The Author(s) *BMC Systems Biology* 2017, **11**(Suppl 3):21
DOI 10.1186/s12918-017-0401-9

BMC Systems Biology

RESEARCH Open Access

Stochastic modeling and simulation of reaction-diffusion system with Hill function dynamics

Minghan Chen†, Fei Li†, Shuo Wang and Young Cao*

From: Third International Workshop on Computational Network Biology: Modeling, Analysis, and Control (CNB-MAC 2016) Seattle, WA, USA. 10 February 2016

Abstract

Background: Stochastic simulation of reaction-diffusion systems presents great challenges for spatiotemporal biological modeling and simulation. One widely used framework for stochastic simulation of reaction-diffusion systems is reaction diffusion master equation (RDME). Previous studies have discovered that for the RDME, when discretization size approaches zero, reaction time for bimolecular reactions in high dimensional domains tends to infinity.

Results: In this paper, we demonstrate that in the 1D domain, highly nonlinear reaction dynamics given by Hill function may also have dramatic change when discretization size is smaller than a critical value. Moreover, we discuss methods to avoid this problem: smoothing over space, fixed length smoothing over space and a hybrid method.

Conclusion: Our analysis reveals that the switch-like Hill dynamics reduces to a linear function of discretization size when the discretization size is small enough. The three proposed methods could correctly (under certain precision) simulate Hill function dynamics in the microscopic RDME system.

Keywords: Reaction diffusion master equation (RDME), Hill function, Stochastic simulation, Hybrid method

Background

Cell reproduction requires elaborate spatial and temporal coordination of crucial events, such as DNA replication, chromosome segregation, and cytokinesis. In cells, protein species are well organized and regulated throughout their life cycles. Theoretical biologists have been using classic chemical reaction rate laws with deterministic ordinary differential equations (ODEs) and partial differential equations (PDEs) to model molecular concentration dynamics in spatiotemporal biological processes. However, wet-lab experiments in single cell resolution demonstrate that biological data present considerable variations from cell to cell. The variations arise from the fact that cells are so small that there exist only one or two copies of genes, tens of mRNA molecules and hundreds or thousands of protein molecules [1–3]. At this scale, the traditional way of modeling molecule "concentration" is not applicable. Noise in molecule populations cannot be neglected, as noise may play a significant role in the overall dynamics inside a cell. Therefore, to accurately model the cell cycle mechanism, discrete and stochastic modeling and simulation should be applied.

A convenient strategy to build a stochastic biochemical model is to break a deterministic model into a list of chemical reactions and simulate them with Gillespie's stochastic simulation algorithm (SSA) [4, 5]. One of the major difficulties in this conversion strategy lies in the propensity calculation of reactions. Gillespie's SSA is well defined for mass action rate laws. However, in many biochemical models, in addition to mass action rate laws, other phenomenological reaction rate laws are often used. For example, the Michaelis-Menten equation [6] and Hill functions [7] are widely used in biological models to model the fast response to signals in regulatory control systems. Although theoretically these phenomenological

*Correspondence: ycao@cs.vt.edu
†Equal contributors
Department of Computer Science, Virginia Tech, Blacksburg, VA 24061, USA

BioMed Central

© The Author(s). 2017 **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.

rate laws may be generated from elementary reactions with mass action rate laws, in practice the detailed mechanisms behind these phenomenological rate laws are not well known and may not be very important. Stochastic modeling and simulation with these phenomenological rate laws are sometimes inevitable.

In recent years, stochastic modeling and simulation for spatiotemporal biological systems, particularly reaction-diffusion systems, have captured more and more attention. Several algorithms and tools [8–11] to model and simulate reaction-diffusion systems have been proposed. These methods can be categorized into two theoretical frameworks: the spatially and temporally continuous Smoluchowski modeling framework [12] and the compartment-based modeling framework, formulated as the spatially discretized reaction-diffusion master equation (RDME) [13, 14]. The Smoluchowski framework [12, 15, 16] stores the exact position of each molecule and is mathematically fundamental, whereas the RDME is coarse-grained and better suited for large scale simulations [17]. In RDME, the spatial domain is discretized into small compartments. Within each compartment, molecules are considered "well-stirred". Under the RDME scheme, diffusion is modeled as continuous time random walk on mesh compartments, while reactions fire only among molecules in the same compartment. Stochastic dynamics of the chemical reactions in each compartment is governed by the chemical master equation (CME) [18, 19]. Yet, the CME is computationally impossible to solve for most practical problems. Stochastic simulation methods were then applied to generate realizations of system trajectories. It has been well established that the discretization compartment size for RDME should be smaller than the mean free path of the reactions for the compartment to be considered as well-stirred [20]. In addition, it has been proved that the RDME of bi-molecular reactions in 3D domain becomes incorrect and yields unphysical results when the discretization size approaches microscopic scale [21–23].

In this paper, we focus on the stochastic modeling of reaction-diffusion systems with reaction rate laws given by Hill functions. In the Results section, we present our numerical analysis on a toy model of reaction-diffusion system with Hill function dynamics. We will show that the RDME framework of the Hill function dynamics has serious simulation defects when the discretization size approach microscopic limit: When the discretization size is small enough, the typical switching pattern of Hill dynamics becomes linear to the input signal (and the discretization size). Later, we propose potential solutions for the discretization of the reaction-diffusion systems with Hill function rate laws. Finally, we conclude this paper with a discussion on RDME for general nonlinear functions and the hybrid method.

Caulobacter modeling

Caulobacter crescentus captures great interest in the study of asymmetric cell division. When a *Caulobacter* cell divides, it produces two functionally and morphologically distinct daughter cells. The asymmetric cell division of *Caulobacter crescentus* requires elaborate temporal and spatial regulations [24–27]. In literature [28–30], four essential "master regulators" of the *Caulobacter* cell cycle, DnaA, GcrA, CtrA and CcrM, have been identified. These master transcription regulators determine the dynamics of around 200 genes. They oscillate temporally to drive the dynamics of cell cycle. Among them, the molecular mechanisms governing CtrA functions have been well studied. The simulation we are concerned with in this paper is also related to this CtrA module. So we give a brief introduction to it.

In swarmer cells, a two-component phosphorelay system (with both CckA and ChpT) phosphorylates the CtrA. Then the chromosomal origin of replication (*Cori*) is bound by the phosphorylated CtrA (CtrAp) to inhibit the initiation of chromosome replication [31]. Later during the swarmer-to-stalked transition period, CtrAp gets dephosphorylated and degraded, allowing the initiation of chromosome replication again. Thus the CtrA has important impact on the chromosome replication in our model, and should be well regulated.

The regulation of CtrA is achieved by the histidine kinase CckA through the following pathway. An ATP-dependent protease, ClpXP, degrades CtrA [32, 33] and is localized to the stalk pole by CpdR. As the nascent stalked cell progresses through the cell cycle, CpdR is phosphorylated by CckA/ChpT, losing it polar localization, and consequently losing its ability to recruit ClpXP protease for CtrA degradation. In addition, CtrA is reactivated through CckA/ChpT phosphorylation [34]. Moreover, the regulatory network of the histidine kinases CckA is influenced by a non-canonical histidine kinase, DivL [35]. DivL promotes CckA kinase, which then phosphorylates and activates CtrA in the swarmer cell. During the swarmer-to-stalked transition period, DivL activity is down-regulated, thereby inhibiting CckA kinase activity. As a result, dephosphorylation and degradation of CtrA trigger the initiation of chromosome replication.

In order to study the regulatory network in *Caulobacter crescentus*, Subramanian et al. [26, 27] developed a deterministic model with six major regulatory proteins. The deterministic model provides robust switching between swarmer and stalked states. Figure 1 (left) demonstrates the total population change during the *Caulobacter crescentus* cell cycle with this deterministic model. In the swarmer stage (from $t = 0$ to 30 min), the CtrA is phosphorylated at a high population level, which inhibits the initiation of chromosome replication. During the swarmer-to-stalked transition period (from $t = 30$ to

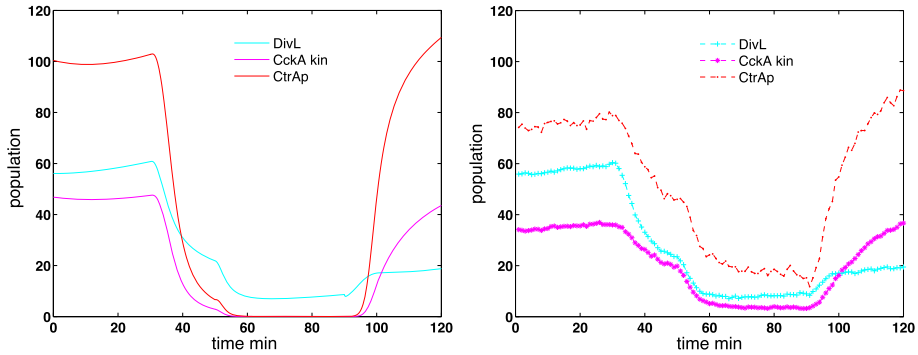


Fig. 1 The population oscillation of CtrAp during *Caulobacter crescentus* cell cycle. Left figure shows the simulation result of deterministic model and the right figure shows the stochastic simulation result. In the swamer stage ($t = 0 \sim 30$min), the CtrA is phosphorylated and at high population level state, which inhibits the initiation of chromosome replication. During swarmer-to-stalked transition ($t = 30 \sim 50$min), the CtrAp population quickly switch to low state, allowing the consequent initiation of chromosome replication in the stalked stage

50 min), the CtrAp population quickly drops to a low level, allowing the consequent initiation of chromosome replication in the stalked stage.

In stochastic simulation of the spatiotemporal model of this regulatory network, the phosphorylated CtrA (CtrAp) population switch from a high level in swarmer stage to a low level in stalked stage is not as sharp as expected, shown in Fig. 1 (right). On the other hand, the DivL population level from the stochastic simulation seems similar to that from the deterministic simulation. A simple analysis suggests that the Hill function dynamics, which models the up regulation of CckA kinase activity by DivL, might be the culprit. Further investigation leads to the discovery of the Hill function limitation at small discretization sizes, as analyzed in the next section.

Methods

Reaction diffusion master equation

Before we plunge into Hill functions in reaction-diffusion systems, we will first briefly review mathematical modeling and simulation methods of spatially inhomogeneous stochastic systems.

The dynamics of a spatially inhomogeneous stochastic system has been considered as governed by the reaction-diffusion master equation (RDME), developed in an early work of Gardiner [13]. The RDME framework partitions the spatial domain into small compartments, such that molecules within each compartment can be considered well-stirred. Assume a biochemical system of N species $\{S_1, S_2, \ldots, S_N\}$ and M reactions within a spatial domain Ω, which is partitioned into K grids V_k, $k = 1, 2, \ldots, K$. For simplicity, we assume that the space Ω is one dimensional (1D). Each species population, as well as the reactions in the system will have a local copy for each compartment. The state of the reaction-diffusion system at any time t is represented by the vector state vector $X(t) = \{X_{1,1}(t), X_{1,2}(t), \ldots, X_{1,K}(t), \ldots, X_{n,k}(t), \ldots, X_{N,K}(t)\}$, where $X_{n,k}(t)$ denotes the molecule population of species S_n in the grid V_k at time t. Reactions in each compartment is governed by the Chemical Master Equation (CME), while diffusion is modeled as random walk across neighboring compartments. Each reaction channel R_j in any compartment k can be characterized by the *propensity function* $a_{j,k}$ and the *state change vector* $\nu_j \equiv (\nu_{1j}, \nu_{2j}, \ldots, \nu_{Nj})$. The dynamics of the diffusion of species S_i from compartment V_k to V_j is formulated by the *diffusion propensity function* $d_{i,k,j}$ and the *diffusion state change vector* $\mu_{k,j}$ similarly. $d_{i,k,j}(x)dt$ gives the probability that, given $X_{i,k}(t) = x$, one molecule of species S_i at grid V_k diffuses into grid V_j in the next infinitesimal time interval $[t, t + dt)$. If $j = k \pm 1$, then $d_{i,k,j}(x) = \frac{D}{h^2}x$, where D is the diffusion rate coefficient and h is the characteristic length, also called discretization size, of a grid; Otherwise $d_{i,k,j} = 0$. The state change vector $\mu_{k,j}$ is a vector of length K with -1 in the k-th position, 1 in the j-th position and 0 everywhere else.

With the reaction-diffusion propensity functions and state change vectors, the RDME completely depicts the dynamics of the system:

$$\begin{aligned}\frac{\partial P(\mathbf{x}, t|\mathbf{x_0}, t_0)}{\partial t} &= \sum_{k=1}^{K}\sum_{j=1}^{M}\left(a_{j,k}(\mathbf{x} - \nu_{j,k})P(\mathbf{x} - \nu_{j,k}, t|\mathbf{x_0}, t_0) - a_{j,k}(\mathbf{x})P(\mathbf{x}, t|\mathbf{x_0}, t_0)\right) \\ &+ \sum_{i=1}^{N}\sum_{k=1}^{K}\sum_{j=1}^{K}\left(-d_{i,k,j}(x_{ik})P(\mathbf{x}, t|\mathbf{x_0}, t_0)\right. \\ &\left. + d_{i,k,j}(X_{ik} - \mu_{k,j})P(X_{11}, \ldots, X_{ik} - \mu_{k,j}, \ldots, X_{N,K}, t|\mathbf{x_0}, t_0)\right), \end{aligned} \quad (1)$$

where $P(\mathbf{x}, t|\mathbf{x_0}, t_0)$ denotes the probability that the system state $X(t) = \mathbf{x}$, given that $X(t_0) = \mathbf{x_0}$. The RDME is a set

of ODEs that gives one equation for every possible state. It is both theoretically and computationally intractable to solve RDME for practical biochemical systems due to the huge number of possible combinations of states. Instead of solving RDME for the time evolution of the probabilities, we can construct numerical realizations of $\mathbf{X}(t)$. A popular method to construct the trajectories of a reaction-diffusion system is to simulate each diffusive jumping and chemical reaction event explicitly. With enough trajectory realizations, we can derive the distribution of each state vector at different times.

The RDME model have been used as an approximation of the Smoluchowski framework in the mesoscopic scale. Furthermore, researches have discovered that in the microscopic limit, bimolecular reactions may be eventually lost when the grid size becomes infinitely small in the three dimensional domain [21, 23]. The RDME framework requires that the two reactant molecules for a bimolecular reaction must be in the same compartment in order to fire a reaction. Intuitively, we may realize that with more discrete compartments, it is less likely for the two molecules to encounter each other at the same compartment in a high dimensional domain. In order to model the reaction-diffusion system with RDME in the microscopic limit, Radek and Chapman [22] derived a formula of mesh-dependent reaction propensity correction for bimolecular reactions when the discretization size h is larger than a critical size h_{crit}. This reaction propensity correction formula fails when the discretization size h is smaller than this critical value. Recently, Isaacson [36] proposed a convergent RDME framework (cRDME). In the cRDME framework the diffusion is modeled exactly as in the RDME, while the bimolecular reaction occurs with a nonzero propensity, as long as the distance of the two reactant molecules is less than the reaction radius as defined in the Smoluchowski framework.

In conclusion, the discretization size for the RDME framework should be small enough to avoid discretization error. Yet when the mesh size is less than a critical value, the RDME may become inaccurate for the loss of bimolecular reactions in high dimensional domains. In this paper we will demonstrate that discretization size in space also has great influence on Hill function dynamics in reaction-diffusion systems. The switch-like Hill dynamics breaks even in a 1D domain when the discretization size is small.

Hill function

The Hill function [7], as well as the Michaelis-Menten function [6] are widely used in enzyme kinetics modeling. In molecular biology, enzymes catalyze biochemical substrates into products, while remaining unchanged. The enzyme kinetics reactions are usually formulated as

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P \tag{2}$$

Leonor Michaelis and Maud Leonora Menten proposed the "quasi-steady state" assumption and formulated the reaction rate equation for the enzyme kinetics, which is mostly referred to as the "Michaelis-Menten" equation. With the conservation law and the quasi-steady state assumption, the Michaelis-Menten equation is given as

$$\frac{d[P]}{dt} = V_{max}\frac{[S]}{K_M+[S]}, \tag{3}$$

with $V_{max} = k_2[E]_0$ being the maximum reaction rate and $K_m = \frac{k_{-1}+k_2}{k_1}$ being the Michaelis constant.

Sometimes one substrate molecule can have several enzyme binding sites and multiple bindings (cooperative binding) with enzymes are required to activate the substrate.

$$S + nE \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} SE_n \xrightarrow{k_2} nE + P \tag{4}$$

In real biological models, the binding of the n enzyme molecules to a substrate does not take place at once but in a succession of steps. Using the quasi-steady state assumption and conservation laws, the Hill function that formulates the reaction dynamics is given as

$$\frac{d[P]}{dt} = V_{max}\frac{[E]^n}{{K_m}^n+[E]^n}, \tag{5}$$

with V_{max} as the maximum reaction rate, K_m as the Michaelis constant, and n as the Hill coefficient. The Hill function is widely used to model "step-regulated" reaction as an activity switch.

Results

To simplify the analysis, a toy model of a reaction-diffusion system in one dimension is constructed. As demonstrated in Fig. 2, in the toy model an enzyme species E (typically a transcription factor) is constantly synthesized and degraded. The enzyme E further upregulates the DNA expression of a product P. The synthesis rate of P is formulated as a Hill function.

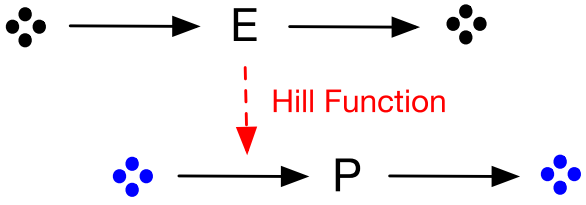


Fig. 2 A simple toy model of Hill function dynamics in 1D domain. Enzyme E is constantly synthesized and upregulates the synthesis of product P

Assume a spatial domain of size L is equally partitioned into K compartments with size $h = L/K$ for each. The list of reactions and reaction propensities in each compartment are given as

$$\begin{aligned}
\emptyset &\rightarrow E_i, \quad a_1 = k_s \cdot h;\\
E_i &\rightarrow \emptyset, \quad a_2 = k_d \cdot E_i;\\
\emptyset &\xrightarrow{E_i} P_i, \quad a_3 = k_{syn} \cdot h \frac{E_i^4}{(K_m \cdot h)^4 + E_i^4};\\
P_i &\rightarrow \emptyset, \quad a_4 = k_{deg} \cdot P_i;\\
E_i &\rightarrow E_{i\pm1}, \; a_5 = 2\frac{D_E}{h^2} E_i;\\
P_i &\rightarrow P_{i\pm1}, \; a_6 = 2\frac{D_P}{h^2} P_i;
\end{aligned} \tag{6}$$

The parameters k_s, k_d are the synthesis, degradation rates, respectively, for enzyme species E, and similarly k_{syn}, k_{deg} are those for product P. K_m is the Michaelis constant in the Hill function.

In the one-dimensional domain, the enzyme E is constantly synthesized and degraded. At the equilibrium state, the distribution of the total population of E is given by the Poisson distribution,

$$P_E(n) = \frac{\alpha^n}{n!} e^{-\alpha}, \tag{7}$$

where $\alpha = \frac{k_s}{k_d} L$ denotes the mean value of the total number of enzyme E molecules in the domain.

For an individual compartment (bin), consider the probability $P_E^{(i)}(n)$ that an individual bin i contains n molecules of enzyme E. At the equilibrium state, enzyme E is homogeneously distributed in the system. The probability that each molecule of E stays in a certain bin i is given by $p = 1/K$. The probability that, of all the E molecules in the domain, none is in bin i is approximated by

$$\begin{aligned}
P_E^{(i)}(0) &= P_E(0) + P_E(1)(1 - \tfrac{1}{K}) + P_E(2)(1 - \tfrac{1}{K})^2 \\
&\quad + \ldots + P_E(N)(1 - \tfrac{1}{K})^N + \ldots \\
&= \sum_{n=0}^{N} e^{-\alpha} \tfrac{\alpha^n}{n!} \left(1 - \tfrac{1}{K}\right)^n \\
&= e^{-\alpha/K}.
\end{aligned} \tag{8}$$

The other probability terms are not important in the analysis.

With the distribution of the enzyme molecular population, the mean reaction propensity for the synthesis of protein P in the i-th bin is

$$\langle a_{syn}^i \rangle = k_{syn} h \sum_{n=0}^{\infty} \frac{n^4}{(K_m \cdot h)^4 + n^4} P_E^{(i)}(n). \tag{9}$$

Notice that when $n = 0$, the Hill function is zero, and when the discrete bin size h is small, the Hill function approaches one quickly if $n \geq 1$. For example, when $K_m \cdot h \leq 0.5$ the Hill function $\frac{n^4}{(K_m \cdot h)^4 + n^4} \geq 0.94$ for $n \geq 1$. Therefore, upper and lower bounds for the product P synthesis propensity, when $k_m \cdot h \leq 0.5$, are

$$0.94 k_{syn} \cdot h \sum_{n=1}^{\infty} P_E^{(i)}(n) \leq \langle a^{syn} \rangle \leq k_{syn} \cdot h \sum_{n=1}^{\infty} P_E^{(i)}(n). \tag{10}$$

Hence, when the discretization size h is small enough, the propensity for the product P synthesis reaction can be approximated as

$$\begin{aligned}
\langle a_{syn}^{(i)} \rangle &\approx k_{syn} \cdot h \cdot \sum_{n=1}^{\infty} P_E^{(i)}(n) \\
&= k_{syn} \cdot h \cdot (1 - P_E^{(i)}(0)) \\
&= k_{syn} \cdot h \cdot (1 - e^{-\alpha/K}).
\end{aligned} \tag{11}$$

When the discretization size h is small and K is large, the mean reaction propensity can be further approximated as

$$\langle a_{syn}^{(i)} \rangle \approx k_{syn} \cdot h \cdot \alpha / K. \tag{12}$$

Notice that α/K is the mean population of enzyme E in the i-th bin. The Hill function of the product P synthesis is now reduced to a *linear function* of the enzyme E population in the i-th bin.

Furthermore, from (12) the mean population of product P in the bin i is

$$\langle P^{(i)} \rangle = \frac{k_{syn} \cdot h}{k_{deg}} \frac{\alpha}{K}, \tag{13}$$

and the total product P population in all K bins is

$$\begin{aligned}
\langle P \rangle &= \tfrac{k_{syn} \cdot L}{k_{deg}} \tfrac{k_s \cdot L}{k_d} \tfrac{1}{K} \\
&= \tfrac{k_{syn}}{k_{deg}} \cdot \alpha \cdot h.
\end{aligned} \tag{14}$$

Equation 14 shows that the total population of product P is a linear function of α, the mean population of E and $h = L/K$, the discretization size. With finer discretization, less product P is produced. Figure 3 shows the histograms and the mean values of the product P population with different discretization sizes. The histograms show that with finer discretization, the population histograms shift further to the left.

The log-log plot (Fig. 3, right) shows that when the discretization size is small enough, the total product P population is a linear function of discretization size. The slope of the log-log plot is about 1.0 at small discretization size h, regardless of K_m.

Moreover, simulation results show that when the mean enzyme E population is less than the constant K_m in the Hill function ($K_m > \alpha$), the population of product P increases slightly before the Hill function dynamics breaks at small discretization sizes. Note that the Hill function dynamics show a concave shape with respect to enzyme E population when the enzyme E population is smaller than the Michaelis constant K_m. Therefore, it is reasonable that the product P population in this reaction-diffusion model

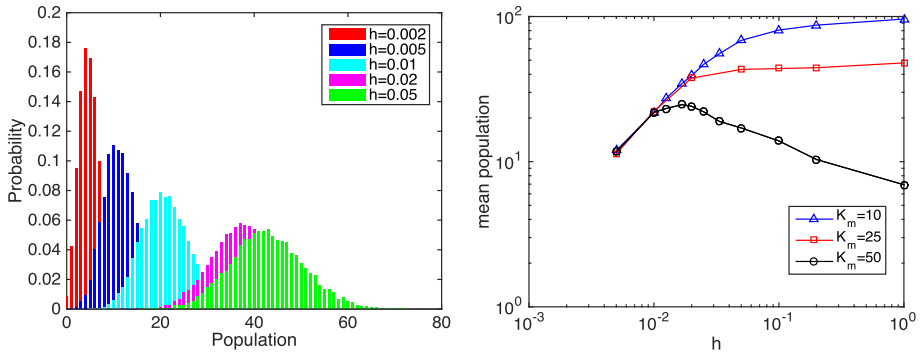


Fig. 3 The histogram (*left*) and mean (*right*) population of product P with different discretization. Parameters: $D_e = 1.0$, $k_s = 2.5$, $k_d = 0.1$, $k_{syn} = 5.0$, $k_{deg} = 0.05$, system size $L = 1.0$. For the histogram figure, $K_m = 25.0$. The log-log plot shows the mean total product population under different discretization and different parameter sets

increases slightly when the Michaelis constant K_m is larger than the mean enzyme E population α.

The numerical analysis above makes two approximations

$$\begin{cases} \frac{n^4}{(K_m \cdot h)^4 + n^4} \approx 1, \text{for } n \geq 1; \\ e^{-\alpha/K} \approx 1 - \alpha/K. \end{cases} \tag{15}$$

to get the linear relation. Assuming an error tolerance of 5%, the two approximations can be simplified to

$$\begin{cases} K_m \cdot h < 0.5, \\ \alpha/K < 1/3. \end{cases} \tag{16}$$

Hence, when the discretization bin number

$$K > \max\{2LK_m, \quad 3\alpha\}, \tag{17}$$

the Hill dynamics reduce to a linear function.

Equivalently, in order for the Hill function dynamics to work well, the discretization number K should be less than or equal to this threshold. However, the coarse discretization from a small K leads to spatial error. Two potential solutions to this discretization dilemma are proposed next.

Discussion

From the previous analysis, the Hill dynamics in RDME systems reduces to a linear function due to the lack of intermediate states — the discrete population in each individual bin yields an integer value (0 or 1) for the Hill function. Thus a natural solution to it is to generate intermediate states by a smoothing technique that averages the population over neighboring bins when calculating the reaction propensity.

To model a RDME system in high dimensions with fine discretization, previous studies [21] have suggested relaxing the same-compartment reaction assumption and allowing reactions within neighboring compartments. The next subsection shows that allowing reactions within neighboring compartments is equivalent to smoothing over neighboring compartments.

Smooth over neighboring bins

A natural technique that bridges the discrete and continuous models is to smooth the spatial population by taking the average of neighboring bins. Consider first smoothing the enzyme E population within the neighboring m bins (including the bin itself) when calculating the reaction propensity.

Following previous analysis, the reaction probability for the synthesis of product P in the i-th bin is

$$\begin{aligned} \langle \hat{a}_{syn}^{(i)} \rangle &= k_{syn} \cdot h \sum_{\hat{n}=0}^{\infty} \left(\frac{(n/m)^4}{(K_m \cdot h)^4 + (n/m)^4} P_E^{(i)}(n;m) \right) \\ &= k_{syn} \cdot h \sum_{n=0}^{\infty} \left(\frac{n^4}{(m \cdot K_m \cdot h)^4 + n^4} P_E^{(i)}(n;m) \right), \end{aligned} \tag{18}$$

where $P_E^{(i)}(n;m)$ denotes the probability that the m neighboring bins of the i-th bin have a total enzyme E population of n. The interpretation of this equation is that the synthesis reaction in the i-th bin is interacting with the m neighboring bins and the propensity is calculated based on the total enzyme E population of all the neighboring bins. By probability theory,

$$P_E^{(i)}(0;m) = e^{-\alpha m/K}. \tag{19}$$

As before, only the term $P_E^{(i)}(0;m)$ is important.

In Eq. (18), for any fixed integer $m \geq 0$, there exists an $h \geq 0$, such that $m \cdot K_m \cdot h < 0.5$ and the Hill function is still approximately one. With such a discretization size h, the product P synthesis propensity can be approximated as

$$\begin{aligned} \langle \hat{a}_{syn}^{(i)} \rangle &\approx k_{syn} \cdot h \sum_{n=1}^{\infty} P_E^{(i)}(n;m), \\ &= k_{syn} \cdot h(1 - P_E^{(i)}(0;m)), \\ &= k_{syn} \cdot h(1 - e^{-\alpha m/K}) \\ &\approx k_{syn} \cdot h \cdot \alpha \cdot m/K. \end{aligned} \tag{20}$$

Again, with a fixed smoothing bin number m, the synthesis reaction propensity becomes linear in the mean enzyme E population $\alpha m/K$ of the m bins, and the mean population of product P in the system is

$$\langle P \rangle = \frac{k_{syn} \cdot L}{k_{deg}} \frac{k_s \cdot L}{k_d} \frac{m}{K}, \tag{21}$$

which is linear in m/K and the mean total enzyme E population α. The linear function can be achieved with an h such that

$$\begin{cases} m \cdot K_m \cdot h < 0.5, \\ m \cdot \alpha / K < 0.33. \end{cases} \tag{22}$$

Figure 4 plots the mean population of product P in the toy model with the smoothing technique and $m = 5$. Numerical results show that smoothing over a fixed number m of compartments gives a good solution for a certain range of discretization sizes. However, there always exists a small enough critical discretization size h_{crit} such that the Hill function dynamics reduce to a linear function when the discretization size is smaller than this h_{crit}. Moreover, fixed length smoothing, in the scenarios where the Michaelis constant K_m is larger than the mean enzyme E population α, gives a result closer to that of the deterministic simulation when the discretization sizes are not too small.

Convergent hill function dynamics in reaction-diffusion systems

The previous subsection demonstrates that a sufficiently small discretization size h will still break the Hill dynamics even with the strategy of smoothing over a fixed number of bins, thus the number of bins needs to vary with the discretization size.

Inspired by the convergent-RDME framework [36], a remedy for the failure of Hill function dynamics in reaction-diffusion systems is to smooth the population over bins within a certain distance.

From the analysis, a small smoothing length would cause the failure of the Hill function dynamics and a large smoothing length would degrade the spatial accuracy of the model. Based on the criteria of failure for the Hill function dynamics with fixed m, Eq. (22), we can choose the smallest m that would not result in failure for the Hill function dynamics, i.e., m such that neither of the two assumptions in the previous analysis are valid. This choice is

$$m = \lceil \max \left\{ \frac{0.5}{K_m \cdot h}, \frac{0.33 \cdot L}{\alpha \cdot h} \right\} \rceil. \tag{23}$$

Following the terminology in the convergent-RDME framework [36], the "reaction radius ρ" of the Hill function dynamics is defined as $\rho = m \cdot h$, where m is given in (23).

Figure 5 shows numerical results for the toy model in the reaction-diffusion system with different discretization sizes and with the convergent smoothing technique (m and h related by (23)). It is clear that the convergent smoothing technique gives very good simulation results for all h values.

Applying the fixed length smoothing technique to the DivL-CckA Hill function model in the *Caulobacter crescentus* cell cycle results in a sharp CtrAp population change during swarmer-to-stalked transition. Figure 6 shows the CtrAp trajectories from the deterministic model and stochastic model simulation results. The fixed length smoothing technique yields more CtrAp in the swarmer stage and less CtrAp in the stalked stage, which yields a sharp CtrAp population change during the swarmer-to-stalked transition as expected.

Conclusions

Motivated by the misbehavior of DivL-CckA dynamics in the stochastic simulation of the *Caulobacter crescentus* cell cycle model, a study of the Hill function dynamics in reaction-diffusion systems reveals that when the discretization size is small enough, the switch-like behavior

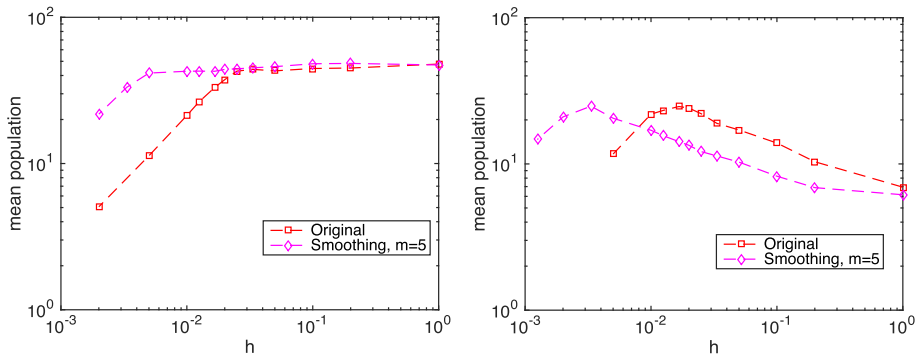


Fig. 4 The total population of product P with different discretization. Parameters: system size $L = 1.0$, $D_e = 1.0$, $k_s = 2.5$, $k_d = 0.1$, $k_{syn} = 5.0$, $k_{deg} = 0.05$. For the left figure $K_m = 25.0$, while for the right figure $K_m = 50.0$

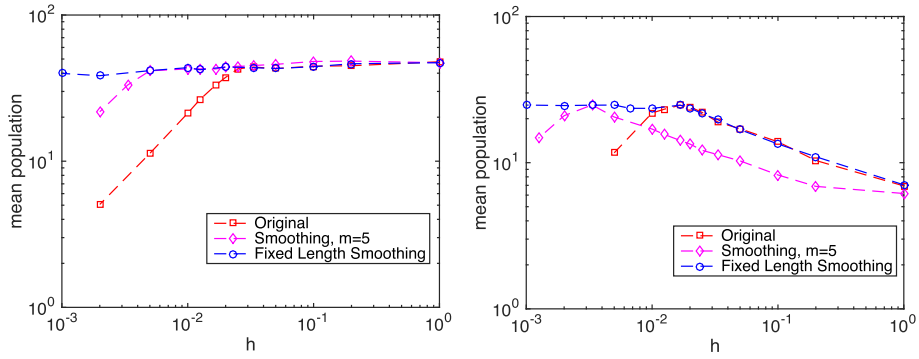


Fig. 5 The total population of product P with different discretization. Parameters: system size $L = 1.0$, $D_e = 1.0$, $k_s = 2.5$, $k_d = 0.1$, $k_{syn} = 5.0$, $k_{deg} = 0.05$. For the left figure $K_m = 25.0$, while for the right figure $K_m = 50.0$

of Hill function dynamics reduces to a linear function of input signal and discretization size. A proposed fixed length smoothing method, which allows chemical reactions to occur with reactant molecules within a distance of fixed length, the "reaction radius"of the Hill function dynamics, seems to give a very good remedy to this problem.

It is known that in high dimensions bimolecular reactions are lost with the RDME in the microscopic limit [21]. This work shows that one-dimensional Hill function dynamics in a RDME framework gives a similar challenge when the discretization size is small enough. The conjecture is that the problem lies in the RDME requirement that reactions only fire with the reactant molecules in the same discrete compartment.

Furthermore, this defect in RDME at the microscopic limit is believed to be a common scenario for all highly nonlinear reaction dynamics. Theoretical biologists have developed many highly nonlinear reaction dynamics that need special attention when converted to stochastic models.

Here we will extend our analysis and discuss a general situation in stochastic simulation of reaction diffusion systems. Suppose that we have a species X, whose population is represented by state variable x, and there is a particular reaction R:

$$\emptyset \xrightarrow{X} P, \tag{24}$$

in which X serves as an enzyme to produce P and the propensity function is represented by $f(x)$. For each X molecular, it can diffuse in a 1D domain with a small length L and with a diffusion coefficient D. Suppose the 1D domain is partitioned into K bins, thus the discretization size is $h = \frac{L}{K}$. The system can then be represented as a chain reaction

$$X_1 \underset{d}{\overset{d}{\rightleftharpoons}} X_2 \underset{d}{\overset{d}{\rightleftharpoons}} \cdots \underset{d}{\overset{d}{\rightleftharpoons}} X_K, \tag{25}$$

where $d = \frac{D}{h^2}$ is the jump rate corresponding to diffusion. The concerned reaction R could fire in any of the bins with propensity $f(x_i)$. Assume that L is small enough such that $\frac{D}{L^2}$ is very large and $d \gg \sum_{i=1}^{K} f(x_i)$ regardless of K. In that

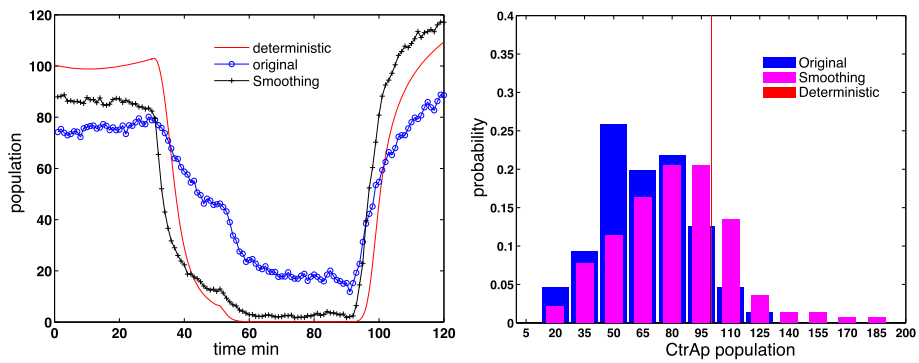


Fig. 6 The Comparison of CtrAp of deterministic model and the stochastic simulation results. *Left*: CtrAp population oscillation trajectory during *Caulobacter crescentus* cell cycle. *Right*: The histogram of CtrAp population in the swarmer cells ($t = 30min$). For model parameters, please refer to [27]

case, the chain reaction system (25) can be considered as a virtual fast system and the slow scale SSA [37] can be applied here. As a result, if the total population of X is n, in each bin, the mean value of x_i is given by

$$\langle x_i \rangle = \frac{n}{K}. \tag{26}$$

Then based on the theory of slow scale SSA, the propensity of the corresponding synthesis reaction (24) should be

$$\langle f(x_i) \rangle = \sum_{j=0}^{\infty} f(j) P(x_i = j), \tag{27}$$

where $P(\cdot)$ is the probability under the distribution when the virtual fast system (25) converges to stochastic partial equilibrium [37].

However, the propensity function converted directly from the deterministic model has a different form as $f(\langle x_i \rangle)$. Note that for a nonlinear function, such as the Hill function or the Michaelis Menten function,

$$\langle f(x_i) \rangle \neq f(\langle x_i \rangle). \tag{28}$$

(28) highlights the mismatch between the RDME framework and the deterministic model.

Hybrid method

In order to have a stochastic model that is consistent with its deterministic counterpart, the propensity function should take the form $f(\langle x_i \rangle)$. This motivates us to adopt the hybrid ODE/SSA method [38] and apply it to the reaction diffusion systems. This hybrid method was a simple idea. It was originally presented by Haseltine and Rawlings [38] and our implementation has some modification to make it fit better with the root finding function used in LSODAR [39]. Consider a system of N species (denoted by $\{S_1, \ldots, S_N\}$) and M reactions (denoted by $\{R_1, \ldots, R_M\}$). For each reaction R_j, there is a propensity function $a_j(x)$ and a state-change vector ν_j. We partition these M reactions into two subsets. The subset S_{slow} contains slow reactions, with index 1 to M_S, and is simulated by the SSA. The subset S_{fast} contains fast reactions, with index M_S+1 to M, and is formulated and solved by ODEs. The simulation of these two subsets is then combined as described below. Let τ be the jump interval of the next slow (stochastic) reaction, and μ be its reaction index. Set $t = 0$. The hybrid method simulate the system as follows:

1) Two uniform random numbers, r_1 and r_2 in $U(0,1)$, are generated.
2) Solve the ODE system for S_{fast} and find the root τ for the integral equation:
$$\int_t^{t+\tau} a_{tot}(\mathbf{x}, s) ds + \log(r_1) = 0, \tag{29}$$
where $a_{tot}(\mathbf{x}, t)$ is the sum of propensities of all reactions in S_{slow}. Because $\mathbf{x}$ varies with t in the ODE system, $a_{tot}(\mathbf{x}, t)$ is a function of t as well.
3) μ is selected as the smallest integer satisfying
$$\sum_{i=1}^{\mu} a_i(\mathbf{x}, t) > r_2 a_{tot}(\mathbf{x}, t). \tag{30}$$
4) Update $\mathbf{x} \leftarrow \mathbf{x} + \nu_\mu$.
5) Return to step 1) if stopping condition is not reached.

Note that our implementation is different from Haseltine and Rawling's original method in step 2. Suppose that the ODE system is given by

$$\mathbf{x}' = f(\mathbf{x}). \tag{31}$$

We add an integration variable z and the following equation to the ODE system.

$$z' = a_{tot}(\mathbf{x}), \quad z(t) = \log(r_1), \tag{32}$$

where we note that $\log(r_1)$ is negative and a_{tot} is always nonnegative. In the hybrid simulation, for each step we start from the current time t and numerically [39]

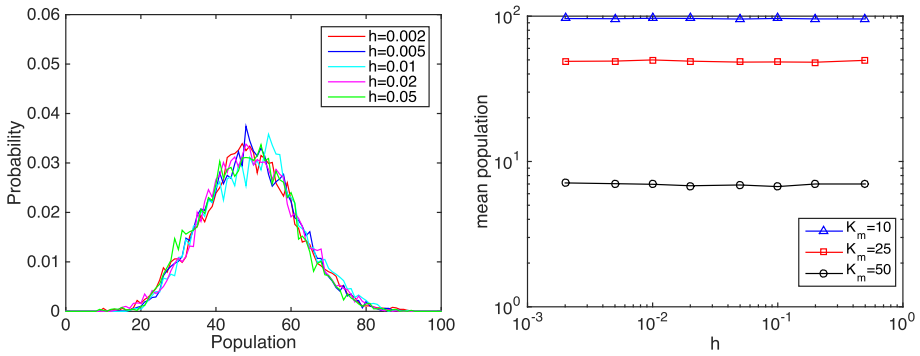


Fig. 7 The histogram (*left*) and mean (*right*) population of product P with different discretization, simulated by the hybrid method. Parameters: $D_e = 1.0$, $k_s = 2.5$, $k_d = 0.1$, $k_{syn} = 5.0$, $k_{deg} = 0.05$, system size $L = 1.0$. For the histogram figure, $K_m = 25.0$. The log-log plot shows the mean total product population under different discretization and different parameter sets

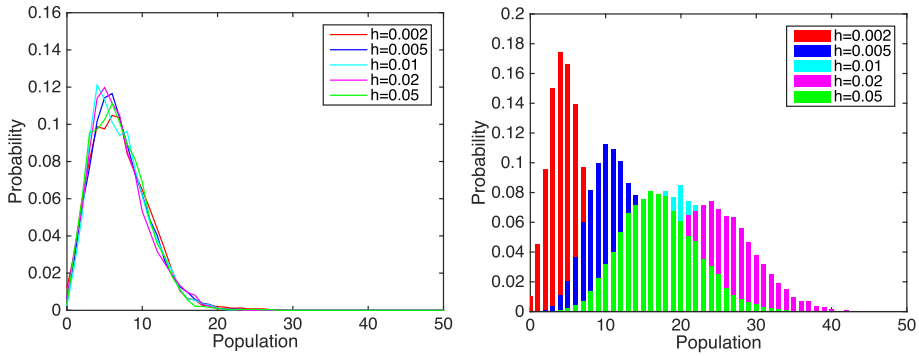


Fig. 8 The distribution of product P with different discretization sizes, simulated by the hybrid method (*left*) and SSA (*right*). Parameters: $K_m = 50.0$, the rest remains the same

integrate the original ODEs (31) and the extra integral Eq. (32). The integration stops when $z(t + \tau) = 0$. As a result, τ is the solution to (29). This procedure can be numerically simulated using standard ODE solvers combined with root-finding functions, such as the LSODAR [39]. Note that since z is an integration variable, one may choose to omit it from the error control mechanism [40]. Adding this extra variable will not greatly affect the efficiency.

We applied the hybrid method to the toy model (6). In our simulation, all diffusion events are partitioned into fast systems and solved by the ODE solver LSODAR, while chemical reactions are simulated by SSA under the hybrid framework described above. We test cases when $K_m = 10, 25, 50$ and Figs. 7 and 8 show the corresponding numerical results. It is obvious that in all three cases, the mean population remains horizontal even when the bin size decreased to the magnitude of 10^{-3}. In Fig. 8, the mean molecule of product P centers around seven under different discretization sizes, while results from SSA shift to the left as discretization size decreases.

Numerical results certainly suggest that the hybrid method has great potential in stochastic simulation of RD systems. We would like to note that great details still need to be studied, but that is not the focus for this paper.

Abbreviations
CME: Chemical master equation; ODE: Ordinary differential equation; PDE: Partial differential equation; RDME: Reaction diffusion master equation; SSA: Stochastic simulation algorithm

Funding
This work was partially supported by the National Science Foundation award DMS-1225160, CCF-0953590, CCF-1526666, and MCB-1613741. In particular, the publication of this paper is directly funded from CCF-1526666, and MCB-1613741.

Availability of data and materials
Not applicable.

Authors' contributions
FL initially realized the simulation error described in this paper. FL, MC and YC then designed the toy model and analyzed the numerical error caused by Hill function simulation. Later SW joined to help the implementation of hybrid simulation. FL and MC together drafted the manuscript, and YC gave critical revisions on the writing. All authors have read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Consent for publication
All authors consent to publish this work through BMC'17.

Ethics approval and consent to participate
Not applicable.

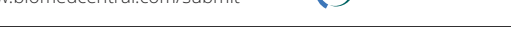
About this supplement
This article has been published as part of *BMC Systems Biology* Volume 11 Supplement 3, 2017: Selected original research articles from the Third International Workshop on Computational Network Biology: Modeling, Analysis, and Control (CNB-MAC 2016): systems biology. The full contents of the supplement are available online at http://bmcsystbiol.biomedcentral.com/articles/supplements/volume-11-supplement-3.

Published: 14 March 2017

References
1. McAdams H, Arkin A. Stochastic mechanisms in gene expression. Proc Natl Acad Sci. 1997;94(3):814–9. http://www.pnas.org/content/94/3/814.
2. Fedoroff N, Fontana W. Small numbers of big molecules. Science. 2002;297(5584):1129–31. http://www.sciencemag.org/content/297/5584/1129.
3. Samoilov M, Plyasunov S, Arkin AP. Stochastic amplification and signaling in enzymatic futile cycles through noise-induced bistability with oscillations. Proc Natl Acad Sci U S A. 2005;102(7):2310–5. http://www.pnas.org/content/102/7/2310.
4. Gillespie DT. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. J Comput Phys. 1976;22(4):403–34.
5. Gillespie DT. Exact stochastic simulation of coupled chemical reactions. J Phys Chem. 1977;81(25):2340–61.
6. Michaelis L, Menten ML. Die Kinetik der Invertinwirkung. Biochem Z. 1913;49:333–69.
7. Hill AV. The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves. J Physiol. 1910;40(Suppl):iv–vii.
8. Hattne J, Fange D, Elf J. Stochastic reaction-diffusion simulation with MesoRD. Bioinformatics. 2005;21(12):2923–4.
9. Andrews SS, Bray D. Stochastic simulation of chemical reactions with spatial resolution and single molecule detail. Phys Biol. 2004;1:137–51.
10. van Zon JS, ten Wolde PR. Green's-function reaction dynamics: A particle-based approach for simulating biochemical networks in time and space. J Chem Phys. 2005;123(23):234910. http://scitation.aip.org/content/aip/journal/jcp/123/23/10.1063/1.2137716.

11. Novère NL, Shimizu TS. StochSim: modelling of stochastic biomolecular processes. Bioinformatics. 2001;17:575–6.
12. von Smoluchowski M. Zur kinetischen Theorie der Brownschen Molekularbewegung und der Suspensionen. Annalen der Physik. 1906;326(14):756–80. http://dx.doi.org/10.1002/andp.19063261405.
13. Gardiner CW, McNeil KJ, Walls DF, Matheson IS. Correlations in stochastic theories of chemical reactions. J Stat Phys. 1976;14:307–31.
14. Nicolis G, Prigogine I. Self-organization in nonequilibrium systems : from dissipative structures to order through fluctuations. New York: A Wiley-Interscience Publication; 1977. http://opac.inria.fr/record=b1078628.
15. doi M. Stochastic theory of diffusion-controlled reaction. J Phys A Math General. 1976;9(9):1479. http://stacks.iop.org/0305-4470/9/i=9/a=009.
16. Keizer J. Nonequilibrium statistical thermodynamics and the effect of diffusion on chemical reaction rates. J Phys Chem. 1982;86(26):5052–67. http://dx.doi.org/10.1021/j100223a004.
17. Fange D, Berg OG, Sjöberg P, Elf J. Stochastic reaction-diffusion kinetics in the microscopic limit. Proc Natl Acad Sci. 2010;107(46):19820–5. http://www.pnas.org/content/107/46/19820.
18. McQuarrie DA. Stochastic approach to chemical kinetics. J Appl Probab. 1967;4(3):413–78. http://www.jstor.org/stable/3212214.
19. Gillespie DT. A rigorous derivation of the chemical master equation. Physica A: Stat Mech Appl. 1992;188(1–3):404–25.
20. Baras F, Mansour MM. Reaction-diffusion master equation: a comparison with microscopic simulations. Phys Rev E. 1996;54:6139–48. http://link.aps.org/doi/10.1103/PhysRevE.54.6139.
21. Isaacson SA. The reaction-diffusion master equation as an asymptotic approximation of diffusion to a small target. SIAM J Appl Math. 2009;70(1):77–111.
22. Erban R, Chapman SJ. Stochastic modelling of reaction-diffusion processes: algorithms for bimolecular reactions. Phys Biol. 2009;6(4):046001. http://stacks.iop.org/1478-3975/6/i=4/a=046001.
23. Hellander S, Hellander A, Petzold L. Reaction-diffusion master equation in the microscopic limit. Phys Rev E. 2012;85:042901. http://link.aps.org/doi/10.1103/PhysRevE.85.042901.
24. Li S, Brazhnik P, Sobral B, Tyson JJ. A Quantitative Study of the Division Cycle of *Caulobacter crescentus*, Stalked Cells. PLoS Comput Biol. 2008;4(1):e9. http://dx.plos.org/10.1371%2Fjournal.pcbi.0040009.
25. Li S, Brazhnik P, Sobral B, Tyson JJ. Temporal controls of the asymmetric cell division cycle in *Caulobacter crescentus*. PLoS Comput Biol. 2009;5(8):000463. http://dx.doi.org/10.1371%2Fjournal.pcbi.1000463.
26. Subramanian K, Paul MR, Tyson JJ. Potential role of a bistable histidine kinase switch in the asymmetric division cycle of *Caulobacter crescentus*. PLoS Comput Biol. 2013;9(9):003221. http://dx.doi.org/10.1371%2Fjournal.pcbi.1003221.
27. Subramanian K, Paul MR, Tyson JJ. Dynamical localization of DivL and PleC in the asymmetric Division cycle of *Caulobacter crescentus*: a theoretical investigation of alternative models. PLoS Comput Biol. 2015;11(7):004348.
28. Collier J, Murray SR, Shapiro L. DnaA couples DNA replication and the expression of two cell cycle master regulators. EMBO J. 2006;25(2):346–56.
29. Collier J, Shapiro L. Spatial complexity and control of a bacterial cell cycle. Curr Opin Biotechnol. 2007;18(4):333–40. http://www.sciencedirect.com/science/article/pii/S0958166907000894.
30. Holtzendorff J, Hung D, Brende P, Reisenauer A, Viollier PH, McAdams HH, et al. Oscillating global regulators control the genetic circuit driving a bacterial cell cycle. Science. 2004;304(5673):983–7. http://www.sciencemag.org/content/304/5673/983.
31. Quon KC, Yang B, Domian IJ, Shapiro L, Marczynski GT. Negative control of bacterial DNA replication by a cell cycle regulatory protein that binds at the chromosome origin. Proc Natl Acad Sci. 1998;95(1):120–5. http://www.pnas.org/content/95/1/120.
32. McGrath PT, Iniesta AA, Ryan KR, Shapiro L, McAdams HH. A dynamically localized protease complex and a polar specificity factor control a cell cycle master regulator. Cell. 2006;124(3):535–47. http://www.sciencedirect.com/science/article/pii/S0092867406000663.
33. Jenal U, Fuchs T. An essential protease involved in bacterial cell-cycle control. EMBO J. 1998;17(19):5658–69.
34. Iniesta AA, McGrath PT, Reisenauer A, McAdams HH, Shapiro L. A phospho-signaling pathway controls the localization and activity of a protease complex critical for bacterial cell cycle progression. Proc Natl Acad Sci. 2006;103(29):10935–40. http://www.pnas.org/content/103/29/10935.
35. Tsokos C, Perchuk B, Laub M. A dynamic complex of signaling proteins uses polar localization to regulate cell-fate asymmetry in *Caulobacter crescentus*. Developmental Cell. 2011;20(3):329–41.
36. Isaacson SA. A convergent reaction-diffusion master equation. J Chem Phys. 2013;139(5):054101. http://scitation.aip.org/content/aip/journal/jcp/139/5/10.1063/1.4816377.
37. Cao Y, Gillespie DT, Petzold LR. The slow-scale stochastic simulation algorithm. J Chem Phys. 2005;122(1):014116.
38. Haseltine EL, Rawlings JB. Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics. J Chem Phys. 2002;117(15):6959–69.
39. Hindmarsh AC. ODEPACK, a systematized collection of ODE solvers. IMACS Trans Sci Comput Amsterdam. 1983;1:55–64.
40. Petzold LR. A Description of DASSL: a differential/algebraic system solver, proceeding of the 1st IMACS World Congress. Montreal. 1982;1:65-68.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit