochastic focusing in the spatial setting, the other case treats the optimization of a small biochemical network.

1 Introduction

In the classical case of non-spatial stochastic modeling of chemical kinetics, the *reaction rates* are understood as transition intensities in a continuous-time Markov chain $X_{t\geq 0}$. When spatial variability is important, space may be discretized in voxels. Between voxels, *diffusion*-, or more generally, *transport rates* become transition intensities in a Markov chain which now takes place in a much larger state space.

This is the point of view taken in the software framework URDME [2, 5] where fairly large-scale spatial stochastic reaction-diffusion models can be simulated. We have developed a solver for sensitivity analysis which allows us to compare single trajectories under arbitrary perturbations of input data and opens up for computing stability estimates as well as optimizing models under various conditions.

P. Bauer • S. Engblom (✉)
Division of Scientific Computing, Department of Information Technology, Uppsala University, Uppsala, Sweden
e-mail: pavol.bauer@it.uu.se; stefane@it.uu.se

A. Abdulle et al. (eds.), *Numerical Mathematics and Advanced Applications - ENUMATH 2013*, Lecture Notes in Computational Science and Engineering 103, DOI 10.1007/978-3-319-10705-9_51

For a given parameter perturbation $c \to c+\delta$ the task is to characterize the mean effect on some function of interest,

$$E[f(X(T, c+\delta)) - f(X(T, c))], \tag{1}$$

for example, by computing a sample average. As a prototypical application, c is a rate constant and f a measure of the molecular population X_t.

The obvious way to carry out this is to conduct two Monte Carlo simulations using independent random numbers and generating N trajectories each of $f(X(T, c+\delta))$ and $f(X(T, c))$, and then taking the average. Two factors can lead to unsatisfactory results with this approach. Firstly, with independent samples, the variance of $f(X(T, c+\delta))$ and $f(X(T, c))$ can be large compared to the difference $f(X(T, c+\delta)) - f(X(T, c))$. This is the *variance reduction problem* which has been discussed in the well-stirred setting by others [12]. Secondly, a slightly more subtle point has been pointed out in [12]; solver algorithms related to Gillespie's Direct Method [8] are not suitable to compute the difference between two processes $X(T, c+\delta)$ and $X(T, c)$ as their coupling is simply not the intended one. This would be a problem for instance, if (1) were to be replaced by

$$E[f(X(T, c+\delta) - X(T, c))], \tag{2}$$

and f some nonlinear function. In fact, a popularly used algorithm for solving spatial stochastic models, the Next Subvolume Method (NSM) [3], belongs to this class of algorithms and can therefore not be used.

In this paper we present the "All Events Method"; a variant of the so-called Common Reaction Path method [12], extended for spatial models in URDME and meeting both the criteria above for an efficient and sound estimation of (1)–(2). In Sect. 2 we give a brief overview of the modeling involved, in the non-spatial as well as in the fully spatial setting, and we also sketch a theory for perturbations, including some implementation aspects of our AEM-solver. In Sect. 3 we discuss two applied examples and show how this solver can be conveniently used in the sense of both forward- and backward formulations.

2 A Viable "All Events Method"-Implementation

After a brief review of stochastic reaction-diffusion modeling we will here summarize the logic behind URDMEs AEM-solver.

2.1 Spatial Stochastic Chemical Kinetics

According to classical *well-stirred* stochastic modeling of chemical kinetics, reactions are transitions between states $x \in \mathbf{Z}_+^D$, counting the number of molecules of each of D distinct species. The transition intensity defines the probability per unit of time for the transition from the state x to $x + S_r$;

$$x \xrightarrow{w_r(x)} x + S_r, \tag{3}$$

where the transition vector $S_r \in \mathbf{Z}^D$ is the rth column in the *stoichiometric matrix* S. Equation (3) defines a continuous-time Markov chain $X_{t\geq 0}$ on $\mathbf{Z}_+^D$.

For spatially extended problems, a stochastic model can be defined by first discretizing space in voxels. Molecular transport can then be handled as a "reaction" which brings a molecule of the lth species from voxel i to j,

$$X_{li} \xrightarrow{a_{ij}\mathbf{x}_{li}} X_{lj}, \tag{4}$$

where $\mathbf{x}_{li}$ is the number of molecules of species l in subvolume i. When space is discretized by general unstructured meshes, suitable rate constants can be obtained by a numerical discretization of the diffusion equation. The consistency in this approach hinges on the fact that the expected value of the concentration converges to the deterministic numerical solution [5].

2.2 Path-Wise Analysis of Perturbations

Without loss of generality, we consider the well-stirred case (3). Let the *state* $X(t) \in \mathbf{Z}_+^D$ count the number of molecules of the D species. The associated Markov chain can be written in the convenient jump SDE form

$$dX_t = S\boldsymbol{\mu}(dt), \tag{5}$$

with counting measure $\boldsymbol{\mu} = [\mu_1, \ldots, \mu_R]^T$. According to this compact notation the time to the arrival of the next reaction of type r is exponentially distributed with intensity $w_r(X_{t-})$. A perhaps more familiar notation is Kurtz's *random time change representation* [6, Chap. 6.2], in which the path is characterized in terms of unit-rate Poisson processes Π_r,

$$X_t = X_0 + \sum_{r=1}^{R} S_r \Pi_r \left(\int_0^t w_r(X_{s-})\, ds \right). \tag{6}$$

This naturally gives rise to the term *operational time* for the argument to each of the R Poissonian processes.

Let a trajectory $Y(t)$ be a *perturbed* version of $X(t)$ in the sense that the former is driven by modified rates $v_r(Y_t)$, but otherwise has an identical reaction topology S. To compare the two trajectories we write

$$dX_t = S\left[\boldsymbol{\mu}^{(0)}(w(X_{t-}), v(Y_{t-});\ dt) + \boldsymbol{\mu}^{(\delta)}(w(X_{t-}), v(Y_{t-});\ dt)\right], \tag{7}$$

$$dY_t = S\left[\boldsymbol{\mu}^{(0)}(w(X_{t-}), v(Y_{t-});\ dt) + \boldsymbol{\mu}^{(\delta)}(v(Y_{t-}), w(X_{t-});\ dt)\right], \tag{8}$$

in terms of the *base* (superscript 0) and *remainder* counting measures (superscript δ), respectively. The intensities for $\mu_r^{(0)}$ and $\mu_r^{(\delta)}$ are given by

$$w_r(x) \wedge v_r(y) \text{ and } w_r(x) - (w_r(x) \wedge v_r(y)). \tag{9}$$

As indicated in the order of the arguments in (7) and (8), there is an asymmetry in the remainder measure.

To analyze $Z_t := \|X_t - Y_t\|^2$ we apply a form of Itô's formula [1, Chap. 4.4.2],

$$dZ_t = 2(X_{t-} - Y_{t-})^T\, S[\boldsymbol{\mu}_{w,v}^{(\delta)} - \boldsymbol{\mu}_{v,w}^{(\delta)}](dt) + S^2[\boldsymbol{\mu}_{w,v}^{(\delta)} + \boldsymbol{\mu}_{v,w}^{(\delta)}](dt). \tag{10}$$

Taking expectation values and ignoring the martingale part we get, after determining the drift parts of the relevant measures,

$$d/dt\, EZ_t = E\left[2(X_t - Y_t)^T\, S[w(X_t) - v(Y_t)] + S^2|w(X_t) - v(Y_t)|\right]. \tag{11}$$

At this point we need some assumption on the dynamics of the process and on the perturbation. Let the rates be *locally Lipschitz* and let the magnitude of the *relative* perturbation be δ. Then for $\|x\| \vee \|y\| \le P$,

$$\|w(x) - v(y)\| \le \|w(x) - w(y)\| + \|w(y) - v(y)\| \tag{12}$$

$$\le L_P\|x - y\| + \delta\|w(y)\| \le C_P(\delta + \|x - y\|). \tag{13}$$

Working similarly, we find from (11) that for some constant C_P,

$$d/dt\, E\|X_t - Y_t\|^2 \le E\, C_P(\delta + \|X_t - Y_t\|^2), \tag{14}$$

where we used the simple observation that for integers n, $\|n\| \le \|n\|^2$. From Grönwall's inequality, assuming $X_0 = Y_0$, we get under a stopping time $t \le \tau_P := \inf_{t\ge 0}\{\|X_t\| \vee \|Y_t\| > P\}$ that

$$E\|X_t - Y_t\|^2 \le \delta(\exp(C_P\, t) - 1). \tag{15}$$

Thus, for *bounded* systems, (15) predicts a RMS perturbation which behaves as $\delta^{1/2}$. For *unbounded* systems, the only immediate generalization is that the limit as $\delta \to 0$ is zero, see [4] and the references therein.

2.3 Simulation Using Consistent Poisson Processes

To motivate our approach to evolving two or more trajectories which can be pathwise compared, consider first the *diffusion approximation* of (5),

$$dX_t = Sw(X_t)\,dt + Sw(X_t)^{1/2}\,d\boldsymbol{W}_t. \tag{16}$$

Two comparable replicas of (16) can clearly be constructed using the *same* Wiener process $\boldsymbol{W}(t)$. In discrete time this boils down to using the same sequence of normal random numbers. This idea can be transferred to the current setting by simply using the same sequence of random numbers when simulating different trajectories, and it leads to the Common Random Numbers method [9].

However, we see from the representation (6) that two trajectories formed by identical Poisson processes are stronger candidates to being similar than any dependency on identical random numbers may generate. This is the motivation behind the Common Reaction Path method [12]. Here all R reaction channels access their own stream of random numbers such that a consistent operational time in the sense of (6) is continuously well-defined. In practise we implement this by storing generator seeds s_i for every channel i and use these for every update of the corresponding Poisson process. For the current case of *spatial* models this implies that all reaction events and all transport events must be associated with a consistent Poisson process. This in contrast to the NSM [3] where only a 'total event' process per voxel is available.

A remark on continuity is made in [12, Appendix B]. When a zero rate is encountered a discontinuity typically forms which is due to the fact that in most implementations, a zero rate will lead to discarding the previous operational time. A new, *uncorrelated* waiting time is drawn whenever the rate becomes non-zero again. In our implementation we circumvent this problem by storing the pre-zero operational time τ^{inf} and associated non-zero rate w^{inf}. When the channel is reactivated we compute the next waiting time τ^{new} using the rescaling (essentially proposed in [7]),

$$\tau^{\text{new}} = t_{\text{current}} + \left(\tau^{\text{inf}} - t_{\text{current}}\right) w^{\text{inf}}/w^{\text{new}}. \tag{17}$$

For more information on implementation of solvers in URDME, consult [2].

3 Sample Applications

We shall now consider two sample applications of our URDME solver; one example in the 'forward' mode, i.e. propagating a definite perturbation, and one example in the 'backward' (or inverse) setting. Due to the computational complexity involved,

the inverse problem we choose to consider is non-spatial. However, it is clearly possible to, at an increased computational cost, also target fully spatial formulations.

3.1 Spatial Stochastic Focusing

As a basic but informative example we consider the following enzymatic law,

$$C + E \xrightarrow{k\,c\cdot e} P + E, \tag{18}$$

in which E is an enzyme and C an intermediate complex which matures into a product P. The model is completed by adding the in- and outflow laws

$$\emptyset \underset{\beta_C c}{\overset{\alpha_C}{\rightleftharpoons}} C, \qquad \emptyset \underset{\beta_E e}{\overset{\alpha_E}{\rightleftharpoons}} E, \qquad P \xrightarrow{\beta_P p} \emptyset. \tag{19}$$

Stochastic focusing [11] is a non-linear stochastic effect under which an input signal is strongly amplified, and notably much more effectively so than for the corresponding mean field model. In the present case this effect can be observed in the response of the number of intermediate complexes C when the birth rate α_E is perturbed according to $\alpha_E \to \alpha_E(1-\delta)$ (Fig. 1, left). A *spatial* version of (18) and (19) can be defined in the geometry $\Omega = [0,1]$ with diffusion of the species. We generate an 'unperturbed' trajectory $C_1(t)$ for which $\alpha_E = c$ is constant and a 'perturbed case' $C_2(t)$ for which α_E is replaced by the space dependent function $\alpha_E(x) = c(1/2 + x)$. Note that this preserves the total production rate in the sense that

$$\int_\Omega \alpha_E(x)dV = c. \tag{20}$$

We combine the reactions with varying diffusion ε and observe a phenomenon which can be referred to as *Spatial stochastic focusing* (Fig. 1, center/right).

In the table below we determine at two different perturbations δ and for several values of TOL, the number of realizations $N = 10, 20, \ldots$ needed to bring the standard Monte Carlo error estimate $\text{std}/\sqrt{N}$ below TOL. This for the case of estimating $E[C_2(1) - C_1(1)]$ using either the Next Subvolume Method [3] or the solver proposed by us.

	NSM				AEM			
$\delta\backslash$ TOL	1/4	1/8	1/16	1/32	1/4	1/8	1/16	1/32
1/2	1,480	3,470	6,990	33,010	30	600	3,630	12,870
1/32	1,350	2,780	5,630	14,970	10	20	60	3,190

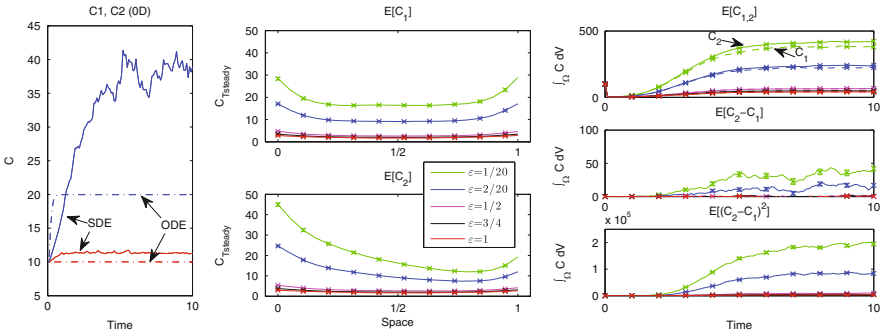

Fig. 1 *Left*: ODE and mean SDE solutions of the unperturbed (C_1, *red, lower*) and perturbed (C_2, *blue, upper*) model (18)–(19) in the well-stirred case ($\delta = 1/2$). *Center*: SDE solutions with spatial perturbation and varying diffusion. Each point represents the mean of C at steady-state. From top to bottom as in legend (Colors online). *Right*: traces of C_1 (*dashed*) and C_2 (*solid*) integrated over space and plotted over time. All SDE solutions are averages of $N = 10^4$ trajectories, error bars are std/$\sqrt{N}$

3.2 *Enzymatic Control*

Consider again the model (18)–(19) but with the enzyme E under *control*,

$$\emptyset \underset{\beta_E e}{\overset{s(t)}{\rightleftharpoons}} E, \tag{21}$$

with $s(t)$ a time-dependent signal. We define a *payoff function* $\varphi(P)$ by

$$\varphi(P) = (P - c_-)[c_- < P \leq C_+] + (C_+ - c_-)[C_+ < P] \tag{22}$$

with c_-/C_+ suitable cutoff values. Reasonable constraints are that $\|s(t)\|_\infty$ and $\|s(t)\|_1$ are bounded. After adding a regularization term the target functional becomes

$$\mathscr{M}[P] := \int_0^T \varphi(P_t)\, dt + \epsilon[s(t)]_{0 \leq t \leq T}, \tag{23}$$

with $[\cdot]$ the total variation. Thus the overall formulation is "*Find* $s(t)$ *such that in expectation,* $\mathscr{M}[P]$ *attains it maximum subject to the constraints*". Here $P = P(t)$ is the solution to (18)–(19) and (21) with $P(0) = E(0) = 0$. We solve the optimization problem both in the deterministic ODE setting and in the stochastic setting using URDME/AEM. The optimization algorithm applied was the Nelder-Mead simplex method [10] and results for two sets of cutoff-values are shown in Fig. 2.

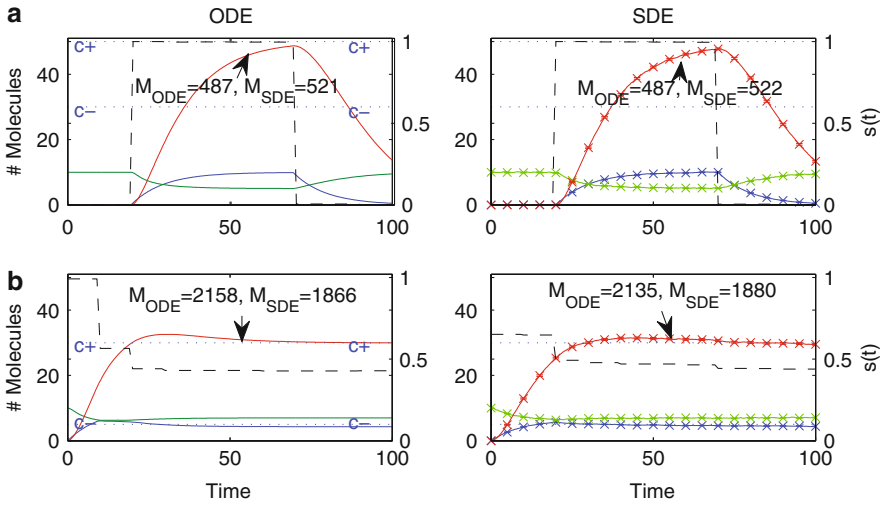

Fig. 2 (**a**) Optimal solution $s(t)$ for cutoff values $c_- = 30$ and $C_+ = 50$, for the case of a deterministic ODE (*left*), and an SDE (*right*). These cutoff values yield similar "all-or-nothing" optimal strategies in both cases. (**b**) Here $c_- = 5$ and $C_+ = 30$, and the optimal solutions are clearly different for the two cases. *Legend*: From top to bottom (colors online), P (*red*), C (*green*), E (*blue*), signal $s(t)$ (*black, dashed*). Values of the target functional for the optimal $s(t)$ are also indicated

Conclusions

We have presented a viable simulation algorithm for continuous-time Markov chains which relies upon a self-consistent use of Poisson processes. This computationally intensive technique enables perturbations in the input parameters to be propagated and opens up for several relevant applications. To the best of our knowledge none of the applications considered here have been addressed previously.

Through straightforward perturbation calculations in the 'forward' mode we have reported results for *spatial stochastic focusing*, where the strong focusing effect can be *uniquely* attributed to the spatial dimension. In a nutshell, the existence of a gradient implies an increase of outgoing products which cannot be explained through well-stirred and/or deterministic analysis.

As an example of an interesting inverse formulation we studied a simple chemical network and defined an arguably quite open criterion for optimality. By wrapping our simulator with a simple external optimizing routine we were able to find optimal control signals which realizes this optimality. In one case the signals found clearly differ from their deterministic versions.

(continued)

While developing computational algorithms simultaneously with challenging applications requires some care, it is our hope that this report shows the value of this approach.

References

1. D. Applebaum, *Lévy Processes and Stochastic Calculus*. Studies in Advanced Mathematics, vol. 93 (Cambridge University Press, Cambridge, 2004)
2. B. Drawert, S. Engblom, A. Hellander, URDME: a modular framework for stochastic simulation of reaction-transport processes in complex geometries. BMC Syst. Biol. **6**, 76 (2012)
3. J. Elf, M. Ehrenberg, Spontaneous separation of bi-stable biochemical systems into spatial domains of opposite phases. Syst. Biol. **1**, 2 (2004)
4. S. Engblom, On the stability of stochastic jump kinetics. Available at http://arxiv.org/abs/1202.3892 (2012)
5. S. Engblom, L. Ferm, A. Hellander, P. Lötstedt, Simulation of stochastic reaction–diffusion processes on unstructured meshes. SIAM J. Sci. Comput. **31**, 3 (2009)
6. S.N. Ethier, T.G. Kurtz, *Markov Processes: Characterization and Convergence*. Wiley Series in Probability and Mathematical Statistics (Wiley, New York, 1986)
7. M.A. Gibson, J. Bruck, Efficient exact stochastic simulation of chemical systems with many species and many channels. J. Phys. Chem. **104**(9), (2000)
8. D.T. Gillespie, Exact stochastic simulation of coupled chemical reactions. J. Phys. Chem. **81**, 25 (1977)
9. P. Glasserman, D.D. Yao, Some guidelines and guarantees for common random numbers. Manage. Sci. **38**, 6 (1992)
10. J.A. Nelder, R. Mead, A simplex method for function minimization. Comput. J. **7**, 4 (1965)
11. J. Paulsson, O.G. Berg, M. Ehrenberg, Stochastic focusing: fluctuation-enhanced sensitivity of intracellular regulation. Proc. Natl. Acad. Sci. U.S.A. **97**, 13 (2000)
12. P.S.M. Rathinam, M. Khammash, Efficient computation of parameter sensitivities of discrete stochastic chemical reaction networks. J. Chem. Phys. **132**, 3 (2010)