RES MILITARIS
REVUE EUROPEENNE D ETUDES EUROPEAN JOURNAL OF MILITARY STUDIES
Social Science Journal

MULTISCALE STOCHASTIC REACTION-DIFFUSION ALGORITHMS COMBINING MARKOV CHAIN MODELS WITH STOCHASTIC PARTIAL DIFFERENTIAL EQUATIONS

Dr.DANDE SRINIVAS
Assistant Professor
Department of Mathematics
Sree Chaitanya College Of Engineering,Karimnagar

Abstract

We analyze two multiscale methods for reaction-diffusion process stochastic simulations. They can be used in systems that contain areas where the concentrations of molecules differ noticeably. Both approaches split an interest region into two subsets, one for stochastic partial differential equations (SPDEs) and the other for continuous-time Markov chain models. The first approach considers a pseudo compartment (also known as an overlap or handshaking zone) in the SPDE portion of the computational domain immediately next to the interface in order to associate Markov chain (compartment-based) models with reaction-diffusion SPDEs. There is no usage of an overlap zone in the second algorithm. Additional developments of both schemes are showcased, encompassing the scenario of an adaptively selected boundary separating distinct modeling methodologies.

Keywords: Gillespie algorithm, multiscale modeling, chemical reaction networks, Markov chain, stochastic reaction-diffusion systems, and stochastic partial differential equations

1 Introduction

Stochastic models of well-mixed chemical systems are traditionally formulated in terms of continuous time Markov chains, which can be simulated using the Gillespie stochastic simulation algorithm (SSA) [42] or its equivalent formulations [12, 41, 60]. These algorithms provide statistically exact sample paths of stochastic chemical models described by the corresponding chemical master equation (CME). However, they can be computationally expensive for larger chemical systems, because they explicitly simulate each occurrence of each chemical reaction. A number of approaches have been developed in the literature to decrease the computational intensity of SSAs. Taking into account separation of time scales, chemical reaction networks can be simplified by model reduction before they are simulated [51–54, 58]. The idea of model reduction can also be used to develop computational methods which efficiently estimate quantities of interest from stochastic simulations [10, 11, 13, 26]. Another approach is to describe the molecular populations in terms of their concentrations that change continuously (rather than treating them as discrete random variables). This can be achieved by the chemical Langevin equation, which is a stochastic differential equation (SDE) acting as a bridge between discrete SSAs and deterministic reaction rate equations [43, 61, 62]. Efficient algorithms which make use of the SDE approximations have been developed for the simulation of chemical systems especially when they include processes

occurring on different time scales [15, 44, 46, 71]. More recently, the SDE approximations have been extensively used to develop hybrid algorithms which use boths SSAs and SDEs for different components of the studied systems [3, 19, 38, 65]. The chemical Fokker-Planck equation corresponding to the chemical Langevin equation can also be used to efficiently estimate quantities of interest from stochastic models [14, 16, 27, 64]. In this paper, we consider spatially-distributed (reaction-diffusion) models which can be described in terms of the reaction-diffusion master equation (RDME) [25]. A spatial domain is discretized into compartments (which are assumed to be well-mixed) and diffusion is modelled as a jump process between neighbouring compartments [24, 48, 56]. In the literature, the RDME approach has been adapted to model and simulate spatially-distributed systems using uniform meshes (equivalently, subvolumes or compartments) [22, 63, 75], nonuniform meshes [8] or complex geometries [49]. The resulting compartment-based model can be simulated by the Gillespie SSA. Compartment-based reaction-diffusion approaches have been used to model several intracellular processes, including Min oscillations in E. coli [5, 31], ribosome biogenesis [20, 21], actin dynamics in filopodia [30, 80] and pattern formation in morphogen signaling pathways [55]. They have also been implemented in a number of software packages including MesoRD [47], URDME [23], STEPS [78], SmartCell [4], Lattice Microbes [68] and Smoldyn [69]. As in the case of the simulation of well-mixed systems, the Langevin approach provides an approximation of the compartment-based model which can reduce the computational intensity of simulations. Spatial Langevin approaches [9, 40, 50] and stochastic partial differential equations (SPDEs) [1, 2, 6, 18, 57] have been suggested to model stochastic reaction-diffusion systems. A hybrid method has also been introduced using the Langevin approximation for diffusion coupled with the compartment-based model for reactions [66]. In the thermodynamic limit (of large populations), compartment-based models lead to reaction-diffusion partial differential equations (PDEs) which are written in terms of spatio-temporal concentrations of chemical species. This property can be exploited to design multiscale (hybrid) algorithms which use the compartmentbased Markov chain model in a subset of the simulated system and apply reaction-diffusion PDEs in other parts [32, 45, 50, 76, 79]. Other hybrid methods have also been developed in the literature including methods which couple more detailed Brownian dynamics (molecular-based) approaches with the compartment-based method [17, 33, 34, 59] or with reaction-diffusion PDEs [7, 36, 73]. In this paper, we analyze two multiscale algorithms which couple compartment-based models with suitably discretized SPDEs. They can be used when a large number of molecules of some species are located in parts of the computational domain. In the region with a small number of molecules, we use a compartmentbased model written as a continuous-time Markov chain. In other regions, we use SPDEs derived from the Markov process. The goal of this multiscale methodology is to get an approximation of the spatio-temporal statistics which we would obtain by running the underlying Markov chain model in the entire computational domain. The paper is organized as follows. In Section 2, we present the derivation of the SPDE description from the compartment-based model. In Section 3, two multiscale schemes are derived. An illustrative example with a static boundary between the SPDE and Markov chain subdomains is studied in Section 4. The algorithm is extended to a time-dependent interface in Section 5. In Section 6, we discuss an example with multiple species.

2 From continuous-time Markov jump processes to stochastic partial differential equations

We consider a system of N chemically reacting species S1, S2, . . . , SN, which are diffusing (with diffusion constants Di , i = 1,2,...,N) in the bounded domain Ω ⊂ R 3 . We use a compartment-based stochastic reaction-diffusion model [25], i.e. we divide the domain Ω into K compartments Ck , k = 1,2,...,K, and model the diffusion as a jump process between neighbouring compartments. In order to simplify the analysis, we consider that Ω is an elongated pseudo-one-dimensional domain Ω = [0,Kh]×[0,hy]×[0,hz], where h,hy,hz > 0, as shown in Figure 1(a). Compartments are rectangular cuboids with the volume hhyhz where

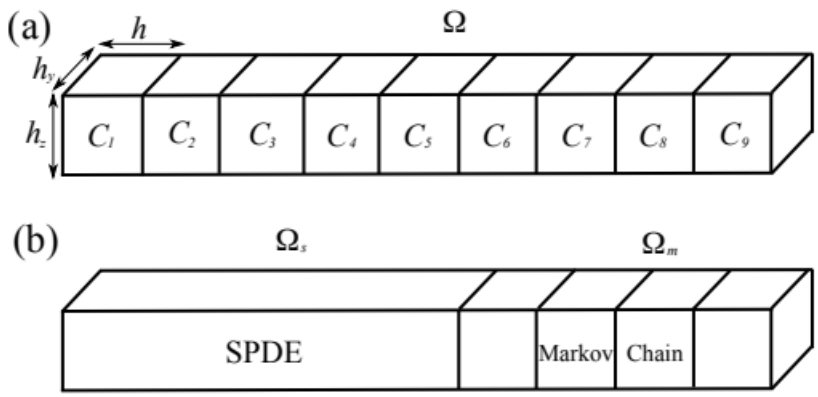

Fig. 1 (a) A schematic illustration of the elongated domain Ω for K = 9. (b) A schematic illustration of the multiscale setup.

Ck = [(k−1)h, kh]×[0,hy]×[0,hz] for k = 1,2,...,K. Let Z k i (t), i = 1,2...,N, k = 1,2,...,K, be the number of molecules of the i-th chemical species in the k-th compartment at time t. Then Z k (t) is an N-dimensional column vector with each component representing the number of molecules of the corresponding species in the k-th compartment at time t. We define

$$\mathbf{Z}(t) = \left(\mathbf{Z}^1(t)^{\mathrm{T}}, \mathbf{Z}^2(t)^{\mathrm{T}}, \ldots, \mathbf{Z}^K(t)^{\mathrm{T}}\right)^{\mathrm{T}},$$

which is a KN-dimensional column vector and T denotes the transpose of a vector. We assume that the chemical system is subject to M chemical reactions with ζ j , j = 1,2,...,M, being the corresponding Ndimensional stoichiometric vector. Let ζ k j , j = 1,2...,M, k = 1,2,...,K, be a KN-dimensional stoichiometric vector which gives a net molecule change during each occurrence of the j-th reaction in the k-th compartment. Let ν k −,i (resp. ν k +,i), i = 1,2,...,N, k = 1,2,...,K be a KN-dimensional stoichiometric vector which gives a net molecule change during diffusion of the i-th species from the k-th compartment to the (k −1)-th (resp. (k +1)-th) compartment. Let

$$\lambda_j^k : [0,\infty)^N \to [0,\infty), \qquad j = 1,2\ldots,M, \quad k = 1,2,\ldots,K,$$

be the propensity function of the j-th chemical reaction in the k-th compartment, i.e. λ k j (Z k (t))dt is the probability that the j-th reaction occurs in the k-th compartment during the time [t,t + dt) given that the current state at time t is Z k (t). We denote by R k j (t), j = 1,2...,M, k = 1,2,...,K, a random process which counts the number of times the j-th reaction occurs in the k-th compartment up to time t. Then

$$R_j^k(t) = Y_j^k\left(\int_0^t \lambda_j^k(\mathbf{Z}^k(s))\,\mathrm{d}s\right), \tag{2.1}$$

where Y k j are independent unit Poisson processes. We define R k −,i (t) (resp. R k +,i (t)), i = 1,2,...,N, k = 1,2,...,K, random processes counting the numbers of times that one molecule of the i-th species in the k-th compartment diffuses to the (k −1)-th compartment (resp. to the (k +1)-th compartment) up to time t. Then,

$$R_{\pm,i}^k(t) = Y_{\pm,i}^k\left(\int_0^t \frac{D_i}{h^2} Z_i^k(s)\,\mathrm{d}s\right), \tag{2.2}$$

where Y k ±,i are independent unit Poisson processes. The governing equation for the state vector Z(t) i

$$\mathbf{Z}(t) = \mathbf{Z}(0) + \sum_{k=1}^{K}\sum_{j=1}^{M} R_j^k(t)\,\boldsymbol{\zeta}_j^k + \sum_{k=2}^{K}\sum_{i=1}^{N} R_{-,i}^k(t)\,\boldsymbol{\nu}_{-,i}^k + \sum_{k=1}^{K-1}\sum_{i=1}^{N} R_{+,i}^k(t)\,\boldsymbol{\nu}_{+,i}^k. \tag{2.3}$$

When the propensities are large [62], the counting processes in Equations (2.1)–(2.2) can be approximated as

$$R_j^k(t) \approx \int_0^t \lambda_j^k(\mathbf{Z}^k(s))\,\mathrm{d}s + \int_0^t \sqrt{\lambda_j^k(\mathbf{Z}^k(s))}\,\mathrm{d}W_j^k(s),$$
$$R_{\pm,i}^k(t) \approx \int_0^t \frac{D_i}{h^2} Z_i^k(s)\,\mathrm{d}s + \int_0^t \sqrt{\frac{D_i}{h^2} Z_i^k(s)}\,\mathrm{d}W_{\pm,i}^k(s),$$

where Wk j and Wk ±,i are standard Brownian motions. Using ν k +,i = −ν k+1 −,i for k = 1,2,...,K −1 and changing the index (k+1) → k in the last term of Equation (2.3), the governing equation (2.3) can be approximated by the following SDE [43, 62]

$$\begin{aligned}\mathbf{Z}(t) = \mathbf{Z}(0) &+ \sum_{k=1}^{K}\sum_{j=1}^{M}\left(\int_0^t \lambda_j^k(\mathbf{Z}^k(s))\,\mathrm{d}s + \int_0^t \sqrt{\lambda_j^k(\mathbf{Z}^k(s))}\,\mathrm{d}W_j^k(s)\right)\boldsymbol{\zeta}_j^k \\ &+ \sum_{k=2}^{K}\sum_{i=1}^{N}\left\{\int_0^t \frac{D_i}{h^2}\left(Z_i^k(s) - Z_i^{k-1}(s)\right)\mathrm{d}s + \int_0^t \sqrt{\frac{D_i}{h^2} Z_i^k(s)}\,\mathrm{d}W_{-,i}^k(s) - \int_0^t \sqrt{\frac{D_i}{h^2} Z_i^{k-1}(s)}\,\mathrm{d}W_{+,i}^{k-1}(s)\right\}\boldsymbol{\nu}_{-,i}^k.\end{aligned} \tag{2.4}$$

Since Wk −,i and Wk−1 +,i terms always appear together in Equation (2.4), and since the sum of independent normal random variables is normally distributed, Equation (2.4) can be rewritten as

$$\begin{aligned}\mathbf{Z}(t) = \mathbf{Z}(0) &+ \sum_{k=1}^{K}\sum_{j=1}^{M}\left(\int_0^t \lambda_j^k(\mathbf{Z}^k(s))\,\mathrm{d}s + \int_0^t \sqrt{\lambda_j^k(\mathbf{Z}^k(s))}\,\mathrm{d}W_j^k(s)\right)\boldsymbol{\zeta}_j^k \\ &+ \sum_{k=2}^{K}\sum_{i=1}^{N}\left\{\int_0^t \frac{D_i}{h^2}\left(Z_i^k(s) - Z_i^{k-1}(s)\right)\mathrm{d}s + \int_0^t \sqrt{\frac{D_i}{h^2}\left(Z_i^k(s) + Z_i^{k-1}(s)\right)}\,\mathrm{d}W_i^{k-1}(s)\right\}\boldsymbol{\nu}_{-,i}^k,\end{aligned} \tag{2.5}$$

where Wk−1 i is a standard Brownian motion. Let Vh = hhyhz be the volume of each compartment, and define c(t) = Z(t)/Vh as a concentration vector for species at time t. Define

$$\widehat{\lambda}_j^{k,h}(\mathbf{c}^k(t)) = \frac{\lambda_j^k(\mathbf{Z}^k(t))}{V_h}, \qquad \text{where} \quad \mathbf{c}^k(t) = \frac{\mathbf{Z}^k(t)}{V_h}.$$

Dividing Equation (2.5) by V_h, we get

$$\begin{aligned}\mathbf{c}(t) = \mathbf{c}(0) &+ \sum_{k=1}^{K}\sum_{j=1}^{M}\left(\int_0^t \widehat{\lambda}_j^{k,h}(\mathbf{c}^k(s))\,\mathrm{d}s + \int_0^t \frac{1}{\sqrt{V_h}}\sqrt{\widehat{\lambda}_j^{k,h}(\mathbf{c}^k(s))}\,\mathrm{d}W_j^k(s)\right)\boldsymbol{\zeta}_j^k \\ &+ \sum_{k=2}^{K}\sum_{i=1}^{N}\left\{\int_0^t \frac{D_i}{h^2}\left(c_i^k(s) - c_i^{k-1}(s)\right)\mathrm{d}s + \int_0^t \frac{1}{\sqrt{V_h}}\sqrt{\frac{D_i}{h^2}\left(c_i^k(s) + c_i^{k-1}(s)\right)}\,\mathrm{d}W_i^{k-1}(s)\right\}\boldsymbol{\nu}_{-,i}^k,\end{aligned} \tag{2.6}$$

where c k i (t) = Z k i (t)/Vh. The second part of Equation (2.6) is consistent with the discretized Langevin scheme for a diffusion equation, as studied in [1]. We rewrite Equation (2.6) using the fact that reaction happens among species in the same compartment and that diffusion occurs between neighbouring compartments. Differentiating Equation (2.6), the concentration of the chemical species in the k-th compartment satisfy

$$\begin{aligned}\mathrm{d}\mathbf{c}^k(t) = &\sum_{j=1}^{M}\left(\widehat{\lambda}_j^{k,h}(\mathbf{c}^k(t))\,\mathrm{d}t + \frac{1}{\sqrt{V_h}}\sqrt{\widehat{\lambda}_j^{k,h}(\mathbf{c}^k(t))}\,\mathrm{d}W_j^k(t)\right)\boldsymbol{\zeta}_j \\ &+ \frac{\mathbf{D}}{h^2}\left\{\left(\mathbf{c}^{k+1}(t) - \mathbf{c}^k(t)\right)\chi_{\{k\neq K\}} - \left(\mathbf{c}^k(t) - \mathbf{c}^{k-1}(t)\right)\chi_{\{k\neq 1\}}\right\}\mathrm{d}t \\ &+ \frac{1}{\sqrt{V_h}}\frac{1}{h}\left\{\mathrm{d}\mathbf{W}^k(t)\sqrt{\mathbf{D}\left(\mathbf{c}^{k+1}(t) + \mathbf{c}^k(t)\right)}\chi_{\{k\neq K\}} - \mathrm{d}\mathbf{W}^{k-1}(t)\sqrt{\mathbf{D}\left(\mathbf{c}^k(t) + \mathbf{c}^{k-1}(t)\right)}\chi_{\{k\neq 1\}}\right\},\end{aligned} \tag{2.7}$$

where Wk (t) are N × N diagonal matrices with Wk i (t) on its diagonal for i = 1,2,...,N, k = 1,2,...,K − 1 and χ{·} is an indicator function. In Equation (2.7), ζ j is an N-dimensional stoichiometric vector of the j-th reaction for j = 1,2,...,M, and D is a N × N diagonal matrix which has diffusion constants of individual species on its diagonal, i.e.

$$\mathbf{D} = \begin{bmatrix} D_1 & 0 & \dots & 0 \\ 0 & D_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & D_N \end{bmatrix}.$$

We approximate white noise processes in Equation (2.7) using spatio-temporal white noise processes as

$$\frac{1}{\sqrt{h}}\frac{\mathrm{d}W_j^k(t)}{\mathrm{d}t} \approx \eta_j(x,t), \qquad \frac{1}{\sqrt{h}}\frac{\mathrm{d}\mathbf{W}^k(t)}{\mathrm{d}t} \approx \boldsymbol{\xi}(x,t),$$

where ηj(x,t), j = 1,2,...,M, are spatio-temporal white noise processes [77], i.e.

$$\int_{x'}^{x'+\Delta x}\int_{t'}^{t'+\Delta t}\eta_j(x,t)\,\mathrm{d}x\,\mathrm{d}t, \quad x' \in [0,Kh],\ t' \in [0,\infty),$$

is normally distributed with zero mean and variance ΔxΔt. Matrices ξ (x,t) are diagonal N×N matrices where diagonal entries are independent spatio-temporal white noise processes. Then Equation (2.7) is a solution of a discretized version of a SPDE in space which can be formally written in the following form

$$\frac{\partial \mathbf{c}(x,t)}{\partial t} = \sum_{j=1}^{M} \hat{\lambda}_j(\mathbf{c}(x,t),x)\,\boldsymbol{\zeta}_j + \mathbf{D}\frac{\partial^2 \mathbf{c}(x,t)}{\partial x^2} + \sum_{j=1}^{M} \sqrt{\frac{\hat{\lambda}_j(\mathbf{c}(x,t),x)}{h_y h_z}}\,\boldsymbol{\zeta}_j\,\eta_j(x,t) + \frac{\partial}{\partial x}\left[\boldsymbol{\xi}(x,t)\sqrt{\frac{2\mathbf{D}\,\mathbf{c}(x,t)}{h_y h_z}}\right], \tag{2.8}$$

where c(x,t) is a spatio-temporal concentration related to c k (t) by

$$\frac{1}{h}\int_{(k-1)h}^{\kappa n} \mathbf{c}(x',t)\,dx' \approx \mathbf{c}^k(t).$$

The reaction term $\hat{\lambda}_j : [0,\infty)^N \times [0,Kh] \to [0,\infty)$ in Equation (2.8) is related to $\hat{\lambda}_j^{k,h}$ by

$$\frac{1}{h}\int_{(k-1)h}^{kh} \hat{\lambda}_j(\mathbf{c}(x',t),x')\,dx' \approx \hat{\lambda}_j^{k,h}(\mathbf{c}^k(t)).$$

Note that Equations (2.6)-(2.7) are discretized versions of Equation (2.8), but the compartment-based model in (2.3) breaks down as h → 0 as discussed in Section 2.2 of [23]. The SPDE in Equation (2.8) is consistent to the ones in the previous work (Equation (1) in [57] and Equation (3.24) in [18]). For more details, see derivations of the SPDE for diffusion in Section 3.1 of [18] and the general version (Equations (8.2.54)- (8.2.56)) in Sections 8.1-8.2 of [39].

3 Multiscale algorithms combining compartment-based models with SPDEs

In this section, we present a multiscale approach which uses both SPDEs and Markov chain models. We develop two algorithms, denoted Scheme 1 and Scheme 2 in what follows, which are applied to illustrative examples in Sections 4, 5 and 6. Considering the same set up as in Section 2, we study a system of N chemically reacting species S1, S2, . . . , SN, which are diffusing (with diffusion constants Di , i = 1,2,...,N) in an elongated domain Ω = [0,L]×[0,hy]×[0,hz], where L = Kh, given in Figure 1. The domain Ω is divided into K compartments (rectangular cuboids) with Ck = [(k −1)h, kh]×[0,hy]×[0,hz] for k = 1,2,...,K. The main goal of this paper is to replace the Markov chain description in a part of the computational domain by the corresponding SPDEs. Let us consider that we use the SPDE in Equation (2.8) in the domain

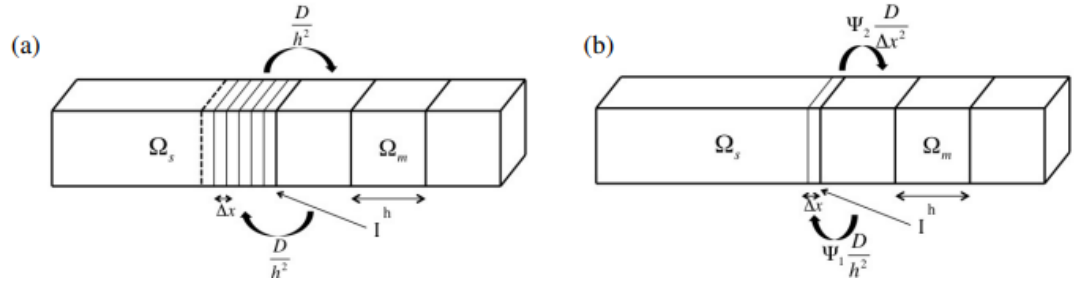

Fig. 2 Schematic diagrams of (a) Scheme 1 and (b) Scheme 2 describing molecular transfer between Ωs and Ωm. Note that the size of a virtual compartment in Ωs is h in panel (a).

Ωs = [0,I]×[0,hy]×[0,hz] where I = Ksh and K > Ks ∈ N; i.e. we consider that the first Ks compartments are described by a suitable discretization of the SPDE in Equation (2.8), see

Figure 1(b). We only use the Markov chain model for the remaining Km = K −Ks compartments, i.e. in subdomain Ωm = [I,L]×[0,hy]×[0,hz]. In this section, we develop an appropriate boundary condition on the interface I between Ωs and Ωm. In order to design the numerical scheme, the SPDE in Equation (2.8) needs to be appropriately discretized. We denote by Δx the mesh size used in the discretization of the SPDE. There are two important cases: (i) Δx > h and (ii) Δx ≤ h. In this section, we focus on case (ii), because we are interested in coupling the SPDE in Equation (2.8) with Markov chain models. The case (i) is important when one uses discretized SPDEs to design efficient multiscale schemes, but this introduces additional discretization errors. We will discuss case (i) in Section 7. In Ωs , each compartment of size h is discretized into α grid points (α ∈ N) with each grid size equal to Δx. In the remaining part of the computational domain Ωm, the compartment-based model is used. The state of the system of the multiscale model is described by vectors X k (t), k = 1,2,...,Ksα +Km. The vector X k (t) for k = 1,2,...,Ksα represents species 'numbers' in the mesh interval [(k −1)Δx, kΔx] in the SPDE region Ωs , i.e. it is related to spatio-temporal concentration c(x,t) used in the SPDE description by

$$\mathbf{X}^k(t) \approx h_y h_z \int_{(k-1)\Delta x}^{k\Delta x} \mathbf{c}(x,t)\,\mathrm{d}x.$$

The vector X Ksα+k (t) for k = 1,2,...,Km represents species numbers in CKs+k = [I + (k −1)h,I +kh] in the Markov chain region Ωm, i.e. it is related to the state vector Z k (t) used in the Markov chain description by X Ksα+k (t) = Z Ks+k (t). We consider two different schemes to describe transfer of molecules near the interface I coupling discretized SPDEs and the Markov chain model, as shown in Figure 2. Without loss of generality, both schemes are introduced for diffusion, because the description of reactions does not influence the transfer of molecules across the interface I. In Scheme 1, we assume that there is a virtual compartment, CKs = [I − h,I], in Ωs , where the molecules are partially treated using a compartment-based approach. Such overlap (handshaking) regions are common in many multiscale methodologies, including coupling molecular dynamics with Brownian dynamics simulations [28,29], Brownian dynamics with PDEs [36], or in atomistic to continuum coupling methods [67]. We define a state vector

$$\mathbf{X}(t) = \left(\mathbf{X}^1(t)^{\mathrm{T}}, \mathbf{X}^2(t)^{\mathrm{T}}, \ldots, \mathbf{X}^{K_s\alpha+K_m}(t)^{\mathrm{T}}\right)^{\mathrm{T}},$$

which is a (Ksα + Km)N-dimensional column vector. Scheme 1 is described using the following evolution equation for state vector X(t):

$$
\begin{aligned}
\mathbf{X}(t) = \mathbf{X}(0) + \sum_{k=2}^{K_s\alpha}\sum_{i=1}^{N}\Bigg\{ & \int_0^t \frac{D_i}{\Delta x^2}\left(X_i^k(s) - X_i^{k-1}(s)\right) \mathrm{d}s \\
& + \int_0^t \sqrt{\frac{D_i}{\Delta x^2}\left(X_i^k(s) + X_i^{k-1}(s)\right)\chi_{\{X_i^k(s)+X_i^{k-1}(s)\geq 0\}}}\, \mathrm{d}W_i^{k-1}(s)\Bigg\}\boldsymbol{\nu}_{-,i}^k \\
& + \sum_{k=K_s\alpha+2}^{K}\sum_{i=1}^{N} Y_{-,i}^k\left(\int_0^t \frac{D_i}{h^2} X_i^k(s)\,\mathrm{d}s\right)\boldsymbol{\nu}_{-,i}^k + \sum_{k=K_s\alpha+1}^{K-1}\sum_{i=1}^{N} Y_{+,i}^k\left(\int_0^t \frac{D_i}{h^2} X_i^k(s)\,\mathrm{d}s\right)\boldsymbol{\nu}_{+,i}^k \\
& + \underbrace{\sum_{\ell=1}^{\alpha}\sum_{i=1}^{N} Y_{-,i}^{K_s\alpha+1}\left(\int_0^t \frac{D_i}{h^2} X_i^{K_s\alpha+1}(s)\,\mathrm{d}s\right)\mathscr{I}_{-,i}^{\ell}(t)\,\boldsymbol{\eta}_{-,i}^{\ell}}_{\text{Markov chain} \to \text{SPDE}} \\
& + \underbrace{\sum_{\ell=1}^{\alpha}\sum_{i=1}^{N} Y_{+,i}^{K_s\alpha}\left(\int_0^t \frac{D_i}{h^2}\sum_{j=1}^{\alpha} X_i^{(K_s-1)\alpha+j}(s)\chi_{\left\{\sum_{j=1}^{\alpha} X_i^{(K_s-1)\alpha+j}(s)\geq 1\right\}}\,\mathrm{d}s\right)\mathscr{I}_{+,i}^{\ell}(t)\,\boldsymbol{\eta}_{+,i}^{\ell}}_{\text{SPDE} \to \text{Markov chain}},
\end{aligned}
\tag{3.1}
$$

where the first term on the right hand side represents diffusion in Ωs (compare with Equation (2.5) replacing h by Δx). Note that the indicator function χ {X k i (s)+X k−1 i (s)≥0} is used to make sure the term inside the square root not being negative. Here the symbols ν k ±,i describe (Ksα + Km)N-dimensional stoichiometric vectors. The second and third terms represent diffusion in the compartment-based region, Ωm, where Y k ±,i are independent unit Poisson processes (compare with Equation (2.2)). The last two terms represent transition of a molecule from Ωm to Ωs and from Ωs to Ωm, respectively. A molecule in Ωm in the boundary compartment, CKs+1, jumps to the SPDE domain with a rate Di/h 2 . A molecule which jumps is placed to one of the mesh points in the overlap compartment, CKs . To describe this process in Equation (3.1), we have defined indicator function

$$
\mathscr{I}_{\pm,i}^{\ell}(t) = \chi_{\{U_{\pm,i}(t)\in[I-h+(\ell-1)\Delta x, I-h+\ell\Delta x]\}}, \qquad \text{for} \quad \ell = 1,2,\ldots,\alpha,
$$

where U±,i(t) are independent uniform variables for each t and i. Stoichiometric vectors, η ` ±,i for ` = 1,2,...,α, i = 1,2,...,N, give changes due to the diffusion of the i-th species between the `-th SPDE discretization point in CKs and the compartment CKs+1 across the interface I. Transition of a molecule from Ωs to Ωm is described by the last term of Equation (3.1) using time-changed Poisson processes. A molecule, anywhere in the overlap compartment CKs , can be transferred with a rate Di/h 2 . The corresponding molecule is then randomly subtracted from one of α discretization grid points which are in CKs . Note that the molecular copy number, ∑ α j=1 X (Ks−1)α+j i (s), in the last term of Equation (3.1) can be non-integer value due to the non-integer concentration in CKs . To prevent X (Ks−1)α+j i (s) being negative due to the molecular transfer from Ωs to Ωm, another indicator function is used in the last term of Equation (3.1) to set the propensity as zero if the total molecular copy number in CKs is less than 1

Scheme 2 is described in terms of two unknown parameters, denoted Ψ1 and Ψ2, by the following evolution equation for the state vector X(t):

$$\mathbf{X}(t) = \mathbf{X}(0) + \sum_{k=2}^{K_s\alpha}\sum_{i=1}^{N}\left\{\int_0^t \frac{D_i}{\Delta x^2}\left(X_i^k(s) - X_i^{k-1}(s)\right)\mathrm{d}s + \int_0^t \sqrt{\frac{D_i}{\Delta x^2}\left(X_i^k(s) + X_i^{k-1}(s)\right)\chi_{\{X_i^k(s)+X_i^{k-1}(s)\geq 0\}}}\,\mathrm{d}W_i^{k-1}(s)\right\}\boldsymbol{\nu}_{-,i}^k + \sum_{k=K_s\alpha+2}^{K}\sum_{i=1}^{N} Y_{-,i}^k\left(\int_0^t \frac{D_i}{h^2}X_i^k(s)\,\mathrm{d}s\right)\boldsymbol{\nu}_{-,i}^k + \sum_{k=K_s\alpha+1}^{K-1}\sum_{i=1}^{N} Y_{+,i}^k\left(\int_0^t \frac{D_i}{h^2}X_i^k(s)\,\mathrm{d}s\right)\boldsymbol{\nu}_{+,i}^k \quad (3.2)$$
$$+ \underbrace{\sum_{i=1}^{N} Y_{-,i}^{K_s\alpha+1}\left(\int_0^t \Psi_1\frac{D_i}{h^2}X_i^{K_s\alpha+1}(s)\,\mathrm{d}s\right)\boldsymbol{\nu}_{-,i}^{K_s\alpha+1}}_{\text{Markov chain}\to\text{SPDE}} + \underbrace{\sum_{i=1}^{N} Y_{+,i}^{K_s\alpha}\left(\int_0^t \Psi_2\frac{D_i}{\Delta x^2}X_i^{K_s\alpha}(s)\chi_{\left\{X_i^{K_s\alpha}(s)\geq 1\right\}}\,\mathrm{d}s\right)\boldsymbol{\nu}_{+,i}^{K_s\alpha}}_{\text{SPDE}\to\text{Markov chain}}.$$

The first three terms in Equation (3.2) are identical to those in Equations (3.1). The fourth and fifth terms describe molecular transfer between the last grid point in Ωs and the boundary compartment CKs+1. A molecule in Ωm in the boundary compartment, CKs+1, jumps to the last grid point of the SPDE domain with rate Ψ1Di/h 2 , and the transfer rate in the opposite direction is Ψ2Di/Δx 2 . Note that X Ksα i (s) in the fifth term of Equation (3.2) can be non-integer value due to the non-integer concentration in Ωs . To prevent X Ksα i (s) being negative due to the molecular transfer from Ωs to Ωm, we use an indicator function to set the propensity as zero if the molecular copy number in the last grid point in Ωs is less than 1. To determine parameters Ψ1 and Ψ2 of Scheme 2, we use the discretization of the 1-dimensional partial differential equation for diffusion using a finite volume approximation [8]. It gives the jump coefficient of the i-th species from the j-th compartment to the neighbouring j 0 -th compartment as Di/(hj |aj −aj 0|), where hj is the length of the j-th compartment and aj and aj 0 are the centers of the j-th and j 0 -th compartments, respectively. Considering the size of the domain allowed for molecule transfer across the interface in Scheme 2, we set |aj −aj 0| = (Δx+h)/2. We take hj = Δx for the jump coefficient from Ωs to Ωm and hj = h for the jump coefficient from Ωm to Ωs . Then, we match the jump coefficients to the rate constants for jump across the interface given in Equation (3.2) to derive the following formula for the parameters of Scheme

$$\Psi_1 = \frac{2h}{\Delta x + h} \quad \text{and} \quad \Psi_2 = \frac{2\Delta x}{\Delta x + h}.$$

The multiscale algorithm for Scheme 1 for the case of diffusion only is given in Table 1. We denote a propensity of diffusion of the i-th species in the (Ks + k)-th compartment in Ωm to the left (resp. right) as a k −,i (t) = Di/h 2X Ksα+k i (t), for k = 1,2,...,Km, (resp. a k +,i (t) = Di/h 2X Ksα+k i (t), for k = 1,2,...,Km −1) for i = 1,2,...,N. This definition also includes the propensity of a diffusive jump of the i-th species from the Markov chain domain, given as a 1 −,i (t). We denote a propensity of diffusive jump of the i-th species from the SPDE domain b

$$a_{+,i}^0(t) = \frac{D_i}{h^2}\sum_{j=1}^{\alpha} X_i^{(K_s-1)\alpha+j}(t).$$

Then, we define total propensity in Ω

$$a_0 = \sum_{i=1}^{N}\sum_{k=1}^{K_m} a^k_{-,i} + \sum_{i=1}^{N}\sum_{k=0}^{K_m-1} a^k_{+,i}. \qquad (3.3)$$

Total propensity a0 is used in steps [A] and [B] in the pseudocode in Table 1 to select time when the next event occurs in Ωm. The pseudocode denotes the time of the next update in each subdomain as ts and tm, an

[A] Set $t = 0$ and $t_s = \Delta t$. Initialize species numbers, $\mathbf{X}(0)$, in Ω_m and Ω_s. Then, generate random numbers r_1 and r_2 uniformly distributed in $(0,1)$. Set τ so that $\tau = -a_0^{-1}\log(r_1)$, where a_0 is defined in Equation (3.3). Set the next time when the diffusion occurs in Ω_m as $t_m = \tau$.

[B] If $t_m \le t_s$,

- Set $t = t_m$.
- Use r_2 to determine which diffusive jump occurs. Each diffusive jump to the left (resp. to the right) has the probability $a^k_{-,i}/a_0$ (resp. $a^k_{+,i}/a_0$) to occur.
- If the selected diffusive jump only includes internal compartments in Ω_m, update species numbers in the corresponding compartments.
- If the diffusion occurs across the interface from Ω_m to Ω_s, update the species number in C_{K_s} by transferring one molecules from C_{K_s+1} to the corresponding grid point in Ω_s.
- If the diffusion occurs across the interface from Ω_s to Ω_m, update the species number in C_{K_s+1} by adding one molecule and subtracting one from the corresponding grid point in Ω_s.
- Generate random numbers r_1 and r_2 uniformly distributed in $(0,1)$. Set τ so that $\tau = -a_0^{-1}\log(r_1)$, where a_0 is defined in Equation (3.3). Set the next time when the diffusion occurs in Ω_m as $t_m = t + \tau$.

[C] If $t_s \le t_m$,

- Set $t = t_s$.
- Use Equation (3.4) to update the SPDE part of the system from t to $t + \Delta t$.
- Set the next time of the update of the SPDE part as $t_s = t + \Delta t$.

[D] Repeat steps **[B]**–**[C]** until the simulation ends.

Table 1 Pseudocode for the multiscale reaction-diffusion algorithm with Scheme 1 applied to simulation of diffusion.

the current time as t. In step [B], we update the compartment-based part of the system. In step [C], we update the SPDE part of the system by

$$\mathbf{X}(t+\Delta t) = \mathbf{X}(t) + \sum_{k=2}^{K_s\alpha}\sum_{i=1}^{N}\left\{\frac{D_i}{\Delta x^2}\left(X_i^k(t) - X_i^{k-1}(t)\right)\Delta t + \sqrt{\frac{D_i\Delta t}{\Delta x^2}\left(X_i^k(t) + X_i^{k-1}(t)\right)}\,\zeta_i^{k-1}\right\}\mathbf{v}^k_{-,i}, \qquad (3.4)$$

where ζ k−1 i are independent normally distributed random numbers with zero mean and unit variance

4 Application: static boundary

In this section, we apply the multiscale approach to examples in which we know a priori the position of the boundary I between Ωs and Ωm. Generalization to a more complicated case with a moving boundary is presented in Section 5. 4.1 A morphogen gradient model We consider a morphogen gradient model in Ω = [0,L]×[0,hy]×[0,hz]. It consists of one chemical species S, i.e. Z k (t) is a scalar describing the number of molecules of S in Ck . The

state of the Markov chain model is described by the K-dimensional column vector Z(t) = Z 1 (t),Z 2 (t),...,Z K(t) T . Morphogen is subject to diffusion which is described by Equation (2.2). There are also two reactions in the system. Morphogen, S, is produced in the first compartment with rate J, i.e. the propensity is λ 1 1 Z 1 = J. Morphogen degrades everywhere with rate δ, i.e. with propensity λ k 2 Z k = δZ k for k = 1,2,...,K. In all stochastic simulation

Notations	Description	Values
L	Length of the domain	$20\,\mu$m
D	Diffusion coefficient	$0.8\,\mu\text{m}^2\,\text{s}^{-1}$
δ	Degradation rate	$0.05\,\text{s}^{-1}$
J	Production rate	$25\,\mu\text{m}^{-2}\,\text{s}^{-1}$
Δx	Spatial discretization in Ω_s	$0.2\,\mu$m
h	Spatial discretization in Ω_m	$1\,\mu$m
h_y, h_z	Spatial discretization in y and z directions	$1\,\mu$m
Δt	Time discretization for SPDE	0.0005 s

Table 2 Parameter values in the morphogen gradient model studied in Section 4.1.

the morphogen gradient model, we assume that 500 morphogen molecules are initially uniformly distributed in the half of the domain, Ωs = [0,L/2]×[0,hy]×[0,hz]. The parameters are given in Table 2. Denoting c(x,t) the morphogen density at point x and time t, the deterministic model can be written as PDE

$$\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2} - \delta c, \qquad \text{with boundary conditions} \qquad -D\frac{\partial c}{\partial x}\bigg|_{x=0} = J, \quad D\frac{\partial c}{\partial x}\bigg|_{x=L} = 0,$$

where D is the diffusion constant of S. We apply the multiscale approach using both schemes developed in Section 3. Since the morphogen is produced at the left end, the morphogen has a decreasing gradient as it goes towards L. Therefore, we split the spatial domain in half, and set the left half as Ωs and the right half as Ωm, i.e. I = L/2. The (Ksα +Km)-dimensional state vector of the multiscale model is denoted X(t) = X 1 (t),X 2 (t),...,X Ksα+Km (t) T . Note that morphogens are produced only in the first discretization mesh point with size Δx in Ωs . In Figure 3, we simulate the morphogen gradient model using Scheme 1 of the multiscale algorithm. We calculate 104 realizations of the sample paths of the stochastic process, and present mean and standard deviations of the morphogen numbers in Ω at different times, t = 0,2,5,20 s. Morphogen numbers in α grid points of Ωs are summed so that they can be compared to the numbers in the underlying Markov chain model. We compare the results with mean and standard deviations of the morphogen numbers which we calculate analytically using matrix analysis for reaction-diffusion Markov chain models [37, 56]. In Figure 3, morphogen numbers in Ωs (resp. in Ωm) are expressed as green bars (resp. blue bars). Error bars represent one standard deviations from the mean number of morphogens in each compartment. Mean and standard deviations of the morphogen numbers from the analytic solution are drawn as a red line and blue dotted lines. The results using the multiscale algorithm match perfectly to the ones from the exact solution. In Figure 4, we present relative errors of the means and standard deviations of the number of molecules between the Markov chain model and multiscale model. The analytic solution is used for the statistics of the Markov chain model, and both schemes are applied numerically for the multiscale approaches. Errors are defined as

$$e_m(k) = \begin{cases} 1 - \dfrac{\mathrm{E}\left[\sum_{\ell=1}^{\alpha} X^{(k-1)\alpha+\ell}\right]}{\mathrm{E}[Z^k]}, \\ 1 - \dfrac{\mathrm{E}\left[X^{k+K(\alpha-1)/2}\right]}{\mathrm{E}[Z^k]}, \end{cases} \quad e_v(k) = \begin{cases} 1 - \dfrac{\sigma\left[\sum_{\ell=1}^{\alpha} X^{(k-1)\alpha+\ell}\right]}{\sigma[Z^k]}, & \text{for } k = 1, 2, \ldots, \dfrac{K}{2}, \\ 1 - \dfrac{\sigma\left[X^{k+K(\alpha-1)/2}\right]}{\sigma[Z^k]}, & \text{for } k = \dfrac{K}{2} + 1, \ldots, K, \end{cases} \tag{4.1}$$

where E[·] and σ[·] represent a mean and standard deviation. In Figure 4(a), red and green lines represent em(k) and ev(k) at time t = 50 s using Scheme 1, respectively, and blue and purple lines are for Scheme 2. We observe that the relative errors in Equation (4.1) are less than 4% in the entire simulation domain. In Figure 4(b), we compare the maximum absolute values of the relative errors defined in Equation (4.1) with α = 1, 5, 10, 25 and fixed compartment size h where α = h/Δx. In both schemes, the relative errors are in a range of less than 4% except for the case when α = 25 with Scheme 2. The relative errors in the mean and standard deviation become significantly larger when we apply the multiscale algorithm using Schem

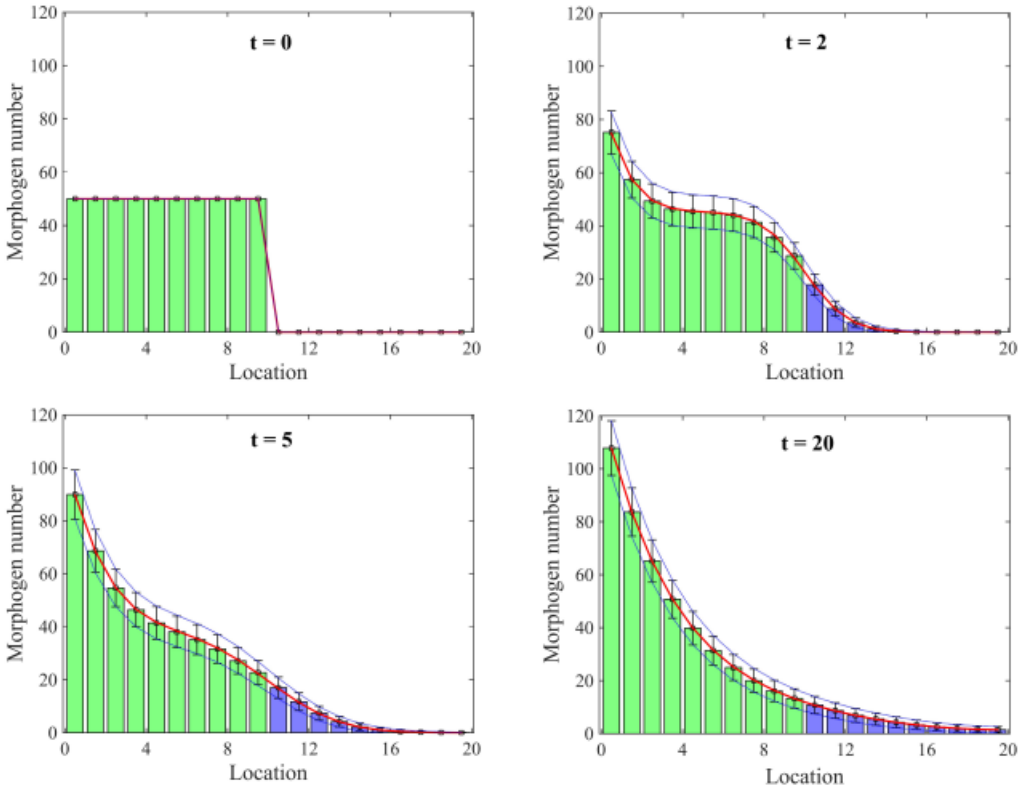

Fig. 3 Comparison between mean numbers of morphogens and their standard deviations from the mean using the analytic solution (red lines and blue dotted lines) and Scheme 1 (green bars and blue bars for the means in Ωs and Ωm, respectively, and error bars for the standard deviations).

2 with α = 25. In this case, the mean in CKs+1 gets larger than the mean in CKs which shows a bias in the method for larger values of α (the exact mean number of molecules decreases along the x-axis). We provide an explanation of this phenomenon in the next section.

4.2 A diffusion model with two compartments

In Section 4.1, we have observed that the error of Scheme 2 increases when we decrease the ratio of the numerical discretization in Ωs and the compartment size. When α = 25, the mean number of molecules of the morphogen does not have a decreasing gradient across the interface I in Scheme 2. Therefore, we set a two-compartment model with diffusion to see what causes this numerical error. The setting is similar to the one in Section 4.1, but we set J = δ = 0 so that there is no flux or degradation of the morphogen. Set L = 2h and I = h (= 1μm) so that Ωs = [0,h]×[0,hy]×[0,hz] and Ωm = [h,2h]×[0,hy]×[0,hz]. Then, each region consists of one compartment, Ks = Km = 1, and X(t) is an (α +1)-dimensional vector. In Figure 5(a), we present simulation results of the two-compartment model using Scheme 1 (red line) and Scheme 2 (green line) with α = 10, 20, 30, 40, 50 and compare them to the simulation result of the Markov chain model using the Gillespie SSA (purple line). The Markov chain model has α + 1 numerical grid points where the first α ones are with size Δx = h/α and the last one with size h. Diffusion of molecules is simulated by jumps from grid points to their nearest neighbours, i.e. the numerical meshes in the Markov chain model are coupled by diffusion in the same way as it is done in Scheme 2. Applying both multiscale algorithms and the Gillespie SSA, we compare the mean morphogen numbers in the second compartment computed from 100 realizations of simulation. Using 50 molecules in total, the exact value of the mean

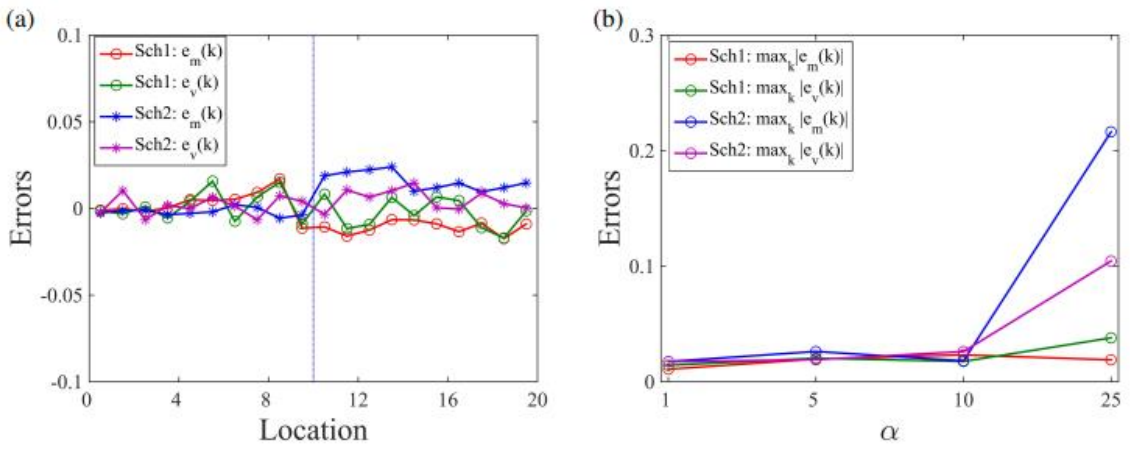

Fig. 4 (a) Errors em(k) and ev(k) given by Equation (4.1) are computed at time 50 s. (b) The maximum absolute values of the errors em(k) and ev(k) given by Equation (4.1) are computed at time 50 s with a static boundary and different values of α. The maximum value of the errors is taken over all region, Ω. Red and green lines are relative errors of the means and standard deviations between the analytic solution of the Markov chain model and Scheme 1. Blue and purple lines are relative errors between the Markov chain model and Scheme 2.

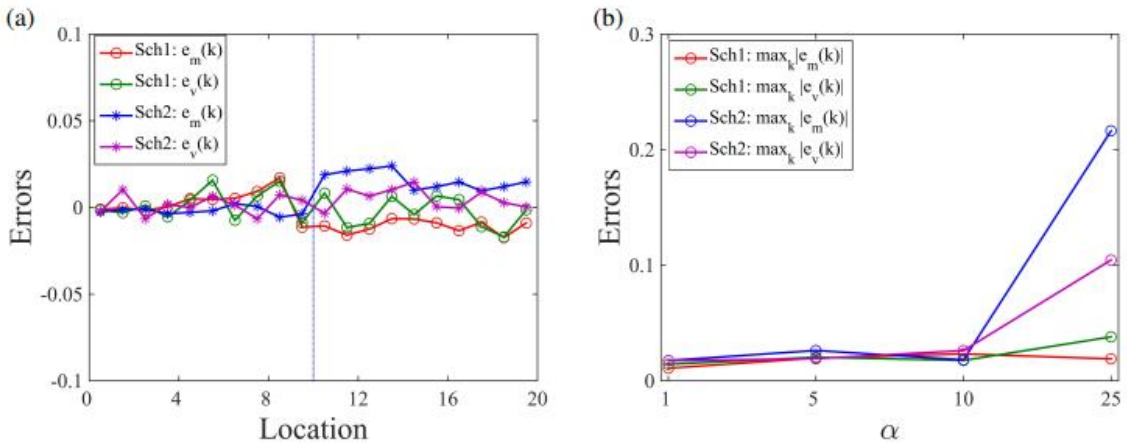

Fig. 5 (a) The mean number of morphogens in C2 at time t = 50 s. Different simulation methods are compared with α = 10, 20, 30, 40, 50 (with a static boundary): the Gillespie SSA with multigrid discretization (α grid points with size Δx and one grid point with size h), Scheme 1, Scheme 2, and Scheme 2 with no noise due to diffusion in Ωs . (b) The probability distribution of the normalized morphogen number in C1 with Scheme 2. The probability distributions are computed for X` (t)/Δx, ` = 1,2,...,α and compared among the cases with α = 10, 50 (with a static boundary) at time t = 0.01, 50 s. Initially, 50 molecules are located in Ωs in panels (a) and (b).

numbers of molecules in Ck , k = 1,2, is 25. Notice that Scheme 1 and the Gillespie SSA with two mesh sizes correctly approximate the means. However, Scheme 2 overestimates the mean morphogen number in C2 as α gets large. To understand where the numerical error arises, we also simulate Scheme 2 without the noise term in the SPDEs (marked as a blue line in Figure 5(a)), i.e. we remove the term with a square root in Equation (3.2). In Figure 5(a), we observe that the mean morphogen number in C2, E[X α+1], is underestimated as α increases when we use Scheme 2 without noise term in the SPDEs. Note that X ` , ` = 1,2,...,α, always have non-negative integer values due to no noise term in Equation (3.2). The molecular transfer from Ωs to Ωm occurs when X α ≥ 1. However, the frequency of this transfer is not sufficient as α gets large, which lowers X α+1 . On the other hand, with noise terms included in Equations (3.1) and (3.2), there are more chances that X ` < 0 for some ` = 1,2,...,α due to large fluctuations with a small number of molecules as α gets large.

[A'] Set $t = 0$, $t_s = \Delta t$ and $k = 1$. Initialize species numbers, $\mathbf{X}(0)$, in Ω_m and Ω_s, and the boundary location, $I(0)$. Then, generate random numbers r_1 and r_2 uniformly distributed in $(0,1)$. Set τ so that $\tau = -a_0^{-1}\log(r_1)$, where a_0 is defined in Equation (3.3). Set the next time when the diffusion occurs in Ω_m as $t_m = \tau$.

[B'] Repeat steps [B]–[C] of algorithm in Table 1 until time $k(n_c\Delta t)$.

[C'] Update position of interface $I(t)$, if necessary, by

- If $\sum_{\ell=1}^{\alpha} X^{(K_s-1)\alpha+\ell} < Q_{lower}$, set $I(t) = I(t - n_c\Delta t) - h$.
- If $\sum_{\ell=1}^{\alpha} X^{(K_s-1)\alpha+\ell} \geq Q_{lower}$ and $X^{K_s\alpha+1} > Q_{upper}$, set $I(t) = I(t - n_c\Delta t) + h$.
- Increase k, the number of updates performed in the interface position $I(t)$, by 1.

[D'] Repeat steps **[B']**–**[C']** until the simulation ends.

Table 3 Pseudocode for the adaptive multiscale reaction-diffusion algorithm with Scheme 1 applied to simulation of diffusion

Then, it is more frequent that X α ≥ 1 due to the fact that ∑ α+1 `=1 X ` = 50 and X ` < 0 for some ` = 1,2,...,α. More frequent molecular transfer from Ωs to Ωm causes overestimation of the mean morphogen number in C2 in Scheme 2. In Figure 5(b), we compare distributions of the morphogen numbers when α = 10 and 50. The distributions are computed from 1000 realizations of simulation when t = 0.01 s and 50 s. Each distribution is computed for all X ` , ` = 1,2,...,α so that we can display an overall range of the morphogen number in each discretization of Ωs . Each X ` is normalized by Δx so that the distributions can be compared for different α's. The normalized mean morphogen number (density) in Ωs decreases significantly in both cases with α = 10 and 50 as time evolves. On the other hand, the variance of the morphogen density is much greater for α = 50 than for α = 10 at t = 0.01 s due to the lower morphogen number in each discretization of Ωs . Therefore, we conclude that the error in Scheme 2 strongly depends on the size of fluctuations close to the interface. On the other hand, the molecular transfer from Ωs to Ωm is decided by ∑ α `=1 X ` in Scheme 1. This setting makes Scheme 1 more robust than Scheme 2 for large values of α since it helps to overcome the errors due to the negative abundance.

5 Application: moving boundary

In some applications [70], it is difficult to decide a position of the interface I a priori. In this section, we extend the presented algorithm to the case when the location of the interface I(t) between Ωs and Ωm moves in time, based on the number of molecules in each location of the domain. The multiscale approach with the adaptive interface is applied to the example introduced in Section 4. The adaptive algorithm is described in Table 3. Following [70], we introduce two thresholds denoted Qupper and Qlower (Qupper ≥ Qlower), and one integer parameter nc. We initialize the position of the interface I(0) = 0 in step [A'], i.e. we initially model the whole domain using the detailed compartment-based approach. We run the original Scheme 1 until time ncΔt. We check whether the interface I(t) should be moved in step [C']. If the number of molecules in the compartment next to the interface in Ωs is smaller than Qlower, a compartment-based model is used in that region. On the other hand, if the number of molecules in the boundary compartment next to interface I(t) in Ωm is larger than threshold Qupper, the corresponding compartment is transferred to the SPDE region where the molecules are redistributed uniformly in α grid points. Due to the uniform redistribution of the molecules, rapid changing of the interface I(t) introduces more errors. Note that in Scheme 1 with a fixed boundary, one molecule has been chosen randomly from α discretizations of CKs in Ωs and transferred to CKs+1 in Ωm. Similarly, we have taken one molecule from CKs+1 and transferred

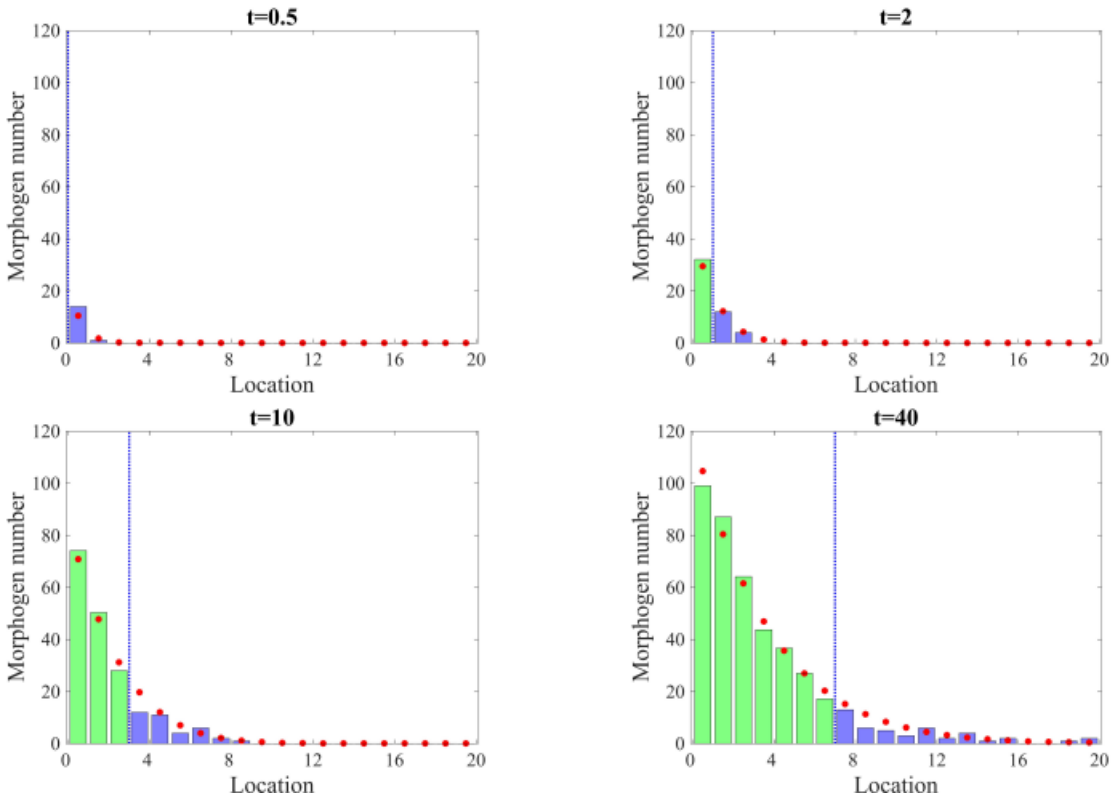

Fig. 6 Comparison between one realization of the number of morphogens using Scheme 1 with a moving interface, given in Table 3 (green bars and blue bars for the morphogen numbers in Ωs and Ωm, respectively) and the analytic solution of the mean (red dots). A blue dotted line represents the location of the interface.

the molecule to the randomly chosen SPDE numerical domain in CKs . However, in the adaptive algorithm, we modify the setting of Scheme 1 so that a molecule is taken uniformly from the entire region of CKs and transferred to CKs+1, i.e. 1/α molecule is subtracted in all α SPDE grid points of CKs . Similarly when the molecule is transferred from CKs+1 to CKs , 1/α molecule is added in all α grid points of CKs . Without this modification of the setting in Scheme 1, the appropriate level of the morphogen gradient is not formed in the next example. The adaptive algorithm [A']–[D'] is applied to the morphogen gradient model introduced in Section 4, and the results are presented in Figure 6. We use Qlower = 15, Qupper = 25 and nc = 10. Other parameters are given in Table 2. Our initial condition is X k (0) = 0, for k = 1,2,...,Ksα +Km, i.e. the system starts with no molecules and the gradient is formed during the simulation. In Figure 6, one realization of the algorithm in Table 3 at different times t = 0.5, 2, 10, 40 s is presented. The green and blue bars represent the numbers of molecules in the corresponding compartments in Ωs and Ωm, respectively. The blue dotted line represents interface I(t) between two regions, and the red circles are the mean numbers of molecules obtained from the analytic solution of the stochastic model. Our results show that the boundary between two regions is moving to the right in time as the molecule numbers increase due to the production on the left. In Figure 7(a) and 7(b), we simulate the adaptive algorithm with fixed thresholds for a range of values of nc = 1, 10, 102 , 103 , 104 , which are the numbers of time steps to check the criterion to move the interface I(t) in step [C']. Two sets of fixed thresholds are chosen, (Qlower,Qupper) = (15,25) in (a) and (Qlower,Qupper) = (20,20) in (b). In Figure 7(c) and 7(d), we simulate the adaptive algorithm with fixed numbers

of time steps, nc, for different values of Qlower and Qupper, which are the threshold values to check before we move the interface I(t) in step [C']. We use the following pairs of the values for the thresholds: (Qlower,Qupper) =

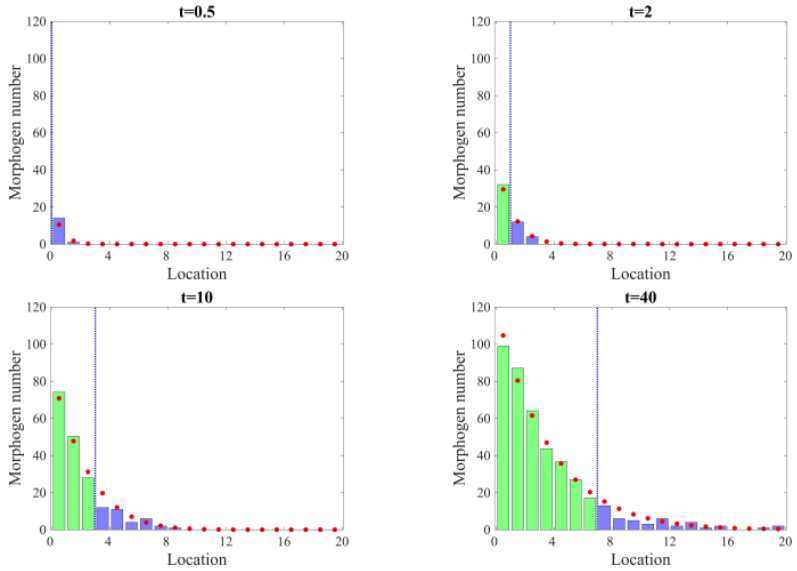

Fig. 7 The maximum absolute values of the relative errors in all locations at time 50s using the multiscale algorithms, with a moving boundary. Different values of nc = 1, 10, 102 , 103 , 104 , are used with fixed thresholds (a) (Qlower,Qupper) = (15,25), and (b) (Qlower,Qupper) = (20,20). Different threshold values (Qlower,Qupper) = (5,55), (10,40), (15,25), (20,20) are used with (c) nc = 1 and (d) nc = 103 . Red and green lines are the maximal relative errors of the means and standard deviations between the analytic solution of the Markov chain model and Scheme 1. Blue and purple lines are maximal relative errors between the Markov chain model and Scheme 2.

(5,55), (10,40), (15,25), (20,20). Two fixed numbers of time steps are used, nc = 1 in (c) and nc = 1000 in (d). As shown in Figure 7(b) and 7(c), we observe that the maximum absolute values of the relative errors increase as the number of time steps, nc, or the size of the threshold window, Qupper −Qlower, gets smaller. This is because the small size of the number of time steps or the threshold window makes the interface location change frequently, which causes additional errors. On the other hand, Figure 7(a) and 7(d) do not show similar pattern since large size of the threshold window (Qupper − Qlower = 10) and the number of time steps (nc = 103) prevents frequent movement of the interface location. Overall, Scheme 2 has slightly smaller errors than Scheme 1. In Figure 7, the maximum absolute values of the relative errors are calculated using 104 realizations of simulation using Scheme 1 or 2 for each value of nc and for each set of values of (Qlower,Qupper) and using the analytic solution of the Markov chain model.

6 Applications:

multiple species In this section, we illustrate the applicability of the multiscale approach to chemical systems with multiple species. Since different chemical species can have very different molecular distributions in the computational domain, the partition of the computational domain into subdomains Ωs and Ωm can be species dependent. We use the pom1p gradient model from Saunders et al. [72] to illustrate a multiscale approach, where

each species has a different partition into Ωs and Ωm depending on its molecular distribution. The model consists

Notations	Description	Values
L	Length of the Domain	$14\,\mu$m
I_1	Left boundary of Ω_s	$3.5\,\mu$m
I_2	Right boundary of Ω_s	$10.5\,\mu$m
D_1	Diffusion coefficient of S_1	$0.02\,\mu\text{m}^2\,\text{s}^{-1}$
D_2	Diffusion coefficient of S_2	$0.2\,\mu\text{m}^2\,\text{s}^{-1}$
a_1	Production parameter of S_1	$1.029\,\text{s}^{-1}$
a_2	Production parameter of S_2	$0.441\,\text{s}^{-1}$
a_3	Fragmentation rate of S_1	$0.26\,\text{s}^{-1}$
a_4	Aggregation rate	$0.049\,\text{s}^{-1}$
a_5	Disassociation rate	$0.035\,\text{s}^{-1}$
a_6	Parameter of production	$0.1089\,\text{s}^{-1}$
Δx	Spatial discretization in Ω_s	$0.035\,\mu$m
h	Compartment size in Ω_m	$0.35\,\mu$m
Δt	Time discretization for SPDE	0.0005 s

Table 4 Parameter values in the two-state model for pom1p gradient.

of two species, slow-diffusing pom1p clusters, denoted S1, and fast-diffusing pom1p particles, denoted S2. We use pseudo 1-dimensional domain Ω as in Figure 1, where L = 14μm, which is divided into K = 40 compartments, Ck , k = 1,2,...,K. Both S1 and S2 are produced in the whole computational domain with space-dependent rates [72], i.e. with propensities

$$\lambda_1^k(\mathbf{Z}^k) = a_1 \exp\left[-a_6\left(k-\frac{K+1}{2}\right)^2\right], \qquad \lambda_2^k(\mathbf{Z}^k) = a_2 \exp\left[-a_6\left(k-\frac{K+1}{2}\right)^2\right],$$

where k = 1,2,...,K, and a1, a2 and a6 are constants given in Table 4. In addition to production, species S1 and S2 are subject to the following reactions which take place in the whole domain

$$S_1 \longrightarrow S_2, \qquad S_1 + S_2 \longrightarrow S_1 + S_1, \qquad S_2 \longrightarrow \varnothing,$$

with the corresponding propensities given by

$$\lambda_3^k(\mathbf{Z}^k) = a_3 Z_1^k, \qquad \lambda_4^k(\mathbf{Z}^k) = a_4 Z_1^k Z_2^k, \qquad \lambda_5^k(\mathbf{Z}^k) = a_5 Z_2^k,$$

where k = 1,2,...,K, and a3, a4 and a5 are constants given in Table 4. In Figure 8, we present an illustrative simulation of pom1p gradient model. We plot spatial distributions of S1 and S2 at times t = 50 s and t = 1000 s. We observe that the spatial distribution of S1 contains a region with high abundance of molecules in the center of the computational domain. The chemical species S2 has low copy numbers in the entire domain. Therefore, we introduce the SPDE region in the middle of the domain by (note that we fix K = 40 in this example)

$$\Omega_s = \bigcup_{k=11}^{30} C_k,$$

where the coarse-graining is only applicable to S1 in Ωs . In particular, we have introduced two interfaces, I1 and I2 between Ωs and Ωm. Diffusion of chemical species S1 is simulated using the algorithm in Table 1. Similarly, production of S1 is implemented using the SPDE and Markov chain model in Ωs and Ωm, respectively, as we did in Equation (3.1). The chemical species S2 is simulated by the Markov chain model in the entire domain, because the average number of molecules of S2 is relatively low. In particular, diffusion, production and degradation of S2 are implemented as in the underlying Markov chain model. The only complications are reactions

$$S_1 \longrightarrow S_2 \quad \text{and} \quad S_1 + S_2 \longrightarrow S_1 + S_1, \tag{6.1}$$

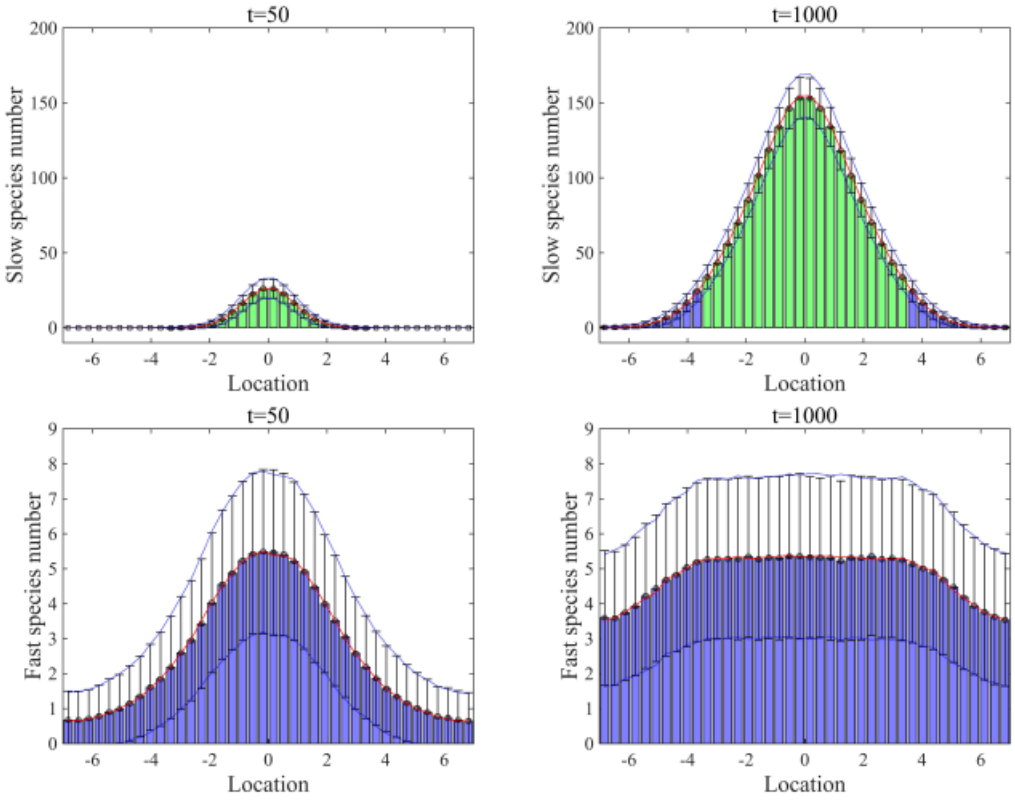

Fig. 8 Mean numbers of the molecules of slow-diffusing pom1p clusters and fast-diffusing pom1p particles, and their standard deviations from the means at t = 50, 1000 s computed from 104 realizations of simulation using the SSA and the multiscale algorithm with Scheme 1.

because they include both species S1 and S2, which are in Ωs described by different modeling approaches. We treat these reactions as time-changed Poisson processes in both subdomains Ωm and Ωs . Discretizing each compartment, Ck , k = 11,12,...,30, into α grid points, the state of S1 variable is described by vector, X(t) = (X 1 ,X 2 ,...,X 20(α+1)) where X 1 ,X 2 ,...,X 10 (resp. X 20α+11 ,X 20α+12 ,...,X 20(α+1)) are the numbers of molecules of S1 in the left (resp. right) part of Ωm. The values of SPDE description in compartment Ck , k = 11,12,...,30, are given by X 10+(k−11)α+` , ` = 1,2,...,α. The state of S2 variable is described by vector, Y(t) = (Y 1 ,Y 2 ,...,Y 40) where Y k is the number of molecules of S2 in

compartment Ck , k = 1,2,...,K. The propensity of the first reaction in (6.1) of the multiscale model is given by

$$\lambda_3^k(\mathbf{X}) = \begin{cases} a_3 X^k, & \text{for } k = 1,2,\ldots,10, \\ a_3 \sum_{\ell=1}^{\alpha} X^{10+(k-11)\alpha+\ell}, & \text{in } \Omega_s \text{ (i.e. for k=11,12,...,30),} \\ a_3 X^{20(\alpha-1)+k}, & \text{for } k = 31,32,\ldots,40. \end{cases} \quad (6.2)$$

The propensity of the second reaction in (6.1) of the multiscale model is given by

$$\lambda_4^k(\mathbf{X}, Y^k) = \begin{cases} a_4 X^k Y^k, & \text{for } k = 1,2,\ldots,10, \\ a_4 Y^k \sum_{\ell=1}^{\alpha} X^{10+(k-11)\alpha+\ell}, & \text{in } \Omega_s \text{ (i.e. for k=11,12,...,30),} \\ a_4 X^{20(\alpha-1)+k} Y^k, & \text{for } k = 31,32,\ldots,40. \end{cases} \quad (6.3)$$

We simulate reactions in (6.1) as time-changed Poisson processes with propensities in Equations (6.2)–(6.3). If the first of these reactions occurs in Ck , k = 11,12,...,30, we subtract 1/α from each X 10+(k−11)α+` ,` = 1,2,...,α, and we add one to Y k . If the second reaction in (6.1) occurs in Ck , k = 11,12,...,30, we add 1/α to each X 10+(k−11)α+` , ` = 1,2,...,α, and we subtract one from Y k . Note that the conversion of S1 in Ck , k = 11,12,...,30, is applied equally to the entire α grid points of Ck rather than to one randomly chosen grid point in Ck as we do for diffusion across the interfaces. In Figure 8, green bars and blue bars represent the mean numbers of molecules of the pom1p clusters and particles in Ωs and Ωm using the multiscale algorithm with Scheme 1. Error bars represent one standard deviation from the mean in the multiscale approach. Red lines and blue dotted lines are the mean numbers and their standard deviations from the means computed by the Gillespie SSA simulating the compartmentbased approach in the entire domain. Both statistics using the compartment-based approach and the multiscale algorithm are computed from the 104 realizations of the simulations for each case.

7 Discussion

A Markov chain model (compartment-based model) has been widely used to describe the discrete nature of the molecular copy numbers and inherent stochasticity in reaction-diffusion systems, but it can be computationally intensive. A possible approach to increase efficiency of simulations is to approximate a part of the model by some coarse-grained methods. In this paper, we have introduced two multiscale algorithms coupling the SPDEs and the Markov chain model, which provide good approximations to the solutions obtained by the Markov chain model applied in the entire spatial domain. Two coupling methods of the Markov chain model and the SPDEs across the interface have been studied. In this section, we compare the presented approach with methods in the literature. Several Langevin formulations have been introduced to model fluctuating hydrodynamics for chemically reactive species [9] and stochastic reaction-diffusion systems [40, 50]. In particular, the spatial chemical Langevin equation was applied to the Gray-Scott model, and its pattern formation was compared to the ones obtained by the reaction-diffusion master equation and PDEs [40]. The spatial chemical Langevin equation consists of a system of stochastic differential equations, and it corresponds to Equation (2.4) in Section 2. On the other hand, several approaches using SPDEs [1, 2, 6, 18, 57] have been introduced to model stochastic reaction-diffusion systems. In [6], the SPDE

was derived for reaction-diffusion systems, and discretization of PDEs and stochastic fields was discussed. Unlike Equation (2.5), the stochastic fields in the discretized SPDEs account for fluctuations due to diffusion but not for reaction. In [57], the SPDE for reaction-diffusion systems was derived which is consistent with Equation (2.8). In their formulation, diffusion was implemented by the SPDE while the reaction was simulated using the exact or modified SSA. In [79], two hybrid algorithms are suggested for coupling a compartment-based model and a PDE model when the size of the PDE discretization is less than or equal to the compartment size. Both algorithms extend the PDE approach to the systems with low copy numbers of molecules in a part of the computational domain. The first algorithm considers the PDE solution as the probability density to find a particle within the region and is applied to both cases of low and high copy numbers of molecules in the PDE region. The second algorithm is a simplified and more efficient version of the first one when the PDE region involves the high copy number of molecules. Like in this paper, both algorithms implement a pseudo-compartment with size h in the PDE region where h represents the compartment size. The second algorithm in [79] is similar to Scheme 1 if a discretized version of SPDEs replaces the PDEs. However, the interface between the two modeling regimes is assumed to be fixed in [79]. In [45], a hybrid algorithm is introduced coupling a compartment-based model and PDEs where the size of the PDE discretization is much finer than the compartment size. In the model, an overlap region is defined with two interfaces (corresponding to the pseudo-compartment in Scheme 1) where both modeling regimes are valid, and both cases with fixed and adaptive interfaces are considered. Unlike our pseudo-compartment in Scheme 1, the overlap region can contain multiple compartments if needed. On one interface between the compartment-based model and the overlap region, the population of the PDE solution on the interface is matched to the average of the population in the neighbouring compartments. On the other interface between the PDE region and the overlap region, flux on the interface was matched. The hybrid algorithm in [45] approximates the mean population numbers in the compartment-based model if it was applied over the entire spatial region. The use of the overlap region allows matching the variance between two models in the compartment-based region when the fixed interface is used. On the other hand, the goal of Scheme 1 and Scheme 2 is to approximate the compartment-based model by employing the discretized version of SPDEs in the region with high molecules. Therefore, we can match both the mean and variance of the population numbers computed by our multiscale algorithms to those in the compartment-based model if it was used in the whole spatial domain. This is done for both cases with a fixed or adaptive boundary. Unlike the previous approaches in [45, 79], the presented multiscale algorithms can apply to systems with multiple species as it is shown in Section 6 where each species has a different partition of the spatial domain into subdomains where different models are used, depending on the spatial distribution of molecules of each species. In [76], a hybrid algorithm is presented using a compartment-based model and PDEs, where the size of the compartment and numerical discretization for the PDE model is equal. In this paper, we have discussed the case when the mesh size of the numerical discretization of the SPDEs is smaller (or equal) than the compartment size in the Markov chain model ($h \geq \Delta x$). This case is useful when we add inherent stochasticity in the PDE model where a fine spatial resolution of the PDE solution is required to describe the solution of the SPDE. This case was also discussed in other hybrid algorithms coupling the compartment-based model and the macroscopic PDEs [45, 79]. The other case, $h < \Delta x$, discussed, for example, in the hybrid algorithm coupling a random walk on a lattice and the PDE model [35], is helpful

when the PDE or SPDE model is used as a coarse-grained approximation of the compartment-based model. Such approximation can be used in the region where spatial concentration gradients are not large, so they do not require a fine resolution in space. Although we have focused on the case $h \geq \Delta x$, the presented approach can be extended to $h < \Delta x$ as well. In fact, if $h = \Delta x$, both Scheme 1 and Scheme 2 will be the same. If $h < \Delta x$, we may be able to consider an overlap region (like a pseudo-compartment) in the compartment-based region to extend Scheme 1. The presented SPDE-based approach provides a bridge between the stochastic approach (using the Markov chain compartment-based model) and the deterministic approach (using the macroscopic PDEs) by incorporating a discretized version of SPDEs. The SPDEs can be utilized to build other hybrid models, for example, by coupling them with macroscopic PDEs. Then some approaches used in the hybrid algorithms coupling the compartment-based model with the PDEs [45, 74, 76, 79] will naturally apply to the case with the SPDEs.

Acknowledgements

Thank you to the Leverhulme Trust for the Philip Leverhulme Prize and the Royal Society for the University Research Fellowship. HWK's research trips to Oxford were funded by the prize money. The Seventh Framework Programme (FP7/2007-2013) of the European Community, under ERC grant agreement no. 239870, provided support for the study that led to these findings. Part funding for this work was provided by the University of Maryland Baltimore County under grant UMBC KAN3STRT (HWK), the Mathematical Biosciences Institute, and the National Science Foundation under grants DMS-1440386 and DMS-1620403 (HWK). In order to acknowledge their support and hospitality during the program "Stochastic Dynamical Systems in Biology: Numerical Methods and Applications," where work on this research was conducted, the authors would like to thank the Isaac Newton Institute for Mathematical Sciences, Cambridge. EPSRC grant number EP/K032208/1 funded this work. The Simons Foundation provided funding for this work.

References

[1] F. Alexander, A. Garcia, and D. Tartakovsky, Algorithm refinement for stochastic partial differential equations: I. linear diffusion, Journal of Computational Physics 182 (2002), no. 1, 47–66.

[2] , Algorithm refinement for stochastic partial differential equations: II. correlated systems, Journal of Computational Physics 207 (2005), no. 2, 769–787.

[3] D. Altintan, A. Ganguly, and H. Koeppl, Efficient simulation of multiscale reaction networks: A multilevel partitioning approach, American Control Conference, 2016, 2016, pp. 6073–6078.

[4] M. Ander, P. Beltrao, B. Di Ventura, J. Ferkinghoff-Borg, M. Foglierini, A. Kaplan, C. Lemerle, I. Tomas-Oliveira, and L. Serrano, ´ SmartCell, a framework to simulate cellular processes that combines stochastic approximation with diffusion and localisation: analysis of simple networks, Systems Biology 1 (2004), no. 1, 129–138.

[5] S. Arjunan and M. Tomita, A new multicompartmental reaction-diffusion modeling method links transient membrane attachment of E. coli MinE to E-ring formation, Systems and Synthetic Biology 4 (2010), no. 1, 35–53.

[6] P. Atzberger, Spatially adaptive stochastic numerical methods for intrinsic fluctuations in reaction–diffusion systems, Journal of Computational Physics 229 (2010), no. 9, 3474–3501. 20 Hye-Won Kang, Radek Erban

[7] J. Bakarji and D. Tartakovsky, On the use of reverse Brownian motion to accelerate hybrid simulations, Journal of Computational Physics 334 (2017), 68-80.

[8] D. Bernstein, Simulating mesoscopic reaction-diffusion systems using the Gillespie algorithm, Physical Review E 71 (2005), no. 4, 041103.

[9] A. Bhattacharjee, K. Balakrishnan, A. Garcia, J. Bell, and A. Donev, Fluctuating hydrodynamics of multi-species reactive mixtures, Journal of Chemical Physics 142 (2015), no. 22, 224107.

[10] Y. Cao, D. Gillespie, and L. Petzold, Multiscale stochastic simulation algorithm with stochastic partial equilibrium assumption for chemically reacting systems, Journal of Computational Physics 206 (2005), 395–411.

[11] , The slow-scale stochastic simulation algorithm, Journal of Chemical Physics 122 (2005), no. 1, 14116.

[12] Y. Cao, H. Li, and L. Petzold, Efficient formulation of the stochastic simulation algorithm for chemically reacting systems, Journal of Chemical Physics 121 (2004), no. 9, 4059–4067. [13] S. Cotter, K. Zygalakis, I. Kevrekidis, and R. Erban, A constrained approach to multiscale stochastic simulation of chemically reacting systems, Journal of Chemical Physics 135 (2011), 094102.

[14] S. Cotter, T. Vejchodsky, and R. Erban, Adaptive finite element method assisted by stochastic simulation of chemical systems, SIAM Journal on Scientific Computing 35 (2013), no. 1, B107-B131.

[15] S. Cotter and R. Erban, Error analysis of diffusion approximation methods for multiscale systems in reaction kinetics, SIAM Journal on Scientific Computing 38 (2016), no. 1, B144-B163.