

A numerical method for mesoscopic reaction-diffusion models

Paul Sjöberg

Dept of Information Technology,
Scientific Computing, Uppsala University
P. O. Box 337, SE-751 05 Uppsala, Sweden.
paul.sjoberg@it.uu.se

Quantitative models in molecular biology have traditionally been formulated as ordinary differential equations that describe the time evolution for concentrations of reactants. As this view fails to accurately describe systems with large fluctuations there is an increasing interest in mesoscopic models which make no assumptions about the properties of the variation. The variation is large for systems with low molecular copy numbers and systems with near-critical behavior [1], both cases are common in intracellular kinetics. Mesoscopic models have been constructed and evaluated in a number of cases, for example in models of circadian rhythms [10], plasmid copy number control [8], fluctuations in metabolite pools [1] and stochastic gene expression [3].

Mesoscopic models are usually formulated as master equations [7]. The solution of the master equation is a probability density function $p(\mathbf{x}, t)$ of system states \mathbf{x} . The state of a system with N reactants is represented by a N -dimensional state vector $\mathbf{x} = x_1 \dots x_N$, where element x_i is the copy number of reactant i . Reactions represent transitions between states. Each reaction is defined by a reaction propensity w and a state change $\mathbf{n} = n_1 \dots n_N$. As each reactant in the model adds a dimension to the state space, the number of states grow exponentially as the model is extended by further components. This makes the master equation unsolvable for large copy numbers, even numerically, for more than, say, three reactants.

The standard assumption for mesoscopic models in biochemistry, with recent exceptions in [6] and [2], is that the spatial distribution of reactants is homogeneous. The homogeneity of the component concentrations depend on the diffusion rate in the cell. Experimental results on protein mobility in *E. Coli* indicate that this assumption is often invalid even for small cells [4].

The most common way to handle the vast state space is to use a Monte Carlo method suggested by Gillespie in 1976 [5]. Monte Carlo methods are insensitive to the increasing dimensionality, but have two drawbacks. First, the convergence is slow. This will be particularly severe for problems having very different time scales and for computations of time evolution of the probability distribution. The second drawback is that it is difficult to decide when to stop the simulation.

An alternative to Monte Carlo methods is to mitigate the impact of the growing state space of the master equa-

tion by approximation by the Fokker-Planck equation [7]. Let \mathcal{H} denote the Hessian matrix of second derivatives with respect to $\mathbf{x} \in (\mathbb{R}^+)^N$. Then the equation is

$$\begin{aligned} \frac{\partial p(\mathbf{x}, t)}{\partial t} &= \sum_r \mathbf{n}_r \cdot \nabla_{\mathbf{x}} (w_r(\mathbf{x}) p(\mathbf{x}, t)) + \\ &+ 0.5 \mathbf{n}_r \cdot \mathcal{H} (w_r(\mathbf{x}) p(\mathbf{x}, t)) \mathbf{n}_r = \\ &= \sum_r \left\{ \sum_{i=1}^N n_{ri} \frac{\partial (w_r(\mathbf{x}) p(\mathbf{x}, t))}{\partial x_i} + \right. \\ &\left. + \sum_{i=1}^N \sum_{j=1}^N \frac{n_{ri} n_{rj}}{2} \frac{\partial^2 (w_r(\mathbf{x}) p(\mathbf{x}, t))}{\partial x_i \partial x_j} \right\}, \end{aligned} \quad (1)$$

where \mathbf{x} is the state vector and $w_r(\mathbf{x})$ is the reaction propensity and \mathbf{n}_r is the state change for reaction r .

Equation (1) is a partial differential equation (PDE) which can be solved numerically using substantially fewer unknowns than the master equation. Finite differences are used to approximate the derivatives [9].

The methods of numerical solution of the master equation can be extended to comprise the spatial distribution. The cell is subdivided into subvolumes of size δ . In each subvolume the distribution of reactants is assumed to be uniform. The diffusion is modeled by a master equation for the motion of molecules between subvolumes [7]. By letting $\delta \rightarrow 0$ this corresponds to a diffusion term in the Fokker-Planck equation:

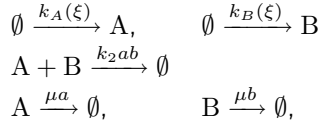
$$\begin{aligned} \frac{\partial p(\mathbf{x}, \xi, t)}{\partial t} &= \sum_r [\mathbf{n}_r \cdot \nabla_{\mathbf{x}} (w_r(\mathbf{x}, \xi) p(\mathbf{x}, \xi, t)) + \\ &+ 0.5 \mathbf{n}_r \cdot \mathcal{H} (w_r(\mathbf{x}, \xi) p(\mathbf{x}, \xi, t)) \mathbf{n}_r] + \\ &+ \nu \left(\sum_{i=1}^N x_i \right) \Delta_{\xi} p(\mathbf{x}, \xi, t), \end{aligned} \quad (2)$$

where ξ is the spatial coordinate, Δ_{ξ} is the Laplace operator and ν is a diffusion constant.

Equation (2) is a PDE in molecule number, space and time. The equation is discretized in the copy number \mathbf{x} and space ξ using second order accurate central differences. The time is discretized using a second order accurate backward differentiation formula. At the lower boundary in copy number space (where $x_i = 0$), the master equation is solved. At the upper copy number boundary (where $x_i = x_N$) $p(\mathbf{x}, \xi, t) = 0$ by assumption. The boundary condition in space is $dp(\mathbf{x}, \xi, t)/d\xi = 0$, since no transport is allowed through the cell membrane.

An example solution with a one-dimensional spatial dimension ξ is shown in Fig. 1. The spatial dimension is discretized using L grid points $\xi_1, \xi_2, \dots, \xi_L$. The reac-

tions are:



where a and b are the copy number for A and B respectively. The space dependent production of metabolites corresponds to a non-homogeneous distribution of synthesizing enzymes. A linear gradient of enzyme activity is introduced for both syntheses. The synthesis of metabolite A is higher at ξ_1 and synthesis of B is higher at ξ_L .

The solution is shown in Fig. 1 with $L = 9$ for no diffusion ($\nu = 0$), intermediate diffusion and high diffusion. The solution for high diffusion converges to the solution for a homogenous cell as is expected. The solution for intermediate diffusion converges towards the high diffusion or the no diffusion solutions as ν is varied.

Acknowledgements

This research is funded by the Swedish Foundation for Strategic Research and the Linnaeus Centre for Bioinformatics.

References

- [1] J. ELF, J. PAULSSON, O. G. BERG, M. EHRENBORG, Near-critical phenomena in intracellular metabolite pools, *Biophys. J.*, 84 (2003), 154–170.
- [2] J. ELF, A. DONČIĆ AND M. EHRENBORG Mesoscopic reaction-diffusion in intracellular signaling, *Proceedings for SPIE's "First International Symposium on fluctuations and noise"* (2003), 114–124.
- [3] M. B. ELOWITZ, A. J. LEVINE, E. D. SIGGIA, P. S. SWAIN, Stochastic Gene Expression in a Single Cell, *Science*, 297 (2002), 1183–1186.
- [4] M. B. ELOWITZ, M. G. SURETTE, P. E. WOLF, J. B. STOCK AND S. LEIBLER, Protein Mobility in the Cytoplasm of *Echerichia Coli*, *J. Bacteriol.*, 180 (1999), 197–203.
- [5] D. T. GILLESPIE, A general method for numerically simulating the stochastic time evolution of coupled chemical reactions, *J. Comput. Phys.*, 22 (1976), 403–434.
- [6] M. HOWARD AND A. D. RUTENBERG, Pattern Formation inside Bacteria: Fluctuations due to the Low Copy Number of Proteins, *Phys. Rev. Lett.*, 90 (2003), 128102.
- [7] N. G. VAN KAMPEN, *Stochastic Processes in Physics and Chemistry*, Elsevier, Amsterdam, 1992.
- [8] J. PAULSSON, M. EHRENBORG, Noise in a minimal regulatory network: plasmid copy number control, *Q. Rev. Biophys.*, 34 (2001), 1–59.
- [9] P. SJÖBERG, P. LÖTSTEDT AND J. ELF, Fokker-Planck approximation of the master equation in molecular biology, Dept. of Information Technology, Scientific Computing, Uppsala University, to appear, will be available at <http://www.it.uu.se/research/reports/>.
- [10] J. M. G. VILAR, H. Y. KUEH, N. BARKAI AND S. LEIBLER, Mechanisms of noise-resistance in genetic oscillators, *Proc. Natl. Acad. Sci. USA*, 99 (2002), 5988–5992.

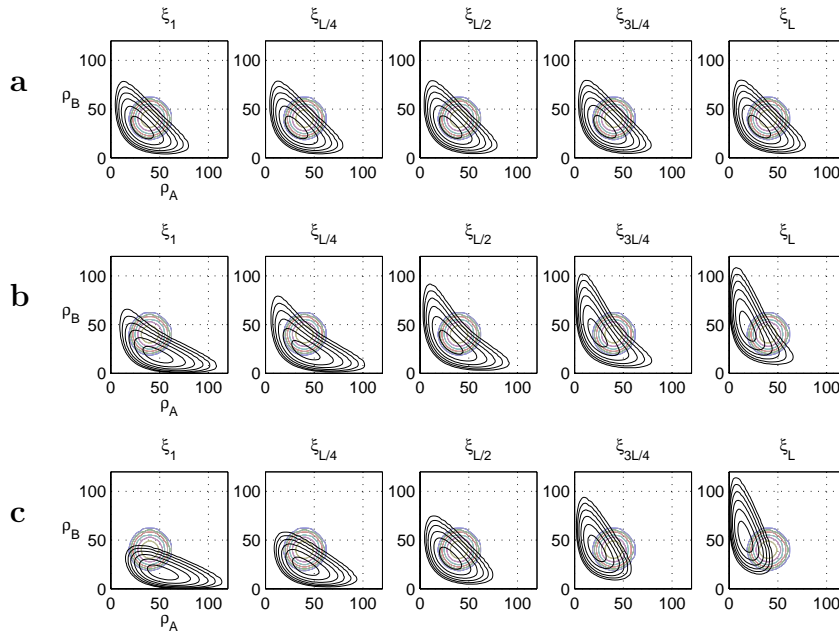


Figure 1: Gray isolines show the initial probability distribution of the molecule densities ρ_A and ρ_B , normalized by the volume δ . Black isolines show the distribution at $t = 100$. **a** shows high diffusion, **b** intermediate diffusion and **c** no diffusion. Model parameters: $k_A(\xi) = (0.8 + 0.4(\xi - \xi_1)/(\xi_L - \xi_1))s^{-1}$, $k_B(\xi) = (1.2 - 0.4(\xi - \xi_1)/(\xi_L - \xi_1))s^{-1}$, $k_2 = 0.001s^{-1}$, $\mu = 0.002s^{-1}$.