The pseudo-compartment method for coupling PDE and compartment-based models of diffusion

Christian A. Yates[1,*], Mark B. Flegg[2]

1 Department of Mathematical Sciences, University of Bath, Claverton Down, Bath, BA2 7AY, United Kingdom
2 School of Mathematical Sciences, Monash University, Wellington Road, Clayton, Victoria, 3800, Australia
∗ E-mail: c.yates@bath.ac.uk

Key index words: hybrid modelling, stochastic reaction-diffusion, pseudo-compartment, multiscale modelling

Abstract

Spatial reaction-diffusion models have been employed to describe many emergent phenomena in biological systems. The modelling technique most commonly adopted in the literature implements systems of partial differential equations (PDEs), which assumes there are sufficient densities of particles that a continuum approximation is valid. However, due to recent advances in computational power, the simulation, and therefore postulation, of computationally intensive individual-based models has become a popular way to investigate the effects of noise in reaction-diffusion systems in which regions of low copy numbers exist.

The specific stochastic models with which we shall be concerned in this manuscript are referred to as 'compartment-based' or 'on-lattice'. These models are characterised by a discretisation of the computational domain into a grid/lattice of 'compartments'. Within each compartment particles are assumed to be well-mixed and are permitted to react with other particles within their compartment or to transfer between neighbouring compartments.

Individual-based stochastic models provide microscopic/mesoscopic accuracy but at the cost of significant computational resources. Models which have regions of both low and high concentrations often necessitate coupled macroscale and microscale modelling paradigms. This is because microscale models are not feasible to simulate at large concentrations and macroscale models are often inappropriate at small concentrations.

In this work we develop two hybrid algorithms in which a PDE in one region of the domain is coupled to a compartment-based model in the other. Rather than attempting to balance average fluxes, our algorithms answer a more fundamental question: 'how are individual particles transported between the vastly different model descriptions?' First, we present an algorithm derived by carefully re-defining the continuous PDE concentration as a probability distribution. Whilst this first algorithm shows very strong convergence to analytic solutions of test problems, it can be cumbersome to simulate. Our second algorithm is a simplified and more efficient implementation of the first, it is derived in the continuum limit over the PDE region alone. We test our hybrid methods for functionality and accuracy in a variety of different scenarios by comparing the averaged simulations to analytic solutions of PDEs for mean concentrations.

Introduction

Often considered to be a stochastic process, diffusion is, at a molecular level, a deterministic process. Diffusion of macromolecules and larger entities is the result of a rapid number of deterministic collisions with solvent molecules. Often, the information about the solvent molecules is not known and the macromolecule (particle) appears to move stochastically. The motion of an individual particle can be simulated in order to generate its random trajectory. Stochastic simulations of this type are known as individual-based simulations and can be subdivided into two types (Erban and Chapman, 2009); those

in which trajectories are defined on continuous domains and those in which trajectories are defined by a random walk on a lattice. Stochastic trajectories of individual particles are appropriate on time-scales which are much larger than the time between solvent molecule collisions and in such circumstances the particle trajectories may resemble a Wiener process (Brownian motion). Coarse-graining the continuous domain into lattice points on which particles undergo a random walk is a simplification which may be appropriate if the lattice spacing is not too large as to impair good spatial resolution. A lower bound may also be implemented on the lattice spacing if the particles interact (react) with each other. We will not concern ourselves with such issues in this paper, but refer the interested reader to the following works Erban and Chapman (2009); Fange et al. (2010) and Isaacson (2009, 2013). Both off-lattice (Erban and Chapman, 2009; Smoluchowski, 1917; van Zon and ten Wolde, 2005) and on-lattice (Baker et al., 2010; Drawert et al., 2010; Engblom et al., 2009; Lampoudi et al., 2009; Yates and Baker, 2013; Yates et al., 2012) individual-based stochastic simulations have remained popular paradigms for stochastic diffusion simulations.

It is common for mathematical models of diffusive processes to treat diffusion using a deterministic continuous partial differential equation (PDE). That is, it is assumed that the copy numbers of particles are so large that a continuous distribution of particles can be defined which evolves deterministically. The effect on the time derivative of the concentration distribution of particles due to the random motion of individuals is represented by a Laplacian term in the governing equation. Other effects, such as particle flow, reactions and other interactions can be incorporated into this mathematical framework by the introduction of additional terms in the governing PDEs. The simplicity with which diffusion can be represented in these deterministic models has, in part, lead to its popularity in modelling complex physical behaviour. This is notably the case in mathematical modelling of biological processes (Adomian, 1995; Field and Noyes, 1974; Schnakenberg, 1979) due to the importance of diffusion of either chemicals and/or cells in many biological environments. Deterministic PDE models have the distinct advantage over stochastic models of diffusion that they are analytically tractable so that the particle dynamics may, in some circumstances, be predicted without numerical simulation. In addition, when numerical solutions are required, there exists a plethora of well-established techniques at hand.

Whilst PDE interpretations of the diffusion mechanisms are convenient and popular they lack the ability to appropriately represent systems in which stochasticity is inherent (Erban et al., 2007). For example, fluctuations in the number of particles in a system are especially important for systems in which one or more of the constituent species is in low abundance. Such situations appear frequently in biological applications due to the small spatial scales that are often involved and can lead to dynamics which are significantly different from a deterministic model which may be postulated based purely on macroscopic considerations and assuming large copy numbers. Examples of such low copy number effects in biology can be seen in genetic transcription (Elowitz et al., 2002; Swain et al., 2002) and protein synthesis (Ghosh et al., 2007). Noise is also important in systems in which random changes in species numbers can cause systems to cross bifurcation thresholds, as evidenced in bistable switching systems such as the lysis/lysogeny genetic regulation of bacteriophage λ (Tian and Burrage, 2004) and oscillatory systems such as the circadian clock (Gonze et al., 2002; Vilar et al., 2002). The effects of stochasticity have also begun to be investigated in spatially extended systems of the type we consider in this paper (Baker et al., 2010; Woolley et al., 2011; Yates et al., 2012). Due to the increase in computer capabilities, stochastic simulation techniques for reaction-diffusion systems have found popularity and have been implemented in a number of freely available simulation packages. As this paper will focus on on-lattice (compartment-based) stochastic diffusion we refer the reader to software packages such as SmartCell (Ander et al., 2004), MesoRD (Hattne et al., 2005), Dizzy (Ramsey et al., 2005), (py)URDME (Drawert et al., 2012) and StochSS (Sanft et al., 2011) (which makes use of pyURDME).

Individual-based simulations (off- and on-lattice stochastic simulations) of diffusion are capable of capturing stochastic effects which PDEs cannot. However, this is at the cost of computational effort which becomes an insurmountable barrier very quickly as copy numbers increase. For extended domains,

in which concentrations of particle species vary dramatically (for example in the formation of morphogen gradients in the early embryo (Bollenbach et al., 2005; Kicheva et al., 2007; Tostevin et al., 2007; Wolpert, 1969)), a computationally feasible model may require localised regions in which diffusion is modelled using a continuous PDE (for regions in which concentration is high) and an individual-based simulation (for regions in which concentration is low).

The importance of efficient and accurate hybrid algorithms for coupling models of diffusion between different regions of space has resulted in a large body of research in the last decade. Most of the models which couple the PDE paradigm with compartment-based stochastic simulations involve the use of significant overlapping regions and/or the averaging of particle flux in order to simulate the necessary boundary condition for the PDE (Alexander et al., 2002; Flekkøy et al., 2001; Moro, 2004; Wagner and Flekkoy, 2004). Ferm et al. (2010) also have a method of coupling a PDE to a compartment-based simulation which employs operator splitting and τ-leaping for the time evolution of the compartment-based model where possible. Recently, hybrid models of stochastic diffusion in which on- and off-lattice models are coupled between different regions of space have been proposed (Erban et al., 2013; Flegg et al., 2012, 2014; Robinson et al., 2014). Central to these models is the ability to treat individual particles as they migrate from one regime to the other. The individual treatment of particles allows for the extension of these coupling algorithms to low copy number regimes. This type of coupling technique was also recently applied the coupling of a PDE with particles undergoing off-lattice Brownian dynamics (Franz et al., 2012, 2013).

In this paper we develop hybrid algorithms which couple a PDE in one region of the domain to a compartment-based model in the other (we will denote these regions as Ω_P and Ω_C respectively). Rather than balancing flux at the interface we use a method which is similar to the ghost-cell method (Flegg et al., 2014). A defining characteristic of this method is the individual treatment of particles as they cross the interface. In our coupling method a small region of the PDE domain (equivalent to one compartment length, h) adjacent to the compartment-based region is allowed to contribute particles to the compartment-based regime as if it were itself a compartment in Ω_C. We will denote this pseudo-compartment as $C_{-1} \subset \Omega_P$. The treatment of this pseudo-compartment is complicated. From the perspective of the PDE region it contains a continuous probability distribution representing the distribution of particles. However, from the perspective of the compartment regime, the pseudo-compartment has compartment-like properties. That is, it contains particles which can jump into their neighbouring compartment according to standard compartment-based rules for diffusion. The pseudo-compartment's duality allows for particles to behave correctly as they cross individually between the two different regimes. When particles cross over the interface a step-function with the mass of a single particle is added to the probability distribution in the region of the pseudo-compartment.

A zero-flux boundary condition is implemented on the PDE at the interface between the two regions. This is because flux of particles over this interface is not governed by a continuous diffusive flow of particles but rather a discrete jumping processes which mimics it. It is important to ensure that there is consistency between the methods by which particles transfer back and forth between regimes since any imbalance can create significant numerical artefacts.

If the size of the PDE region was to be reduced to the size of a compartment (i.e. consisting only of the pseudo-compartment, C_{-1}) then, effectively the PDE region would be exactly equivalent to an extra compartment, whilst if the compartment-based domain disappeared the pseudo-compartment would behave as the rest of the PDE. A one-dimensional graphical representation of this approach is given in Figure 1.

We present two separate methods (which we refer to as pseudo-compartment methods (PCM)) for conceptualising and implementing this coupling scheme. The first method is a necessary for the precise coupling of systems with low copy numbers. The second method is a mathematically justified simplification of the first for situations in which large particulate copy numbers are present within the PDE domain. We will demonstrate these coupling mechanisms in one dimension only but they are easily generalised to
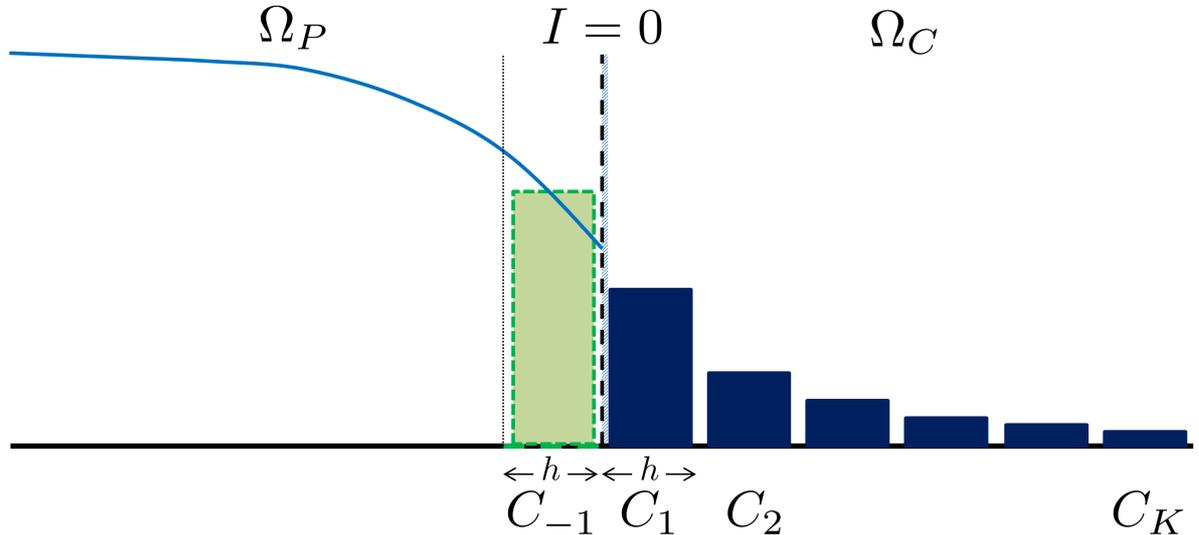
Figure 1. In the left-hand-side of the domain, Ω_p, the distribution of particles is modelled as a PDE. On the right-hand-side, Ω_C, the particle density is modelled with a compartment-based approach. The concentration of particles in each compartment, C_i for $i = 1, \dots, K$, of Ω_C is represented by a dark blue rectangle (numbers of particles can be found by multiplying by the interval width, h). The probability density in Ω_P is represented by the light-blue curve. The two regimes interact through the movement of particles from the left-most compartment in Ω_C, C_1, to the pseudo-compartment, $C{-}1$, (whose average particle concentration is represented by the height of the green box on the right of Ω_P) and vice versa. The height of the green-rectangle is calculated by integrating the PDE solution over the region $[0, h)$ and dividing by h. This 'number of particles' in the pseudo-compartment can be calculated by multiplying the height of the green box by h.

higher dimensional problems with (hyper)-planar interfaces.

Algorithms

Overview of algorithms for deterministic and compartment-based stochastic diffusion

Compartment-based modelling

We will denote the region in which we employ the compartment-based regime Ω_C. In this region we partition the domain into K regular compartments of length h. We choose h to be sufficiently small that particles in each compartment can be considered to be well-mixed, reacting only with particles in their current compartment. We do not keep detailed information about the specific location of each particle, rather we simply store the number of particles of each type in each of these compartments. Whilst this produces a significant saving on computational intensity in comparison to Brownian dynamics models (in which the position of each individual is stored and/or updated at each step), we lose information about the specific location of each of the particles in a compartment and are forced to assume that each particle is distributed uniformly throughout the compartment.

Considering a general system of N species and M chemical reactions, a propensity function, $\alpha_{i,j}(t)$, $i = 1, \ldots, K$ and $j = 1, \ldots, M + 2$, is assigned to each reaction in each compartment (including the two "diffusive reactions" by which particles leave compartment i). The propensity functions are defined such that $a_{i,j}(t)dt$ is the probability that the jth reaction occurs in the ith compartment during the infinitesimally small time interval $[t, t + dt]$. The propensity function for a specific reaction in a particular compartment depends on the number of reactant particles available in that compartment (for all but zeroth order reactions), the method by which they combine, the kinetic rate constant and (for all but first order reactions) the size of the compartment, h (Erban and Chapman, 2009). Similarly, for each species, we can characterise the diffusive jumps of particles between compartment i and its two nearest neighbouring compartments as a set of first order interactions with propensities

$$\alpha_{i,j} = A_i D_A / h^2, \tag{1}$$

for $j = M + 1, M + 2$. Here, D_A denotes the macroscopic diffusion coefficient of species A and A_i is the copy number of species A in compartment i.

There are two predominant mechanisms for simulating the time evolution of the system. In the *time-driven* approach, a fixed time-step Δt is designated and each reaction j in compartment i occurs with probability $\alpha_{i,j}\Delta t$. This method is simple to implement, but less efficient than the alternative *event-driven* algorithm. In order to ensure accuracy in the time-driven algorithm the time-step must be chosen to be sufficiently small that single particles do not undergo multiple diffusion or reaction events in a single step. This can lead to very small choices of Δt and consequently many time-steps in which no reaction/diffusion events occur. In comparison, an event-driven algorithm simulates the evolution of the system by choosing the time-step between reaction/diffusion events from an exponential distribution. The most commonly cited example of such a simulation scheme is the Gillespie's direct method (DM) (Gillespie, 1977) in which the time for the next reaction of any sort, τ, is chosen from an exponential distribution with parameter α_0 (where $\alpha_0 = \sum_{i=1}^{N} \sum_{j=1}^{M+2} \alpha_{i,j}$ is the sum of all the propensity functions). In practice τ is simulated according to the following formula

$$\tau = \frac{1}{\alpha_0} \ln \frac{1}{r_1}, \tag{2}$$

where r_1 is uniformly distributed in $(0,1)$. The event that takes place τ units of time from the current time is chosen from a weighted probability proportional to each of the event propensities $\alpha_{i,j}$.

There are efficient adaptations of Gillespie's original algorithm available depending on context (Cao et al., 2004; Gibson and Bruck, 2000; Li and Petzold, 2006; McCollum et al., 2006; Yates and Klingbeil, 2013). These algorithms are mathematically identical and differ only in the computational bookkeeping required to implement them. An efficient algorithm designed specifically for spatially extended systems is the next sub-volume method (NSM) from Elf and Ehrenberg (2004), and an adaptation of the next reaction method (NRM) from (Gibson and Bruck, 2000) (which is itself based on Gillespie's original first reaction method (FRM) (Gillespie, 1976)). For simplicity and efficiency we choose Gillespie's DM in order to evolve particle numbers in Ω_C, although we note that our algorithms can be trivially modified to incorporate the alternative event-driven simulation methods discussed above or time-driven evolution.

PDE-based modelling

We denote the region in which we solve the PDE as Ω_P. In this region we solve the relevant one-dimensional PDE numerically, but to a high degree of accuracy. In particular we solve the PDE

$$\frac{\partial p}{\partial t} = D \frac{\partial^2 p}{\partial x^2} + R(p), \qquad x \in \Omega_p, \tag{3}$$

where $R(\rho)$ represents a set of, as yet unspecified, reactions.

In the examples that follow, we discretise the PDE-domain, Ω_p, into a regular grid with spacing Δx where $\Delta x \ll h$. In order that the time-step which we take is not restricted by the requirement for stability of the PDE solver we employ the θ-method with $\theta \leq 1/2$ which is unconditionally stable. In particular we choose $\theta = 1/2$ (Crank-Nicolson method) which gives a solution which is second order accurate in space. Note that the coupling algorithm that follows is independent of the method which is used to solve the PDE, providing that method is sufficiently accurate.

Low copy number coupling technique: Algorithm 1

It is common to think of the PDE description of diffusion as the evolution of a continuous concentration. However, at low concentrations this quantity is not well defined. Here we consider a situation in which there may be insufficient particles to satisfy the continuum limit. In this situation, the continuous distribution cannot be interpreted as a concentration but rather as the probability density to find each of the N_P particles that are within the PDE domain (we denote the number of compartment particles N_C such that total number of particles is $N = N_P + N_C$). We specify that in the continuum limit, as $N_P \to \infty$, the PDE should describe the concentration of particles. As such, we define the continuous distribution $p(x,t)$ to be the probability density to find each of the N_P PDE-based particles at position x, scaled by the number of particles, such that $\int_{\omega \subset \Omega_P} p(x,t)\,\mathrm{d}x$ is the expected number of particles in a subset, ω, of Ω_P.

Our first algorithm (which is appropriate for situations in which the copy numbers are low) progresses asynchronously. The compartment-based regime is updated in an event-driven manner whilst the PDE domain is updated in a time-driven manner. We use the model design shown in Figure 1 and describe here how to couple the transition of particles between the PDE and compartment-based regimes using the pseudo-compartment, C_{-1}.

Particle numbers in region Ω_C are updated according to jump propensities given in equation (1). We extend the definition of these jump propensities to include transitions to and from the pseudo-compartment, C_{-1}, that lies inside the PDE domain. When a particle is chosen to jump left from compartment C_{-1} in Ω_C, into the pseudo-compartment, C_{-1}, its mass is incorporated into the PDE solution. The expected number of particles within the pseudo-compartment is increased by one by adding a step function of amplitude $1/h$ to the PDE distribution in $[-h, 0)$.

In Ω_C, the propensity for a jump from compartment to neighbouring compartment is given by equation (1). We modify the propensity to jump from the pseudo-compartment, C_{-1}, into its neighbouring compartment, C_1, in Ω_C, in order to take account of the expected number of particles in C_{-1}. The jump propensity from C_{-1} to C_1 is therefore given by

$$\alpha^* = \frac{D}{h^2} A_{-1}, \tag{4}$$

where $A_{-1} = \int_{C_{-1}} p(x,t)\mathrm{d}x$ is the expected number of particles in the pseudo-compartment. For convenience in what follows we denote the normalised number of particles in C_{-1} as $a_{-1} = A_{-1}/N_P$.

Jumps from C_{-1} to C_1 occur as compartment-based events in the algorithm. This means that in each small time interval $(t, t + dt_p)$ in which the PDE is updated, none of the PDE-based particles can jump into Ω_C. Since $p(x,t)$ is the scaled probability distribution describing the probability of finding these particles in the PDE regime, the new distribution must be conditioned on the event that none of these particles jumped into the compartment-based regime during the PDE update interval $(t, t + dt_p)$. After running the PDE solver on the distribution $p(x,t)$ a change of $\delta_p(x,t)$ must be made to reflect the fact that none of the remaining N_P particles transferred into C_1 via C_{-1}.

In the small time step dt_p each of the expected A_{-1} particles inside the pseudo-compartment should

be capable of jumping over the interface with a probability of $P_{\text{jump}} = D dt_p / h^2$. By definition

$$
\begin{aligned}
\delta_p(x,t) &= N_P \left(Pr(x,t|\text{no jump}) - Pr(x,t) \right) \\
&= \frac{Pr(\text{no jump}|x,t)p(x,t)}{Pr(\text{no jump})} - p(x,t) \\
&= \frac{Pr(\text{no jump}|x,t)p(x,t)}{1 - P_{\text{jump}}a_{-1} + O(dt_p^2)} - p(x,t),
\end{aligned} \tag{5}
$$

where Pr indicates the probability for a single particle. $Pr(\text{no jump}|x,t)$ depends on whether the particle in question is inside or outside the pseudo-compartment, C_{-1}:

$$
Pr(\text{no jump}|x,t) = \begin{cases} 1, & x \notin C_{-1} \\ 1 - P_{\text{jump}}, & x \in C_{-1} \end{cases}. \tag{6}
$$

Therefore,

$$
\delta_p(x,t) \approx \begin{cases} \dfrac{p(x,t)P_{\text{jump}}a_{-1}}{1 - P_{\text{jump}}a_{-1}}, & x \notin C_{-1} \\[4mm] \dfrac{p(x,t)P_{\text{jump}}(a_{-1} - 1)}{1 - P_{\text{jump}}a_{-1}}, & x \in C_{-1} \end{cases}. \tag{7}
$$

As $dt_p \to 0$, $P_{\text{jump}} \to 0$ and we find that

$$
\lim_{dt_p \to 0} \frac{\delta_p(x,t)}{dt_p} = \begin{cases} \dfrac{Da_{-1}}{h^2}p(x,t), & x \notin C_{-1} \\[4mm] \dfrac{Da_{-1}}{h^2}p(x,t) - \dfrac{D}{h^2}p(x,t), & x \in C_{-1} \end{cases}. \tag{8}
$$

The PDE is therefore solved with an additional source term of size $p(x,t)Da_{-1}/h^2$, throughout the domain and a sink term inside the pseudo-compartment of size $-p(x,t)D/h^2$. These additional terms ensure that the probability distribution inside the PDE domain redistributes correctly as a result of not finding pseudo-compartment particles available for jumping in a given time step.

Changes in the compartment-based regime have an impact on the PDE only when particles cross the interface. We have already discussed the change to the relatively straightforward step-function addition of density to the PDE when particles cross from C_1 to C_{-1}. However, when particles cross from C_{-1} to C_1 it would be erroneous to simply subtract a step function of amplitude $1/h$ from the PDE distribution (reversing the operation for the transfer of particles in the opposing direction). This is because we assume that each of the particles in the PDE region has the same distribution function $p(x,t)/N_P$. Therefore, taking any one of these N_P identical particles across the boundary into Ω_C will result only in a rescaling of this distribution. Recognising that we lose a single particle's worth of mass, the distribution $p(x,t)$ in Ω_P should now represent the probability distribution of $N_P - 1$ identical particles so we can ca lculate the rescaled distribution as

$$
p_{new}(x,t) = \frac{N_P - 1}{N_P}p(x,t).
$$

A cartoon schematic illustrating the algorithm is given in Figure 2. The algorithm is also outlined in detail in Table 1.

In Algorithm 1, changes to the PDE solution are as a result of diffusion, auxiliary source/sink terms for changes in conditional probability, events where a particle moves from the pseudo-compartment, C_{-1}, to the first compartment, C_1, in the compartment-based regime, Ω_C, and events where a particle moves in
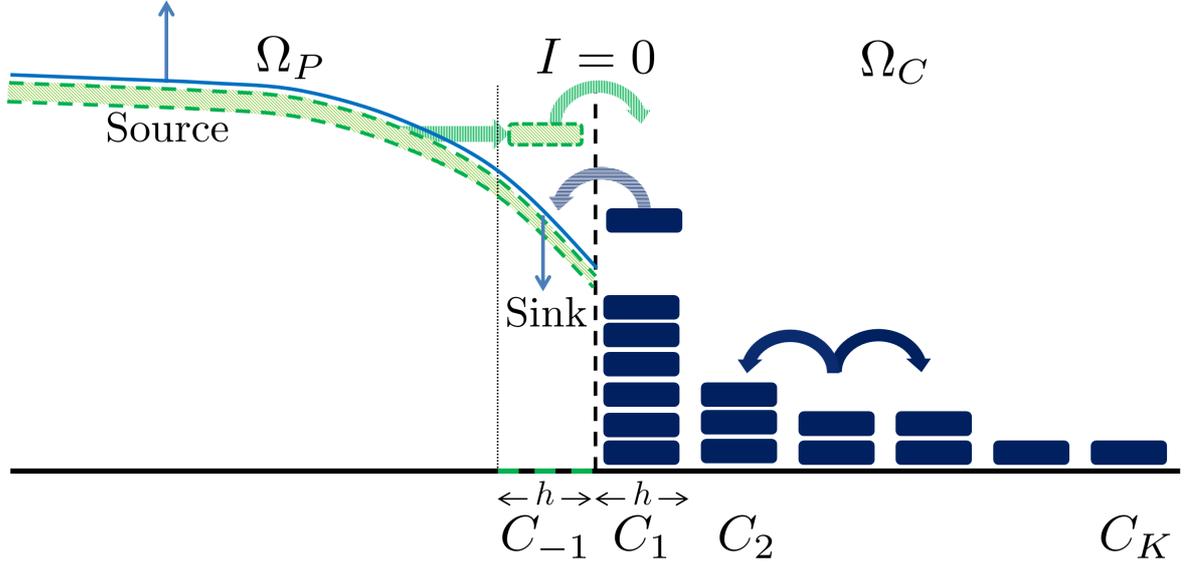
Figure 2. A cartoon schematic illustrating the most important features of Algorithm 1 in one dimension. In Ω_C, to the right of the interface, particles are represented as individuals in compartments (dark blue rectangles) with each individual adding a step function to the density in that compartment. Particles can jump to neighbouring compartments within Ω_C (dark blue block arrows). In Ω_P, to the left of the interface, density is represented by a 'continuous' function (light blue curve). Particles in the left-most compartment in Ω_C (adjacent to the interface) can jump into the PDE regime (blue horizontally striped arrow) where they are added as a step function to the PDE solution in the region $[-h, 0)$ - the pseudo-compartment. The PDE does not include information about the specific location of individuals and so when a particle jumps from the pseudo-compartment into Ω_C it must first be realised in the pseudo-compartment by subtraction of its contribution to the whole PDE followed by an instantaneous jump over the interface. Every instance in which a particle does not jump from the pseudo-compartment into Ω_C, the conditional probability to find particles in the distribution must necessarily be adjusted so that more weight lies outside of the pseudo-compartment. This effect results in a net source of distribution outside of the pseudo-compartment in the PDE and a balancing net sink of distribution inside the pseudo-compartment such that the total number of particles is conserved in the absence of reactions.

the opposite direction from C_1 to C_{-1}. All of these changes may be represented formally in the following PDE (in one dimension)

$$\frac{\partial p(x,t)}{\partial t} = D\frac{\partial^2 p(x,t)}{\partial x^2} + \left(\frac{N_P^+}{N_P^-} - 1\right)p\sum_J \delta(t - t_J) + \frac{Da_{-1}(t)}{h^2}p - \frac{\mathbb{I}_{C_{-1}}(x)}{h}\left(\frac{D}{h}p - \sum_j \delta(t - t_j)\right), \quad (9)$$

where $\mathbb{I}_B(x) = 1$ if $x \in B$ and zero otherwise, t_j are instants when particles jump from compartment to PDE, t_J are instants when particles jump from PDE to compartment and $N_P^- = N_P(t^-)$ and $N_P^+ = N_P(t^+)$ represent numbers of particles in Ω_P immediately before and after times in which the number discontinuously jumps at respective moments t_J (they can be thought of as left and right limits of $N_p(t)$ respectively). Note that $N_P^+ = N_P^- - 1$. As a common sense check we can integrate over the Ω_P to derive

(i) Initialise the time, $t = t_0$ and set the final time, t_{final}. Specify the PDE-update time-step dt_P and initialise the next PDE time-step to be $t_P = t + dt_P$.

(ii) Initialise the number of particles in each compartment in Ω_C, A_i for $i = 1, \ldots, K$, and the distribution of probability density in Ω_P, $p(x, t_0)$, for $x \in \Omega_P$.

(iii) Calculate the propensity functions for diffusion between the compartment regimes as $\alpha_{i,j} = A_i D / h^2$ for $i = 1 \ldots K$ and $j = M + 1, M + 2$ and for reactions as $\alpha_{i,j}$ for $i = 1 \ldots K$ and $j = 1, \ldots, M$ using the usual mass action kinetics.

(iv) Calculate the propensity function for diffusion from the pseudo-compartment, C_{-1}, in Ω_P into the adjacent compartment, C_1, in Ω_C: $\alpha^* = D \int_{C_{-1}} p(x, t) \mathrm{d}x / h^2$ (equation (4)).

(v) Calculate the sum of the propensity functions, $\alpha_0 = \sum_{i=1}^{K} \sum_{j=1}^{M+2} \alpha_{i,j} + \alpha^*$.

(vi) Determine the time for the next 'compartment-based' event $t_C = t + \tau$, where τ is given by equation (2).

(vii) If $t_C < t_P$ then the next compartment-based event occurs:

   (a) Determine which event occurs. Each event occurs with probability proportional to its propensity function (see Gillespie (1977)).

   (b) If the event corresponds to $\alpha_{i,j}$ for $i = 1 \ldots K$ and $j = M + 1, M + 2$ then move a particle from interval i in the direction specified by j. If the particle crosses the interface into pseudo-compartment C_{-1} then add a particle's worth of mass to the region C_{-1} i.e. $p(x, t + \tau) = p(x, t) + \mathbb{I}_{C_{-1}}/h$. Here $\mathbb{I}_{C_{-1}}$ is the indicator function which takes the value 1 in C_{-1} and 0 elsewhere.

   (c) If the event corresponds to propensity function α^* then place a particle in C_1. To remove this particle from the PDE solution, simply rescale the solution i.e. $p(x, t + \tau) = p(x, t)(N_P - 1)/N_P$. Update N_P and N_C to reflect the exchange of particles between regimes.

   (d) Update the current time, $t = t_C$.

(viii) If $t_P < t_C$ the the PDE regime is updated:

   (a) Update the PDE solution according to the numerical method described in Section using $p(x, t)$ as the previous value of the solution.

   (b) Add the auxiliary source and sink terms (see equation (8)) to the numerical solution of the PDE.

   (c) Update the current time, $t = t_P$ and set the time for the next PDE update step to be $t_P = t_P + dt_P$.

(ix) If $t \leq t_{final}$, return to step (iii).

   Else end.

Table 1. Algorithm for the coupling of the compartment-based regime with the PDE-based regime in the case of low copy numbers (Algorithm 1).

an equation for the change in the total number of particles represented by the PDE regime, N_P:

$$\frac{\mathrm{d}N_p}{\mathrm{d}t} = -\sum_J \delta(t - t_J) + \sum_j \delta(t - t_j). \tag{10}$$

As expected, the number of particles changes by integer amounts at the discrete times when particles enter or leave the domain.

High copy number coupling technique: Algorithm 2

The algorithm presented in the previous subsection is accurate for both high and low copy numbers. However, its complexity leads to computational overheads which may prove costly, especially in situations in which copy numbers of species are relatively high and a less sophisticated coupling method would suffice. For this reason we derive the following coupling method (which we refer to as Algorithm 2) as a limiting case of our original algorithm, valid only for the scenario of high copy numbers.

We predict the benefits of these hybrid methods to be greatest for applications that have regions in which there are low copy numbers (requiring individual-based simulation) and other regions of high copy numbers (in which a PDE description will be more efficient). In particular the copy numbers at the interface should be high in order to justify the use of the PDE there. Arguably if the copy numbers are low at the interface then the interface is positioned incorrectly.

In order to derive our simplified algorithm we allow $N_P \to \infty$. This implies that times between transfer events (particles crossing from C_{-1} to C_1 and vice versa) are low: $0 < t_{j+1} - t_j \ll 1$ and $0 < t_{J+1} - t_J \ll 1$. Let us define a small time, δt, such that the number of transfer events which occur between t and $t + \delta t$ is large, but significantly smaller than the total number of particles in the pseudo-compartment. Integrating equation (9) from t to $t + \delta t$ and dividing by δt we find

$$\frac{p(x, t + \delta t) - p(x, t)}{\delta t} = D\frac{\partial^2 p(x,t)}{\partial x^2} - \frac{p}{N_P \delta t}\int_t^{t+\delta t}\sum_J \delta(t' - t_J)\mathrm{d}t' + \frac{Da_{-1}(t)}{h^2}p \tag{11}$$

$$- \frac{D\mathbb{I}_{C_{-1}}(x)}{h^2}p + \frac{\mathbb{I}_{C_{-1}}(x)}{h\delta t}\int_t^{t+\delta t}\sum_j \delta(t' - t_j)\mathrm{d}t'.$$

For each $C_{-1} \to C_1$ transfer event we approximate $(N_P^+/N_P^- - 1) = -1/N_P^- \approx -1/N_P$, since N_P is large. It should be noted that in order to obtain equation (11) from equation (9) one requires δt be sufficiently small that temporally varying dependant variables are constant on $(t, t + \delta t)$, that is $\int_t^{t+\delta t} p(x,t')dt' \approx p(x,t)\delta t$ and similarly for $a_{-1}(t)$ and $1/N_P(t)$. The large number of transfer events that occur between t and $t + \delta t$ mean that the fluctuations in the number of transfer events are negligible and, as such, they can be predicted deterministically using the propensities given in the previous section (equations (1) and (4)):

$$\int_t^{t+\delta t}\sum_J \delta(t' - t_J)\mathrm{d}t' = \frac{D\delta t A_{-1}}{h^2}, \qquad \text{and} \tag{12}$$

$$\int_t^{t+\delta t}\sum_j \delta(t' - t_j)\mathrm{d}t' = \frac{D\delta t A_1}{h^2}, \tag{13}$$

Substituting equation (12) into equation (11) and recalling that $A_{-1} = a_{-1}N_P$, we find the second and third terms on the right-hand side of equation (11) cancel each other. Consequently, we see that changes in the PDE as a result of hybridisation with the discrete subdomain (i.e. not including the Laplacian

diffusion operator) are confined to within the pseudo-compartment.

$$\frac{p(x,t+\delta t) - p(x,t)}{\delta t} = D\frac{\partial^2 p(x,t)}{\partial x^2} - \frac{D\mathbb{I}_{C_{-1}}(x)}{h^2}p + \frac{\mathbb{I}_{C_{-1}}(x)}{h\delta t}\int_t^{t+\delta t}\sum_j \delta(t'-t_j)\mathrm{d}t'. \tag{14}$$

Substituting equation (12) into the second term on the right hand side of equation (14) we obtain (correct to $O(h^2)$ accuracy; as accurate as the compartment-based simulation)

$$\frac{p(x,t+\delta t) - p(x,t)}{\delta t} = D\frac{\partial^2 p(x,t)}{\partial x^2} + \frac{\mathbb{I}_{C_{-1}}(x)}{h\delta t}\int_t^{t+\delta t}\sum_j \delta(t'-t_j) - \sum_J \delta(t'-t_J)\mathrm{d}t'. \tag{15}$$

In order to find equation (15) it should be noted that

$$\mathbb{I}_{C_{-1}}(x)p(x,t) = \mathbb{I}_{C_{-1}}(x)\left(p(x_0,t) + p_x(x_0,t)(x-x_0) + \ldots\right) = \frac{\mathbb{I}_{C_{-1}}(x)}{h}\left(A_{-1} + O(h^2)\right), \tag{16}$$

where x_0 is at the centre of C_{-1} and $p_x(x_0,t) = O(h)$ (due to the no flux boundary condition placed on the PDE at the interface between the two subdomains). Multiplying equation (15) by δt, differentiating with respect to δt and taking the limit as $\delta t \to 0$ we obtain

$$\frac{\partial p(x,t)}{\partial t} = D\frac{\partial^2 p(x,t)}{\partial x^2} + \frac{\mathbb{I}_{C_{-1}}(x)}{h}\left(\sum_j \delta(t-t_j) - \sum_J \delta(t-t_J)\right). \tag{17}$$

If the number of particles within the pseudo-compartment, A_{-1}, becomes large then it will have a small relative variance and hence can effectively be assumed to be deterministic (albeit not necessarily integer valued).

Equation (17) is the motivating representation for Algorithm 2, which is valid for large N_P. Transfer events from Ω_P and Ω_C are calculated in the same way as Algorithm 1, however, their effect on the PDE domain is different. The source and sink terms in Algorithm 1, required for conditional probability changes in the PDE distribution as a result of periods of 'no jump', are no longer part of Algorithm 2. In Algorithm 2, as particles transfer from C_{-1} to Ω_C particles are withdrawn from the PDE in step functions from the pseudo-compartment in the same way that they are added when particles transfer from from Ω_C to C_{-1} (as indicated in equation (17)).

Unsurprisingly, for large N_P we obtain a near deterministic hybrid simulation in both subdomains. Using equations (12) and (13), equation (15) becomes

$$\frac{\partial p(x,t)}{\partial t} = D\frac{\partial^2 p(x,t)}{\partial x^2} + \frac{\mathbb{I}_{C_{-1}}(x)D}{h^2}\left(A_1 - A_{-1}\right). \tag{18}$$

The last two terms in equation (18) represents a lattice approximation to a free diffusive flux over the interface, a desired outcome of Algorithm 2.

A cartoon schematic illustrating the simplified algorithm for large copy numbers (Algorithm 2) in one dimension is given in Figure 3.

Table 2 describes the algorithm used to interface the two regimes. Note that when removing a pseudo-particle from the PDE regime in step (vii) (c) we enforce the condition that $p(x,t) > 1/h$ for all $x \in C_{-1}$ to ensure that when we remove a step-function corresponding to a particle's worth of mass that we do not make the PDE solution negative. In practise this condition will rarely be enforced since the PDE region should be used to describe regions of the domain that have sufficiently high density that the individual-based model becomes inefficient. We note, however, that with a fixed interface this will not always be possible and that this restriction may lead to slight inaccuracies in the solution.

We have shown that, theoretically, Algorithm 1 simplifies to Algorithm 2 in the case of large N_P. In the following section, we will demonstrate this empirically on a number of simple test problems. We will also demonstrate the strong agreement between the hybrid simulations of both algorithms and known exact analytic solutions.
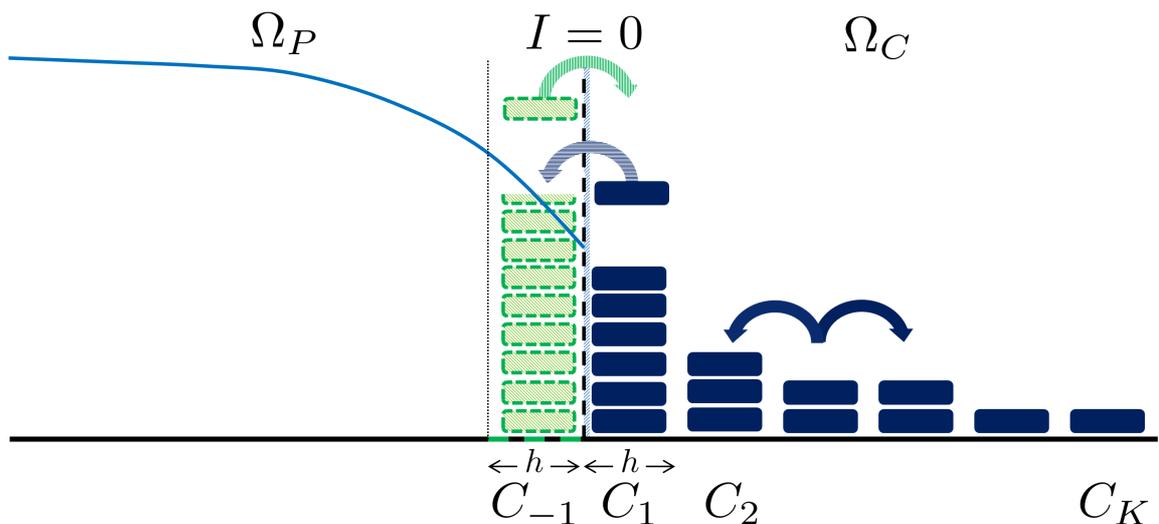
Figure 3. A cartoon schematic illustrating the most important features of Algorithm 2 in one dimension. In Ω_C, to the right of the interface, particles are represented as individuals in compartments (dark blue rectangles) with each individual adding a step function to the density in that compartment. Particles can jump to neighbouring compartments within Ω_C (dark blue block arrows). In Ω_P, to the left of the interface, density is represented by a 'continuous' function (light blue curve). Particles in the left-most compartment in Ω_C (adjacent to the interface) can jump into the PDE regime (blue horizontally striped arrow) where they are added as a step function to the PDE solution in the region $[-h, 0)$ - the pseudo-compartment (green dashed line). Pseudo-particles (hatched green rectangles with dashed borders) corresponding to the mass of the PDE in the pseudo-compartment are allowed to jump right (green vertically striped arrow) into the first compartment in Ω_C. When a particle leaves the pseudo-compartment a step function of mass corresponding to one particle is removed from the PDE in that region.

Results

In this section we assesses the accuracy of our proposed algorithms in simulating three straightforward test problems of increasing complexity.

(i) Initialise the time, $t = t_0$ and set the final time, t_{final}. Specify the PDE-update time-step dt_P and initialise the next PDE time-step to be $t_P = t + dt_P$.

(ii) Initialise the number of particles in each compartment in Ω_C, A_i for $i = 1, \ldots, K$, and the distribution of probability density in Ω_P, $p(x, t_0)$, for $x \in \Omega_P$.

(iii) Calculate the propensity functions for diffusion between the compartment regimes as $\alpha_{i,j} = A_i D/h^2$ for $i = 1 \ldots K$ and $j = M + 1, M + 2$ and for reactions as $\alpha_{i,j}$ for $i = 1 \ldots K$ and $j = 1, \ldots, M$ using the usual mass action kinetics.

(iv) Calculate the propensity function for diffusion from the pseudo-compartment, C_{-1}, in Ω_P into the adjacent compartment in Ω_C: $\alpha^* = D \int_{C_{-1}} p(x, t) \mathrm{d}x / h^2$ (equation (4)).

(v) Calculate the sum of the propensity functions, $\alpha_0 = \sum_{i=1}^{K} \sum_{j=1}^{M+2} \alpha_{i,j} + \alpha^*$.

(vi) Determine the time for the next 'compartment-based' event $t_C = t + \tau$, where τ is given by equation (2).

(vii) If $t_C < t_P$ then the next compartment-based event occurs:

(a) determine which event occurs according to the method described in the text (see Gillespie (1977)).

(b) If the event corresponds to $\alpha_{i,j}$ for $i = 1 \ldots K$ and $j = M + 1, M + 2$ then move a particle from interval i in the direction specified by j. If the particle crosses the interface into pseudo-compartment C_{-1} then add a particle's worth of mass to the region C_{-1} i.e. $p(x, t + \tau) = p(x, t) + \mathbb{I}_{C_{-1}}/h$. Here $\mathbb{I}_{C_{-1}}$ is the indicator function which takes the value 1 in C_{-1} and 0 elsewhere.

(c) If the event corresponds to propensity function α^* and $p(x, t) > 1/h$ for all $x \in C_{-1}$ then place a particle in C_1. Remove a particle's worth of mass from the PDE solution in the region C_{-1} i.e. $p(x, t + \tau) = p(x, t) - \mathbb{I}_{C_{-1}}/h$.

(d) update the current time, $t = t_C$.

(viii) If $t_P < t_C$ the the PDE regime is updated:

(a) Update the PDE solution according to the numerical method described in Section using $p(x, t)$ as the previous value of the solution.

(b) Update the current time, $t = t_P$ and set the time for the next PDE update step to be $t_P = t_P + dt_P$.

(ix) If $t \leq t_{final}$, return to step (iii).
Else end.

Table 2. Algorithm for the coupling of the compartment-based regime with the PDE-based regime in the case of high copy numbers.

Test Problem 1: Uniform particle distribution

In Test Problem 1 we demonstrate that our proposed algorithms are capable of maintaining a steady state distribution of particles. We consider the diffusion of non-interacting particles on a domain with

reflective boundaries which will maintain a steady state distribution. We assume that all N particles are initially uniformly distributed across the domain and move with diffusion coefficient $D = 1/400$, in non-dimensionalised units. The PDE that we solve in Ω_P is the diffusion equation with zero flux boundary conditions at either end.
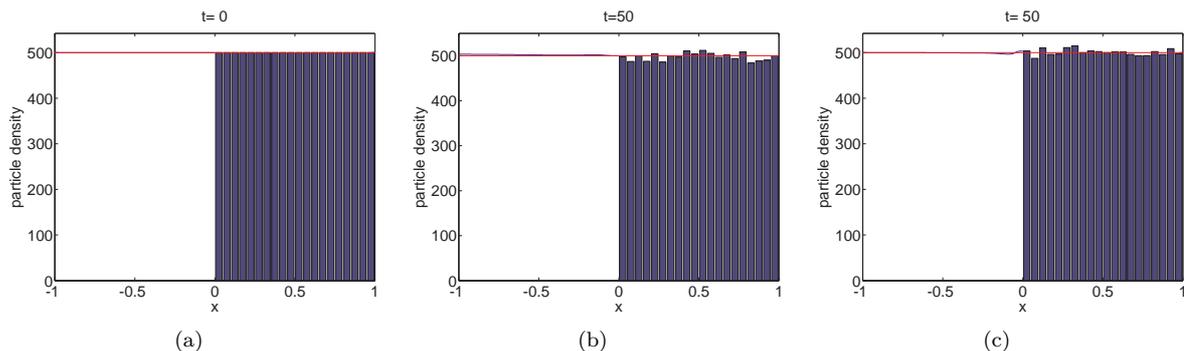

Figure 4. Maintenance of (a) an initially uniform steady state by (b) Algorithm 1 and (c) Algorithm 2. In each subfigure the blue curve is the solution of the PDE in Ω_P, the blue histograms represent the concentration of particles in each of the compartments in Ω_C and the red curve represents the analytic solution of the mean-field model. Note that the blue curve is difficult to see since, for the most-part, it lies underneath the analytic solution. The number of particles is $N = 1000$. The simulation results are averaged over 100 repeats.

Figure 4 demonstrates that a uniform steady state is maintained by both algorithms. The disparity between the PDE solution and the analytic solution close to the interface in Figure 4 (c) is transient and oscillates either side of the analytic solution as time progresses (see Supplementary Material (SM) Movies M1 and M2 corresponding to Algorithms 1 and 2, respectively).

Test Problem 2: Particle redistribution

As a further example of the propriety of our algorithms we consider a situation in which particles are initially uniformly distributed in either Ω_P or in Ω_C with no particles in the complementary region. We note that the situation in which all the particles are in Ω_C and hence the concentration in the compartment-based region is much higher than that in the PDE region is not the ideal situation in which to employ our algorithm. However, it is important to show that our algorithm is robust to such situations. In physical terms each of these simulations corresponds to particles, artificially kept in one half of a box, being allowed to diffuse into the other half of the box.

The mean-field model corresponding to this situation can be formulated as an initial-boundary-value problem:

$$\frac{\partial \rho}{\partial t} = D\frac{\partial^2 \rho}{\partial x^2}, \text{ for } x \in [-1, 1], \tag{19}$$

with zero-flux boundary conditions

$$\frac{\partial \rho}{\partial x}\bigg|_{x=-1,1} = 0, \tag{20}$$

and initial conditions

$$\rho(x, 0) = H(-x) \text{ for } x \in [-1, 1], \tag{21}$$

or

$$\rho(x,0) = H(x) \text{ for } x \in [-1,1], \tag{22}$$

depending on which side of the box we wish to initialise the particles. This equation can be solved using separation of variables and Fourier series. For initial condition (21) the solution is given by:

$$\rho(x,t) = \frac{1}{2} - \sum_{n=1}^{\infty} \frac{2}{(2n-1)\pi}(-1)^n \cos\left((2n-1)\pi\left(\frac{x+1}{2}\right)\right) \exp\left(-\frac{(2n-1)^2\pi^2 Dt}{4}\right). \tag{23}$$

For initial condition (22) the minus in front of the sum is replaced with a plus.

In Figure 5 we compare the analytically derived density to the density produced using our algorithms given initial condition (21). We see a favourable comparison between the solution according to our algorithms and the mean-field solution. A comparison of the densities through time for Algorithm 1 and 2 can be seen in Movies M3 and M4 of the SM, respectively. In Movies M5 and M6 of the SM we present the same comparison but for a single realisation of Algorithms 1 and 2 respectively. In Figure 1 of the SM we compare the particle density of the two algorithms against the analytic solution of the mean-field model for initial condition (22). The comparison is again favourable.
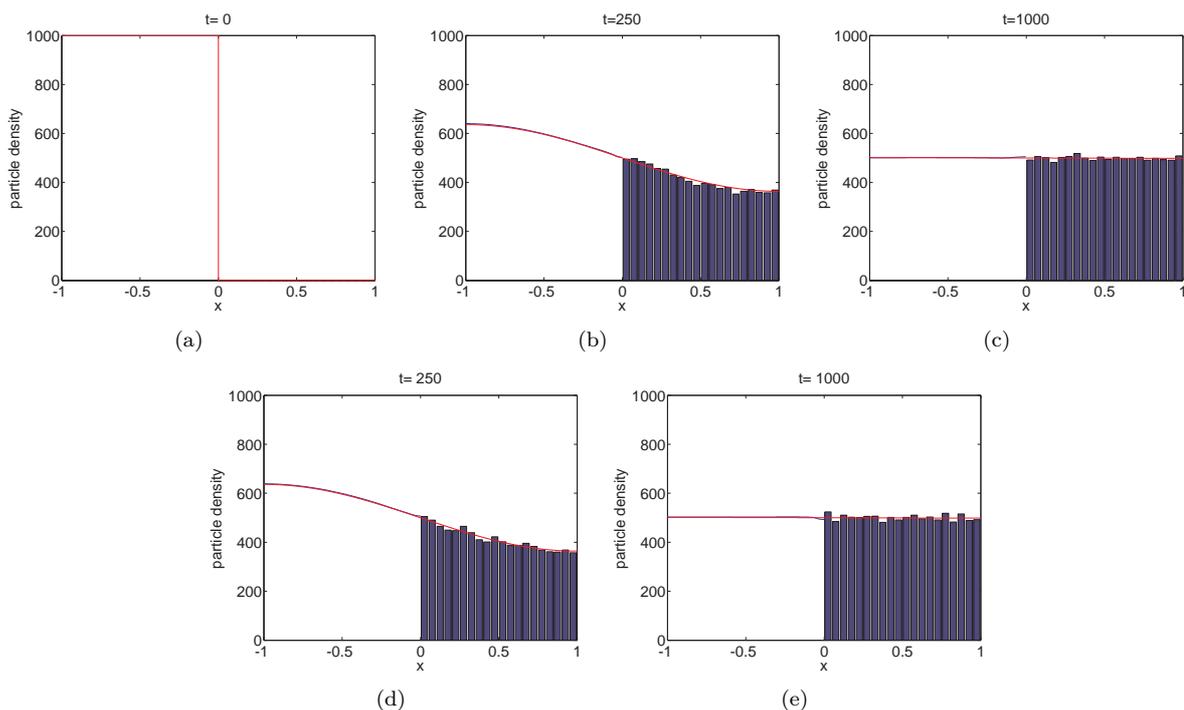

Figure 5. Particles initially in the left-hand half of the domain (a) spread diffusively (with diffusion coefficient $D = 1/400$) throughout the domain eventually reaching steady state. By $t = 1000$ the system is assumed to have reached steady state. Panels (b) and (c) show the particle density profile according to Algorithm 1 at times 250 and 1000 respectively. Panels (d) and (e) show the same comparison for Algorithm 2. Figure descriptions and particle numbers are as in Fig 4. The simulation results are averaged over 100 repeats.

Test Problem 3: A morphogen gradient formation model

In this final test problem we model the formation of a morphogen gradient by considering basic first order reactions. Morphogen is produced at $x = -1$ with rate $D\lambda$. We implement this production as a flux boundary condition in the PDE region of the model:

$$\frac{\partial p}{\partial x}\Big|_{x=-1} = -\lambda. \tag{24}$$

We also allow morphogen to decay uniformly across the domain with rate μ and we implement a zero-flux boundary condition at $x = 1$. These zeroth order (production) and first order (degradation) reactions allow for the description of the mean dynamics of the fully stochastic model by a PDE:

$$\frac{\partial \rho}{\partial t} = D\frac{\partial^2 \rho}{\partial x^2} - \mu\rho. \tag{25}$$

This is the mean-field PDE which we expect the density to satisfy across the whole domain. The analytic solution for the expected evolution of the density for this system (with boundary condition analogous to equation (24)) can therefore be found explicitly as

$$\rho(x,t) = \rho_{st}(x) + \rho_u(x,t), \tag{26}$$

where

$$\rho_{st}(x) = \lambda\sqrt{\frac{D}{\mu}}\frac{\cosh\left(\sqrt{\mu/D}(1-x)\right)}{\cosh\left(2\sqrt{\mu/D}\right)}, \tag{27}$$

gives the steady state of the morphogen gradient and

$$\rho_u(x,t) = -\frac{\lambda D}{2\mu}\tanh\left(2\sqrt{\frac{\mu}{D}}\right)\left\{\exp(-\mu t) + 2\sum_{n=1}^{\infty}\frac{\cos(n\pi(x+1)/2)}{1 + D(n\pi)^2/4\mu}\exp\left(-\left(\frac{D(n\pi)^2}{4} + \mu\right)t\right)\right\}. \tag{28}$$

Since equation (25) is the PDE we would expect the mean-field density of the fully stochastic model to satisfy, we use this as the PDE we solve in Ω_P in our hybrid model. In Ω_C we implement the usual diffusive 'reactions' and additional mono-molecular decay reactions in each compartment and a zero flux boundary condition at $x = 1$.

In Figure 6 we compare the analytically derived density to that produced using our algorithms given an initially empty domain. We increase the value of the diffusion coefficient by a factor of 10 (in comparison to previous simulations) to $D = 10 \times h^2 = 1/40$ (in order to achieve a sensible steady state profile). We choose $\lambda = 1000$ and $\mu = 0.1$, giving a rate of influx of particles into the domain of $D\lambda = 25$. As in previous examples, we see a favourable comparison between the solution according to our hybrid algorithms and the expected mean-field solution given by equations (26)-(28). A comparison of the densities through time for Algorithms 1 and 2 can be seen in Movies M7 and M8 of the SM, respectively.

A more quantitative study of the similarities between the solution of our algorithms and the analytic solutions is presented in the next section.

Error analysis

In this section we revisit the comparison between the mean particle densities predicted by our hybrid algorithm and the expected values given by the corresponding deterministic model. It is our aim in developing our algorithms that the error between the hybrid methods and the deterministic solution should be of the same order of magnitude as the accuracy associated with the individual numerical
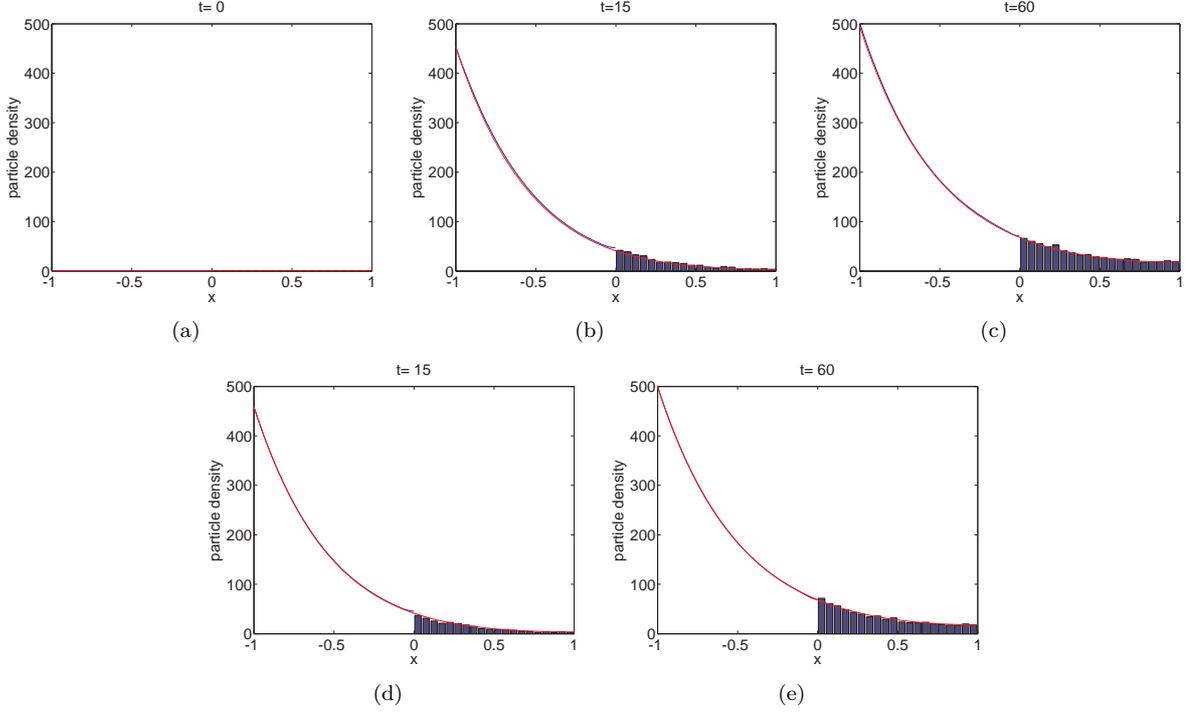
Figure 6. A morphogen gradient forms though a combination of influx at the left-hand boundary and decay of morphogen particles throughout the domain. Panel (a) shows the initially empty domain. By $t = 60$ the system is assumed to have reached an approximately steady state. Panels (b) and (c) show the particle density profile according to Algorithm 1 at times 15 and 60 respectively. Panels (d) and (e) show the same comparison for Algorithm 2. Figure descriptions are as in Fig 4. The simulation results are averaged over 100 repeats.

simulation techniques. In particular we developed our algorithms in order that the stochastic model should not 'see' (i.e. be influenced by) the interface. Since it is well established that the descriptions of particle density in Ω_P and Ω_C are valid representations of diffusion, we address the question of 'whether our algorithms can correctly simulate the flux over the interface?' In order to test this we compare

$$N_{DC}(t) = \int_0^1 \rho(x,t)\mathrm{d}x, \tag{29}$$

the expected number of particles in Ω_C in the deterministic model to

$$N_{HC}(t) = \sum_{i=1}^K \bar{A}_i, \tag{30}$$

the expected number of particles in Ω_C in the hybrid models. Here $\bar{A}_i$ represents the mean number of particles in the ith compartment averaged over 100 repeats of the hybrid algorithms. For completeness we also compare

$$N_{DP}(t) = \int_{-1}^0 \rho(x,t)\mathrm{d}x, \tag{31}$$

the expected number of particles in Ω_P in the deterministic model to

$$N_{HP}(t) = \int_{-1}^{0} p(x,t)\mathrm{d}x, \tag{32}$$

the expected number of particles in Ω_P in the hybrid models.

The results for test problem 1 are shown in Figure 7. In simulations generated by our hybrid Algorithm 2, the mass in the PDE-based regime is very similar to the expected mass given by the uniform solution of the diffusion equation. The same is true of the mass in the compartment-based regimes. Results are quantitatively indistinguishable (excepting for expected stochastic variation) for Algorithm 1, although for brevity we do not explicitly show the corresponding figures.
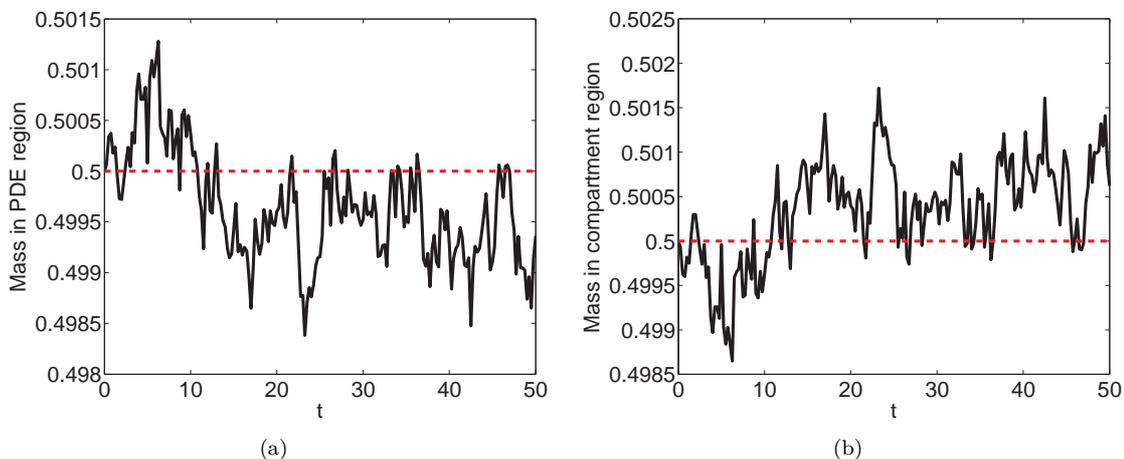

(a)   (b)

Figure 7. Error plots for the uniform particle distribution maintenance example. In panel (a) we display the normalised (i.e. divided by the total number of particles in the domain) expected value of the total mass in Ω_P given by the analytic solution (red, dashed line, $N_{DP}(t)/(N_{DP} + N_{DC}(t))$) and the normalised average mass in Ω_P according to our hybrid Algorithm 2 (black line, $N_{HP}(t)/(N_{HP} + N_{HC}(t))$). In panel 7(b) we display the same comparison but for the mass in the compartment-based regime (i.e. $N_{DC}(t)/(N_{DP} + N_{DC}(t))$ (red, dashed line) and $N_{HC}(t)/(N_{HP} + N_{HC}(t))$ (black line)). The masses generated by our hybrid algorithm are averaged over 100 repeats as in previous density comparison figures.

In Figure 8 we display similar comparisons for test problem 2, in which the particles are initialised uniformly in Ω_P only. As the simulation evolves we see that the expected mass predicted by the hybrid algorithm matches closely the expected mass given by the analytic solution in both Ω_P and Ω_C. Further, the insets demonstrate that the relative error between the masses generated by the analytic solution and the hybrid model are low and have no systematic bias, but instead fluctuate about their expected values. Similar results hold for the initial condition in which the particles are initialised uniformly in Ω_C only (see Figure 2 of the SM). As in Figure 7, since the results of the simulations of our two hybrid algorithms are quantitatively similar, we present only the simulation results from our second hybrid algorithm.

Finally in Figure 9 we display similar comparisons between the mass in each of the regimes for *both* of the hybrid algorithms we have developed when simulating the formation of a morphogen gradient (test problem 3). Again, a comparison of the results from the hybrid algorithm and the analytic solution shows good agreement, especially for long times. However, there is a clear disparity between the mass in Ω_C for Algorithm 2 and the expected mass in that region (see Figure 9 (d)). In particular the hybrid algorithm
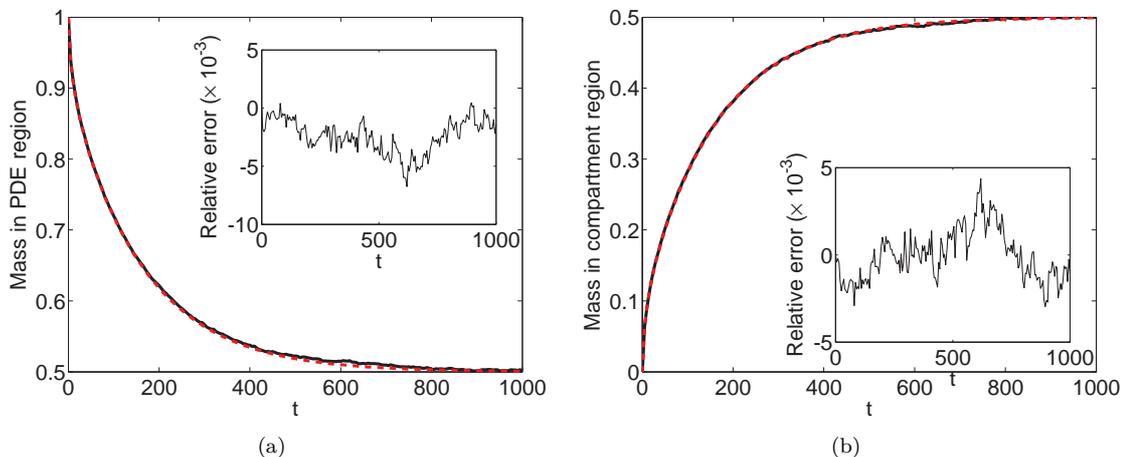
Figure 8. Error plots for test problem 2, in which all particles are initially in Ω_P. In (a) we display the normalised expected value of the total mass in Ω_P given by the analytic solution (red, dashed line, $N_{DP}/(N_{DP} + N_{DC})$ and the normalised average mass in Ω_P (black line, $N_H P/(N_{DP} + N_{DC})$ according to our second hybrid algorithm. In (b) we display the same comparison but for the compartment-based regime. In each panel, the inset displays the evolution of the relative error (i.e. $(N_{HP} - N_{DP})/N_{DP}$ or $(N_{HC} - N_{DC})/N_{DC}$, respectively) throughout the simulation. Again, the masses generated by our hybrid algorithm are averaged over 100 repeats.

under-represents the expected mass in the compartment-based regime for early times. Although it is not clearly visible in Figure 9 (c) due to the scale of the axes, there is a corresponding over-representation of the expected mass in Ω_P at early times for Algorithm 2. This is not the case for Algorithm 1 (see Figures 9 (a) and (b)).

The reason for this disparity is a technical one. Recall that in Algorithm 2, in order to avoid negative solution values, a particle is not allowed to be removed from the pseudo-compartment, C_{-1}, until the PDE solution across that compartment has value at least $1/h$. This avoids the possibility that the removal of a pseudo particle may send the value of the PDE solution in C_{-1} negative. As the domain is initially empty and fills from the PDE region there is a significant period of time in which mass should be flowing over the interface, but is restricted from doing so by an artificial barrier in our approximate Algorithm 2. The flow of particles into Ω_C is, therefore, retarded in comparison to the analytic solution for which, of course, no such barrier exists. In Algorithm 1, however, we see no such problem (*cf* Figure 9 (b)), because mass which is moved across the interface is removed from the whole of Ω_P, rather than just the pseudo-compartment. This justifies our assertion that Algorithm 2 is suitable for simulations in which the number of particles at the interface are high, whereas Algorithm 1 is suitable for situations in which there may be low copy numbers at the interface. We explore the ramifications of these results further in the discussion.

We also analysed the error in each of the algorithms as we decrease the value of h, the compartment width in Ω_C. The order of magnitude of the error was was indistinguishable from the error associated with the compartment-based model.
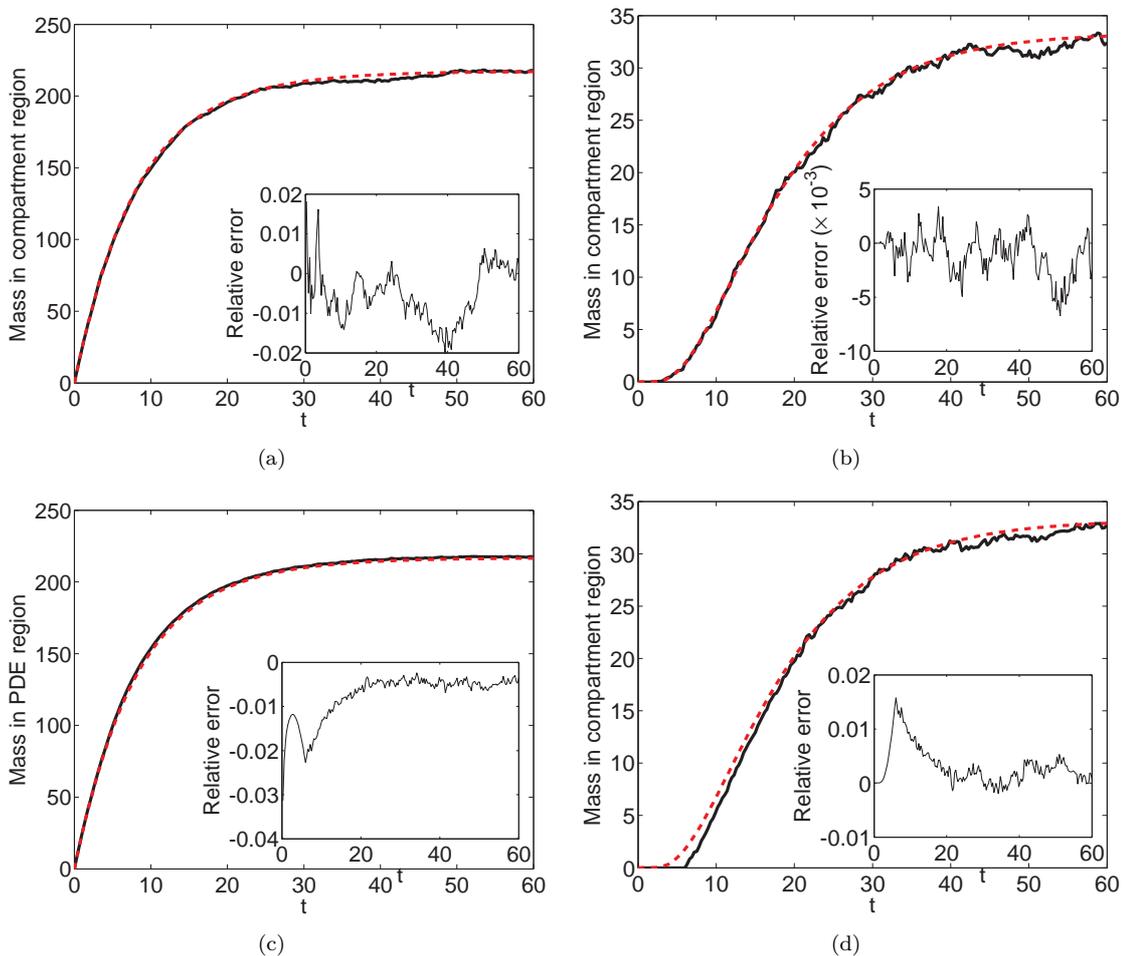
Figure 9. Error plots for the morphogen gradient formation example. Figure descriptions are as in Figure 8

Discussion

We have introduced two algorithms which allow a coupling between a continuum PDE description of particle density in one region of the domain and a discrete stochastic compartment-based description in another. Using a *pseudo-compartment* in the PDE regime, which has properties of both descriptions, we were able to couple the two simulation methodologies together in such a way that the expected flux across the interface is maintained. We evidenced this though a series of examples, demonstrating a good correspondence between the expected mass in each of the modelling regions (averaged over several repeats of the hybrid simulation algorithms) and the mass in those regions determined by the corresponding fully-deterministic mean-field model.

Our method relies on a knowledge of the deterministic PDE description to which the mean of the compartment-based simulations corresponds. Whilst this is easy to obtain in the examples we have studied, when higher order reactions are introduced to the stochastic model there is no longer an exact closed-form representation of the expected particle density as a PDE in terms of the first moment alone.

In order to obtain a self-contained equation we must make moment closure assumptions, which means that the PDE description will no longer correspond exactly to the mean behaviour of the compartment-based system. However, in situations where particle numbers are sufficiently large the closed PDE description will often be a close match to the expected densities of the compartment-based model and, as such, we expect that our algorithms will still be applicable. Furthermore, large particle numbers are precisely the regime in which we want to make use of the PDE description, further improving the applicability of our algorithms.

It is precisely this consideration which will provide motivation for our next work on this subject to develop a hybrid method with an adaptive interface. In order to make the hybrid algorithm as efficient as possible, the interface should be able to move so that only regions with sufficiently low concentrations are represented by the compartment-based model and regions with high concentration are consistently represented with the PDE-based model. This will increase the computational efficiency of the algorithm, since we will no longer be simulating the individual particle movements in the discrete model when the continuum description can be tolerated. There is also scope to interface the algorithms we have developed in this manuscript with algorithms designed to couple compartment-based models and molecular-based models (Flegg et al., 2012, 2014) in order to build a "three-regime method". Another important extension of this work is to consider compartment-based models on irregular lattices which will allow us to capture more complex and realistic domain geometries.

For the sake of clarity in this manuscript, we presented the algorithms themselves and the illustrative examples in a single dimension, however, the algorithm is easily adapted for higher dimensional problems with (hyper)-planar interfaces. We therefore expect our algorithms to be highly applicable to stochastic simulations of systems in which molecular populations are high in some regions and low in others, meaning that neither a fully-deterministic-continuum nor a fully-stochastic-individual-based model is appropriate. Our algorithms provide significant increases in speed in comparison to fully stochastic models and more accurate and representative modelling in comparison to fully-deterministic models.

References

G. Adomian. The diffusion-Brusselator equation. *Comput. Math. Appl.*, 29(5):1–3, 1995.

F.J. Alexander, A.L. Garcia, and D.M. Tartakovsky. Algorithm refinement for stochastic partial differential equations: I. linear diffusion. *J. Comput. Phys.*, 182(1):47–66, 2002.

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

R.E. Baker, C.A. Yates, and R. Erban. From microscopic to macroscopic descriptions of cell migration on growing domains. *Bull. Math. Biol.*, 72(3):719–762, 2010.

T. Bollenbach, K. Kruse, P. Pantazis, M. González-Gaitán, and F. Jülicher. Robust formation of morphogen gradients. *Phys. Rev. Lett.*, 94(1):018103:1–4, 2005.

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

B. Drawert, M.J. Lawson, L. Petzold, and M. Khammash. The diffusive finite state projection algorithm for efficient simulation of the stochastic reaction-diffusion master equation. *J. Chem. Phys.*, 132(7): 074101, 2010.

B. Drawert, S. Engblom, and A. Hellander. URDME: a modular framework for stochastic simulation of reaction-transport processes in complex geometries. *BMC Syst. Biol.*, 6(1):76, 2012.

J. Elf and M. Ehrenberg. Spontaneous separation of bi-stable biochemical systems into spatial domains of opposite phases. *Syst. Biol.*, 1(2):230–236, 2004.

M.B. Elowitz, A.J. Levine, E.D. Siggia, and P.S. Swain. Stochastic gene expression in a single cell. *Sci. Signal.*, 297(5584):1183, 2002.

S. Engblom, L. Ferm, A. Hellander, and P. Lötstedt. Simulation of stochastic reaction-diffusion processes on unstructured meshes. *SIAM J. Sci. Comput.*, 31(3):1774–1797, 2009.

R. Erban and S.J. Chapman. Stochastic modelling of reaction–diffusion processes: algorithms for bimolecular reactions. *Phys. Biol.*, 6(4):1–18, 2009.

R. Erban, S.J. Chapman, and P. Maini. A practical guide to stochastic simulations of reaction-diffusion processes. *arXiv preprint arXiv:0704.1908*, 2007.

R. Erban, M.B. Flegg, and G.A. Papoian. Multiscale stochastic reaction-diffusion modelling: application to actin dynamics in filopodia. *preprint*, 2013.

D. Fange, O.G. Berg, P. Sjöberg, and J. Elf. Stochastic reaction-diffusion kinetics in the microscopic limit. *Proc. Natl. Acad. Sci. USA*, 107(45), 2010.

L. Ferm, A. Hellander, and P. Lötstedt. An adaptive algorithm for simulation of stochastic reaction-diffusion processes. *J. Comput. Phys.*, 229(2):343–360, 2010.

R.J. Field and R.M. Noyes. Oscillations in chemical systems. iv. limit cycle behavior in a model of a real chemical reaction. *J. Chem. Phys.*, 60:1877, 1974.

M.B. Flegg, S.J. Chapman, and R. Erban. The two-regime method for optimizing stochastic reaction–diffusion simulations. *J. Roy. Soc. Interface.*, 9(70):859–868, 2012.

M.B. Flegg, S. Hellander, and R. Erban. Convergence of methods for coupling of microscopic and mesoscopic reaction-diffusion simulations. *arXiv preprint arXiv:1304.7393*, 2014.

E.G. Flekkøy, J. Feder, and G. Wagner. Coupling particles and fields in a diffusive hybrid model. *Phys. Rev. E*, 64(6):066302, 2001.

B. Franz, M.B. Flegg, S.J. Chapman, and R. Erban. Multiscale reaction-diffusion algorithms: Pde-assisted brownian dynamics. *SIAM J. Appl. Math.*, 2012.

B. Franz, C. Xue, K.J. Painter, and R. Erban. Travelling waves in hybrid chemotaxis models. *Bull. Math. Biol.*, 2013.

S. Ghosh, P. Ghosh, K. Basu, and S.K. Das. Modeling the stochastic dynamics of gene expression in single cells: A birth and death markov chain analysis. In *IEEE International Conference on Bioinformatics and Biomedicine*, pages 308–316. IEEE, 2007.

M.A. Gibson and J. Bruck. Efficient exact stochastic simulation of chemical systems with many species and many channels. *J. Phys. Chem. A.*, 104(9):1876–1889, 2000.

D.T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.*, 22(4):403–434, 1976.

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

D. Gonze, J. Halloy, and A. Goldbeter. Robustness of circadian rhythms with respect to molecular noise. *Proc. Natl. Acad. Sci. USA*, 99(2):673, 2002.

J. Hattne, D. Fange, and J. Elf. Stochastic reaction-diffusion simulation with MesoRD. *Method. Biochem. Anal.*, 21(12):2923–2924, 2005.

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

S.A. Isaacson. A convergent reaction-diffusion master equation. *J. Chem. Phys.*, 139(5):054101, 2013.

A. Kicheva, P. Pantazis, T. Bollenbach, Y. Kalaidzidis, T. Bittig, F. Jülicher, and M. González-Gaitán. Kinetics of morphogen gradient formation. *Science*, 315(5811):521–525, 2007.

S. Lampoudi, D.T. Gillespie, and L.R. Petzold. The multinomial simulation algorithm for discrete stochastic simulation of reaction-diffusion systems. *J. Chem. Phys.*, 130(9):094104, 2009.

H. Li and L. Petzold. Logarithmic direct method for discrete stochastic simulation of chemically reacting systems. Technical report, University of California Santa Barbara, 2006.

J.M. McCollum, G.D. Peterson, C.D. Cox, M.L. Simpson, and N.F. Samatova. The sorting direct method for stochastic simulation of biochemical systems with varying reaction execution behavior. *Comput. Biol. Chem.*, 30(1):39–49, 2006.

E. Moro. Hybrid method for simulating front propagation in reaction-diffusion systems. *Phys. Rev. E*, 69(6):060101, 2004.

S. Ramsey, D. Orrell, and H. Bolouri. Dizzy: stochastic simulation of large-scale genetic regulatory networks. *J. Bioinform. Comput. Biol.*, 3(2):415–436, 2005.

M. Robinson, M. Flegg, and R. Erban. Adaptive two-regime method: application to front propagation. 2014.

K.R. Sanft, S. Wu, M. Roh, J. Fu, R.K. Lim, and L.R. Petzold. Stochkit2: software for discrete stochastic simulation of biochemical systems with events. *Bioinformatics*, 27(17):2457–2458, 2011.

J. Schnakenberg. Simple chemical reaction systems with limit cycle behaviour. *J. Theor. Biol.*, 81(3): 389–400, 1979.

M. Smoluchowski. Versuch einer mathematischen theorie der koagulationskinetik kolloider lösungen. *Z. Phys. Chem.*, 92(129-168):9, 1917.

P.S. Swain, M.B. Elowitz, and E.D. Siggia. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc. Natl. Acad. Sci. USA*, 99(20):12795, 2002.

T. Tian and K. Burrage. Bistability and switching in the lysis/lysogeny genetic regulatory network of bacteriophage [lambda]. *J. Theor. Biol.*, 227(2):229–237, 2004.

F. Tostevin, P. R. Ten Wolde, and M. Howard. Fundamental limits to position determination by concentration gradients. *PLoS. Comput. Biol.*, 3(4):e78, 2007.

J.S. van Zon and P.R. ten Wolde. Simulating biochemical networks at the particle level and in time and space: Green's function reaction dynamics. *Phys. Rev. Lett.*, 94(12):128103, 2005.

J.M.G. Vilar, H.Y. Kueh, N. Barkai, and S. Leibler. Mechanisms of noise-resistance in genetic oscillators. *Proc. Natl. Acad. Sci. USA*, 99(9):5988–5992, 2002.

G. Wagner and E.G. Flekkoy. Hybrid computations with flux exchange. *Phil. Trans. R. Soc. A*, 362: 1655–1666, 2004.

L. Wolpert. Positional information and the spatial pattern of cellular differentiation. *J. Theor. Biol.*, 25 (1):1–47, 1969.

T.E. Woolley, R.E. Baker, E.A. Gaffney, and P.K. Maini. Power spectra methods for a stochastic description of diffusion on deterministically growing domains. *Phys. Rev. E*, 84(2):021915, 2011.

C.A. Yates and R.E. Baker. Importance of the voronoi domain partition for position-jump reaction-diffusion processes on non-uniform rectilinear lattices. *Phys. Rev. E*, 88(5):054701, 2013.

C.A. Yates and G. Klingbeil. Recycling random numbers in the stochastic simulation algorithm. *J. Chem. Phys.*, 138(9):094103, 2013.

C.A. Yates, R.E. Baker, R. Erban, and P.K. Maini. Going from microscopic to macroscopic on non-uniform growing domains. *Phys. Rev. E*, 86(2):021921, 2012.