Society for
Mathematical
Biology

ORIGINAL ARTICLE

Stochastic Analysis of Reaction–Diffusion Processes

Jifeng Hu · Hye-Won Kang · Hans G. Othmer

Abstract Reaction and diffusion processes are used to model chemical and biological processes over a wide range of spatial and temporal scales. Several routes to the diffusion process at various levels of description in time and space are discussed and the master equation for spatially discretized systems involving reaction and diffusion is developed. We discuss an estimator for the appropriate compartment size for simulating reaction–diffusion systems and introduce a measure of fluctuations in a discretized system. We then describe a new computational algorithm for implementing a modified Gillespie method for compartmental systems in which reactions are aggregated into equivalence classes and computational cells are searched via an optimized tree structure. Finally, we discuss several examples that illustrate the issues that have to be addressed in general systems.

1 Introduction

Reaction–diffusion (RD) equations are at the core of many models of biological processes, ranging from the molecular level, at which they are used to describe signaling, metabolic processes or gene control, to the population level, at which they are used to

J. Hu
School of Mathematics, University of Minnesota, Minneapolis, MN 55455, USA

H.-W. Kang
Mathematical Biosciences Institute, Ohio State University, Columbus, OH, USA

H.G. Othmer (✉)
School of Mathematics and Digital Technology Center, University of Minnesota, Minneapolis, MN 55455, USA
e-mail: othmer@math.umn.edu

⚫ Springer

describe birth–death processes and random movement. When coupled with other processes that describe directed movement, the resulting transport models can account for a wide variety of observed spatio-temporal patterns at the population level, ranging from those in bacterial colonies to those in large mammals. In this paper spatial pattern formation in developmental biology serves as an example of how RD equations are used, and we focus on stochastic effects in such systems, how one properly discretizes a deterministic RD system, and some of the issues that arise in comparing deterministic and stochastic solutions of nonlinear RD systems.

A classical model of RD equations applied to pattern formation in biology is due to Alan Turing (Turing 1952). In Turing's model the uniform state of a spatially distributed system can be unstable to some non-uniform perturbations if the kinetic interactions and the diffusion constants are chosen appropriately, and such instabilities can lead to a steady or a time-periodic spatially non-uniform distribution of chemicals called morphogens. A widely applied example of a Turing mechanism that produces stable non-uniform spatial patterns is an activator-inhibitor system in which activation is short-range due to slow diffusion of the activator, and inhibition is long-range due to rapid diffusion of the inhibitor (Gierer and Meinhardt 1972). Turing's model is based on the spontaneous emergence of pattern in a homogeneous system, but a second class of RD-based models of pattern formation postulates that there is an imposed heterogeneity in the form of a gradient of a morphogen that is produced at a boundary of the system. In such 'positional information' (PI) models cells read the level of morphogen, and thus determine their 'position' in the system and develop accordingly. PI models and their descendants have been more widely used than Turing models because model developmental systems such as *Drosophila* are well-characterized systems in terms of the molecular components involved in patterning, and these have localized morphogen sources that drive pattern formation (Umulis and Othmer 2013).

The French flag model serves as a paradigm of PI models (cf. Fig. 1) and captures some essential aspects of pattern formation. The question is how to divide an initially uniform one-dimensional domain into three equally sized sub-domains, and in a PI model this is done by setting thresholds in the morphogen to define the boundary of different types. The deterministic version of how a French flag or three cell types can be produced in 1D, using suitable thresholds in a linear morphogen distribution, is shown in Fig. 1(A).

Most studies of pattern formation have used a deterministic model, but stochastic effects may be important in gene expression and spatial pattern formation if the copy number of some important components is low. For example, in dorsal-ventral patterning of *Drosophila*, a primary signaling molecule called Dpp is present at concentrations of 10^{-10} M to 10^{-9} M (Shimmi and O'Connor 2003), and at these concentrations there are on average fewer than 10 signaling molecules per nucleus in the embryo. To illustrate the effect this may have on patterning, consider the French flag problem. In Fig. 1(B) we show one realization of a stochastic model in which the input flux of morphogen is fixed at the left, and it diffuses into the domain and is degraded by a first order reaction. The solid line shows the steady-state mean of the distribution, which can be computed directly since the equations are linear (Gadgil et al. 2005), and this also represents the steady-state distribution for the corresponding deterministic system. Since each embryo is a single realization of the patterning

process, the results illustrate the uncertainty in fixing the location of the boundaries between cell types. Of course the downstream interpretation of the signal, as well as feedback from downstream steps may play a role in reducing the effect of signaling noise, and some examples of this are given in Kang et al. (2012).

In the following section we discuss several different transport processes that occur on different time and space scales which can lead to a diffusion-based description of transport. In later sections we describe several methods for estimating an appropriate cell or compartment size for simulations of stochastic RD equations, we discuss a new computational algorithm for solving discretized RD equations, and in the last section we illustrate some of the issues involved in applying the theoretical analysis to an RD example.

2 Routes to the Diffusion Process

2.1 Classical Diffusion

The motion of point particles under deterministic or random forcing is described by Newton's law, which we write in the form

$$d\mathbf{x}_i = \mathbf{v}_i \, dt + d\mathbf{X}_i, \tag{1}$$

$$m_i \, d\mathbf{v}_i = \mathbf{F}_i \, dt + d\mathbf{V}_i. \tag{2}$$

Here $(\mathbf{x}_i, \mathbf{v}_i) \in \mathbb{R}^n$, $n = 1, 2, 3$ are the positions and velocities of the ith particle and m_i is the mass. If the imposed forces $\mathbf{X}$ and $\mathbf{V}$ are smooth deterministic forces they can be written as $d\mathbf{X}_i = \mathbf{X}_i \, dt$, and similarly for $d\mathbf{V}$, and (1) and (2) are the standard Newton equations for particles. When $\mathbf{X}_i$ and $\mathbf{V}_i$ are random forces these are stochastic differential equations, the integral forms of which are interpreted in the Ito sense (Arnold 1974; Capasso and Bakstein 2005). The random forcing terms that are most widely used are Gaussian white noise and compound Poisson processes, both of which are stable Lévy processes (Sato 1999; Applebaum 2004), i.e., stochastic-continuous processes having independent, stationary increments and sample paths that are right-continuous and have left limits.

A general description of Brownian motion of a heavy particle in a fluid is based on (1) and (2), with the assumption that the forcing on position is zero, the random forcing on velocity is Gaussian white noise $\mathbf{W}$, and velocity-dependent frictional forces are admitted. This leads to

$$d\mathbf{x} = \mathbf{v} \, dt \tag{3}$$

$$m \, d\mathbf{v} = -m\zeta \mathbf{v} \, dt + \mathbf{F} \, dt + \sqrt{2\zeta m k_B T} \, d\mathbf{W}(t) \tag{4}$$

where ζ is the friction coefficient, k_B is Boltzmann's constant, and T is the temperature. Under the assumption that the fluid motion relaxes on a much shorter time scale than the motion of the particle, the hydrodynamic forces appear both via the deterministic friction force and the random forces. Under these assumptions (3) and (4) are equivalent to a partial differential equation for the conditional probability density $p(\mathbf{x}, \mathbf{v}, t | \mathbf{x}', \mathbf{v}', t')$ of finding the particle at $(\mathbf{x}, \mathbf{v})$ in phase space at time t, having started at $(\mathbf{x}', \mathbf{v}')$ at time t', namely,

$$\frac{\partial p}{\partial t} + \nabla_{\mathbf{x}} \cdot \mathbf{v} p + \nabla_{\mathbf{v}} \cdot \left(\left(-\zeta \mathbf{v} + \frac{\mathbf{F}}{m} \right) p \right) = \frac{\zeta k_B T}{m} \nabla_{\mathbf{v}} \cdot \nabla_{\mathbf{v}} p. \tag{5}$$

This is known as the Fokker–Planck–Kramers–Klein equation (Wilemski 1976), or simply the Fokker–Planck equation. The evolution of p is driven by both drift-diffusion in $\mathbf{v}$ and drift in $\mathbf{x}$ due to the external force, but when the friction coefficient ζ is large, the velocity relaxes on a time scale $\mathcal{O}(\zeta^{-1})$, and then (5) reduces to the Smoluchowski equation

$$\frac{\partial n}{\partial t} = D \nabla_{\mathbf{x}} \cdot \left(\nabla_{\mathbf{x}} n - \frac{1}{k_B T} \mathbf{F} n \right) \tag{6}$$

for the number density $n(\mathbf{x}, t) = \int p \, d\mathbf{v}$, where the diffusion coefficient is defined by the relation

$$D = \frac{k_B T}{\zeta} = \frac{k_B T}{6\pi \eta r}.$$

The latter equality is the Stokes–Einstein relation for a spherical particle of radius r in a fluid of viscosity η. In the absence of an external force the drift term vanishes and the standard diffusion equation results.

However, this approach only applies when the concentration of the diffusing particle is sufficiently low that inter-particle interactions can be neglected. A more general

formulation begins with the ansatz that, other processes and external fields being absent, a spatially non-uniform particle distribution evolves so as to minimize the Gibbs or Helmholtz free energy, depending on the constraints on the system (Callen 1960). Thus a more general description of molecular diffusion leads to the equation

$$\frac{\partial n}{\partial t} = \nabla_{\mathbf{x}} \cdot n M \nabla_{\mathbf{x}} \mu \tag{7}$$

where μ is the chemical potential of the mobile species and M its mobility. This is a more realistic description of diffusion in the complex cytoplasm of a cell, since in general the free energy and hence μ incorporates like–like interactions and like–unlike interactions (Prigogine and DeFay 1954; Othmer 1976).

2.2 Standard and Anomalous Diffusion as the Limit of Space-Jump Processes

An alternate route, which is not predicated on the molecular origin of diffusion, begins with a jump process, either on a lattice or of a more general kind. Now the forces are impulses given, for example, by a Poisson forcing function acting on the position. Thus the particle or walker moves by a sequence of jumps, distributed in time according to a Poisson law, and the position is given by

$$\mathbf{x}_i(t) = \sum_{k=1}^{N(t)} Y_k H(t - t_k). \tag{8}$$

The amplitudes Y_k are independent random variables, H is the step function, and $N(t)$ is a homogeneous Poisson counting process with parameter λ that counts the number of jumps in $[0, t]$, assuming that $N(0) = 0$ with certainty. A generalization of this allows coupling between the amplitudes of the impulses and their temporal occurrence, and can be defined by a random measure $M(\mathrm{d}t, \mathrm{d}Y)$ that gives the number of jumps in $((t, t + \mathrm{d}t) \times (y, y + \mathrm{d}y))$.

More general waiting time distributions for the jumps can be treated as well. Let $\mathcal{T}$ be the waiting time between jumps, assume they are independent and identically distributed, and let $\phi(t)$ be the probability density function (PDF) for the waiting time distribution (WTD). If a jump has occurred at $t = 0$ then

$$\phi(t)\, \mathrm{d}t = \Pr\{t < \mathcal{T} \leq t + \mathrm{d}t\}.$$

In general the jumps in space may depend on the waiting time, and conversely, the WTD may depend on the size of the preceding jump, but to make the analysis tractable, we assume that the spatial redistribution that occurs at jumps is independent of the WTD. Let $T(\mathbf{x}, \mathbf{y})$ be the PDF for a jump from $\mathbf{y}$ to $\mathbf{x}$, i.e., given that a jump occurs at T_i,

$$T(\mathbf{x}, \mathbf{y})\, \mathrm{d}\mathbf{x} = \Pr\{\mathbf{x} \leq X(T_i^+) \leq \mathbf{x} + \mathrm{d}\mathbf{x} \mid X(T_i^-) = \mathbf{y}\}, \tag{9}$$

where the superscripts $\pm$ denote limits from the right and left, respectively. If the underlying medium is spatially non-homogeneous and anisotropic, the transition probability depends on $\mathbf{x}$ and $\mathbf{y}$ separately, while in a homogeneous medium

$T(\mathbf{x}, \mathbf{y}) = \tilde{T}(\mathbf{x} - \mathbf{y})$, where $\tilde{T}$ is the unconditional probability of a jump of length $|\mathbf{x} - \mathbf{y}|$. We further assume that T is sufficiently smooth and that for any fixed $\mathbf{y}$ the first two $\mathbf{x}$-moments of T are finite—the effect of infinite moments will be discussed later.

Let $P(\mathbf{x}, t|0) \, d\mathbf{x}$ be the probability that a walker who begins at the origin at $t = 0$ is in the interval $(\mathbf{x}, \mathbf{x} + d\mathbf{x})$ at time t, and let $\hat{\Phi}(t)$ be the probability of no jump up to time t. Then $P(\mathbf{x}, t|0)$ satisfies the renewal equation

$$P(\mathbf{x}, t|0) = \hat{\Phi}(t)\delta(\mathbf{x}) + \int_0^t \int_{\mathbb{R}^n} \phi(t - \tau) T(\mathbf{x}, \mathbf{y}) P(\mathbf{y}, \tau|0) \, d\mathbf{y} \, d\tau, \qquad (10)$$

and many of the standard jump processes can be derived from this by particular choices of ϕ and T (Othmer et al. 1988). For instance, if the WTD is exponential, one obtains the continuous time random walk

$$\frac{\partial P}{\partial t}(\mathbf{x}, t|0) = -\lambda P(\mathbf{x}, t|0) + \lambda \int_{\mathbb{R}^n} T(\mathbf{x}, \mathbf{y}) P(\mathbf{y}, t|0) \, d\mathbf{y} \qquad (11)$$

which has been widely studied (cf. Weiss 1994 and references therein). If we rewrite this as

$$\frac{\partial P}{\partial t}(\mathbf{x}, t|0) = \lambda \int_{\mathbb{R}^n} \left[T(\mathbf{x}, \mathbf{y}) P(\mathbf{y}, t|0) - T(\mathbf{y}, \mathbf{x}) P(\mathbf{x}, t|0) \right] d\mathbf{y} \qquad (12)$$

it has the form of a master equation, which will be discussed later.

It follows from the central limit theorem applied to the sum of the IID steps taken in the walk that any jump process with (i) a WTD that has a finite first moment and (ii) a jump distribution that has a finite second moment evolves like a standard Brownian motion for large t. Thus a standard diffusion process emerges from a wide variety of space jump processes under these two conditions. However, when the large-time limit of the mean-square displacement grows either sub- or super-linearly the process is said to exhibit anomalous diffusion. For example, if

$$\langle x^2(t) \rangle \sim \gamma t^\beta$$

for $\beta \neq 1$ and $t \to \infty$, it is called *subdiffusion* if $\beta < 1$ and *superdiffusion* if $\beta > 1$ (Metzler and Klafter 2000). Subdiffusion occurs when particles spread slowly, whether because they rest or are trapped for a long time, and in particular, if the mean waiting time between jumps is infinite. For example, if the WTD is given by

$$\phi(t) = \frac{1}{(1 + t)^2},$$

then the moments M_k are infinite for all $k \geq 1$, and one finds from the Laplace transform that

$$\langle x^2(t) \rangle \sim \log t,$$

and thus the process is subdiffusive (Othmer and Xue 2013). The superdiffusive case arises when the walk is highly persistent in time. For example, if the walker never changes direction this leads to a wave equation for which the mean square displacement scales as t^2.

2.3 Diffusion as the Limit of Poisson-Driven Velocity-Jump Processes

A second class of transport processes that can be described as a diffusion process on certain space and time scales is called the velocity-jump process (Othmer et al. 1988). Again we use position and velocity as the primary variables, but instead of jumps in space we introduce jumps in velocity at random instants. Let $p(\mathbf{x}, \mathbf{v}, \mathbf{y}, t)$ be the density function for individuals in a $(2n + m)$-dimensional phase space with coordinates $(\mathbf{x}, \mathbf{v}, \mathbf{y})$, where $\mathbf{x} \in \mathbb{R}^n$ is the position, $\mathbf{v} \in \mathbb{R}^n$ is the velocity, and $\mathbf{y} \in \mathbb{R}^m$ is the set of intracellular state variables involved in cell movement. As in the case of space jumps, we suppose that the waiting time between jumps and the changes in velocity are independent, and that the WTD is exponential. As a result, the turning can be described by the turning rate λ, and the turning kernel $T(\mathbf{v}, \mathbf{v}')$, which defines the probability of a change in velocity from $\mathbf{v}'$ to $\mathbf{v}$, given that a reorientation occurs. $T(\mathbf{v}, \mathbf{v}')$ is non-negative and normalized so that $\int T(\mathbf{v}, \mathbf{v}') \, d\mathbf{v} = 1$, and at present we assume that it is independent of time and space. The evolution of p is then governed by the equation

$$\frac{\partial p}{\partial t} + \nabla_{\mathbf{x}} \cdot (\mathbf{v} p) + \nabla_{\mathbf{v}} \cdot (\mathbf{F} p) + \nabla_{\mathbf{y}} \cdot (\mathbf{f} p) = -\lambda p + \lambda \int T(\mathbf{v}, \mathbf{v}') p(\mathbf{x}, \mathbf{v}', t) \, d\mathbf{v}', \quad (13)$$

and the underlying stochastic process is called a *velocity jump process*. Under suitable assumptions on the turning kernel the parabolic scaling $\tau = \epsilon^2 t$ and $\xi = \epsilon \mathbf{x}$, where ϵ is a small dimensionless parameter (e.g., $\epsilon \sim \mathcal{O}(sT/L)$ with L the domain size, s a characteristic speed, and T a characteristic time), leads to a diffusion approximation of the transport equation. If we neglect the force F and the internal variables, then

$$\epsilon^2 \frac{\partial p}{\partial \tau} + \epsilon \mathbf{v} \cdot \nabla_{\xi} p = -\lambda p + \lambda \int_V T(\mathbf{v}, \mathbf{v}') p(\xi, \mathbf{v}', \tau) \, d\mathbf{v}' \equiv \mathcal{L} p \quad (14)$$

where the subscript on ∇, which we drop hereafter, indicates differentiation with respect to the scaled space variable. The right-hand side of (14) is $\mathcal{O}(1)$ compared with the left-hand side, whatever the magnitude of p, and this leads to a diffusion equation for the lowest order term p_0 of an outer or Hilbert expansion, which we write as

$$p(\xi, \mathbf{v}, \tau) = \sum_{i=0}^{k} p_i(\xi, \mathbf{v}, \tau) \epsilon^i + \epsilon^{k+1} p_{k+1}(\xi, \mathbf{v}, \tau). \quad (15)$$

An approximation result for any order in ϵ that provides a bound on the difference between the solution of the transport equation and an expansion derived from the solution of the associated parabolic diffusion equation has also been proven (Hillen and Othmer 2000).

Theorem 1 *Given suitable assumptions on the turning kernel and the Hilbert expansion* (15), p_0 *solves the parabolic limit equation*

$$\frac{\partial p_0}{\partial \tau} - \nabla \cdot (D \nabla p_0) = 0, \quad p_0(\xi, 0) = \int_V p(\xi, \mathbf{v}, 0) \, d\mathbf{v}, \quad (16)$$

with diffusion tensor

$$D = -\frac{1}{\omega}\int_V \mathbf{v}\mathcal{F}\mathbf{v}\,d\mathbf{v}. \tag{17}$$

In addition, the first two of the higher-order corrections are given by

$$p_1 = \mathcal{F}(\mathbf{v}\cdot\nabla p_0), \qquad p_2 = \mathcal{F}(p_{0,\tau} + \mathbf{v}\cdot\nabla\mathcal{F}\mathbf{v}\cdot\nabla p_0),$$

where $\mathcal{F}$ is the pseudoinverse of the turning operator $\mathcal{L}$, $V \subset \mathbb{R}^n$ is a compact set and $\omega = |V|$. Then for each $\vartheta > 0$ there exists a constant $C > 0$ such that for each $\vartheta/\epsilon^2 < t < \infty$ and each $x \in \mathbb{R}^n$, $q_2 := \sum_{j=0}^{2}\epsilon^j p_j$ satisfies

$$\left\| p(x,.,t) - q_2\big(\epsilon x,.,\epsilon^2 t\big) \right\|_{L^2(V)} \leq C\epsilon^3.$$

The constant C depends on V, D, ϑ and the second eigenvalue μ_2 of $\mathcal{L}$.

Therefore, on suitably chosen time and space scales that in essence guarantee that individuals turn frequently enough and thus do not, on average, move too far, the velocity jump process can be approximated by a diffusion process. Such restrictions are easily understood when it is realized that some individuals may simply continue in their original direction and never turn (Hillen and Othmer 2000).[1] This result applies as well when reactions occur and has led to a better understanding of how individual-level behavior leads to the classical Patlak–Keller–Segel chemotaxis equations (Erban and Othmer 2005, 2007), and to macroscopic, individual-based models of bacterial pattern formation (Xue and Othmer 2009; Xue et al. 2011).

The general conclusion is that diffusion processes arise from a variety of more fundamental processes at the molecular or higher level, and in general are an approximation valid only on appropriate time and space scales.

3 The Stochastic Description of Spatially Discrete RD Systems

A master equation is essentially a balance equation that expresses the conservation of probability for a Markovian stochastic process. The form of a master equation depends on whether the state space is discrete or continuous, and whether time is discrete or continuous. An example of a continuous-time, continuous state space process is the space-jump process that led to (12). All such equations can be derived from the fundamental description of a Markovian process given by the Chapman–Kolmogorov equation. In this section we introduce a master equation for reaction–diffusion processes in which the underlying physical space is discretized into N_c compartments of equal volume h^3.

Suppose that in each cell there are s distinct species and N_r reactions amongst them. In the stochastic description of reactions and transport the evolution of the system is modeled as a continuous-time Markov jump process. Let $N(t) =$

---

[1] Other scalings may be applicable in other regimes, but would lead to different evolution equations (Hillen and Othmer 2000).

$(N^1(t), N^2(t), \ldots, N^{N_c}(t))$ be the vector of random variables whose kth vector component represents the number of molecules of species in the kth compartment. Let $P(n, t)$ be the probability that $N(t) = n$, i.e., the joint probability that $N^1(t) = n^1 = (n^1_1, n^1_2, \ldots, n^1_s)$, $N^2(t) = n^2 = (n^2_1, n^2_2, \ldots, n^2_s)$, ..., $N^{N_c}(t) = n^{N_c} = (n^{N_c}_1, n^{N_c}_2, \ldots, n^{N_c}_s)$. Then $P(n, t) \equiv P(n^1, n^2, \ldots, n^k, \ldots, n^{N_c}, t)$. For any reaction network one defines a set of complexes that represent combinations of molecules that react as a unit in a reaction, a map $\mathcal{E}$ that encodes the topology of the reaction network, and a map v that converts complex concentrations into species concentrations. One can show that the ℓth reaction $C(k) \to C(k')$ induces a change $\Delta n^{(\ell)} = v\mathcal{E}_{(\ell)}$ in the number of molecules of all species after one step of the reaction, where subscript ℓ denotes the ℓth column of $\mathcal{E}$ and $C(k)$ denotes the kth complex (Gadgil et al. 2005). Therefore the state $m = n - v\mathcal{E}_{(\ell)}$ is a source or predecessor to n under one step of the ℓth reaction. Similarly, states of the form $m = n + v\mathcal{E}_{(\ell)}$ are reachable from n in one step of the ℓth reaction. Once the graph of the network and the stoichiometry are fixed one can easily derive the explicit form of the master equation for a given network of reactions from the corresponding deterministic equations (Gadgil et al. 2005).

In the presence of diffusive coupling between compartments, the flux of the ith species from k to one of its neighbors k' is assumed to be given by

$$J^{kk'}_i = \frac{D_i}{h^2} n^k_i, \tag{18}$$

while the reverse flux is

$$J^{k'k}_i = \frac{D_i}{h^2} \left(n^{k'}_i + 1 \right).$$

The transfer of one molecule in the first step involves the change $n^k_i \to n^k_i - 1$, $n^{k'}_i \to n^{k'}_i + 1$ whereas the reverse flux involves the change $n^{k'}_i + 1 \to n^{k'}_i$, $n^k_i - 1 \to n^k_i$. The general form of the master equation for both reaction and diffusion processes is

$$
\begin{aligned}
\frac{d}{dt} P(n, t) = \sum_{k=1}^{N_c} \Bigg\{ & \sum_{k' \in \mathcal{N}(k)} \Bigg[\sum_{i=1}^{s} \frac{D_i}{h^2} \left(n^{k'}_i + 1 \right) P\left(n^1, \ldots, n^k - e_i, \ldots, n^{k'} \right. \\
& + e_i \ldots, n^{N_c}, t \Big) - \frac{D_i}{h^2} n^k_i P\left(n^k, t \right) \Bigg] \\
& + \sum_{\ell=1}^{N_r} \mathcal{R}_\ell \left(n^k - v\mathcal{E}_{(\ell)} \right) \cdot P\left(n^1, \ldots, n^k - v\mathcal{E}_{(\ell)}, \ldots, n^{N_c}, t \right) \\
& - \sum_{\ell=1}^{N_r} \mathcal{R}_\ell \left(n^k \right) \cdot P\left(n^k, t \right) \Bigg\}
\end{aligned}
\tag{19}
$$

where $\mathcal{N}(k)$ is the set of all neighbors of k in the reaction graph for the cellular network and $e_i = (0, 0, \ldots, 1, \ldots, 0)^T$ has a 1 in the ith position and zeroes elsewhere. The formulation in (18) and (19) is based on the assumption that the domain

is decomposed into equal-size compartments. If the compartments have side lengths that are different in different directions then the flux at (18) must be scaled by the h appropriate for each direction, and the sum over neighbors in (19) must be changed to reflect this.

3.1 The Maximal Compartment Size

Given the reaction network and the network of connections between cells, one must next determine the appropriate cell size. The question as to how large it can be has been addressed by many, and a variety of criteria have been proposed (Kang et al. 2012). Most are based on first estimating the smallest reaction time scale and then choosing a compartment size to ensure that the diffusion time scale is less than this. Since the diffusion time in a cube of side h scales as h^2/D, one has to compare the kinetic rate for the fastest step with the slowest diffusion rate. Alternatively one could generalize this by treating each reaction separately and first applying the criterion to the species involved in a given reaction, and then comparing the estimates for all reactions. Later we will discuss a result that determines a suitable time scale for the entire network and use this to set the cell size.

Typically the accuracy of a numerical solution of deterministic continuum RD equations increases as the grid size is reduced, but this not true for RD master equations such as (19) without further analysis. To see this, consider a closed system of fixed size that contains N molecules, and suppose that the system is discretized into N_c compartments that are connected by diffusion. The stationary distribution is spatially uniform and multinomial (Gadgil et al. 2005), with mean and variance given by

$$M_i = \frac{N}{N_c}, \quad \sigma_i^2 = M_i\left(1 - \frac{1}{N_c}\right) = \frac{N}{N_c}\left(1 - \frac{1}{N_c}\right), \quad i = 1, \ldots, N_c.$$

An appropriate measure of the fluctuations is the coefficient of variation, $CV = \sigma_i/M_i$, and this is

$$CV = \sqrt{\frac{N_c - 1}{N}}, \tag{20}$$

which is zero for $N_c = 1$ and which grows as $\sqrt{N_c}$ for large N_c. Thus this measure of the fluctuations tends to infinity as $N_c \to \infty$. However, as is shown in Kang et al. (2012), a suitably scaled CV defined later stabilizes as $N_c \to \infty$ and the evolution equation for the corresponding scaled mean value converges to the continuum RD equations for linear reactions. Later the onset of the plateau in this scaled CV as a function of N_c is used to set the compartment size. If bimolecular reactions are involved there are also lower bounds on the size because each compartment must be large enough to contain both reacting species, and the continuum limit of the RD master equation does not capture such reactions (Isaacson 2009). To circumvent this one has to allow reactions between molecules in adjacent compartments (Agbanusi and Isaacson 2013).

An alternate criterion for determining the maximal compartment size is based on the deterministic RD equations for a reaction network. The basic premise is that the

cell size must be small enough so that if one were to impose no-flux boundary conditions on the cell boundary, the solution would relax to a spatially uniform solution, whether stationary or periodic. This requires that diffusion dominates reaction in an appropriate sense, and that was first made precise in Othmer (1977), Ashkenazi and Othmer (1978) and generalized in Conway et al. (1978). Let c be the vector of species concentrations let $\bar{c}$ be the spatial average of c, and let δ be the smallest diffusion coefficient and α_1 the largest non-zero eigenvalue of the Laplacian with homogeneous Neumann conditions on $\partial\Omega$. Then it was shown that under the assumption that all species diffuse, $\|c(\mathbf{x}, t) - \bar{c}(t)\|_{L_2} \to 0$ exponentially in t in any bounded domain $\Omega \in \mathbb{R}^n$ if $|\alpha_1\delta| > \hat{r}$, where $\hat{r}$ is the maximum Euclidean norm, taken over an appropriate invariant set C_∞, of the Jacobian of the reaction terms. This result was extended in Kang et al. (2012) to allow non-diffusing species and reactions at the boundary, as described below. We briefly describe this result here and later use it to compute a maximal compartment size for a discretized reaction–diffusion system. We discuss conditions for the existence of an invariant set there.

Let $\Omega \subset \mathbb{R}^3$ be a domain with a smooth boundary $\partial\Omega$, and let $u(\mathbf{x}, t) \in \mathbb{R}^m$ and $v(\mathbf{x}, t) \in \mathbb{R}^n$ denote the concentrations of diffusing and non-diffusing species, respectively, that react in Ω at time t. We redefine the vector $c = [u^T, v^T]^T$ of species concentrations and assume that c is contained in a closed, invariant rectangle C_∞ at all times. Then, we write the governing equations and the boundary conditions for the species as follows:

$$\frac{\partial u(\mathbf{x}, t)}{\partial t} = \mathcal{D}\Delta u(\mathbf{x}, t) + \mathcal{R}\big(u(\mathbf{x}, t), v(\mathbf{x}, t)\big), \quad \mathbf{x} \in \Omega,$$

$$\frac{\partial v(\mathbf{x}, t)}{\partial t} = \mathcal{S}\big(u(\mathbf{x}, t), v(\mathbf{x}, t)\big), \quad \mathbf{x} \in \Omega, \tag{21}$$

$$\frac{\partial u(\mathbf{x}, t)}{\partial n} = 0, \quad \mathbf{x} \in \partial\Omega.$$

We assume that $\mathcal{R}$ and $\mathcal{S}$ are C^1 in the non-negative cone of (u, v)-space, and in general both are nonlinear functions of (u, v).

The Jacobians of a function $f : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^m \times \mathbb{R}^n$ with respect to u and v are denoted

$$D_u f(u, v) \equiv \frac{\partial f(u, v)}{\partial u}, \qquad D_v f(u, v) \equiv \frac{\partial f(u, v)}{\partial v}.$$

We assume that for all $c \in C_\infty$, $\sigma(D_v\mathcal{S}(c)) \subset LHP$, where $\sigma(A)$ denotes the spectrum of A. Then there exists a Lyapunov matrix W_c for which $W_c D_v\mathcal{S}(c)$ is uniformly negative. Letting $\lambda_m(W_c)$ and $\lambda_M(W_c)$ be the smallest and the largest eigenvalues of W_c, we define

$$\lambda_m \equiv \inf_{c \in C_\infty} \lambda_m(W_c),$$

$$\lambda_M \equiv \sup_{c \in C_\infty} \lambda_M(W_c), \tag{22}$$

$$w \equiv \frac{1}{2} \sup_{U \in C_\infty} \left\| \frac{\partial W_c}{\partial t} \right\|_E.$$

Here $\| \cdot \|_E$ denotes the Euclidean norm, and $U = (u, v, \bar{v})$, where $\bar{v}$ is a solution of

$$\frac{d\bar{v}(t)}{dt} = \mathcal{S}\big(\bar{u}(t), \bar{v}(t)\big),$$

and $\bar{u}$ is the spatial average of u. The Lyapunov matrix is needed because some species do not diffuse—details are given in Kang et al. (2012). We define measures of kinetic sensitivity via the Jacobians of the reaction terms as

$$
\begin{aligned}
\hat{r}_u &\equiv \sup_{c \in C_\infty} \big\| D_u \mathcal{R}(c) \big\|_E, & \hat{r}_v &\equiv \sup_{c \in C_\infty} \big\| D_v \mathcal{R}(c) \big\|_E, \\
\hat{s}_u &\equiv \sup_{c \in C_\infty} \big\| D_u \mathcal{S}(c) \big\|_E, & \check{s}_v &\equiv \inf_{c \in C_\infty} \big| \sigma \big(W_c D_v \mathcal{S}(c) \big) \big|
\end{aligned}
\tag{23}
$$

where $| \cdot |$ denotes the absolute value.

Conditions for exponential convergence in time of $c(\mathbf{x}, t)$ to a spatially uniform solution under homogeneous Neumann boundary conditions are derived in Kang et al. (2012). In essence these require that the smallest non-zero diffusion coefficient should be large enough compared to some function of the kinetic sensitivities, $\hat{r}_u$, $\hat{r}_v$, $\hat{s}_u$, and $\check{s}_v$, and of the constants, λ_m, λ_M, and w, defined by the positive operator W_c. There it was shown that $\| c(\mathbf{x}, t) - \bar{c}(t) \|_{L_2} \to 0$ exponentially in t if

$$
\begin{aligned}
&\text{(i)} \quad \sigma \big(D_v \mathcal{S}(c) \big) \subset LHP \quad \text{for all } c \in C_\infty, \\
&\text{(ii)} \quad \check{s}_v > w, \\
&\text{(iii)} \quad |\alpha_1 \delta| > \hat{r}_u + \frac{\hat{r}_v \hat{s}_u}{\check{s}_v - w} \cdot \frac{\lambda_M^2}{\lambda_m},
\end{aligned}
\tag{24}
$$

where $\delta = \min_i \mathcal{D}_i$ and α_1 is the largest non-zero eigenvalue of the scalar problem

$$
\begin{aligned}
\Delta \phi(\mathbf{x}) &= \alpha \phi(\mathbf{x}), & \mathbf{x} &\in \Omega, \\
\frac{\partial \phi}{\partial n} &= 0, & \mathbf{x} &\in \partial \Omega.
\end{aligned}
\tag{25}
$$

If all species diffuse this reduces to the previous condition $|\alpha_1 \delta| > \hat{r}_u$. A kinetic scheme that satisfies the conditions for a Turing instability shows that the conditions can be violated on a general domain, but they can be met on a sufficiently small domain, even for such mechanisms. This criterion will be used later to estimate an appropriate cell size in an example.

3.2 A Generalized Measure of Fluctuations

To assess the fluctuations in the network one can use one of several forms of a coefficient of variation. The first measure results from defining the noise component- and compartment-wise, which may be appropriate when assessing the effect of fluctuations in a morphogen used to define the boundary between two tissue types. In this case one computes the standard deviation of the number of molecules for the ith

species in the kth compartment divided by its mean, viz.,

$$\left(CV(t)\right)^{\eta} \equiv \frac{\sqrt{E[N^{\eta}(t)^2] - E[N^{\eta}(t)]^2}}{E[N^{\eta}(t)]} \tag{26}$$

where $\eta = (k, i)$. This reflects the fluctuations of each species in each compartment, and thereby leads to a total of sN_c measures. However, for most purposes a single global measure that averages over all species and compartments is more appropriate. Among the many ways to do this, we later use a measure based on a normalized covariance matrix, which is defined as (Tomioka et al. 2004)

$$\left[\Sigma_0(t)\right]_{\mathbf{i}(\eta), \mathbf{i}(\zeta)} = \frac{E[N^{\eta}(t)N^{\zeta}(t)] - E[N^{\eta}(t)]E[N^{\zeta}(t)]}{E[N^{\eta}(t)]E[N^{\zeta}(t)]}$$

where $\zeta = (q, j)$ and the index function $\mathbf{i}$ labels the component. We then define a generalized coefficient of variation as the square root of the maximum eigenvalue of $\Sigma_0(t)$, and denote it

$$CV^*(t) \equiv \sqrt{\lambda_{\max}\left(\Sigma_0(t)\right)}. \tag{27}$$

Since the covariance matrix is symmetric and at least positive semi-definite, this is well-defined. The long-time behavior of this is defined by $\Sigma_{0,\infty} \equiv \lim_{t \to \infty} \Sigma_0(t)$, and the steady-state CV^* is given by

$$CV^* = \sqrt{\lambda_{\max}(\Sigma_{0,\infty})}.$$

In certain classes of linear reaction networks one can prove that CV^* is an increasing function of the number of cells, as in the pure diffusion case. In particular, if there are no catalyzed inputs to an open system, or if in a closed system the kinetic matrix has exactly one zero eigenvalue, then CV^* is the square root of the reciprocal of the smallest mean molecule number in all cells and tends to infinity as $N_c \to \infty$ (Kang et al. 2012). However, it can be shown that a scaled CV^* defined as

$$\overline{CV}^* \equiv \sqrt{\frac{V}{N_c}} CV^*, \tag{28}$$

where V is the total volume of the system, stabilizes as the cell number increases under appropriate conditions (Kang et al. 2012). Moreover, it was shown there that the estimate of cell size derived from the PDE criterion agrees quite well with the cell size at which the scaled CV stabilizes.

In nonlinear problems an analytic expression for the scaled CV has not been established, but whether or not it stabilizes as the cell size decreases can always be checked computationally, and we show in a later section that it provides a reasonable estimate for a nonlinear system.

4 An Exact Algorithm for Simulating Stochastic RD Systems

4.1 A Brief Description of the Original and an Optimized Direct Method

Numerous stochastic simulation algorithms (SSAs) have been proposed since the original exact algorithms, called the first reaction method and the direct method, were formulated by Gillespie (1977). Later Gibson and Bruck (2000) designed the next reaction method in which random numbers generated for the direct method are re-used, and a priority tree is used to accelerate the search for the next reaction. In this section we develop a new algorithm for RD systems, and to set it in context we first describe the original direct method. A comparison with SSAs for RD equations will be done after developing the new algorithm.

We first suppose that the system volume V is well-mixed and that there are N_r reactions amongst s species. The reaction rates can be written

$$\mathcal{R}_\ell = c_\ell h_{j(\ell)}(n) \tag{29}$$

where c_ℓ is the probability per unit time that the molecular species in the jth complex reacts, $j(\ell)$ denotes the reactant complex for the ℓth reaction, and $h_{j(\ell)}(n)$ is the number of independent combinations of the molecular components in this complex, viz.,

$$c_\ell = \frac{k_\ell}{(\mathcal{N}_A V)^{\sum_i v_{ij(\ell)} - 1}}, \qquad h_{j(\ell)} = \prod_i \binom{n_i}{v_{ij(\ell)}}. \tag{30}$$

Here k_ℓ is the rate constant for the deterministic rate equations, written in terms of the number density, and $v_{ij(\ell)}$ is the stoichiometric coefficient of species i in the $j(\ell)$th reaction.

In Gillespie's notation, the probability density for reaction type μ is

$$p_\mu(\tau) = h_\mu c_\mu e^{-\sum_{i=1}^{N_r} h_i c_i \tau} = a_\mu e^{-a_0 \tau} \tag{31}$$

where

$$a_\mu = c_\mu h_\mu, \qquad a_0 = \sum_{i=1}^{N_r} a_i.$$

In the Monte Carlo simulation algorithm of the direct method, a basic reaction cycle comprises three steps: first, determine the waiting time for the next reaction; second, determine which reaction will occur; and lastly, update the system state to reflect changes in the species involved as reactants or products in the reaction that has occurred. Accordingly, during each cycle two random numbers $R_1, R_2 \in URN[0, 1]$ are generated, one of which is used to calculate the waiting time according to

$$\tau = -\frac{\ln(R_1)}{a_0} \tag{32}$$

Fig. 2 Original (*bottom*) and optimized (*top*) Gillespie's direct method applied to reaction diffusion systems
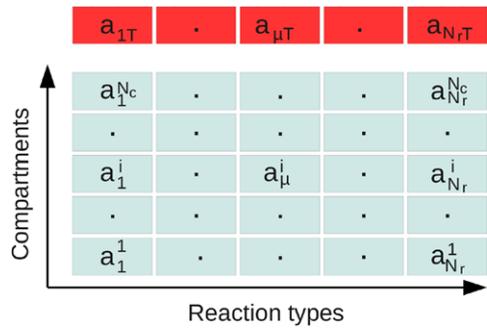
and the other of which is used to determine the next reaction type μ according to

$$\sum_{i=1}^{\mu-1} a_i < R_2 \cdot a_0 \leq \sum_{i=1}^{\mu} a_i. \tag{33}$$

Cao et al. proposed several potential ways of optimizing the original direct method algorithm by reducing the cost of specific steps (Cao et al. 2004), such as preordering the reactions according to their firing frequency and recomputing only the propensities of reactions which are affected by the current reaction.

If we apply the direct method to a discretized volume of N_c equal-sized cells, as described by the RD master equation (19), there are a total of $R_x = N_r \cdot N_c$ reactions, which can be large for 3D problems, but an enhanced direct method used for polymer dynamics (Matzavinos and Othmer 2007) can be used here. In this algorithm, reactions of the same type in different cells are lumped into an equivalence class, and the propensity of an equivalence class is simply the sum of the propensities of all individual reactions that belong to that class. For example, if the 'reaction' is a diffusion step, then diffusive jumps in *all directions* among all cells are considered as one class for that species. To determine the next reaction, one first searches for the next reaction type among all equivalence classes, and then, given the reaction type, searches for the cell in which the reaction occurs. The comparison between this algorithm and the direct method is illustrated in Fig. 2, where

$$a_{\mu T} = \sum_{i=1}^{N_c} a_\mu^i = \sum_{i=1}^{N_c} c_\mu \, h_\mu^i, \tag{34}$$

and where the superscript i denotes the computational cell. In the original Gillespie method one searches for the next reaction among $N_r \times N_c$ cells, while in the optimized algorithm one searches for the reaction type μ among N_r equivalence classes, and then in the column labeled μ searches for the target cell i. The average number of search steps are estimated as $\mathcal{O}(N_r \times N_c)$ and $\mathcal{O}(N_r + N_c)$, respectively.[2] The ratio of search times implies that the improved algorithm leads to a lower search time

---

[2]These estimates are predicated on a uniform distribution of species in the compartments. They could be very different if there are spatial gradients, but estimates are difficult in that case.

Fig. 3 The tree structure used in searching for the current reacting cell. In this tree, there are a total of $(N_L + 2)$ levels, where the root is the 0th level, and the nodes at the lowest (i.e., $(N_L + 1)$th) layer are the computational cells. We assume each node in the ith level $(i \in \{0, 1, \ldots, N_L\})$ has m_{i+1} daughters
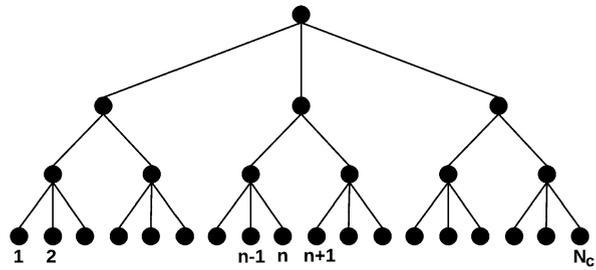
whenever $N_c > N_r/(N_r - 1)$, and if for example, there are 10 reactions in 1000 cells, the speedup is approximately ten-fold. However, one can do better, as is shown in the next section.

4.2 A New Search Algorithm for Discretized RD Equations

Earlier we showed how to predict a maximal compartment size, and later we will show in an example that this size may be so small that N_c may easily be of the order of thousands or tens of thousands in a 3D system. When N_c is large, the search time for the target cell as described above will dominate the total computational cost, and could present a significant bottleneck for simulations.

Therefore we developed an optimized search algorithm that is based on a subdivision of the computational cells into blocks, and which incorporates the previously defined equivalence classes. Associated with each block is the sum of propensities for each reaction type within individual cells contained in that block, and, instead of searching for target cells sequentially among the N_c cells, one first determines which block the target cell belongs to, and then searches for the cell within that block. Suppose we group all cells into m disjoint blocks, of which each includes N_c/m cells.[3] One needs $\mathcal{O}(m)$ steps to find the target block, and within that block, one needs another $\mathcal{O}(N_c/m)$ steps to determine the target cell. Accordingly, one needs only $\mathcal{O}(m + N_c/m)$ search steps instead of $\mathcal{O}(N_c)$.

One can generalize the above procedure by defining a tree structure as shown in Fig. 3. The root is a node that reflects the propensities of all reaction types, which in turn are the sum of propensities of corresponding reaction types in all daughter nodes linked to the root, and similarly for each level in the tree. In the lowest level the nodes are simply the computational cells, and based on this construction, the total propensity of a reaction type is simply the sum of the propensities of that reaction among all the computational cells. Searching for the target cell in the tree is a straightforward top down search, as in a decision tree.

We call the resulting algorithm for a tree with $(N_L + 2)$-levels an N_L-partition algorithm—the detailed steps are given later. The root is designated as the 0th level, and the lowest level, which contains all computational cells, as the $(N_L + 1)$th level. For the sake of simplicity, we assume an equal number of daughter nodes for any node in the same level. Let each node in the ith level have m_{i+1} daughter nodes; then

---

[3]Here and hereafter we assume that block sizes are chosen so that N_c is a multiple of m.

$N_c = \prod_{i=0}^{N_L} m_{i+1}$. In such a tree, searching for the target cell through $N_L + 1$ levels takes

$$\sum_{i=0}^{N_L} \mathcal{O}(m_{i+1}) \tag{35}$$

steps whereas updating the tree structure after the current reaction takes

$$(N_L + 2) \sum_{j=1}^{N_c} g_j \tag{36}$$

steps, where g_j denotes the number of reaction propensities at the jth cell affected by the current reaction step. Note that if the current step is a diffusive jump, there will be more than one non-zero g-term, whereas when the current reaction is a non-diffusion reaction, there will be only one non-zero g-term. The value of each g-term indicates the complexity of the inter-connection between reactions, and thus (36) encodes both information on diffusion and reactions in the system.

4.3 The Algorithm in Detail

To explain the algorithm in detail, we define the following additional quantities.

- h_μ^i: number of reactant combinations of reaction type μ in the ith cell, $i = 1, \ldots, N_c$.
- a_μ^i: reaction rate of reaction type μ in the ith cell, $i = 1, \ldots, N_c$.
- $h_{\mu T}$: total number of reactant combinations of type μ among all cells: $\sum_{i=1}^{N_c} h_\mu^i$.
- $a_{\mu T}$: total rate of reaction type μ among all cells: $\sum_{i=1}^{N_c} a_\mu^i$.
- a_0: total reaction rate of all types in all cells: $\sum_{\mu=1}^{N_r} a_{\mu T} = \sum_{i=1}^{N_c} \sum_{\mu=1}^{N_r} c_\mu h_\mu^i$.

1. *Initialization steps*
   (a) Partition the system domain into N_c computational cells. Each cell is associated with a data structure with information local to the cell, such as the numbers of each molecular species, the rate constants of reactions (c_μ), the reactant combinations (h_μ^i), etc.
   (b) Create a tree structure with nodes as in Fig. 3. Assign the propensity vector of all reaction types to each node, using the fact that for the interior levels the propensity of a reaction type at a node is the sum of the propensities of that type in all daughter nodes attached to the node.
   (c) Initialize the computational cells and reaction rate tree corresponding to the initial condition of the system.
2. *The time stepping procedure*
   Suppose that the simulation time is t_n after n reaction steps. Let the state of the system evolve as follows.
   (a) Generate a random number R_1, and compute the waiting time τ_n as

$$\tau_n = -\frac{\ln(R_1)}{a_0}$$

where $a_0 = \sum_{\mu=1}^{N_r} a_{\mu T}$ and $a_{\mu T}$ is stored in the propensity vector associated with the root node of the tree.

(b) Generate a second random number (R_2) to determine the type and cell location of the next reaction.

(i) The reaction type μ is such that $1 \leq \mu \leq N_r$ and

$$\sum_{j=1}^{\mu-1} a_{jT} < R_2 \cdot a_0 \leq \sum_{j=1}^{\mu} a_{jT}.$$

(ii) To determine the target cell, first compute the random number $Q \in URN[0, a_0]$ given by

$$Q = R_2 \cdot a_0 - \sum_{j=1}^{\mu-1} a_{jT}.$$

The target cell d in which the next reaction occurs is found by searching top down through the reaction tree until

$$\sum_{i=1}^{d-1} a_\mu^i < Q \leq \sum_{i=1}^{d} a_\mu^i. \tag{37}$$

In detail, first examine the m_1 nodes at the first layer of the propensity tree structure in Fig. 3, and find the smallest $m_{1\mu}$ such that $m_{1\mu} \leq m_1$

$$Q < \sum_{i=1}^{m_{1\mu}} a_\mu^{i,m_1} \tag{38}$$

where a_μ^{i,m_1} is the propensity of reaction type μ associated with the ith node at the first layer. After that, replace Q by

$$Q \leftarrow Q - \sum_{i=1}^{m_{1\mu}-1} a_\mu^{i,m_1} \tag{39}$$

and examine the m_2 nodes connected with the $m_{1\mu}$st node at the first layer, and find the smallest $m_{2\mu}$ such that $m_{2\mu} \leq m_2$ and

$$Q < \sum_{i=1}^{m_{2\mu}} a_\mu^{i,m_2}. \tag{40}$$

Repeat the above procedure until the lowest layer is reached and the target cell is found. A path from the root to the target cell will be represented by the vector $(m_{1\mu}, m_{2\mu}, \ldots, m_{(N_L+1)\mu})$, which means that the path passes through the $m_{1\mu}$th node connected with the root, and $m_{(i+1)\mu}$ signifies that the path passes through the $m_{(i+1)\mu}$th node connected with the previous node of the path.

Table 1 Comparison of searching and updating steps in three algorithms

Algorithm	N_{search}	N_{update}	Optimal parameter	Optimal N_{search}
G_1	$\frac{m_1+m_2+2}{2}$	$2\sum_{i=1}^{N_c} g_i$	$m_1 = m_2 = \sqrt{N_c}$	$\sqrt{N_c}+1$
G_2	$\frac{m_1+m_2+m_3+3}{2}$	$3\sum_{i=1}^{N_c} g_i$	$m_1 = m_2 = m_3 = \sqrt[3]{N_c}$	$\frac{3\sqrt[3]{N_c}+3}{2}$
G_{binary}	$\log_2(N_c)$	$\log_2(N_c)\sum_{i=1}^{N_c} g_i$		$\log_2(N_c)$

(c) Update the state of the affected cells based on result (b). First, one changes the molecular numbers and reactant combinations of the affected reactions in those cells. To update the tree, suppose that at the ith cell reaction type j changes its reactant combinations from $a_j^{i,\text{old}}$ to $a_j^{i,\text{new}}$. Let $\delta_j^i = a_j^{i,\text{new}} - a_j^{i,\text{old}}$; then for any branch points on the path from the computation cell to the tree root, add δ_j^i to the reactant combination at that point.

3. Advance the time from t_n to $t_{n+1} = t_n + \tau_n$, and repeat step (2) until the desired time is reached.

4.4 Estimates of the Cost of Searching and Updating in Three Algorithms

In the following we will focus on three specific partition algorithms depending on the number of levels in the tree. First is a one-level partition algorithm (G_1) obtained by dividing N_c cells into m_1 groups. In reference to the tree in Fig. 3, the second level contains m_1 nodes, each of which links to N_c/m_1 cells. The second algorithm is a two-level partition algorithm (G_2) obtained by making a tree structure with a root connecting to m_1 daughter nodes, each of which in turn connects to m_2 nodes. The latter nodes each connect to $N_c/(m_1 \cdot m_2)$ computational cells. The third algorithm is the extreme case of a binary tree (G_{binary}) in which each node in the tree has two daughter nodes. Later the efficiency of these algorithm is compared when they are applied to two simple examples.

When N_c is large, searching for the target reacting cell through the tree structure and subsequent updating could be costly. In searching for the target cell, one goes through the sequential comparison as shown in the detailed procedure in (38) and (40). The number of comparisons made in finding the target cell, which involve both vertical steps and horizontal comparisons, is called the *search depth*. In updating the tree after a reaction occurs, the changed propensity in an affected cell must propagate from that cell to the root node, and therefore ($N_L + 1$) updating steps are needed. The average searching and updating steps in a reaction cycle for a uniformally distributed system are summarized in Table 1 for different algorithms. The optimal parameter for each algorithm is predicted under the present assumption that N_c is large.

4.5 A Numerical Comparison of Three Algorithms

Next we apply the three algorithms described in last section to two examples.

Fig. 4 The search depth and CPU cost of algorithm G_1 in pure diffusion of one species
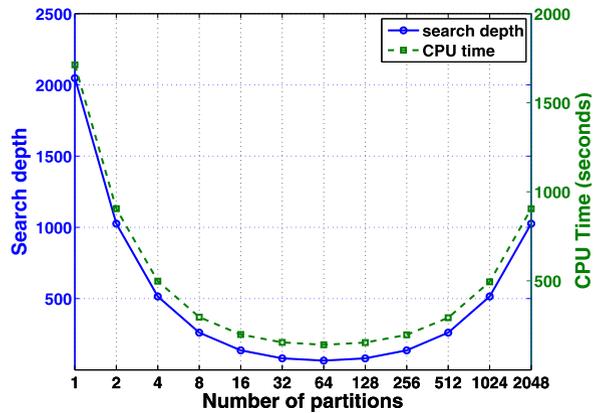
4.5.1 Pure Diffusion of One Species

The first example is a comparison of the efficiency of the three algorithms when simulating pure diffusion of one species on a quasi-2D domain. We use a domain of size 6.4 μm × 6.4 μm divided into 64×64 rectangular compartments of size 0.1 μm × 0.1 μm, each of height 0.1 μm. The 5.0 μM concentration, which is equivalent to 12329 molecules or approximately 3 molecules per cell, is uniformly distributed in the domain, and the diffusion coefficient is 2.0 μm²/s. The run time of the test of the algorithms is based on 50 seconds of dynamics.

First, we test the CPU cost using algorithm G_1. The search depth and total CPU cost for different cell segmentation is summarized in Fig. 4. The numerical simulations show that the total CPU cost is dependent on how the computational cells are partitioned into groups (i.e., m_1). Both the search depth and the CPU cost are convex functions of the partition number. The minimum search depth N_{search} and the CPU cost occur at the same optimal partition, which is predicted to occur when the partition number $m_1 = \sqrt{N_c}$. In the current simulation, $N_c = (L_x/\Delta h_x) \cdot (L_y/\Delta h_y) = 4096$, which gives the optimal $m_1 = 64$. The optimal algorithm with $m_1 = 64$ is more than ten-fold faster than that without partitioning i.e., when $m_1 = 1$.

According to step (2) in the detailed algorithm, the CPU time for each reaction type consists of the cost for the following operations: generation of two random numbers, determination of the target reaction type, searching for the target reacting cell, and updating the tree and the system configuration. For G_1 algorithms with various cell partitions, the costs are constant for the above operations except for the cost of searching for the target cell. Thus, it is reasonable to postulate that the CPU cost (C_{CPU}) in a typical reaction cycle is linearly related to the average search depth (*ASD*), i.e.,

$$C_{\text{CPU}} = C_0 + ASD \times C_1 \tag{41}$$

where C_0 is the total machine cost for all operations except for the searching step, and C_1 is the CPU time cost for each search step. Note that in the detailed algorithm, a typical search path for the target cell, say $(m_{1\mu}, m_{2\mu}, \ldots, m_{(N_L+1)\mu})$, will have a search depth of $\sum_{i=1}^{(N_L+1)} m_{i\mu}$. The numerical results confirm the linear dependence

Fig. 5 The linear dependence of CPU time per reaction cycle on the average search depth in algorithm G_1. The search depth of the circled points corresponds to cell partitions $m_1 = 64, 32, 16, 8, 4, 2, 1$
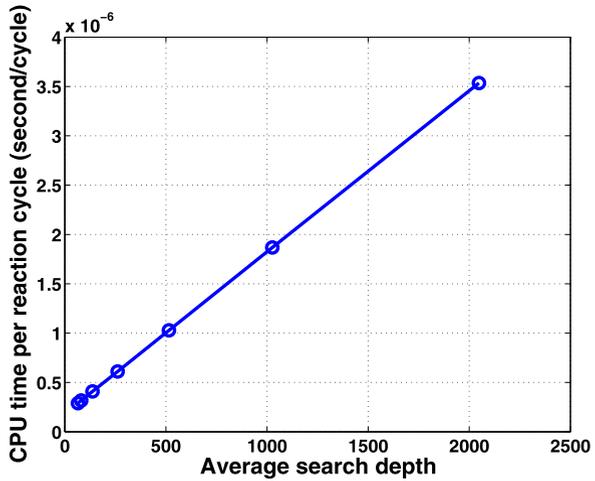

of the average CPU cost on the average search depth, as shown in Fig. 5. In particular, for pure diffusion of one species, the numerical computation in Fig. 5 shows that $(C_0, C_1) = (1.84 \times 10^{-7}, 1.64 \times 10^{-9})$ given the data structures in the current implementation. Therefore, the sum of the updating and other costs is equivalent to $1.84 \times 10^{-7}/1.64 \times 10^{-9} \approx 112$ searching steps. One thus predicts that the original Gillespie direct algorithm without optimization of the searching step can be accelerated by $\frac{112+(N_c+1)/2}{112+\sqrt{N_c}+1}$-fold using the G_1 algorithm, which is approximately 12 with $N_c = 4096$ in the current example.

We also ran the simulation using algorithm G_2. As previously described for the tree structure in Fig. 3, this algorithm requires three parameters, m_1, m_2 and m_3, but only two are independent since their product is N_c. We fixed the total number of computational cells to be $N_c = 4096$, and tried different values for (m_1, m_2, m_3). For uniform species distribution, the theoretical prediction of the average search depth for a reaction cycle in Table 1 is confirmed by the stochastic simulations, as shown in Fig. 6 (top). As in algorithm G_1, the time for updating the tree structure is negligible compared with the time for the searching steps. The average CPU cost per reaction cycle is also a linear function of the average search depth as in (41), and the function fitted in Fig. 6 (bottom) gives $(C_0, C_1) = (2.06 \times 10^{-7}, 1.33 \times 10^{-9})$. Thus the cost of other operations in an average reaction cycle is approximately equal to that of 155 searching steps. The minimal average search depth occurs when $m_1 = m_2 = m_3 = \sqrt[3]{N_c} = 16$, where the CPU cost is least.

We also tested the efficiency of algorithm G_{binary} on this example. For a fixed array of N_c cells, the tree structure associated with this algorithm is unique, unlike the previous two algorithms, which have variable parameters. Therefore, the search depth is $\log_2(N_c)$, and the number of updating steps is $2\log_2(N_c)$ in this example. A comparison between the three algorithms is summarized in Fig. 7. We discover that the optimal G_2 algorithm is faster than optimal G_1 algorithm by reducing the searching steps efficiently. Although the searching and updating steps are both logarithmic in N_c for G_{binary}, this algorithm is not as efficient as G_2. To understand this,

Fig. 6 The average search depth (*top*) and the dependence of CPU time on search depth (*bottom*) for four G_2 algorithms, each characterized by a choice of three parameters (m_1, m_2, m_3)
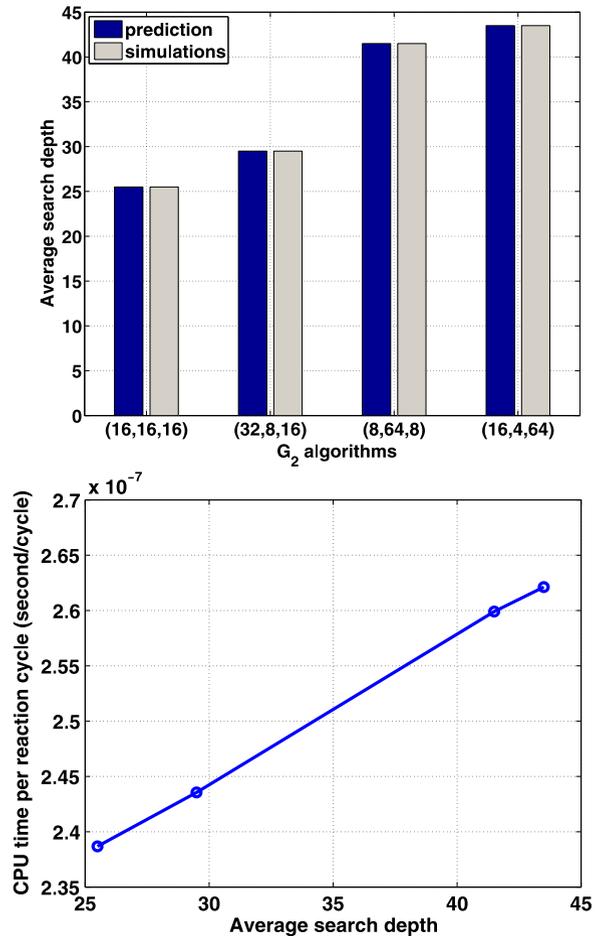
we plot the CPU cost per reaction cycle as a function of the average search depth for algorithm G_{binary} in Fig. 8. One sees there that the CPU cost is a highly nonlinear function of *ASD* over the range of search depths shown. The authors postulate that the searching/updating of the tree structure between levels is costly, whereas the searching amongst the lateral nodes within each level is cheaper given the data structure we are using in the simulation. Therefore, even though algorithm G_{binary} requires less searching and updating steps, algorithm G_2 uses more lateral searching rather than vertical searching in the tree and thus outperforms G_{binary}. Thus for large N_c, the most efficient algorithm with our optimization strategy is one with a tree structure that optimizes the combined lateral searching and vertical searching, which suggests tree structures with a small number of levels, rather than the binary tree used for G_{binary}. When the reaction network is complicated, species become more interconnected, and a reaction firing may affect the propensity of many other reactions, thereby lengthening the updating time in G_{binary}.

Fig. 7 The average CPU cost per reaction cycle (*left*) and numbers of searching/updating steps (*right*) are compared among the algorithms of G_1, G_2 and G_{binary}
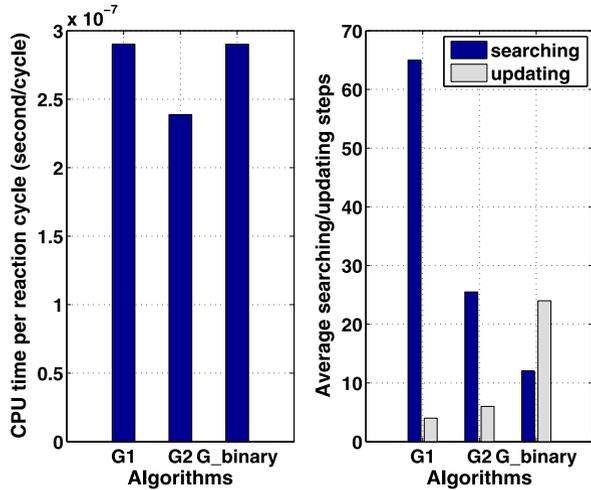
Fig. 8 The dependence of the average CPU cost on the average search depth for the G_{binary} algorithm
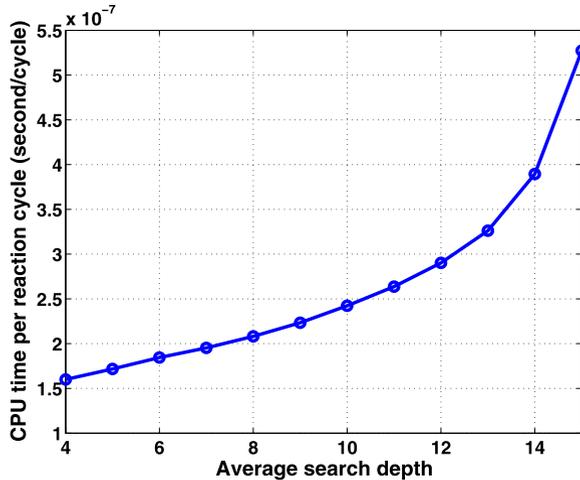
4.5.2 A Three Species Reaction–Diffusion System

In this example, molecules A and B bind reversibly and produce C. Only A and B can diffuse freely on the domain of size $6.4\ \mu m \times 6.4\ \mu m \times 0.1\ \mu m$. The domain is divided into cubes of side length $0.1\ \mu m$. The initial concentrations of A, B and C are 5.0, 2.0, and $0.0\ \mu M$, respectively. The on- and off-rates of the reaction are $1.0\ \mu M^{-1}s^{-1}$ and $1.0\ s^{-1}$. The diffusion coefficients for A and B are 2.0 and $1.0\ \mu m^2/s$, respectively.

As shown in Fig. 9, algorithm G_1 is optimal when partitioning the cells into $m_1 = \sqrt{N_c} = 64$ groups, as we predicted in the previous section. The CPU cost is also linear in the average search depth.

As for the pure diffusion problem, we compare the efficiency of three algorithms in simulating the three species system, and the results are summarized in Fig. 10. As in last example, algorithm G_2 is the most efficient among these three. However,

Fig. 9 The search depth and CPU cost of algorithm G_1 in a three species RD system
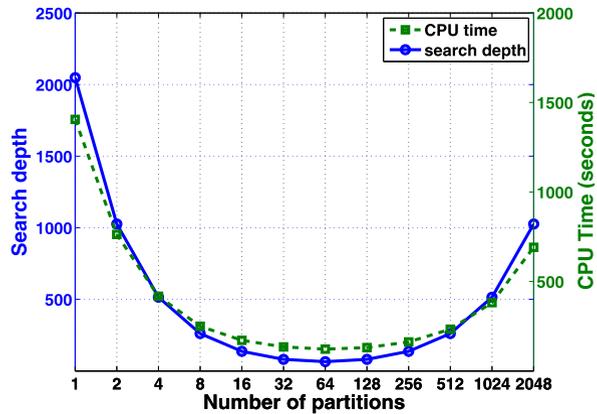
Fig. 10 The average CPU cost per reaction cycle (*left*) and the number of searching/updating steps (*right*) are compared among the optimal algorithms for G_1, G_2 and G_{binary}
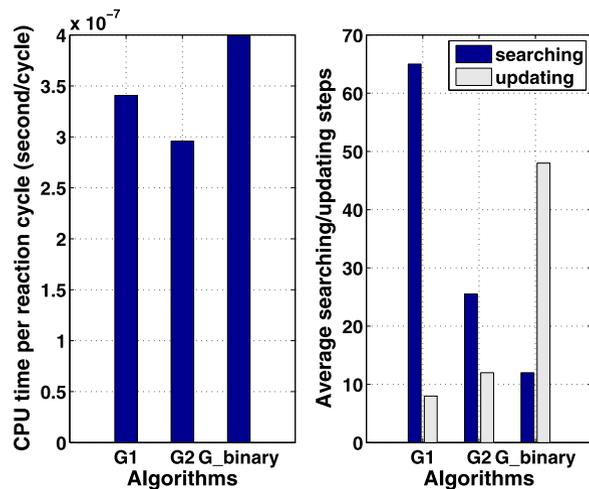
with more reactant interactions in the system, the occurrence of one reaction affects more reactions in the next round, with the result that the updating of G_{binary} is more costly, and this leads to slower performance of G_{binary} than that of G_1. As predicted in Table 1, the update cost is proportional to $\sum_{i=1}^{N_c} g_i$, which is the total affected reactions in all cells after the current reaction step. The current example gives rise to a higher average of total affected reactions due to the more complex reaction network, and the updating steps increase, as seen in a comparison of Fig. 7 and 10.

4.6 Other SSAs for RD Equations

There have been numerous applications of the direct method and next reaction method to spatially inhomogeneous systems by compartmentalizing the domain and regarding reactions in different compartment as distinct reactions (Ander et al. 2004; Cao et al. 2004; Li and Petzold 2006). Cao et al. (2004) carefully examined and compared the computational performance of the direct method and next reaction method.

They showed that an optimized direct method (ODM) could outperform the next reaction method except for a class of very loosely coupled systems. This is due to the high computational cost in maintaining the inherent data structure in the next reaction method, although the method has less computational operations. Li and Petzold (2006) improved the ODM further by proposing the logarithmic direct method (LDM) with a binary search tree. In the LDM, an array of N_c accumulated partial sums of propensities among all cells is created, on which a binary search is conducted. Even though searching for the target cell could be as fast as logarithmic in N_c, the updating of the array could be expensive. In that algorithm, the array is updated upward from the first partial sum which was changed by the current reaction. Thus, if it is the first of the array of partial sums that changes, the remaining $N_c - 1$ partial sums must be updated. For large N_c this again may become the bottleneck for this algorithm.

Another algorithm, called the next subvolume method (NSM), was proposed by Elf and Ehrenberg (2004). The next subvolume method (NSM) combines the next reaction method for determining the reacting subvolume in which the next reaction or diffusion occurs, with Gillespie's direct method for determining which reaction or diffusion step occurs in that subvolume. There is no optimization in the NSM algorithm.

In the algorithm developed here a tree structure defined on the set of computational cells is used to facilitate the searching and updating procedures. There are two major parameters associated with the tree—the number of levels in the tree and the number of daughter nodes emanating from a parent node at each level. One can optimize the performance of this algorithm by testing benchmark problems with varying parameters. As analyzed above, for a fixed tree depth there may be an optimal partition within levels. When the tree depth becomes larger, the search depth decreases whereas the updating steps increase. There may be crossover in the performance of this algorithm when the tree depth increases, due to the increased updating cost. Of course the crossover will depend on the complexity of the reaction network. Our experience suggests that for large problems one should experiment with these numbers to determine the best choices. For instance, one could with tree depths from one to five, and compare the performance with the G_{binary} algorithm. But a more rigorous investigation into the optimization problem of the algorithm remains to be done.

Aside from the optimization of partitioning of the compartment, the primary difference between our method and the NSM lies in the order in which operations are performed. We first sum the reaction rates of each reaction type across the subvolumes to decide where reaction occurs (using the direct method) and then search the tree to determine where it occurs. In contrast, the NSM sums all reaction rates in each subvolume, and first determines in which subvolume reaction occurs (using the next reaction method), and then decides the reaction type within the subvolume (using the direct method). Our method is the most highly optimized one using the direct method, and in particular, it eliminates the bottleneck in the LDM method by using a tree structure, which speeds up both the searching and updating steps. However, at present we cannot claim that the algorithm is best for all cases, and a performance comparison between our method and others is needed.

5 Computational Analysis of a Nonlinear RD System

With an efficient computational algorithm for stochastic RD systems at hand, we turn to the analysis of a nonlinear RD system with the objective of illustrating when the deterministic and stochastic descriptions agree and when their predictions differ. While no hard and fast rules can be stated, some cautionary notes on the interpretation of stochastic simulation results emerge. We analyze the deterministic behavior of the system using bifurcation analysis of the PDEs, and compare the predicted dynamics with that from a stochastic description.

Consider the simplified model of the key steps in the glycolytic pathway in a distributed system given in Ashkenazi and Othmer (1978):

$$\frac{\partial c_1}{\partial t} = \mathcal{D}_{c1}\Delta c_1 + k_0 - k_1 c_1 c_2^2 - k_2 c_1,$$
$$\frac{\partial c_2}{\partial t} = \mathcal{D}_{c2}\Delta c_2 + k_1 c_1 c_2^2 + k_2 c_1 - k_3 c_2,$$
(42)

with Neumann boundary conditions. The polynomial form of the reaction kinetics is a reduced form of a more complex model of enzyme activations and is used here for ease of analysis (Othmer and Aldridge 1978). In general trimolecular kinetics are not physically realistic.

In order to obtain spatially non-uniform solutions, we assume that $\mathcal{D}_{c1} > \mathcal{D}_{c2}$, which is necessary for a Turing instability. To rewrite the equations in a dimensionless form, define the dimensionless variables

$$\tau = k_3 t, \qquad \kappa = \frac{k_2}{k_3}, \qquad \delta = \frac{k_0}{k_3}\sqrt{\frac{k_1}{k_3}},$$
$$u = \sqrt{\frac{k_1}{k_3}}c_1, \qquad v = \sqrt{\frac{k_1}{k_3}}c_2, \qquad \mathcal{D}_u = \frac{\mathcal{D}_{c1}}{L^2 k_3}, \qquad \theta = \frac{\mathcal{D}_{c2}}{\mathcal{D}_{c1}},$$
(43)

where L is a measure of the domain size. The dimensionless equations are

$$\frac{\partial u}{\partial \tau} = \mathcal{D}_u \Delta u + \delta - \kappa u - uv^2,$$
$$\frac{\partial v}{\partial \tau} = \theta \mathcal{D}_u \Delta v + \kappa u + uv^2 - v,$$
(44)

where the Laplacian is now in scaled coordinates. For later comparison with stochastic simulations of this system we will regard the domain as a quasi-1D domain that is a rectangular solid of small cross-section in the y and z directions. We focus on the three parameter sets given in Table 2 to illustrate different aspects of the relationship between solutions of the deterministic system (42) and the corresponding stochastic system.

5.1 Bifurcation Analysis

Bifurcation diagrams constructed using AUTO-07P and δ as the bifurcation parameter, with other parameters at the fixed values in Table 2, are shown in Figs. 11 and 12.

Table 2 The three sets of parameters used for the dimensionless glycolytic reaction system

| Parameters	$\mathcal{D}_u$	θ	δ	κ
Set 1	2/9	0.1	2	0.1
Set 2	2/9	0.75	0.8	0.01
Set 3	2/9	0.005	2	0.1
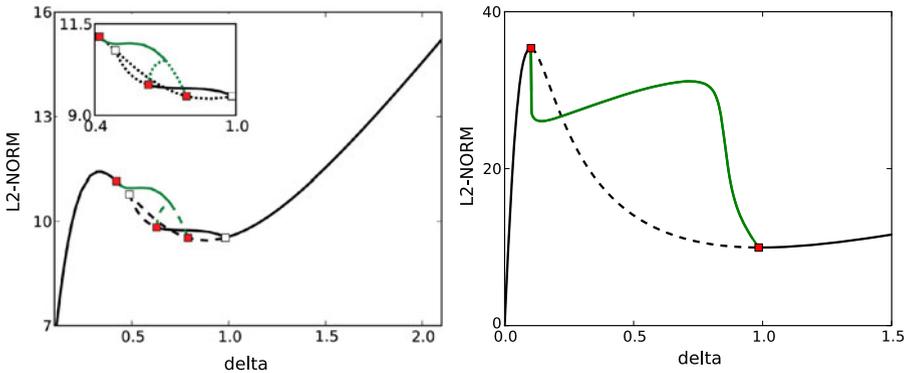
Fig. 11 The bifurcation diagram with respect to δ, using parameters other than δ as in Set 1 (*left*) and Set 2 (*right*) in Table 2. Here and hereafter *solid* (*dashed*) *lines* denote stable (unstable) solutions, *open squares* denote bifurcation points for static solutions, and *filled* (*red*) *squares* denote Hopf bifurcation points
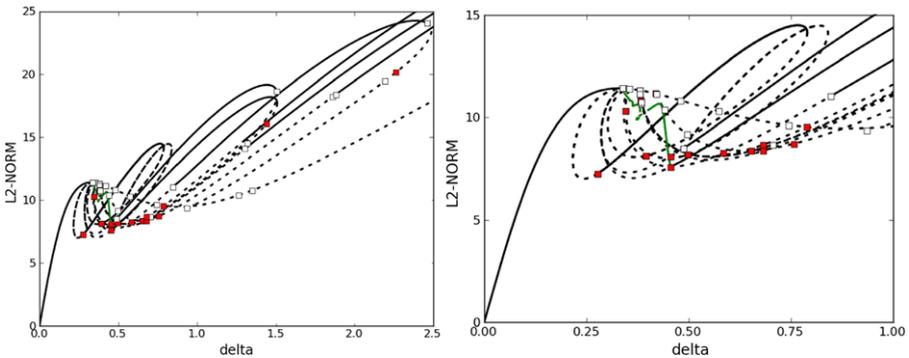
Fig. 12 The bifurcation diagram (**a**) and a blowup of one segment (**b**) with respect to δ using Set 3 in Table 2

One sees in Fig. 11 (left) that for an intermediate ratio of the diffusion coefficients there are multiple bifurcations in the range of 0.25–1.0, but at $\delta = 2$, which is used later, there is a unique stable steady state. At a larger ratio of the diffusion coefficients no Turing instabilities are possible and the only bifurcation is to a uniform stable periodic solution that exists for a range of input fluxes (Fig. 11 (right)).

When θ is small enough, as in the third set of parameters, there are many bifurcations and one finds coexistence of multiple stable non-uniform steady states in Fig. 12(a) and in the blowup shown in Fig. 12(b). One also sees that the interaction
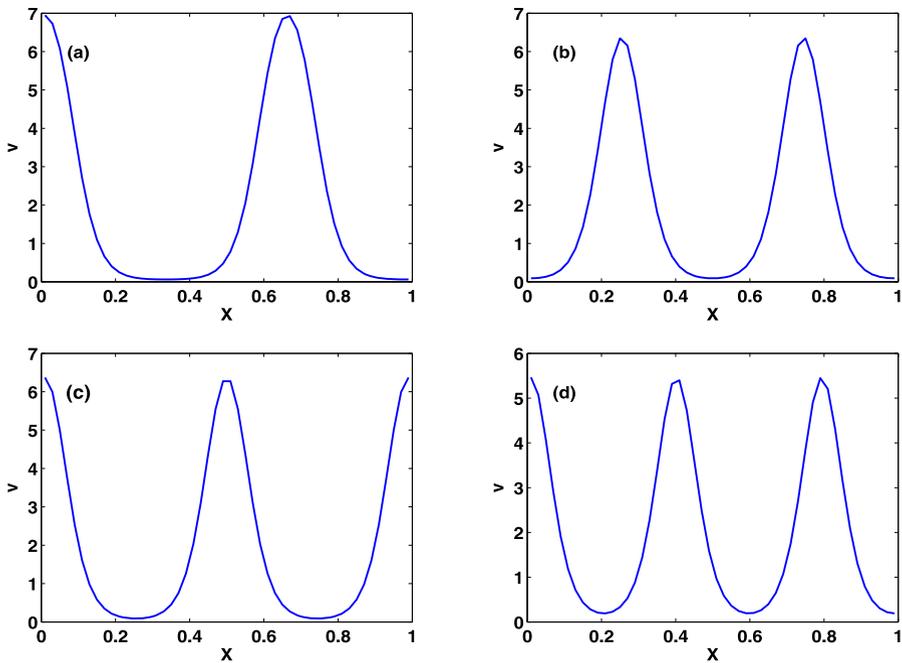
Fig. 13 Coexistence of multiple stable non-uniform steady-state solutions at $\delta = 2$

of the stationary and Hopf branches is very complex, and we have not attempted to resolve all the details. At $\delta = 2$, in Fig. 12, which is the base value in Table 2, there are several unstable and stable steady-state solutions. Some of the stable steady-state solutions are shown in Fig. 13, and there are also symmetry pairs for solutions that are symmetric about the mid-point. In the deterministic simulation, the initial condition uniquely determines the type of the steady-state solution, but later we show that stochastic simulations migrate from one type of solution to another type in time.

5.2 An Estimate of the Domain Size

In Sect. 3 we introduced two methods to estimate an upper bound for the maximal compartment size in the stochastic simulation of RD systems. In the first approach using the PDE-driven criterion, one must compute the appropriate reaction sensitivities over a positively invariant region of the local dynamics. With diagonal diffusion, as in the model, that region must be a rectangle (Chueh et al. 1977). However, the polynomial form does not have an invariant rectangle, and therefore we consider the modified system

$$
\begin{aligned}
\frac{du}{d\tau} &= \delta - \kappa u - \frac{uv^2}{1 + \epsilon v^2}, \\
\frac{dv}{d\tau} &= \kappa u + \frac{uv^2}{1 + \epsilon v^2} - v.
\end{aligned}
\tag{45}
$$

We have

$$\frac{du}{d\tau} = \delta - \kappa u - \frac{uv^2}{1 + \epsilon v^2} \leq \delta - \kappa u,$$

$$\frac{dv}{d\tau} = \kappa u + \frac{uv^2}{1 + \epsilon v^2} - v = \left(\kappa + \frac{v^2}{1 + \epsilon v^2}\right)u - v \qquad (46)$$

$$\leq \left(\kappa + \frac{1}{\epsilon}\right)u - v.$$

Therefore, there exists an invariant rectangle, $\Sigma = [0, u_{\max}] \times [0, v_{\max}]$, where

$$u_{\max} = \frac{\delta}{\kappa} \quad \text{and} \quad v_{\max} = \left(\kappa + \frac{1}{\epsilon}\right)\frac{\delta}{\kappa}. \qquad (47)$$

Next we derive the maximal cell size, but to simplify the calculations, we do it for the polynomial system. We will see later that the invariant rectangle gives a very conservative estimate of the size. We have

$$\mathcal{R}(u, v) = \begin{bmatrix} \delta - \kappa u - uv^2 \\ \kappa u + uv^2 - v \end{bmatrix},$$

and therefore

$$D_{(u,v)}\mathcal{R}(u, v) = \begin{bmatrix} -\kappa - v^2 & -2uv \\ \kappa + v^2 & 2uv - 1 \end{bmatrix}.$$

We compute $\|D_{(u,v)}\mathcal{R}(u, v)\|_E = \sqrt{\lambda_{\max}(D_{(u,v)}\mathcal{R}(u, v)^* D_{(u,v)}\mathcal{R}(u, v))}$ where $*$ is the conjugate transpose, and we obtain

$$D_{(u,v)}\mathcal{R}(u, v)^* D_{(u,v)}\mathcal{R}(u, v) = \begin{bmatrix} 2(\kappa + v^2)^2 & (\kappa + v^2)(4uv - 1) \\ (\kappa + v^2)(4uv - 1) & (2uv)^2 + (2uv - 1)^2 \end{bmatrix}.$$

The eigenvalues of $D_{(u,v)}\mathcal{R}(u, v)^* D_{(u,v)}\mathcal{R}(u, v)$ satisfy

$$\lambda^2 - \left\{2(\kappa + v^2)^2 + (2uv)^2 + (2uv - 1)^2\right\}\lambda + (\kappa + v^2)^2 = 0,$$

and therefore

$$\lambda = \frac{A(u, v) \pm \sqrt{A(u, v)^2 - B(u, v)^2}}{2},$$

where

$$A(u, v) = 2(\kappa + v^2)^2 + (2uv)^2 + (2uv - 1)^2,$$

$$B(u, v) = 2(\kappa + v^2).$$

Then

$$\hat{r}_{(u,v)} = \sup_{(u,v)\in C_\infty} \sqrt{\lambda_{\max}}$$

$$= \sup_{(u,v)\in C_\infty} \frac{\sqrt{A(u,v) + B(u,v)} + \sqrt{A(u,v) - B(u,v)}}{2}.$$

For a one-dimensional system of length 1, $\alpha_1 = -\pi^2$, and therefore the convergence condition becomes

$$\theta \mathcal{D}_u \pi^2 > \sup_{(u,v)\in C_\infty} \frac{\sqrt{A(u,v) + B(u,v)} + \sqrt{A(u,v) - B(u,v)}}{2}. \tag{48}$$

This provides a restriction on the scaled diffusion coefficient $\mathcal{D}_u$ of the system, and to convert it to a system size we use (43), and set $k_1 = 1/\mu M^2$ s and $k_3 = 1 \text{ s}^{-1}$. This leads to the following criterion on the allowable size of the system:

$$L < \sqrt{\frac{2\mathcal{D}_{c2}\pi^2}{\sup_{(u,v)\in C_\infty}\left(\sqrt{2(\kappa + v^2 + \frac{1}{2})^2 + 2(2uv - \frac{1}{2})^2} + \sqrt{2(\kappa + v^2 - \frac{1}{2})^2 + 2(2uv - \frac{1}{2})^2}\right)}}. \tag{49}$$

The minimum values of the $u_{\max}$ and $v_{\max}$ obtained for the three standard parameter sets are shown in the second column in Table 3.

These are rough estimates, and one can do better if one restricts attention to a certain class of initial conditions. To obtain estimates in this way, we numerically solve (44) with parameters given in Table 2, considering only initial data for which both species are zero, and obtain the maximal values given in the third column of Table 3. With these refined estimates we obtain the upper bounds on compartment size shown in the third column of Table 4.

The estimates in the fourth column of Table 4 are obtained by determining the value of N_c at which the CV stabilizes, using 1000 realizations of the system to

Table 3 $u_{\max}$ and $v_{\max}$ derived from the invariant region and from the numerical simulation using three sets of parameters in Table 2

Parameters	Invariant region	Numerical solutions
Set 1	$(20, 2 + 20/\epsilon)$	$(4, 4)$
Set 2	$(80, 0.8 + 80/\epsilon)$	$(8, 8)$
Set 3	$(20, 2 + 20/\epsilon)$	$(4, 8)$

Table 4 An upper bound of the domain size in the system with three sets of parameters as given in Table 2

| Parameters	Invariant region	Numerical solutions	CVs	Other estimates
Set 1	0.04 μm	0.2 μm	0.3 μm	0.08 μm
Set 2	0.03 μm	0.3 μm	0.6 μm	0.08 μm
Set 3	0.008 μm	0.03 μm	0.06 μm	0.009 μm
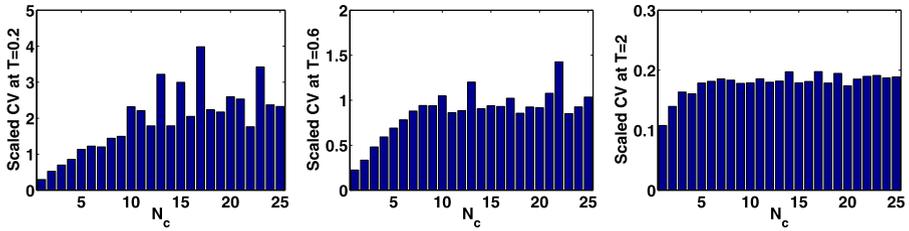
Fig. 14 $\overline{CV}^*(t)$ at $t = 0.2$ s, $t = 0.6$ s, and $t = 2.0$ s with different numbers of compartments using Set 1 in Table 2 with spatially non-uniform initial values
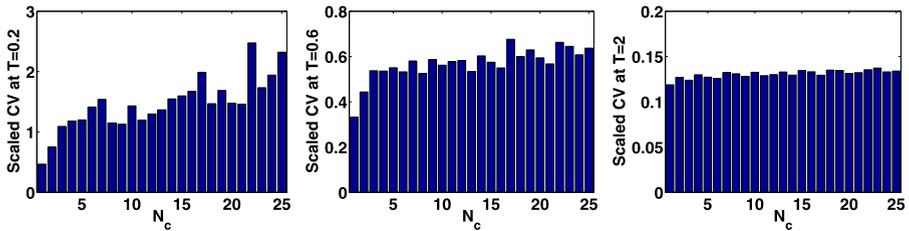

Fig. 15 $\overline{CV}^*(t)$ at $t = 0.2$ s, $t = 0.6$ s, and $t = 2.0$ s with different numbers of compartments using Set 2 in Table 2 with spatially non-uniform initial values

compute the scaled coefficient of variation in time scaled defined as

$$\overline{CV}^*(t) \equiv \sqrt{\frac{V}{N_c}} CV^*(t),$$

where $CV^*(t)$ is defined in (27). Stabilization of the stationary value $\overline{CV}^*$ of this was used in Kang et al. (2012) as a criterion to determine a lower bound for N_c, but this does not provide a lower bound when the system has a spatially uniform solution at the steady state. Therefore we consider the transient evolution, and because the reaction system in (42) is spatially homogeneous, we assume that all cells but the first are initially empty, and that the first contains 1000 molecules of each species. Under these conditions $\overline{CV}^*(t)$ can capture how fast spatial non-uniformity diffuses across the domain.

We show $\overline{CV}^*(t)$ at three times for different compartment numbers in Figs. 14–16. In Figs. 14 and 15, $\overline{CV}^*(t)$ stabilizes at $N_c \sim 10$ and $N_c \sim 5$, respectively, when $t = 0.6$ s, which leads to the upper bounds

$$h < \frac{3\ \mu\text{m}}{10} = 0.3\ \mu\text{m}, \tag{50}$$

$$h < \frac{3\ \mu\text{m}}{5} = 0.6\ \mu\text{m}. \tag{51}$$

One sees that the compartment size for parameters in Set 1 is smaller than that with Set 2, because the larger θ and smaller κ in Set 2 imply faster spread of species in the domain.
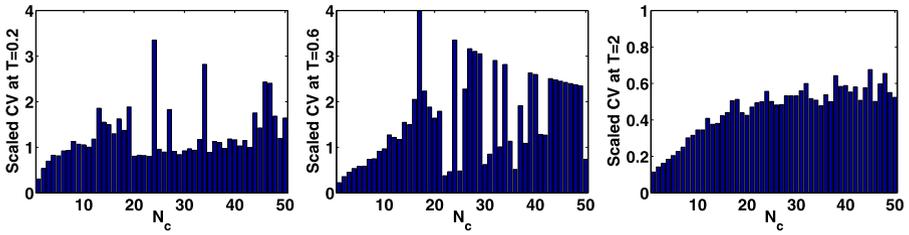
Fig. 16 $\overline{CV}^*(t)$ at $t = 0.2$ s, $t = 0.6$ s, and $t = 2.0$ s with different numbers of compartments using Set 3 in Table 2 with spatially non-uniform initial values

In Fig. 16 the result is less definitive. One might use $N_c = 30$ as a minimal number but there are still significant variations in the scaled CV, and if we use $N_c = 50$ we predict an upper bound of

$$h < \frac{3\ \mu m}{50} = 0.06\ \mu m. \tag{52}$$

The smaller size compared with the other estimates results from the smaller θ in Set 3. In this example, the compartment sizes estimated by CV stabilization as in the fourth column of the Table 4 are similar to those estimated by the PDE-driven criterion given in the third column of Table 4.

The last column of Table 4 gives an upper bound of the domain size based on the criterion suggested earlier (Bernstein 2005), which requires that the fastest reaction time scale should be slower than the slowest diffusion time scale. In the system (44), the nonlinear reaction has the largest propensity among the reactions based on the parameters in Table 2 and $(u_{\max}, v_{\max})$ given in the third column in Table 3. This criterion requires that

$$\frac{1}{2\theta \mathcal{D}_u} < \frac{1}{u_{\max} v_{\max}^2},$$

and using $\mathcal{D}_u = \mathcal{D}_{c1}/(L^2 k_3)$ and setting $\mathcal{D}_{c1} = 2\ \mu m^2/s$ and $k_3 = 1\ s^{-1}$, we obtain the upper bounds given in the last column in Table 4.

Table 4 shows that the invariant region method and Bernstein's method give comparable estimates of maximal compartment sizes, and both are significantly smaller than estimates based on solving the PDEs for a restricted class of initial data, or computing the CV from stochastic simulations. One reason for the rough estimation using the invariant region method is because the negative nonlinear term is omitted in the estimation in (46), so the interaction between u and v cannot be captured in the estimates. The numerical solutions of PDEs give finer estimates but they use both $u_{\max}$ and $v_{\max}$ of the PDE solution in (49), although it may not be true that $u_{\max}$ and $v_{\max}$ occur at the same time point of the PDE solution. In any case, the results suggest that various criteria should be tested when beginning analysis of a new reaction network.
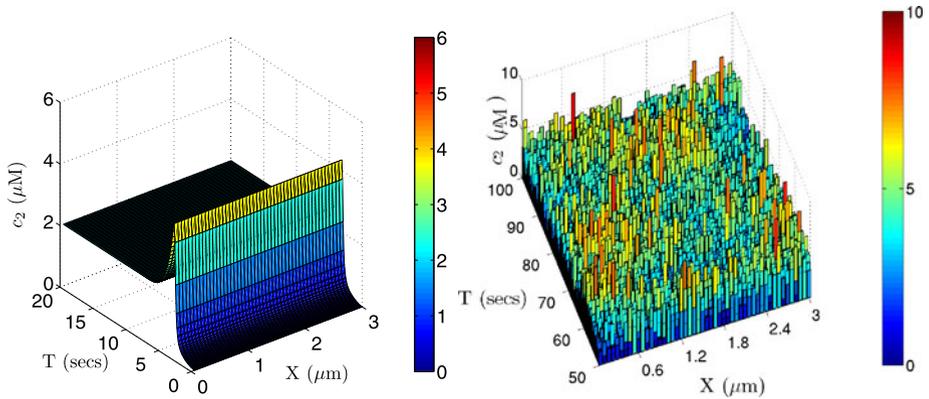
Fig. 17 The solutions for the second component of the glycolytic reaction with parameter set 1. The PDE solution with zero initial condition is shown *on the left*, and *on the right* is one realization of a stochastic simulation of the system using 50 cells along x-direction of the domain of size $(3, 0.2, 0.2)$ µm

5.3 Comparison of the Deterministic and Stochastic Dynamics

In this section we compare the deterministic and stochastic behavior of the model for the glycolytic reactions. The three parameter sets given in Table 2 are used here, the first of which leads to a unique uniform steady state, the second to a unique uniform periodic solution, and the third to one of several stable steady states.

The bifurcation analysis indicates that the RD system corresponding to the first set of parameters has a unique and stable steady-state solution. Starting with uniform zero conditions, the system evolves to the unique stable steady state as shown in Fig. 17 (left). The PDE is solved using 50 grid points in the 1D 3 µm domain on the x-axis. To simulate the dynamics of this 1D system stochastically, we use a rectangular domain of size $(3, 0.2, 0.2)$ µm in the x-, y- and z-directions, respectively. The width in the y- and z-directions is chosen such that it is within the maximal compartment size as estimated previously in the *CV* analysis. We partition the domain into 50 compartments in the x-direction, and use the optimized search algorithm as described in the previous section to simulate the dynamics. The stochastic simulation indicates that the system approaches a uniform steady state at large times, but due to the small volume of the computational cells, the fluctuation in number of the second species is significant, and the average concentration in individual compartments can be several-fold higher than the steady-state prediction of the PDE. In a typical realization there are about 360 molecules in the entire system.

Similarly, we compare the solution of the deterministic PDE and the predictions of the stochastic system (Fig. 18), using parameter set 2 for both. Starting from zero initial conditions, the deterministic system evolves to a periodic solution, which the bifurcation analysis done earlier shows is the only stable solution. The stochastic dynamics also evolve to a periodic solution with similar period (Fig. 18, right), but the fluctuations in the peak concentrations can be large. Despite this, the long-term evolution of the stochastic dynamics reflect those of the PDE prediction with the first two parameter sets, due to the fact that the system has a unique and stable solution in each case.
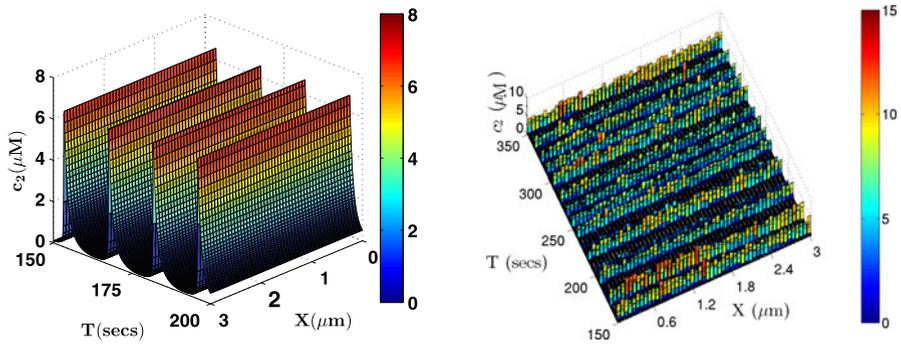
Fig. 18 The solutions of the glycolytic RD system with parameter set 2. The periodic solution of the PDE is displayed *on the left* with 50 grid points on the 1D 3.0 μm domain. *On the right* is one realization of a stochastic simulation of the system with 50 compartments along the x-axis of the domain of size $(3, 0.3, 0.3)$ μm
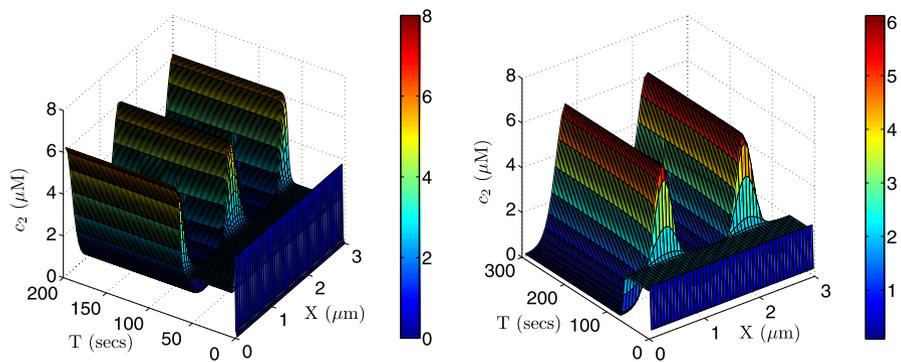

Fig. 19 The PDE solution of the glycolytic RD system with parameter set 3. With the same 50-compartment discretization of the 1D 3 μm domain, different initial conditions lie in the domain of attraction of different steady states. In the *left panel* the initial condition is uniform $(c_1, c_2) = (0, 0)$ μM, whereas *on the right* the initial condition is uniform $(c_1, c_2) = (0.1, 0.05)$ μM

In contrast, the evolution of the RD system using the third parameter set is much more complicated. The bifurcation analysis done earlier based on 50-grid-point discretization of the PDE shows that there are several coexisting uniform and non-uniform steady-state solutions, but only the non-uniform steady states are stable (Fig. 13). As a result, different initial conditions are required to begin in the domain of attraction of different steady states, as shown in Fig. 19. In addition, we find that different discretizations of the PDE may lead to distinct steady states, starting with the same zero initial conditions, as shown in Fig. 20. In all cases the maxima of the solutions are comparable, ranging from 5.5–7.0 μM as was found in the bifurcation analysis.

To simulate the stochastic dynamics of this system, we partition the domain into 50 compartments along the x-axis, which yields a maximal compartment size consistent with the *CV* analysis. We also choose different cross-sectional sizes in the y–z plane
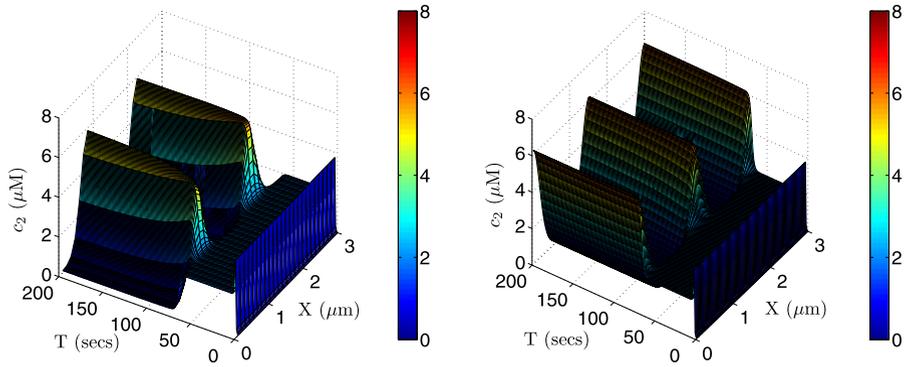
Fig. 20 Starting with the same zero initial condition, discretizations of the 1D 3 μm domain into 25 (*left*) and 100 (*right*) cells lead to different stable steady states
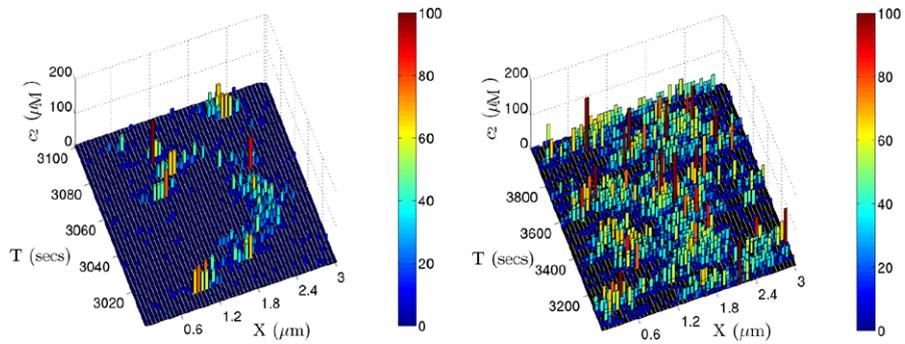

Fig. 21 The evolution of one realization of a stochastic simulation of the glycolytic RD system with parameter set 3. The domain $(3, 0.05, 0.05)$ μm is divided into 50 cells along the x-axis. The dynamics for 100 seconds and 1000 seconds are shown on *left* and *right*, respectively

but do not further discretize them, and we compare the resulting dynamics. We first use a cross section that lies within the maximal compartment size, which is 0.06 μm. An example of the dynamics in this discretization indicates that the stochastic simulation results are quite different from those of the deterministic prediction of the PDE. As shown in Fig. 21, the evolution of the stochastic simulation shows no clear spatial pattern, and in particular, it is far from the distinct three-peak non-uniform steady state shown in Fig. 19 (left).

However, if we increase the cross-sectional area of the compartment in the y–z plane to 0.5×0.5 μm^2, and scale the kinetic coefficients appropriately for the increase in volume, the system evolution shows clearly identifiable pattern formation with spatial distributions corresponding to different steady states predicted by the bifurcation and PDE analysis. Different initial random seeds for the optimized Gillespie algorithm result in completely different behavior of both dynamic and long-term behavior of the system, as indicated in Fig. 22. These simulations show at least two features of the stochastic dynamics of this nonlinear system. Firstly, depending on the initial seed, the intrinsic fluctuations in the stochastic system may lead
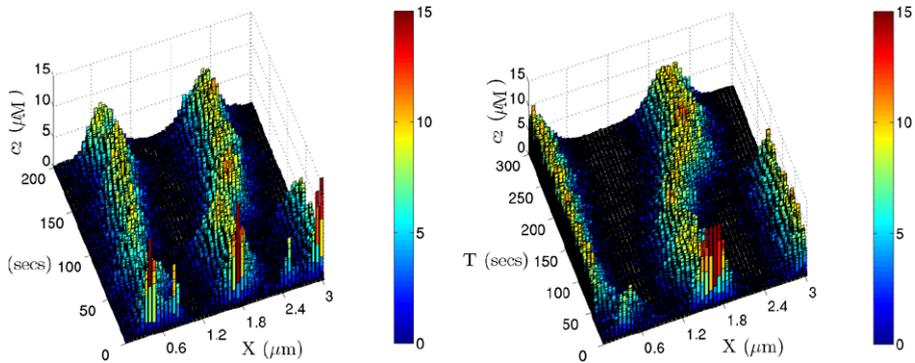
Fig. 22 The evolution of two realizations of the glycolytic RD system with parameter set 3. The domain (3, 0.5, 0.5) µm is divided into 50 cells along the x-axis. Different random number seeds lead to different steady states: that corresponding to Fig. 13(b) *on the left* and Fig. 13(a) *on the right*

Fig. 23 The evolution of one realization of the RD system with parameter set 3. The domain of size (3, 0.9, 0.9) µm is divided into 50 cells along the x-axis, and with this choice the stochastic simulations always converge to the same distribution, which is also a deterministic solution of the PDE
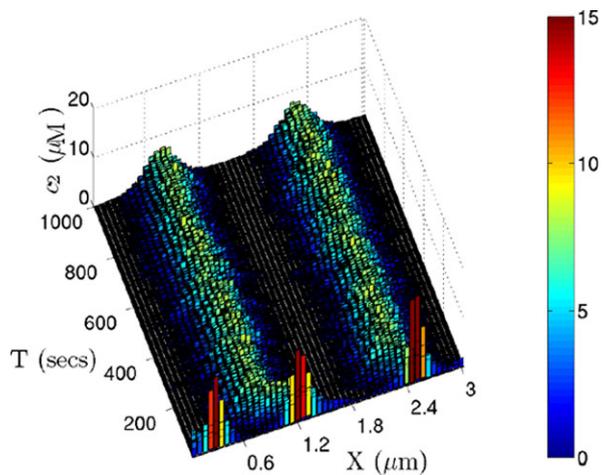
to stationary distributions that are very different from the stable steady states for the PDE, and secondly, during the early dynamics, the system may migrate from one stable steady state of the PDE to another. However, the results suggest that the final stationary distribution that the system evolves to has a large enough domain of attraction that the fluctuations do not drive the system from this distribution.

If we further increase the cross-section to 0.9×0.9 µm^2, the stochastic simulation always approaches the same stable non-uniform steady state as shown in Fig. 23. The deterministic prediction in Fig. 19 shows that different initial conditions drive the system into different steady states, but the stochastic simulation always evolves to a distribution corresponding to that obtained from the PDE with non-zero initial condition. This indicates that even with zero initial conditions the intrinsic fluctuations are large enough to drive the system to this stationary distribution.

Fig. 24 The temporal evolution of the total number of second component of the glycolytic system for three different cross section sizes of the domain as in Figs. 21–23. The mean of the total number of molecules in the whole domain for three cross sections from least to largest are 9.01, 902.4, 2925.4, respectively, whereas the variances are 0.647, 0.037, 0.02
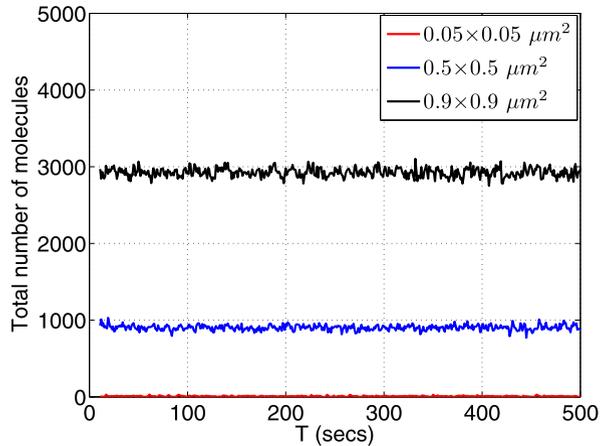
In these simulations we have assumed that the distributions on a cross-section are uniform, and we find that as the cross-section increases the solution of the stochastic problem resembles a solution of the PDE more reliably. On the smallest cross section of $0.05 \times 0.05 \ \mu m^2$, the dynamics of the system wanders 'aimlessly', whereas on a larger cross section of $0.5 \times 0.5 \ \mu m^2$ we observe more persistent wandering between limited numbers of stationary distributions. With an even larger cross section of $0.9 \times 0.9 \ \mu m^2$, the dynamics always leads to one stationary distribution. The underlying reason for this progression is simple: we scale coefficients with the volume, which leads to a larger amount of molecules, and thus the stochastic simulation corresponds more closely with the PDE version of the system dynamics. Particularly for the glycolytic system, which has several coexisting steady states, the smaller the cross section is, the larger the fluctuations in the numbers of molecules in individual compartments. The evolution of the total number of molecules for these different cross sections is depicted in Fig. 24.

One limitation of the larger cross-section is that the compartment width in the y- and z-directions may exceed the maximal compartment size, and thereby invalidate the stochastic simulation. In other words, the real dynamics in the full 3D system with the same large cross section may show a non-uniform distribution of species in the y–z plane, under which the assumption of a well-mixed cross section is no longer appropriate. To demonstrate this, we used the optimized algorithm to do a stochastic simulation of the full 3D system similar to the above, but further divided the $0.9 \times 0.9 \ \mu m^2$ cross section into 20×20 compartments. First, we examine the averaged concentration of the cross section along the x-axis. The result in Fig. 25 shows a temporally changing dynamics in the pattern which is quite different from that in Fig. 23. Interestingly, the averaged cross-sectional concentrations shown there occasionally resemble the non-uniform steady states predicted by the PDE, but detailed plots of the species concentration on various cross sections along the x-axis show that the cross sections are clearly not well-mixed (cf. Fig. 26). Thus the assumption of well-mixed species for the large cross-sections is not valid in this system. We also extracted the dynamics of the compartments with same y–z position along the x-axis as depicted in Fig. 27, and found that the dynamics along these fibers show no clear
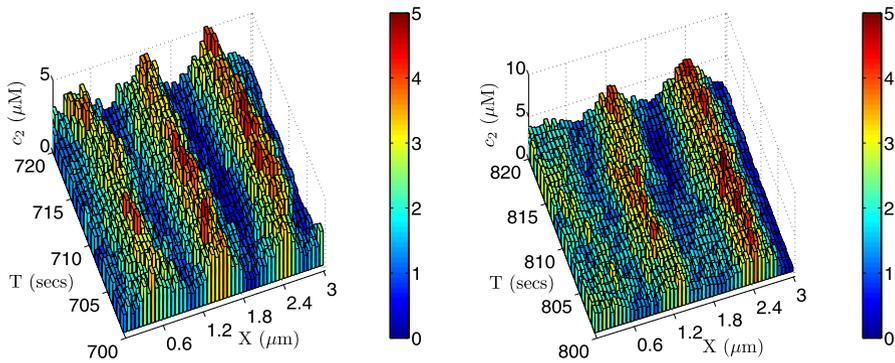
Fig. 25 The evolution of one realization of the glycolytic RD system with parameter set 3. The domain of size $(3, 0.9, 0.9)$ μm has 50, 20 and 20 divisions in the x-, y- and z-directions, respectively. Shown are the dynamics in different time intervals, and the concentration in any y−z plane is averaged and projected to the x-axis
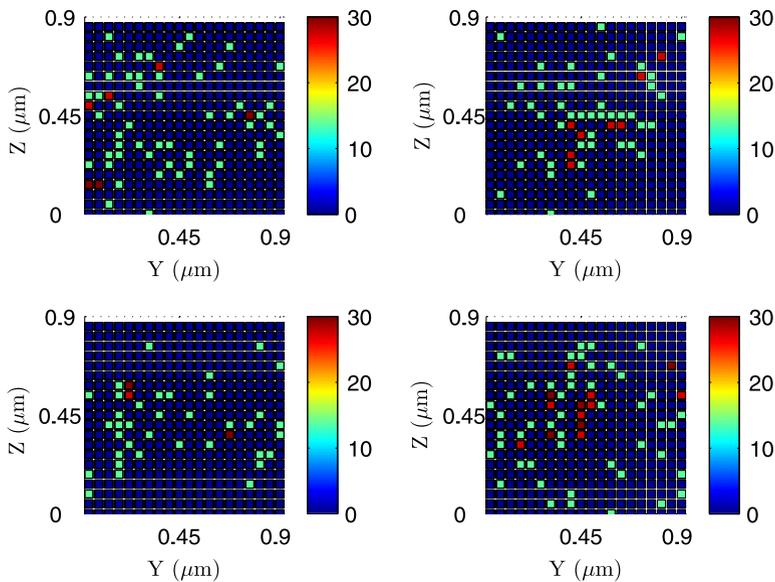

Fig. 26 The dynamics on four cross-sections along the x-axis extracted from the full 3D dynamics at $t = 1000$ s. The x-coordinates are 0.6, 1.2, 1.8, 2.4 μm, respectively

patterns close to the steady state predicted by the PDE. In fact they resemble the earlier results on a small cross section more closely.

5.4 Conclusion

Stochastic simulations of RD systems raise a number of issues that must be confronted in evaluating the results. First is how one can determine the appropriate compartment size, and given that, how to deal with the increased computational com-
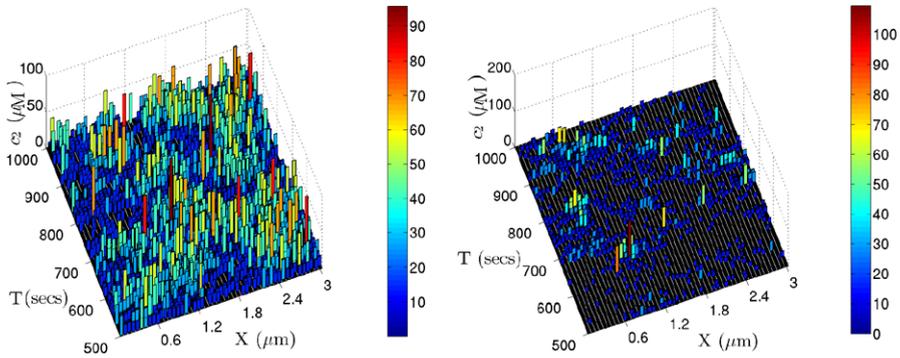
Fig. 27 Dynamics in two fibers of compartments along the x-axis extracted from the full 3D dynamics of the glycolytic reactions. The *left fiber* has coordinates $(y, z) = (0.45, 0.45)$, whereas the *right* is $(y, z) = (0.9, 0.9)$

plexity of distributed systems. We have given several criteria that can be used for the former, and have developed a new computational algorithm that optimizes the search process in a Gillespie algorithm. Our comparison of predictions from the PDE with those from stochastic simulations of the 1D system shows that the choice of the computational cell size is critical, and that results obtained using a sufficiently small size may bear little relation to the deterministic results when there are multiple solutions. In that case the dynamics of the stochastic simulation may wander between these steady states, with the residence time in each depending on the size of the attraction basin. This is less of a problem when the PDE has a unique attractor, but spatial non-uniformity may still lead to significant differences in the predicted behavior when the number of molecules is low.

References

Agbanusi, I. C., & Isaacson, S. A. (2013). *A comparison of bimolecular reaction models for stochastic reaction diffusion models*. arXiv:1301.0547.

Ander, M., Beltrao, P., Di Ventura, B., Ferkinghoff-Borg, J., Foglierini, M., Kaplan, A., Lemerle, C., Tomas-Oliveira, I., & Serrano, L. (2004). SmartCell, a framework to simulate cellular processes that combines stochastic approximation with diffusion and localisation: analysis of simple networks. *Syst. Biol.*, *1*(1), 129–138.

Applebaum, D. (2004). *Lévy processes and stochastic calculus* (Vol. 93). Cambridge: Cambridge University Press.

Arnold, L. (1974). *Stochastic differential equations, theory and applications*. New York: Wiley-Interscience.

Ashkenazi, M., & Othmer, H. G. (1978). Spatial patterns in coupled biochemical oscillators. *J. Math. Biol.*, *5*, 305–350.

Bernstein, D. (2005). Simulating mesoscopic reaction–diffusion systems using the Gillespie algorithm. *Phys. Rev. E, 71(4 Pt 1)*, 041103.

Callen, H. B. (1960). *Thermodynamics*. New York: Wiley.

Cao, Y., Li, H., & Petzold, L. (2004). Efficient formulation of the stochastic simulation algorithm for chemically reacting systems. *J. Chem. Phys.*, *121*(9), 4059–4067.

Capasso, V., & Bakstein, D. (2005). *An introduction to continuous-time stochastic processes: theory, models, and applications to finance, biology, and medicine*. New York: Birkhauser.

Chueh, K. N., Conley, C. C., & Smoller, J. A. (1977). Positively invariant regions for systems of nonlinear diffusion equations. *Indiana University Math. J.*, *26*(2), 373–392.

Conway, E., Hoff, D., & Smoller, J. (1978). Large time behavior of solutions of systems of nonlinear reaction–diffusion equations. *SIAM J. Appl. Math.*, *35*(1), 1–16.

Elf, J., & Ehrenberg, M. (2004). Spontaneous separation of bi-stable biochemical systems into spatial domains of opposite phases. *IET Syst. Biol.*, *1*, 230–236.

Erban, R., & Othmer, H. (2005). From signal transduction to spatial pattern formation in *E. coli*: a paradigm for multi-scale modeling in biology. *Multiscale Model. Simul.*, *3*(2), 362–394.

Erban, R., & Othmer, H. G. (2007). Taxis equations for amoeboid cells. *J. Math. Biol.*, *54*(6), 847–885.

Gadgil, C., Lee, C. H., & Othmer, H. G. (2005). A stochastic analysis of first-order reaction networks. *Bull. Math. Biol.*, *67*, 901–946.

Gibson, M. A., & Bruck, J. (2000). Efficient exact stochastic simulation of chemical systems with many species and many channels. *J. Phys. Chem. A*, *104*, 1876–1889.

Gierer, A., & Meinhardt, H. (1972). A theory of biological pattern formation. *Biol. Cybern.*, *12*(1), 30–39.

Gillespie, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.*, *81*(25), 2340–2361.

Hillen, T., & Othmer, H. G. (2000). The diffusion limit of transport equations derived from velocity jump processes. *SIAM J. Appl. Math.*, *61*, 751–775.

Isaacson, S. A. (2009). The reaction–diffusion master equation as an asymptotic approximation of diffusion to a small target. *SIAM J. Appl. Math.*, *70*(1), 77–111.

Kang, H. W., Zheng, L., & Othmer, H. G. (2012). A new method for choosing the computational cell in stochastic reaction–diffusion systems. *J. Math. Biol.*, *65*, 1017–1099.

Kang, H. W., Zheng, L., & Othmer, H. G. (2012). The effect of the signalling scheme on the robustness of pattern formation in development. *Interface Focus*, *2*(4), 465–486.

Li, H., & Petzold, L. (2006). *Logarithmic direct method for discrete stochastic simulation of chemically reacting systems* (Technical Report). Department of Computer Science, University of California, Santa Barbara.

Matzavinos, A., & Othmer, H. G. (2007). A stochastic analysis of actin polymerization in the presence of twinfilin and gelsolin. *J. Theor. Biol.*, *249*, 723–736.

Metzler, R., & Klafter, J. (2000). The random walk's guide to anomalous diffusion: a fractional dynamics approach. *Phys. Rep.*, *339*(1), 1–77.

Othmer, H. G. (1976). Nonuniqueness of equilibria in closed reacting systems. *Chem. Eng. Sci.*, *31*, 993–1003.

Othmer, H. G. (1977). Current problems in pattern formation. In *Some mathematical questions in biology* (Vol. VIII, pp. 57–85). Providence: Am. Math. Soc.

Othmer, H. G., Dunbar, S. R., & Alt, W. (1988). Models of dispersal in biological systems. *J. Math. Biol.*, *26*, 263–298.

Othmer, H. G., & Aldridge, J. A. (1978). The effects of cell density and metabolite flux on cellular dynamics. *J. Math. Biol.*, *5*, 169–200.

Othmer, H. G., & Xue, C. (2013). The mathematical analysis of biological aggregation and dispersal: progress, problems and perspectives. In M. Lewis, P. Maini, & S. Petrovskii (Eds.), *Dispersal, individual movement and spatial ecology: a mathematical perspective*, Heidelberg: Springer.

Prigogine, I., & DeFay, R. (1954). *Chemical thermodynamics*. New York: Longmans, Green.

Sato, K. I. (1999). *Lévy processes and infinitely divisible distributions*. Cambridge: Cambridge University Press.

Shimmi, O., & O'Connor, M. B. (2003). Physical properties of Tld, Sog, Tsg and Dpp protein interactions are predicted to help create a sharp boundary in Bmp signals during dorsoventral patterning of the Drosophila embryo. *Development*, *130*(19), 4673–4682.

Tomioka, R., Kimura, H., Kobayashi, T. J., & Aihara, K. (2004). Multivariate analysis of noise in genetic regulatory networks. *J. Theor. Biol.*, *229*(4), 501–521.

Turing, A. M. (1952). The chemical basis of morphogenesis. *Philos. Trans. R. Soc. Lond. B, Biol. Sci.*, *237*, 37–72.

Umulis, D. M., & Othmer, H. G. (2013). *Mechanisms of scaling in spatial pattern formation*. Development (to appear).

Weiss, G. H. (1994). *Aspects and applications of the random walk*. Amsterdam: North-Holland.

Wilemski, G. (1976). On the derivation of Smoluchowski equations with corrections in the classical theory of Brownian motion. *J. Stat. Phys.*, *14*(2), 153–169.

Xue, C., & Othmer, H. G. (2009). Multiscale models of taxis-driven patterning in bacterial populations. *SIAM J. Appl. Math.*, *70*(1), 133–167.

Xue, C., Budrene, E. O., & Othmer, H. G. (2011). Radial and spiral stream formation in *Proteus mirabilis* colonies. *PLoS Comput. Biol.*, *7*(12), e1002332.