Two Regime Method for optimizing stochastic reaction-diffusion simulations

Mark B. Flegg*,   S. Jonathan Chapman*,   Radek Erban*

April 21, 2011

Abstract. The computer simulation of stochastic reaction-diffusion processes in biology is often done using either compartment-based (spatially discretized) simulations or molecular-based (Brownian dynamics) approaches. Compartment-based approaches can yield quick and accurate mesoscopic results but lack the level of detail that is characteristic of the more computationally intensive molecular-based models. Often microscopic detail is only required in a small region but currently the best way to achieve this detail is to use a resource intensive model over the whole domain. We introduce the *Two Regime Method* (TRM) in which a molecular-based algorithm is used in part of the computational domain and a compartment-based approach is used elsewhere in the computational domain. We apply the TRM to two test problems including a model from developmental biology. We thereby show that the TRM is accurate and subsequently may be used to inspect both mesoscopic and microscopic detail of reaction-diffusion simulations according to the demands of the modeller.

Introduction

Stochastic spatio-temporal simulations have been successfully used in a number of biological applications, including models of morphogen gradients [1], MAPK pathway [2], signal trasduction in E. Coli chemotaxis [3] and oscillation of Min proteins in cell division [4]. However, detailed stochastic spatio-temporal models are often computationally intensive to simulate. This is one of the reasons why whole cell simulation has been recognised as a "grand challenge of the 21st century" [5]. In this paper, we address this problem using a modelling approach which has computational complexity only in regions of interest.

Spatio-temporal biological processes are often modelled using deterministic reaction-diffusion models which are written in the form of partial differential equations (PDEs) for concentrations of biochemical species [6]. However, cellular or subcellular processes usually take place on very small spatial scales. With such small scales it is not uncommon for populations of biochemical species to be so small that deterministic (PDE-based) approaches are completely inappropriate. Several stochastic reaction-diffusion models have been introduced in the literature. In general, these models can be divided into two very distinct classes [7]. The first class of models is compartment-based which is characterized by a discretization of the spatial domain into compartments [8]. At any particular time the best approximation to the localisation of any individual molecule is the compartment that the molecule is in. Molecules that are in the same compartment and are of the same species are completely indistinguishable. Molecules are free to migrate in the form of discrete jumps from one compartment to another via diffusion. Compartment-based modelling techniques are very popular and are used in a number

---

*Mathematical Institute, University of Oxford, 24-29 St Giles', Oxford OX1 3LB, United Kingdom; e-mails: flegg@maths.ox.ac.uk, chapman@maths.ox.ac.uk; erban@maths.ox.ac.uk
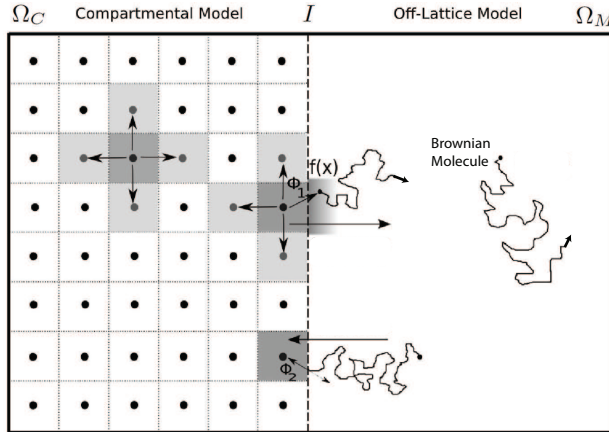
Figure 1: Graphical representation of the TRM.

of available self-contained simulation packages, for example, MesoRD [9] and SmartCell [10]. Whilst compartment-based models do not specifically represent the true noisy trajectory of the molecules, it has been shown that these models can provide accurate results by choosing the mesh spacing carefully [11, 12]. The second class of models is based on Brownian dynamics (molecular-based) simulation. The characterizing property of this method is that each molecule has an exact location on a continuous domain. Molecule diffusion is simulated by calculation of its noisy trajectory. There are a number of simulation packages that implement molecular-based simulation, for example, Smoldyn [13, 14], MCell [15, 16] and GFRD [17].

Brownian dynamics simulation is popular because of its better representation of the microscopic physics. However, if precise information about the trajectories of each molecule is not important then the effort placed on tracking them is a waste of computational resources. As a general rule, if concentrations are really small, tracking every molecule's trajectory is achieveable but becomes less practical as concentrations increase, when compartment-based methods (or even deterministic methods) are preferred. Often it is difficult to choose the most appropriate stochastic model when large spatial concentration variations, specific regions of interest and/or small systems coupled to larger systems are involved. In each of these cases, it would be ideal if a Brownian dynamics model may be used for localized regions of particular interest in which accuracy and microscopic detail is important (such as near the biological cell membrane [18]) and a compartment-based model may be used for other regions in which accuracy may be traded for simulation efficiency.

In this paper, we propose a spatially hybrid model, the *Two Regime Method* (TRM), which includes the best parts of each type of model and therefore optimizes simulation results. We consider a domain Ω divided into two subsets: Ω_C, which is discretized into compartments, and Ω_M, where a molecular-based algorithm (Brownian dynamics) is used to follow trajectories of individual molecules ($\Omega = \Omega_C \cup \Omega_M$). Specific algorithms that are used for each individual model may vary; our goal is to define the set of rules that must be followed on the interface $I \equiv \partial\Omega_C \cup \partial\Omega_M$ for accurate and efficient coupling of the two different modelling paradigms. The situation is schematically illustrated in Figure 1. Here, the compartment-based part Ω_C is on the left and the molecular-based subdomain Ω_M is on the right. Before outlining the implementation of the spatially hybrid algorithm, we first present a general overview of the two different modelling paradigms used respectively in Ω_C and Ω_M.

Compartment-based modelling

The characterizing property of compartment-based stochastic models of reaction-diffusion processes is the fact that the domain is discretized. Molecules do not have exact locations as they do in Brownian dynamics simulations but rather they are only assigned to be located within one of the compartments at a given moment in time. In particular, computer implementations of compartment-based models only store and evolve numbers of molecules in compartments (for each biochemical species). In what follows, we will consider the disretization of the computational domain using the compartments of the same size. This means that, in the case of three-dimensional simulations, compartments will be small cubes with the side equal to h. Similarly, compartments will be squares with the side h (resp. intervals of length h) for two-dimensional (resp. one-dimensional) models. In each case, diffusion is modelled by allowing jumps of molecules between neighboring compartments with rate D/h^2 where D is the diffusion constant [22].

Compartment-based models postulate that chemical reactions only occur if the reactant molecules are in the same compartment [9]. Let us consider a general system of N chemical species which are subject to M chemical reactions in a domain Ω_C which is discretized into K compartments. Then compartment-based models assign a propensity function $\alpha_{i,j}(t)$, $i = 1, 2, \ldots, M$, $j = 1, 2, \ldots, K$, to each chemical reaction in each compartment. Here, the product $\alpha_{i,j}(t)\,\mathrm{d}t$ is the probability that the i-th chemical reaction occurs in the j-th compartment during the infinitesimally small time interval $[t, t + \mathrm{d}t]$. The propensity $\alpha_{i,j}(t)$ depends on the number of available reactants in the compartment, the kinetic rate constant and the compartment size h [7]. In a similar way, one can assign the propensity of diffusive jumps from the j-th compartment, $j = 1, 2, \ldots, K$, to its neighboring compartments to be $\alpha_{i,j}(t) = D/h^2$, where $i = M + 1, \ldots, M + d_j$ and d_j is the number of available compartments adjacent to the j-th compartment. In this way, the reaction events and the diffusive jumps are both contained in the propensity $\alpha_{i,j}$.

Since a real diffusing molecule's trajectory is continuous, if a molecule leaves a compartment it enters a neighboring compartment. This necessarily places a constraint on the time step in which the system is updated. Commonly, there are two separate ways of simulating molecule migration from compartment to compartment. The first is by designation of the time step Δt [19]. We will call this approach *time-driven*. This is a less accurate but simpler method of simulation. Regardless of how small the time step is, there will be a non-zero probability for molecules to move more than one compartment which leads to inaccuracies in the results. However, accuracy is assured so long as the probability for a single molecule to move just one compartment, $D\Delta t/h^2$, is small (which makes the probability for two compartment jumps negligible). Further restrictions on the time step might also be provided by fast chemical reactions.

A more common approach to compartment-based modelling is given by *event-driven* algorithms which simulate the evolution of the system by picking the time step that corresponds to the next diffusive jump or reaction event [20, 21]. Examples include the Gilespie algorithm [20], Next Reaction Method [21] and Next Subvolume Method [9]. In the Next Reaction Method, the time $t + \tau_{i,j}$ for each reaction and diffusion jump is computed putatively by

$$\tau_{i,j} = \frac{1}{\alpha_{i,j}} \ln\left(\frac{1}{r_{i,j}}\right) \tag{1}$$

where $r_{i,j}$ is a uniformly distributed random number in (0,1). In particular, if the compartment size h is reduced, then the propensities D/h^2 of the molecules to migrate between compartments increases which causes the time step between diffusive events to decrease in turn. Thus, in both

time-driven and event-driven algorithms, the time step and the compartment size are inherently linked (either by the constraint of small jumping probabilities or by correlation between the randomly selected time step and the compartment size).

Molecular-based modelling

Computer implementations of molecular-based models in the literature store and evolve the positions of all biomolecules of interest [13]. The trajectories of large biomolecules (proteins) are computed through Brownian dynamics [23]. This is done by the inclusion of a random motion term which models the effect of solvent molecules on the biomolecules of interest without the need to simulate each solvent molecule individually. In the time-driven algorithms, the position $\mathbf{x}_i(t + \Delta t)$ of the i-th molecule at time $t + \Delta t$ is computed from its position $\mathbf{x}_i(t)$ at time t by

$$\mathbf{x}_i(t + \Delta t) = \mathbf{x}_i(t) + \sqrt{2D\Delta t}\,\boldsymbol{\zeta}, \tag{2}$$

where $\boldsymbol{\zeta} = [\zeta_x, \zeta_y, \zeta_z]$ is a vector of three normally distributed random numbers with zero mean and unit variance. Chemical reactions are then performed at each time step according to the probability of each reaction. Probabilites of bimolecular reactions also depend on the distance of possible reactants [13, 7]. Molecules that migrate over domain boundaries are treated depending on whether they are reflective, absorbing or reactive boundaries [18].

A time-driven Brownian model is simple to implement but may not be efficient. Prescribing a time step Δt to an arbitrary simulation can be difficult because if the concentration is too small then many time steps will be calculated before any interactions or interesting features may take place. Sometimes event-driven algorithms are prefered for Brownian dynamical models. An example is the GFRD method [17] which chooses a time step based on the current system configuration. It makes use of the fact that a single particle and two particle problems can be solved analytically using Green's functions, and selects the time step in such a way that the system evolution can be approximated as a collection of two particle (or single particle) problems.

The Two Regime method

The Two Regime method (TRM) is a spatially hybrid model for stochastic reaction-diffusion processes which uses a two-part domain Ω. One part, Ω_C, is discretized into compartments and the other, Ω_M, is a molecular-based subdomain ($\Omega = \Omega_C \cup \Omega_M$). The interface between the subdomains is denoted $I \equiv \partial\Omega_C \cup \partial\Omega_M$. The algorithm is graphically presented in Figure 1. Some details of the implementation of TRM are dependent on the choice of algorithms governing each individual regime, i.e. whether we use a time-driven or event-driven approach in Ω_C and Ω_M. In this section, we focus on the most common case. Other cases are discussed in the Discussion section.

Commonly, compartment-based stochastic reaction-diffusion models are implemented using an event-driven algorithm [9] whilst molecular-based stochastic reaction-diffusion models are implemented using a time-driven approach [13]. We therefore present the TRM which integrates an event-driven compartment-based reaction-diffusion regime (Next Reaction Method [21]) and a time-driven molecular-based reaction-diffusion regime (for example, that used by Smoldyn [13]). The algorithm is summarized in Table 1. Molecules in both Ω_C and Ω_M are simulated according to the rules defined by their particular algorithm. Therefore, when initializing in the step (i), molecules are placed according to compartment in Ω_C and according to coordinate in Ω_M. The propensity for particles to migrate from compartments adjacent to the interface I into

Table 1. The Two Regime Method for stochastic reaction-diffusion simulation

(i) Initialize numbers of molecules in compartments in Ω_C and positions of molecules which are in Ω_M.

(ii) Choose Δt the time between updates of the molecular-based regime (M-events) in Ω_M. Use Eq. 1 to calculate $\tau_{i,j}$, the putative times at which the next reaction or migratory events (C-events) will take place in Ω_C (or are initiated in Ω_c for jumps over the interface I). Set $t_M = \Delta t$ and $t_C = \min(\tau_{i,j})$ where the minimum is taken over all $i = 1, 2, \ldots, M + d_j$ and $j = 1, 2, \ldots, K$.

(iii) If $t_C \leq t_M$, then the next C-event occurs:
- Update current time $t := t_C$.
- Change the number of particles in Ω_C to reflect the specific C-event that has occured. If this event is one in which a particle leaves Ω_C bound for Ω_M, then compute its initial position in Ω_M according to the method described in the text (using Eq. 4) and remove it from the corresponding compartment adjacent to the interface I. Calculate the next putative time for the current C-event by Eq. 1.
- For all propensity functions $\alpha_{i,j}$ that are changed as a result of the C-event, determine the putative times of the corresponding event by Eq. 6. Set $t_C = \min(\tau_{i,j})$.

(iv) If $t_M \leq t_C$, then the next M-event occurs:
- Update current time $t := t_M$.
- Change the locations of all molecules in Ω_M according to Eq. 2.
- Initialize all molecules which migrated from Ω_C to Ω_M during previous C-events at locations computed in the step (iii), using Eq. 4.
- Perform all reaction events in Ω_M.
- Absorb all particles that interact with the interface I from Ω_M (excluding those just initiated) into the closest compartment in Ω_C. The particles moving accross the interface are selected with the help of Eq. 5.
- For all propensity functions $\alpha_{i,j}$ that are changed as a result of the M-event determine the putative times of the corresponding C-event by Eq. 6. Set $t_C = \min(\tau_{i,j})$ if needed.

(v) Repeat steps (iii) and (iv) until the desired end of the simulation.

Ω_M is given by $\Phi_1 D/h^2$ per molecule, where Φ_1 is the change in the propensity of migration (from that of an internal migratory event) that is included to make the molecular flux over the interface I consistent with diffusion. We have determined Φ_1 to be (see Supplemental Material)

$$\Phi_1 = \frac{2h}{\sqrt{\pi D \Delta t}}, \tag{3}$$

where Δt is the fixed time increment defined for updating the molecules in Ω_M. We use Eq. 1 to find the putative times $t + \tau_{i,j}$ for the next reaction/migration events in Ω_C. We call these events as C-events. They include the jumps from Ω_C to Ω_M which have propensities equal to $\Phi_1 D/h^2$ multiplied by the number of molecules in the corresponding compartment. Therefore, we define d_j to be the number of available directions in which a molecule may jump including a jump over the interface I. In the step (ii), we also initialize the time of the next update of molecules in Ω_M as $t_M = \Delta t$ (we call these events as M-events).

The algorithm then proceeds by repeating steps (iii) and (iv) which compute C-events and M-events, respectively. At each C-event, we track the reaction-migratory events that occur in Ω_C. At each M-event, we update the entire system of particles in Ω_M by calculating the new

positions of particles in Ω_M using Eq. 2 and the reactions using the method prescribed by the specific molecular-based algorithm [13, 23]. At each M-event, we also introduce new molecules from Ω_C into Ω_M by placing them at a distance x from the interface I that is sampled from the probability distribution $f(x)$. These new molecules arise from those C-events which have occurred since the previous M-event. We have determined that (see Supplemental Material)

$$f(x) = \sqrt{\frac{\pi}{4D\Delta t}} \text{erfc}\left(\frac{x}{\sqrt{4D\Delta t}}\right), \tag{4}$$

where $x \geq 0$ is the distance from the boundary and $\text{erfc}(x) = 2/\sqrt{\pi} \int_x^\infty \exp(-t^2)\,\mathrm{d}t$ is the complementary error function.

Furthermore, in the step (iv), we also identify all molecules which migrate from Ω_M to Ω_C during an M-event. These include those molecules which have new locations calculated by Eq. 2 in Ω_C. However, we also have to include some molecules which have their new locations calculated by Eq. 2 in Ω_M. We postulate that every molecule has a probability P_m to migrate from Ω_M to Ω_C during an M-event where

$$P_m = \exp\left(\frac{-\Delta x_{\text{old}}\,\Delta x_{\text{new}}}{D\,\Delta t}\right), \tag{5}$$

where Δx_{old} is the distance of the molecule from the interface I before the M-event and Δx_{new} is the calculated distance from the interface I at the current M-event. All molecules which migrate from Ω_M to Ω_C are placed into the corresponding (nearest) compartment in Ω_C. It is worth noting that Eq. 5 is also used in some stochastic reaction-diffusion algorithms for the treatment of boundary conditions [13, 18]. In the TRM, the boundary conditions are implemented for each regime in the usual way, see, for example Ref. [18] where treatment of reflective, absorbing and reactive boundaries is discussed.

For both C-events and M-events, when the propensity $\alpha_{i,j}$ is changed due to a change in Ω_C we must recalculate the putative time $t + \tau_{i,j}$ for the next C-event to take place. Following Gibson and Bruck [21], we update $\tau_{i,j}$ as follows:

$$\tau_{i,j} := t + \frac{\alpha_{i,j}^{\text{old}}(t)}{\alpha_{i,j}^{\text{new}}(t)}\left(\tau_{i,j} - t\right) \tag{6}$$

where $\alpha_{i,j}^{\text{old}}(t)$ is the propensity before the event at time t and $\alpha_{i,j}^{\text{new}}(t)$ is the propensity after the event.

Mathematical justification of the TRM is given in the Supplemental Information. In order to demonstrate the validity of the TRM, we will present results for two test problems for which one can derive analytically the exact solution. We can therefore test our stochastic simulation by increasing the number of realizations and observing convergence between the exact solution and the stochastic simulation. The first test problem has been designed so that a large range of different concentrations and fluxes will exist at some stage on and across the interface so that theoretical agreement will provide more compelling validation of the algorithm in all possible situations. In the second test problem, we have applied the TRM to a particular model of protein concentration gradients in order to demonstrate the usefulness of the TRM to an existing biological model [24].

Illustrative results: diffusion example

We will consider one-dimensional diffusion of N_0 non-interacting molecules with diffusion constant $D = 1$ in the domain $\Omega = [0, 1]$. In this example, all parameters are dimensionless. The

boundaries of the domain Ω are defined to be reflective. We assume that initially all molecules are uniformly distributed in the subinterval $[0, 0.5]$. Physically, this problem is nothing other than the diffusion of molecules squashed into half a room being suddenly allowed to fill the room. This problem can be stated as the following initial boundary value problem for the (normalized) density ϱ of molecules

$$\frac{\partial \varrho}{\partial t} = \frac{\partial^2 \varrho}{\partial x^2}, \quad 0 < x < 1, \tag{7}$$

where $\varrho(x, 0) = 2H(0.5 - x)$, $\varrho_x(0, t) = \varrho_x(1, t) = 0$ and H is the Heaviside function. The solution can be found by applying the separation of variables technique to Eq. 7 as

$$\varrho = 1 - 4 \sum_{i=1}^{\infty} \frac{(-1)^i \cos\left((2i - 1)\pi x\right)}{(2i - 1)\pi} \exp\left(-(2i - 1)^2 \pi^2 t\right). \tag{8}$$

This solution will be used to test the validity of the TRM.

To apply the TRM, we divide Ω into the following compartment-based and molecular-based parts: $\Omega_C = [0, 0.5]$ and $\Omega_M = [0.5, 1]$. Therefore, the interface I is located at $I = \partial\Omega_C \cap \partial\Omega_M = \{0.5\}$. The compartment-based domain Ω_C is discretized into $K = 20$ compartments (subintervals) of the length $h = 0.025$. The compartment-based model will give us the time evolution of a vector of integers $\mathbf{n} = [n_1, n_2, \ldots, n_K]$, where n_i is the number of molecules in the ith compartment, $i = 1, 2, \ldots, K$. Initially all N_0 molecules are placed in Ω_C in random compartments with equal probability. This constitutes the step (i) in Table 1. In the step (ii), we choose the time step Δt as $\Delta t = 10^{-6}$, therefore Φ_1 given by Eq. 3 is approximately 28.

Figure 2 is a plot of the distribution comparing the solution (Eq. 8) and the distribution that is produced using our stochastic algorithm using an initial $N_0 = 2000$ molecules at $t = 0.1$. For comparitive purposes, the molecules in Ω_M are counted and compartmentalized so that we can see their relative concentration to that of the compartments in Ω_C. Whilst good agreement is observed we require a more quantitative comparison. It has been established in the literature that the algorithms that govern the internal processes of Ω_M and Ω_C are consistent with diffusion and therefore we are most interested with the accuracy of the flux over the boundary. For this reason we compare the total probability to find molecules in Ω_C with that predicted by Eq. 8 (i.e. $\int_0^{0.5} \varrho(x, t) \, \mathrm{d}x$). A comparison of a theoretical probability and simulation containing $N_0 = 2000$ molecules shows good agreement over time in Figure 3. Simulation of the TRM with $N_0 = 30000$ overlaps that of the theory in Figure 3. Thus, in Figure 4, we present a plot of the error between the expected probability to find molecules in Ω_C and that of a simulation with $N_0 = 30000$ molecules over time to demonstrate there is no apparent time dependent error that is associated with the algorithm, rather there is simple stochastic white noise around the true expected probability. The test problem therefore provides very strong validation of the proposed algorithm.

Illustrative results: a morphogen gradient model

In the second test problem, we simulate a model of a diffusing morphogen presented by Tostevin, ten Wolde and Howard [24]. Consider a system of no molecules on the domain $\Omega = [0, \infty)$. Morphogen molecules are produced at the origin at a rate $J = 6.25 \ \mu\mathrm{m}^{-1}\mathrm{s}^{-1}$ and diffuse with a diffusion coefficient of $D = 0.5 \ \mu\mathrm{m}^2\mathrm{s}^{-1}$. We implement a zero flux boundary condition at the origin. The molecules undergo a linear death rate of $\mu = 1 \ \mathrm{s}^{-1}$. Intuitively, since molecules are produced only at the origin and die as they diffuse away from the origin there will be a large concentration of molecules near the origin which dissipate as x gets large. If a molecular-based
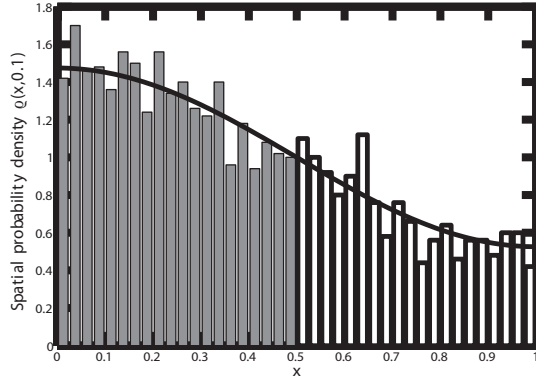
Figure 2: Plot of the distribution of the diffusion example evaluated at $t = 0.1$ according to Eq. 8 (black line) and TRM (compartment molecule probability density shown in dark compartment bars and free molecules counted into compartments and shown in white bars). The compartmental separation is $h = 0.025$.
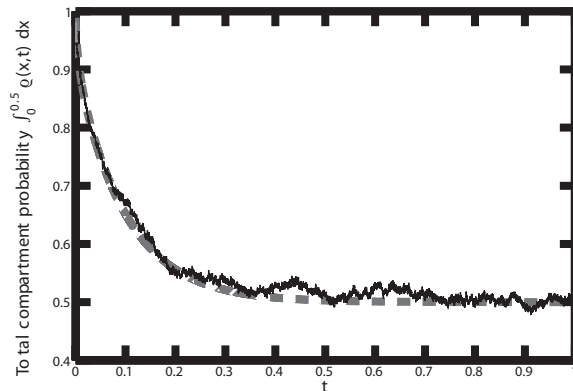

Figure 3: Plot of total probability to find a particle bound to a compartment as a function of time ($\int_0^{0.5} \varrho(x,t)\, \mathrm{d}x$) according to Eq. 8 (gray dashed line) and from simulation of 2000 molecules (solid black line) using the TRM for the diffusion example.

algorithm were to be used for this model, at small values of x, where there are characteristically a lot of molecules, high computational effort will be needed whereas accurate and efficient results are achieveable using a compartment-based model. However, using a compartment-based algorithm for the entire model would mean necessarily truncating the domain and introducing errors associated with small concentrations at large values of x. This type of problem is ideal for the TRM since we would prefer to utilise the efficiency of a compartment-based model for small x and the precision of a molecular-based model for large x. We therefore simulate the model using the TRM and place the interface at an arbitrarily chosen location ($x = 1\ \mu$m). Therefore, $\Omega_C = [0,1)$ and $\Omega_M = [1,\infty)$. In the large J limit, we expect the concentration of particles, $\varrho(x,t)$ to obey the PDE

$$\frac{\partial \varrho}{\partial t} = D\nabla^2 \varrho - \mu \varrho + J\delta(x), \quad x > 0 \tag{9}$$

where $\delta(x)$ is the Dirac delta function. Whilst there is zero flux over the boundary the introduction of the point source term at $x = 0$ gives us a formal boundary condition of $\varrho_x(0,t) = -J/D$.
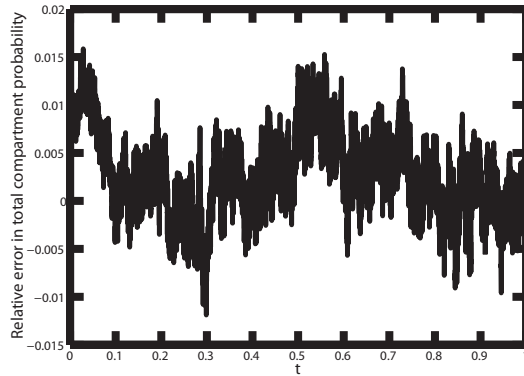
8

Figure 4: Relative stochastic error of the total probability to find a particle bound to a compartment ($\int_0^{0.5} \varrho(x,t)\,\mathrm{d}x$) between simulation of the diffusion example using the TRM with 30000 initial particles and the theoretical value calculated using Eq. 8, as a function of time.

The initial condition is $\varrho(x,0) = 0$. The time dependent analytical solution is given by Ref. [25]

$$\varrho(x,t) = \frac{\lambda J \exp\left(-x/\lambda\right)}{2D} \left[1 - \frac{1}{2}\mathrm{erfc}\left(\frac{2Dt/\lambda - x}{\sqrt{4Dt}}\right) + \right.$$
$$\left. + \frac{-\exp(2x/\lambda)}{2}\mathrm{erfc}\left(\frac{2Dt/\lambda + x}{\sqrt{4Dt}}\right) \right], \tag{10}$$

where $\lambda = \sqrt{D/\mu}$ is the characteristic decay length. Using this analytic solution we can discuss the accuracy of the TRM by combining 5000 simulations of this problem.

We chose the following parameters for our simulation results; $\Delta t = 0.1$ ms for the molecular-based model and $h = 0.05$ μm for the compartment-based model. Figure 5 is a plot of the distribution comparing the analytical solution (Eq. 10) and the distribution that is produced using our stochastic algorithm averaged over 5000 realizations at time $t = 3$ s. In order to test that the interface properties of the TRM are accurate we compare $\int_0^1 \varrho(x,t)\,\mathrm{d}x$ computed by Eq. 10 to the total expected number of compartment-bound molecules over 5000 simulations and $\int_1^\infty \varrho(x,t)\,\mathrm{d}x$ to the total expected number of molecules in Ω_M over the same 5000 simulations. The results are shown in Figure 6. The stochastic and analytic results are in clear agreement.

Discussion and conclusion

In this manuscript we have successfully interfaced a compartment-based model of diffusion to a Brownian dynamics model of diffusion. We have derived and verified, for any process which has diffusion as its dominant form of molecule mobility, very particular rules in which molecules may freely migrate between a compartment-based regime and a free Brownian dynamics regime. The ability to interchange the modelling approach in spatially specific areas allows greater control in simulation and in particular allows a modeller to have the precision of a molecular-based regime in regions where it is needed whilst having the computational efficiency of compartment-based approaches where it is appropriate. In order to interface the two regimes one simply needs to migrate particles between regimes according to the rules described in Table 1 and using the coupling parameters given in Eqs. 3 and 4. Their mathematical justification is given in the Supplemental Material.

It is important to note that while we have focused on presenting the TRM using an event-driven compartment-based regime and a time-driven molecular-based regime, it is possible to use
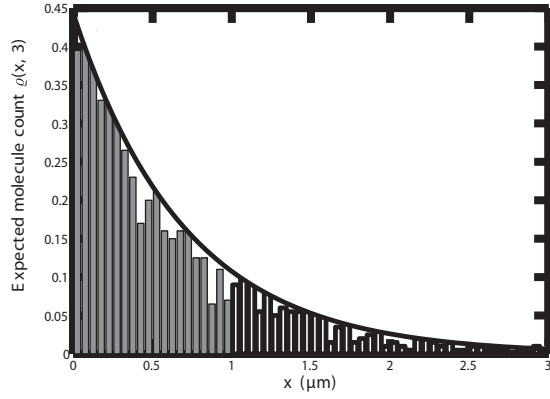
Figure 5: Plot of the expected morphogen molecule count per compartment evaluated at $t = 3$ s according to Eq. 10 (black line) and 200 realizations of the TRM simulation (compartment molecule probability density shown in dark compartment bars and free molecules counted into compartments and shown in white bars).
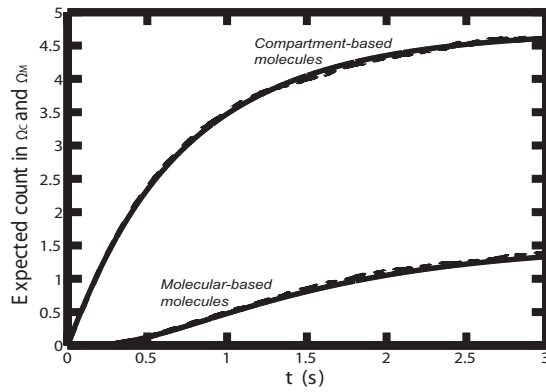

Figure 6: Plot of total expected number of morphogen molecules found in both Ω_C and Ω_M as a function of time for the analytical solution according to Eq. 10 (solid lines) and TRM simulation results averaged over 5000 realizations (dashed lines).

10

any desired combination of compartment-based and molecular-based regimes. In the case where algorithms in both Ω_C and Ω_M are event-driven (for example, using the GFRD [17] in Ω_M), the parameters Φ_1 and $f(x)$ are the same. Care must be taken in this case. Molecule migration from Ω_M to Ω_C occurs at the moment of first contact with the interface I. Furthermore, Eqn. 3 implies that (in the case of an event driven algorithm) Φ_1 is dependent on $\tau_{i,j}$, which in turn determines the propensity for the diffusive events on interfacial compartments in Ω_C and therefore determine $\tau_{i,j}$ itself. The putative time for the migration of a particle from the boundary compartment in Ω_C to Ω_M is calculated from some current time in the following way:

$$\tau_{i,j} = \frac{h^2}{n_b \Phi_1 D} \ln\left(\frac{1}{r}\right), \tag{11}$$

where n_b is the number of molecules in the compartment at the interface, r is a uniformly distributed random number in (0,1). Notice from (3) that $\Phi_1 \sqrt{\tau_{i,j}} = \frac{2h}{\sqrt{\pi D}}$, a constant with respect to $\tau_{i,j}$. Therefore we find $\tau_{i,j}$ using the following formula

$$\tau_{i,j} = \left(\frac{h^2}{n_b(\Phi_1 \sqrt{\tau_{i,j}})D} \ln\left(\frac{1}{r}\right)\right)^2 = \frac{\pi h^2}{4 n_b^2 D} \ln^2 r. \tag{12}$$

The illustrative algorithms which we presented were effectively one-dimensional, but the algorithm naturally extends by symmetry into higher dimensions with flat interfaces. It still remains to extend the work into curved higher dimensional interfaces and systems with advection as a form of molecule mobility. Another important extension is to consider unstructured meshes [26] and complex geometries [27]. The current algorithm has the potential to significantly increase the accuracy and speed at which current stochastic reaction-diffusion simulations are implemented by giving the simulator control over their modelling approaches in regions of interest.

Acknowledgements

References

[1] Erdmann T, Howard M, ten Wolde PR (2009) The role of spatial averaging in the precision of gene expression patterns. *Phys Rev Let* 103:258101.

[2] Takahashi K, Tănase-Nicola S, ten Wolde PR (2010) Spatio-temporal correlations can drastically change the response of a MAPK pathway. *Proc Natl Acad Sci U S A* 107(46):19820-19825.

[3] Lipkow K, Andrews SS, Bray D (2005) Simulated diffusion of phosphorylated CHeY through the cytoplasm of escherichia coli. *J Bacteriol* 187(1):45-53.

[4] Fange D, Elf J (2006) Noise-induced Min phenotypes in E. coli. *PLoS Comput Biol* 2(6):637-648.

[5] Tomita M (2001) Whole-cell simulation: a grand challenge of the 21st century. *Trends Biotechnol* 19(6):205-210.

[6] Murray JD (2003) *Mathematical biology II: Spatial models and biomedical applications*, Springer-Verlag, pp 403-420.

[7] Erban R, Chapman J (2009) Stochastic modelling of reaction-diffusion processes: algorithms for bimolecular reactions. *Phys Biol* 6(4):046001.

[8] Fange D, Berg OG, Sjöberg P, Elf J (2010) Stochastic reaction-diffusion kinetics in the microscopic limit. *Proc Natl Acad Sci U S A* 107(46):19820-19825.

[9] Hattne J, Fange D, Elf J (2005) Stochastic reaction-diffusion simulation with MesoRD. *Bioinformatics* 21(12):2923-2924.

[10] Ander M (2004) SmartCell, a framework to simulate cellular processes that combines stochastic approximation with diffusion and localisation:analysis of simple networks. *Syst Biol* 1(1):129-138.

[11] Isaacson SA (2009) The reaction-diffusion master equation as an asymptotic approximation of diffusion to a small target. *SIAM J Appl Math* 70(1):77-110.

[12] Isaacson SA, Isaacson D (2009) Reaction-diffusion master equation, diffusion-limited reactions, and singular potentials. *Phys Rev E* 80:066106.

[13] Andrews SS, Bray D (2004) Stochastic simulation of chemical reactions with spatial resolution and single molecular detail. *Phys Biol* 1:137-151.

[14] Andrews SS, Addy NJ, Brent R, Arkin AP (2010) Detailed simulations of cell biology with Smoldyn 2.1. *PLoS Comput Biol* 6(3):e1000705.

[15] Bartol TM, Stiles JR, Salpeter MM, Salpeter EE, Sejnowski TJ (1996) Mcell: generalized Monte Carlo computer simulation of synaptic transmission and chemical signaling. *Abstr Soc Neurosci* 22:1742.

[16] Kerr RA, Bartol TM, Kaminsky B, Dittrich M, Chang J-CJ, Baden SB, Sejnowski TJ, Stiles JR (2008) Fast Monte Carlo simulation methods for biological reaction-diffusion systems in solution and on surfaces. *SIAM J Sci Comput* 30(6):3126-3126.

[17] van Zon JS, ten Wolde PR (2005) Simulating biochemical networks at the particle level and in time and space: Green's function reaction dynamics. *Phys Rev Lett* 94:128103.

[18] Erban R, Chapman J (2007) Reactive boundary conditions for stochastic simulations of reaction-diffusion processes. *Phys Biol* 4(1):16-28.

[19] Tomita M et al. (1999) E-Cell software environment for whole-cell simulation. *Bioinformatics* 15(1):72-84.

[20] Gillespie D (1977) Exact Stochastic Simulation of Coupled Chemical Reactions. *J Phys Chem* 81(25):2340-2361.

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

[22] Erban R, Chapman J, Maini P (2007) A practical guide to stochastic simulations of reaction-diffusion processes. http://arxiv.org/abs/0704.1908

[23] Lipkov a J, Zygalakis KC, Chapman SJ, Erban R (2011) Analysis of Brownian dynamics simulations of reversible biomolecular reactions. http://arxiv.org/abs/1005.0698

[24] Tostevin F, ten Wolde PR, Howard M (2007) Fundamental limits of position determination by concentration gradients. *PLoS Comput Biol* 3(4):0763-0771.

[25] Bergmann S et al. (2007) Pre-steady-state decoding of the bicoid morphogen gradient. *PLoS Comput Biol* 5(2):0232-0242.

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

[27] Isaacson SA, Peskin CS (2007) Incorporating diffusion in complex geometries into stochastic chemical kinetics simulations. *SIAM J Sci Comput* 28(1):47-74.

Supplemental Material – the Two Regime method for optimizing stochastic reaction-diffusion simulations

In the supplemental material, we derive equations (3) and (4) from the manuscript, i.e. we derive equations relating the TRM parameters Φ_1 and f(x) and the parameters of compartment-based and molecular-based models. This derivation will also make use of an auxiliary parameter Φ_2 (which does not explicitly appear in the formulation of the TRM, because its value is determined to be equal to 2 in all cases). We also justify equation (5) from the manuscript.

Without loss of generality, consider an interface at $x = 0$. To the left of this interface a compartment-based model is used with fixed compartment size h. To the right of the interface a molecular-based algorithm is used, updated at fixed time increments of Δt. Our aim is to choose the parameters which govern the transfer of molecules across the interface from one model to the other in such a way as to make the interface "invisible" in the solution: the switch of models does not reflect anything in the underlying physical processes, it is simply a mathematical construct to aid the simulation.

To determine these parameters we focus on the purely diffusive problem, since bulk reactions have no affect on boundary conditions [18]. Then the exact description of the underlying physical problem is a set N molecules undergoing Brownian motion. We write $\bar{p}(x,t)$ for the probability distribution function for this process, so that $\bar{p}(x,t)\delta x$ gives the probability of finding a given molecule in the interval $[x, x + \delta x]$ at time t. In the purely diffusive case $\bar{p}(x,t)$ satisfies the diffusion equation.

The stochastic models we are using on either side of the interface at $x = 0$ both provide an approximation to $\bar{p}(x,t)$ (as $h \to 0$ and $\Delta t \to 0$). We need to choose the coupling parameters so that these approximations join together smoothly, that is, they both give the same value of $\bar{p}(0,t)$ and $\bar{p}_x(0,t)$, where the subscript indicates partial differentiation.

We label the compartment $x \in [-kh, -kh + h]$ by k and denote the probability of finding a molecule in it by $p_k(t)h$ (so that $p_k(t)$ approximates the probability density function $\bar{p}(-kh + h/2, t)$ above). We motivate our modification of the transition probabilities for compartment number one (adjacent to the interface) by imagining that a molecule in it which attempts to jump to the right has a probability Φ_1 of crossing onto the molecular-based side, and a probability of $1 - \Phi_1$ of being reflected (in fact, Φ_1 may be greater than 1, so in the end we must think of it as altering the transition rates rather than as a transmission probability). If the molecule is transmitted we have to decide where to place it: we suppose that it is placed at a random position in $x > 0$ with probability distribution function $f(x)$ (so that it is placed in the interval $[x, x + \mathrm{d}x)$ with probability $f(x)\delta x$). Similarly we imagine that a molecule on the molecular-based side $x > 0$ at time t whose new position at time $t + \Delta t$ is in $x < 0$ is added to compartment number 1 with probability Φ_2, and reflected with probability $1 - \Phi_2$ (again, in the end we will find that $\Phi_2 = 2$, so we will later have to change our interpretation of Φ_2).

If we denote by $p(x,t)$ the probability density function of the discrete-time molecular-based algorithm, then these transmission/reflection rules imply

$$
\begin{aligned}
p_1(t + \Delta t) &= \left(1 - \frac{2D\Delta t}{h^2}\right)p_1(t) + \frac{D\Delta t}{h^2}p_2(t) + (1 - \Phi_1)\frac{D\Delta t}{h^2}p_1(t) \\
&\quad + \int_0^\infty \int_0^\infty \frac{\Phi_2 p(y,t)}{h\sqrt{4\pi D\Delta t}} \exp\left(\frac{-(x+y)^2}{4D\Delta t}\right) \mathrm{d}x\,\mathrm{d}y, \quad (13)
\end{aligned}
$$

$$
\begin{aligned}
p(x, t + \Delta t) &= \int_0^\infty \frac{p(y,t)}{\sqrt{4\pi D\Delta t}}\left[\exp\left(\frac{-(x-y)^2}{4D\Delta t}\right) + (1 - \Phi_2)\exp\left(\frac{-(x+y)^2}{4D\Delta t}\right)\right]\,dy \\
&\quad + \frac{D\Delta t\Phi_1}{h}p_1(t)f(x). \quad (14)
\end{aligned}
$$

Note that if $f(x)$ vanishes away from $x = 0$ then equation (14) reduces to the Fokker-Planck equation for finite-time-step Brownian motion and thus to the diffusion equation in the limit $\Delta t \to 0$. However, in the vicinity of $x = 0$ there is a boundary layer of width $O(\sqrt{\Delta t})$. We rescale (13) and (14) using the (dimensionless) boundary layer coordinate $\eta = x/\sqrt{D\Delta t}$. We also denote the probability density and placement function in this boundary layer region by $p_{\mathrm{inner}}(\eta, t) = p(\sqrt{D\Delta t}\,\eta, t)$ and $f_{\mathrm{inner}}(\eta) = \sqrt{D\Delta t}\,f(\sqrt{D\Delta t}\,\eta)$. Note the rescaling of f, which is necessary to satisfy the normalisation condition $\int_0^\infty f(x)\,\mathrm{d}x = 1$ since (as we will see) f vanishes outside the boundary layer. Thus, in the boundary

layer scalings, (13) and (14) become

$$
\begin{aligned}
p_1(t + \Delta t) &= \left(1 - \frac{(1 + \Phi_1)\, D\Delta t}{h^2}\right) p_1(t) + \frac{D\Delta t}{h^2} p_2(t) \\
&\quad + \frac{\Phi_2 \sqrt{D\Delta t}}{h} \int_0^\infty \int_0^\infty p_{\text{inner}}(\xi, t) K(\eta + \xi)\, d\eta\, d\xi,
\end{aligned}
\tag{15}
$$

$$
p_{\text{inner}}(\eta, t + \Delta t) = \int_0^\infty p_{\text{inner}}(\xi, t)\left(K(\eta - \xi) + (1 - \Phi_2) K(\eta + \xi)\right) d\xi + \frac{\sqrt{D\Delta t}\, \Phi_1}{h} p_1(t) f_{\text{inner}}(\eta),
\tag{16}
$$

where $K(x) = (4\pi)^{-1/2} \exp(-x^2/4)$. Now in order for our two models to join smoothly we require on the compartment-based side that

$$
\begin{aligned}
p_1(t) &\sim \bar{p}(-h/2, t) = \bar{p}(0, t) - \frac{h}{2} \bar{p}_x(0, t) + \cdots, \\
p_2(t) &\sim \bar{p}(-3h/2, t) = \bar{p}(0, t) - \frac{3h}{2} \bar{p}_x(0, t) + \cdots,
\end{aligned}
$$

while, for the molecular-based side, we want no rapid variation in the boundary layer, so that

$$
p_{\text{inner}}(\eta, t) \sim \bar{p}(0, t) + \sqrt{D\Delta t}\, (\eta + C)\, \bar{p}_x(0, t) + \cdots.
\tag{17}
$$

We have allowed here for a small shift C in where the molecular-based region "sees" the interface; we will see that this extra degree of freedom is crucial. Equivalently we could have said that $p_k(t)$ approximates $\bar{p}(-kh + h/2 - C\sqrt{D\Delta t}, t)$ rather than $\bar{p}(-kh + h/2, t)$: C is really a small phase shift between the spatial coordinate in the two regions. Substituting these expansions into (15) and (16) and equating coefficients of h^0 and h^1 (with $h \sim \sqrt{\Delta t}$ as h, $\sqrt{\Delta t} \to 0$) gives

$$
\Phi_1 = \frac{h \Phi_2}{\sqrt{D\Delta t}} \int_0^\infty \int_0^\infty K(\eta + \xi)\, d\eta\, d\xi = \frac{h \Phi_2}{\sqrt{D\Delta t \pi}},
\tag{18}
$$

$$
\Phi_1 = 2 - 2\Phi_2 \int_0^\infty \int_0^\infty (\xi + C) K(\eta + \xi)\, d\eta\, d\xi = 2 - \Phi_2 - \frac{2\Phi_2 C}{\sqrt{\pi}},
\tag{19}
$$

$$
\begin{aligned}
1 &= \int_0^\infty \left(K(\eta - \xi) + (1 - \Phi_2) K(\eta + \xi)\right) d\xi + \frac{\sqrt{D\Delta t}\, \Phi_1}{h} f_{\text{inner}}(\eta), \\
&= 1 - \frac{\Phi_2}{2} \text{erfc}\left(\frac{\eta}{2}\right) + \frac{\sqrt{D\Delta t}\, \Phi_1}{h} f_{\text{inner}}(\eta),
\end{aligned}
\tag{20}
$$

$$
\begin{aligned}
\eta + C &= \int_0^\infty (\xi + C)\left(K(\eta - \xi) + (1 - \Phi_2) K(\eta + \xi)\right) d\xi - \frac{\Phi_1}{2} f_{\text{inner}}(\eta), \\
&= \eta + C + (2 - \Phi_2)\left(\frac{e^{-\eta^2/4}}{\sqrt{\pi}} - \frac{\eta}{2} \text{erfc}\left(\frac{\eta}{2}\right)\right) - \frac{C\Phi_2}{2} \text{erfc}\left(\frac{\eta}{2}\right) - \frac{\Phi_1}{2} f_{\text{inner}}(\eta),
\end{aligned}
\tag{21}
$$

where $\text{erfc}(x) = 2/\sqrt{\pi} \int_x^\infty \exp(-t^2)\, dt$ is the complementary error function. From (18) we find

$$
\frac{\Phi_1}{\Phi_2} = \frac{h}{\sqrt{\pi D\Delta t}}.
\tag{22}
$$

Then (20) gives

$$
f_{\text{inner}}(\eta) = \frac{\sqrt{\pi}}{2} \text{erfc}\left(\frac{\eta}{2}\right),
\tag{23}
$$

corresponding to the unscaled

$$
f(x) = \sqrt{\frac{\pi}{4D\Delta t}} \text{erfc}\left(\frac{x}{\sqrt{4D\Delta t}}\right),
\tag{24}
$$

which is denoted as equation (4) in the paper. The normalisation condition on f is automatically satisfied.

Substituting (23) and (19) into (21) gives

$$
0 = (2 - \Phi_2)\left(\frac{e^{-\eta^2/4}}{\sqrt{\pi}} - \frac{\eta}{2} \text{erfc}\left(\frac{\eta}{2}\right) - \frac{\sqrt{\pi}}{4} \text{erfc}\left(\frac{\eta}{2}\right)\right),
$$

from which we find
$$\Phi_2 = 2. \tag{25}$$

Substituting (13) in (10), we derive the formula for Φ_1 which is presented as equation (3) in the paper Finally, substituting this value into (19) gives
$$C = -\frac{h}{2\sqrt{D\Delta t}}. \tag{26}$$

One interpretation of this value of C is that we should think of $p_k(t)$ as approximating $\bar{p}(-kh + h, t)$, so that $p_1(t)$ approximates $\bar{p}(0, t)$ and not $\bar{p}(-h/2, t)$.

We introduced Φ_1 and Φ_2 as probabilities, and yet we have found Φ_2 to be greater than one (as indeed, Φ_1 may be). Thus we need to reinterpret their meaning. For Φ_1, from (15), we see that we simply need to interpret
$$\frac{\Phi_1 D\Delta t}{h^2}$$

as the probability of a molecule in compartment number one jumping to the right. Thus Φ_1 simply represents the relative increase in this propensity function for a boundary compartment over a compartment in the bulk. For Φ_2, from (14), we see that the probability that a molecule at $y > 0$ at time t lies at $x > 0$ at time $t + \Delta t$ is

$$\frac{1}{\sqrt{4\pi D\Delta t}} \left[\exp\left(\frac{-(x-y)^2}{4D\Delta t} \right) - \exp\left(\frac{-(x+y)^2}{4D\Delta t} \right) \right] = \frac{1}{\sqrt{4\pi D\Delta t}} \exp\left(\frac{-(x-y)^2}{4D\Delta t} \right) \left[1 - \exp\left(\frac{-xy}{D\Delta t} \right) \right].$$

We can interpret this as implying that not only are those molecules which end the time step at $x < 0$ removed and placed in compartment number one, but each molecule which ends the time step at a position $x > 0$ is removed with probability

$$P_m = \exp\left(-\frac{xy}{D\Delta t} \right).$$

which is labeled as equation (5) in the manuscript. As shown in [13], P_m is the probability that the particle trajectory crossed the interface $x = 0$ at some point during the time interval $[t, t + \Delta t)$ and then returned to $x > 0$, so that $\Phi_2 = 2$ corresponds in some sense to a maximally absorbent boundary.