# Single molecule simulations in complex geometries with embedded dynamic one-dimensional structures

Stefan Hellander[1]

*Department of Information Technology, Uppsala University, Box 337, SE-75105, Uppsala, Sweden.*[a]

Stochastic models of reaction-diffusion systems are important for the study of bio-chemical reaction networks where species are present in low copy numbers or if reactions are highly diffusion limited. In living cells many such systems include reactions and transport on one-dimensional structures, such as DNA and microtubules. The cytoskeleton is a dynamic structure where individual fibers move, grow and shrink. In this paper we present a simulation algorithm that combines single molecule simulations in three-dimensional space with single molecule simulations on one-dimensional structures of arbitrary shape. Molecules diffuse and react with each other in space, they associate to and dissociate from one-dimensional structures as well as diffuse and react with each other on the one-dimensional structure. A general curve embedded in space can be approximated by a piecewise linear curve to arbitrary accuracy. The resulting algorithm is hence very flexible. Molecules bound to a curve can move by pure diffusion or via active transport, and the curve can move in space as well as grow and shrink. The flexibility and accuracy of the algorithm is demonstrated in four numerical examples.

PACS numbers: 82.37.Np,82.39.-k,87.10.Mn,87.16.Ka

[a] Electronic mail: stefan.hellander@it.uu.se

# I. INTRODUCTION

It has been established that modeling biochemical reaction networks with a macroscopic deterministic model, for instance with ordinary differential equations (ODEs), can be unsuitable when some species are present in low copy numbers or if reactions are highly diffusion limited[1,2]. Such systems may only be accurately modeled by a discrete stochastic model. Usually two levels of stochastic modeling are considered: the mesocopic level and the microscopic level.

The mesoscopic scale is described by the chemical master equation (CME) which is a discrete, stochastic model. Molecules are treated as discrete entities but assumed to be well stirred with a uniform distribution within the simulation volume. The copy number of each species is computed but the position of individual molecules is not tracked. Exact trajectories of such a system can be generated with the stochastic simulation algorithm (SSA) by Gillespie[3].

Systems for which the condition of spatial homogeneity does not hold can be modeled at the mesoscopic level by subdividing the domain into smaller compartments, where in each compartment the well stirred condition is assumed to hold. Diffusion can now be modeled as discrete jumps between the compartments. The governing equation of the system is the reaction-diffusion master equation (RDME) and exact trajectories can be computed with the next subvolume method (NSM)[2], which is an efficient implementation of the SSA in a spatial setting. Several software packages have been developed for simulating systems at the mesoscopic level, including MesoRD[4], URDME[5] and STEPS[6]. The spatial accuracy of the RDME depends on the size of the voxels. If the voxels are too large we might not resolve the spatial dynamics of the system accurately, but on the other hand it has also been shown that when the size of the voxels becomes too small the accuracy deteriorates[7–9]. By allowing the reaction rates in the system to depend on the size of the voxels in a general way, accuracy can be retained, but at most up to a critical lower bound of the voxel size, at which point the model breaks down regardless of how the reaction rates are chosen[8–10]. The mesoscopic model may therefore be unsuitable for systems where a very high spatial accuracy is required to capture important dynamics.

The microscopic level, or microscale, usually refers to models where the positions of single molecules are tracked. Either we compute the continous position in space or let the individual molecules move on a lattice, as in Ref. 11. Reactions occur with some probability per unit time when molecules collide or when they are within some prescribed reaction radius. There are two models that have been used extensively to model systems at the microscale: the Smoluchowski model[12] and the Doi model[13,14]. In this paper we will only be concerned with the Smoluchowski model.

In Ref. 15 they show that fine-grained spatial correlations can affect the macroscopic behaviour of a system. Furthermore, the dynamics of the system cannot be captured on the macroscopic or mesoscopic level but rather require a microscopic level of simulation. This is due to highly diffusion-limited reactions and fast rebindings that the mesoscopic model cannot resolve without hitting the critical lower bound of the voxel size. This motivates development of efficient methods for simulating systems at the microscopic level.

The Green's function reaction dynamics (GFRD) algorithm[16] is an efficient method to generate accurate trajectories of the Smoluchowski model. A full $N$-body problem is divided into many 1- and 2-body problems that can be simulated independently by choosing the time step in the algorithm sufficiently small. For these subproblems the Smoluchowski equation

can be solved analytically or numerically. Improvements of GFRD have been developed in e.g. Refs. 15 and 17. Sampling random numbers from the analytical solution of the Smoluchowski equation is complicated and expensive, but in Ref. 17 this step is simplified by solving the Smoluchowski equation using first order or second order operator splitting schemes. In Ref. 15 the algorithm is made exact in both space and time by combining GFRD with the first-passage kinetic Monte Carlo (FPKMC) algorithm developed in Ref. 18, and this improved algorithm is called eGFRD. The main idea in eGFRD is to introduce protective spheres around single molecules and pairs of molecules and then sample exact times for the molecules to exit the spheres.

Software that implements eGFRD is eCell[19]. Other software packages that simulate systems with the Smoluchowski model as the target model are Smoldyn[20], MCell[21] and Spatiocyte[11]. Both Smoldyn and MCell discretize time while GFRD is exact in time, thus making them less accurate. On the other hand, Smoldyn and MCell can be significantly more efficient for larger systems. In Spatiocyte molecules move on a lattice of densely packed spheres.

Instead of simulating the system with a purely microscopic method one can combine the microscopic level and the mesoscopic level in a hybrid method[22–24]. By restricting the microscale simulations to the species or the parts of space that are necessary for capturing the important dynamics of the system, computational time can be saved without losing a significant amount of accuracy.

In addition to the observed spatial heterogeneity in systems where molecules move only by pure diffusion in free space, there can be spatial effects in systems due to interactions with internal lower dimensional structures in the cells. This could for instance be binding to and reactions on curved surfaces or one-dimensional structures such as DNA or microtubules. In Ref. 22 a method for simulating interactions with curved surfaces at the microscopic level is developed, and in Ref. 25 they develop an FPKMC method for simulating active transport due to chemical potentials on embedded lines of fixed length and with fixed configuration. A mesoscopic method that simulates interactions between molecules in free space and lines is considered in Ref. 26 and a mesoscopic method incorporating active transport is suggested in Ref. 27.

In this paper we have developed an algorithm that combines single molecule simulations in three-dimensional space with single molecule simulations on embedded curves. Molecules in some general three-dimensional volume, for instance a sphere or a cylinder, are simulated with the GFRD algorithm. They can react with the embedded curves according to the Smoluchowski model. Molecules attached to a curve can also diffuse and react with each other according to the one-dimensional Smoluchowski equation. In addition to pure diffusion, we simulate active transport along the curves by letting the molecules bound to the curves move deterministically in one direction. The curves can have an arbitrary shape. This flexibility is obtained by approximating a continuous curve with non-constant curvature by straight linear segments, and this approximation can be made arbitrarily accurate by making the discretization of the curve finer and finer. It is also known that many one-dimensional structures in cells are highly dynamic[28]. A microtubule may for instance both move in space as well as grow and shrink. We take those dynamics into account by letting the embedded curves move according to some general random process as well as letting them grow and shrink over time.

The following sections are organized as follows. In Sec. II we review the GFRD method in one dimension as well as three dimensions. In Sec. III we describe the algorithm for

coupling the one-dimensional simulations with the three-dimensional simulations, as well as simulation of moving lines embedded in space. In Sec. IV we demonstrate the flexibility of the algorithm with four examples. In the first example we verify the accuracy of the method by comparing with mesoscopic simulations and theoretical predictions. In the second example we consider two spirals embedded in a sphere. In the third example we consider several straight lines embedded in a spherical domain. Molecules can bind to the lines and are then moved by active transport along the lines. The lines move in space according to a random process, and we show how the active transport affects the distribution of the molecules in space. In the final example we consider lines that grow and shrink. We show how the number of molecules attached to a line will vary with the total length of the lines.

All parameters are given in SI-units.

## II. SIMULATIONS IN THREE AND ONE DIMENSIONS

We consider a system of $N$ molecules, where the molecules can undergo bimolecular or unimolecular reactions. The evolution of the system is modeled by the Smoluchowski equation with a mixed boundary condition at the reaction radius of the molecules. In principle one could write down the Smoluchowski equation for the full system, but due to the high dimension it would not be tractable to solve. In GFRD[16] the main idea is to divide the full $N$-body problem into smaller subproblems consisting of the simulation of single molecules and pairs of molecules. The subdivision is done in two steps. First we find the nearest neighbor of every molecule. Molecules that have each other as nearest neighbors form pairs and the remaining molecules will be updated as single molecules. Secondly we choose a time step $\Delta t$ such that no molecule in a pair is likely to react with any other molecule except its nearest neighbor and molecules that are not in a pair should be unlikely to react with any other molecule. The full $N$-body problem can now be approximated by smaller, independent, subproblems during the time step $\Delta t$.

### A. Updating positions of single and pairs of molecules

The positions of single molecules are updated by sampling from a normal distribution with standard deviation $\sqrt{2D\Delta t}$ in each direction, where $D$ is the diffusion constant of the molecule. A pair of molecules, where the molecules have the positions $\mathbf{x}_{1n}$ and $\mathbf{x}_{2n}$ at time $t_n$, are updated according to the PDF $p(\mathbf{x}_1, \mathbf{x}_2, t | \mathbf{x}_{1n}, \mathbf{x}_{2n}, t_n)$ obtained by solving the Smoluchowski equation for a pair of molecules, given by:

$$p_t = D_1 \Delta_{\mathbf{x}_1} p + D_2 \Delta_{\mathbf{x}_2} p, \tag{1}$$

where $D_i$ and $\mathbf{x}_i$ are the diffusion coefficient and position of particle $i$. By making a change of variables as suggested in Ref. 16 this equation can be split into two equations: one for the relative position of the molecules and one for a weighted mean of the positions

$$\mathbf{Y} = \sqrt{D_2/D_1}\mathbf{x}_1 + \sqrt{D_1/D_2}\mathbf{x}_2, \ \mathbf{y} = \mathbf{x}_2 - \mathbf{x}_1. \tag{2}$$

Now $p$ can be written as $p(\mathbf{x}_1, \mathbf{x}_2, t | \mathbf{x}_{1n}, \mathbf{x}_{2n}, t_n) = p_{\mathbf{Y}}(\mathbf{Y}, t | \mathbf{Y}_n, t_n) p_{\mathbf{y}}(\mathbf{y}, t | \mathbf{y}_n, t_n)$, and the equations become

$$\partial_t p_{\mathbf{Y}} = D \Delta_{\mathbf{Y}} p_{\mathbf{Y}}, \ \ \partial_t p_{\mathbf{y}} = D \Delta_{\mathbf{y}} p_{\mathbf{y}},$$

where $D = D_1 + D_2$ is the sum of the diffusion coefficients of the molecules. The motion in the $\mathbf{Y}$ direction is in free space and will be described by a normal distribution, but the equation for the $\mathbf{y}$ coordinate will have an additional boundary condition at the reaction radius of the molecules to model associations between molecules. By considering the equation in a spherical coordinate system, $\mathbf{r} = (r, \theta, \phi)$, the boundary condition becomes:

$$4\pi\sigma^2 D \frac{\partial p_{\mathbf{r}}}{\partial r}\Big|_{r=\sigma} = k_r p_{\mathbf{r}}(r = \sigma, t | r_n, t_n), \tag{3}$$

where $\sigma$ is the reaction radius of the molecules, $r = \|\mathbf{r}\|$ and $k_r$ is the intrinsic association rate. The initial condition is given by

$$p_{\mathbf{r}}(\mathbf{r}, t | \mathbf{r}_n, t_n) = \delta(\mathbf{r} - \mathbf{r}_n)$$

and $p_{\mathbf{r}}(\mathbf{r}, t | \mathbf{r}_n, t_n)$ vanishes as $r \to \infty$.

This equation can be solved analytically[16,29], but due to the complexity of the analytical solution we use the method described in Ref. 17. There the full equation is solved in two steps using a first order splitting scheme (or if higher accuracy is needed with a second order splitting scheme). This simplifies the sampling of new positions for the molecules, at the price of an additional error. This error is shown to be small in Ref. 17. The first step of the operator splitting scheme is to solve for the radial part of the equation (in a spherical coordinate system rotated such that $\mathbf{r}_n = (r_n, 0, 0)$), given by:

$$\partial_t p_r = D \left( \frac{\partial^2 p_r}{\partial^2 r} + \frac{2}{r} \frac{\partial p_r}{\partial r} \right), \tag{4}$$

with the initial condition $p_r(r, t_n | r_n, t_n) = \delta(r - r_n)$ and with the boundary condition given by Eq. (3). The analytical solution is derived in Ref. 30. The second step is to solve for the angular part:

$$\partial p_\theta = \frac{D}{r^2} \left( \frac{1}{\sin\theta} \frac{\partial}{\partial\theta} \left( \sin\theta \frac{\partial p_\theta}{\partial^2\theta} \right) + \frac{1}{\sin^2\theta} \frac{\partial^2 p_\theta}{\partial^2\phi} \right), \tag{5}$$

with initial condition given by $p_\theta(\theta, t_n | \theta_n, \phi_n, t_n) = \delta(\theta)/r^2 \sin(\theta)$. The solution to Eq. (5) will be independent of $\phi$, and the distribution of $\phi$ will therefore be uniform on the interval $[0, 2\pi)$. It is possible to solve Eq. (5) analytically and the solution is given in the form of an infinite series, where the terms are products of Legendre functions and exponentials. Efficient strategies to compute the analytical solutions of Eqs. (4) and (5) are discussed in Ref. 31. An efficient method to generate random numbers from the PDF $p_\theta$ is suggested in Ref. 32.

Note that the probability for two molecules to survive until time $t_n + \Delta t$ is given by

$$S(t_n + \Delta t | \mathbf{r}_n, t_n) = 1 - 4\pi\sigma^2 D \int_{t_n}^{t_n+\Delta t} \frac{\partial p_{\mathbf{r}}}{\partial r}\Big|_{r=\sigma} d\tau.$$

From $S$ we can sample the exact time when the pair of molecules react (cf. Ref. 16). Thus, the first step is to check whether the two molecules react before $t_n + \Delta t$. If they do, we execute the reaction at the sampled time, and if they do not react we sample new positions for each molecule by

1. Sample $r_{n+1}$ from the PDF $p_r$.

2. Sample $\theta_{n+1}$ from the PDF $p_\theta$ with $r = r_{n+1}$.

3. Sample $\phi_{n+1}$ from a uniform distribution on the interval $[0, 2\pi)$.

4. Transform $\mathbf{r}_{n+1}$ back to $\mathbf{y}_{n+1}$. Sample $\mathbf{Y}_{n+1}$ from a normal distribution and then solve Eq. (2) to obtain $\mathbf{x}_{1(n+1)}$ and $\mathbf{x}_{2(n+1)}$.

Also molecules in one dimension can react. For pairs of molecules we solve the one-dimensional Smoluchowski equation

$$\partial_t p_s = D\frac{\partial^2 p_s}{\partial s^2}, \tag{6}$$

with the boundary condition at the reaction radius

$$D\frac{\partial p_s}{\partial s}\Big|_{s=\sigma} = k_r p_s(s = \sigma, t | s_n, t_n). \tag{7}$$

to obtain the PDF $p(s, t | s_n, t_n)$ for the distance $s$ at time $t$ between the two molecules, given that the distance was $s_n$ at time $t_n$. The analytical solution to Eq. (6) with boundary condition given by Eq. (7) is derived in Ref. 29.

Both molecules in one dimension and molecules in three dimensions can dissociate. Consider a reaction $A \xrightarrow{k_d} B + C$. The time until an $A$-molecule dissociates is sampled from an exponential distribution with mean $k_d$. After a molecule $A$ has dissociated the $B$- and $C$-molecules are placed at a distance equal to the reaction radius.

## B.  Molecules close to a curved boundary

Most biologically realistic geometries are fairly complex. Cells can be close to spherical or have a shape reminding of a spheroid. Eukaryotes typically have complex internal two-dimensional structures. One flexible and efficient way to deal with such non-regular surfaces is to discretize the surface with an unstructured mesh.

In this paper we do not consider volumes with internal two-dimensional structures, and are consequently only concerned with the outer boundaries of the volume. The outer boundary is approximated by a triangular surface mesh. The vertex of each triangle is also a center vertex in the dual mesh generated from the primal mesh. Consider a center vertex $\psi_i$. This center vertex is the vertex of several triangles $T_1, \ldots, T_M$ in the mesh, and to each dual element we associate one linear plane defined by the center vertex $\psi_i$ and with normal given by

6

the average of all the normals $\mathbf{n}_1, \ldots, \mathbf{n}_M$ to the triangles $T_1, \ldots, T_M$. We have now assigned one linear plane to each dual element in the mesh, and these planes are an approximation of the surface.

Now, assume that a molecule is close to one of the planes approximating the surface. Choose a time step $\Delta t$ small enough such that the interaction between the molecule and the plane can be approximated by the interaction between an infinite flat plane and a molecule during $\Delta t$. Then consider the change of coordinate system where the molecule diffuses parallel to the plane in two directions and along the normal of the plane in the third direction. The molecule can then be updated according to normal diffusion in the first two directions and in the third direction by sampling from a normal distribution and then, in case the molecule ends up on the wrong side of the plane, simply reflecting that coordinate in the plane.

This approach gives great flexibility to the algorithm and enables us to simulate molecules in different complicated geometries. The details of this approach is described in Ref. 22. Another method to simulate molecules close to an absorbing or reflecting boundary has been suggested in Ref. 33.

Note that by making a locally planar approximation we can also include reactions between molecules and surfaces by solving the one-dimensional Smoluchowski equation for molecules that are close to the surface, as done in Ref. 22. In this paper, however, we only consider reflecting outer boundaries, but the methodology for simulating interactions with reactive surfaces (outer boundaries as well as internal surfaces) developed in Ref. 22 can be combined with the method described in this paper. Furthermore, we could also combine this method with the hybrid method coupling microscopic and mesoscopic simulation algorithms, also developed in Ref. 22.

## III.   EMBEDDED ONE-DIMENSIONAL MANIFOLDS

We consider a continuous and differentiable curve $\Gamma$ embedded in space, parametrized by

$$\Gamma(s) = (a(s), b(s), c(s)), \, 0 \leq s \leq 1.$$

$\Gamma$ can be approximated by a piecewise linear curve $\gamma$. Denote the linear segments of $\gamma$ by $\gamma_1, \ldots, \gamma_N$. The segment $\gamma_i$ is defined by the points $\mathbf{p}_i^1$ and $\mathbf{p}_i^2$. The points are chosen such that $\mathbf{p}_i^1 = \mathbf{p}_{i-1}^2$ and thus all segments are connected. The curve $\Gamma$ can be approximated to arbitrary accuracy by letting $N \to \infty$ in such a way that $\|\mathbf{p}_i^1 - \mathbf{p}_i^2\| \to 0$.

We want to simulate reactions between molecules in space and the curve as well as diffusion and reactions between molecules on the curve. Furthermore, we want to simulate moving curves that grow and shrink.

### A.   Interactions with the curve

If a molecule $A$ is close to $\gamma$, we will choose a time step $\Delta t$ such that the molecule is unlikely to interact with more than one of the linear segments during that time step, and such that the linear segment can be approximated, to high accuracy, by an infinitely long straight line during $\Delta t$. The problem of simulating a molecule close to an infinitely long straight line can be reduced to a two-dimensional problem, which can be seen by considering

the position of the molecule at time $t_n$, $\mathbf{x}_n$, in cylindrical coordinates $\mathbf{r} = (r, \theta, z)$, with the $z$-coordinate in the direction of the line. The diffusion of the molecule in the $z$-coordinate is now independent of the diffusion in the $r$ and $\theta$ directions and the interaction with the line.

Assume that the closest linear segment to the molecule is $\gamma_i$. Now consider the sphere $S$ with center at $\mathbf{x}_n$ and radius $R = \min\{\|\mathbf{x}_n - \mathbf{p}_i^1\|, \|\mathbf{x}_n - \mathbf{p}_i^2\|\}$. By choosing the time step $\Delta t$ such that the molecule $A$ is unlikely to diffuse outside of the sphere $S$ during the time step, the requirement that the molecule is unlikely to react with more than one linear segment is satisfied. If the diffusion coefficient of $A$ is $D_A$, the molecule will, on average, diffuse the distance $\sqrt{6 D_A \Delta t}$ during a time step $\Delta t$. Thus, we choose the time step to be

$$\Delta t = \frac{R^2}{K \cdot 6 D_A},$$

where $K$ is a constant chosen sufficiently large in order to ensure that the probability for the molecule to diffuse outside of the sphere is small. Note that if the molecule is really close to $\mathbf{p}_i^1$ or $\mathbf{p}_i^2$ the time step could become very small, making the simulation too slow. To avoid this problem we have to choose a smallest value that $\Delta t$ can take, thus introducing an error. However, with a carefully chosen smallest value of $\Delta t$ that error will be small.

The interaction between the molecule and the linear segment can now be approximated by the two-dimensional Smoluchowski equation during $\Delta t$. The two-dimensional Smoluchowski equation for the relative position (in cylindrical coordinates) is given by:

$$p_t = D \Delta p, \tag{8}$$

with the boundary condition

$$2 \pi \sigma D \frac{\partial p_{\mathbf{r}}}{\partial r} \bigg|_{r=\sigma} = k_r p_{\mathbf{r}}(r = \sigma, t | \mathbf{r}_n, t_n). \tag{9}$$

Again we solve the equation in two steps with a first order operator splitting scheme. The radial part is given by:

$$\partial_t p_r = D \left( \frac{\partial^2 p_r}{\partial r^2} + \frac{1}{r} \frac{\partial p_r}{\partial r} \right) \tag{10}$$

with the boundary condition given by Eq. (9), the initial condition given by $p_r(r, t_n | r_n, t_n) = \delta(r - r_n)$ and with $p_r$ vanishing at infinity. The solution of Eq. (10) with the boundary condition given by Eq. (9) is solved analytically in Ref. 29. The analytical solution is

$$p_r(r, t | r_0, t_n) = \frac{1}{2\pi} \int_0^\infty \exp(-D u^2 (t - t_n)) C(u, r, k, k_r) C(u, r_0, k, k_r) u \, du,$$

where $k = 2\pi\sigma D$ and the function $C$ is defined by

$$C(u, r, k, h) = \frac{J_0(ur)(kuY_1(\sigma u) + hY_0(\sigma u)) - Y_0(ur)(kuJ_1(\sigma u) + hJ_0(\sigma u))}{((kuY_1(\sigma u) + hY_0(\sigma u))^2 + (kuJ_1(\sigma u) + hJ_0(\sigma u))^2)^{\frac{1}{2}}}.$$

Here $J_0$ and $J_1$ are Bessel functions of the first kind and $Y_0$ and $Y_1$ are Bessel functions of the second kind. Despite having the analytical solution available we choose to compute

8

the solution using a finite difference method, as the analytical solution is fairly complex and expensive to evaluate. The solution could also be tabulated as described in Ref. 31. Tabulating the solution is more efficient but has the drawback of introducing an additional error as well as consuming more memory.

The second step is to solve for $\theta$ and $z$

$$\partial p_{\theta z} = D \left( \frac{1}{r^2} \frac{\partial^2 p_{\theta z}}{\partial \theta^2} + \frac{\partial^2 p_{\theta z}}{\partial z^2} \right), \tag{11}$$

with initial condition

$$p_{\theta z}(\theta, z | 0, z_n, t_n) = \frac{\delta(\theta)\delta(z - z_n)}{r}.$$

Again, $p$ vanishes at infinity. The solution of Eq. (11) is a Gaussian in both the $z$-direction and the $\theta$-direction.

Molecules can also dissociate from the lines. The time until a dissociation is assumed to be exponentially distributed with mean equal to the intrinsic dissociation rate, and after a dissociation from a line the molecule is placed at a distance equal to the reaction radius between the molecule and the line. As we are using a finite difference scheme to compute the solution of Eq. (10) we are in fact forced to place the dissociating molecule at a slightly larger distance from the line than the reaction radius, in order to have discretization points between the molecule and the line. This introduces a small error, but it can be made arbitrarily small by placing the molecule closer and closer to the line while increasing the number of discretization points used in the finite difference scheme.

## B.  Diffusion and reactions on moving curves

A new position at time $t + \Delta t$, $\mathbf{x}_{n+1}$, of a single molecule moving by pure diffusion, bound to $\Gamma$ at position $\mathbf{x}_n$ at time $t$, where $\mathbf{x}_n \in \gamma_i$, is computed by sampling a distance $d$ from a one-dimensional normal distribution with standard deviation $\sqrt{2D\Delta t}$ and then updating

$$\mathbf{x}_{n+1} = \mathbf{x}_n + d \frac{(\mathbf{p}_i^2 - \mathbf{p}_i^1)}{\|\mathbf{p}_i^2 - \mathbf{p}_i^1\|}. \tag{12}$$

It is possible that $\mathbf{x}_{n+1}$ is no longer on the segment $\gamma_i$. If this is the case we project $\mathbf{x}_{n+1}$ to the closest point on $\gamma$. The time step $\Delta t$ should thus be chosen such that $\sqrt{2D\Delta t}$ is small in comparison to $\|\mathbf{p}_i^2 - \mathbf{p}_i^2\|$, so that the molecules do not diffuse a long distance compared to the distance of the segments. Since we are projecting the molecule onto the line we are introducing a small error and the effective diffusion will be slightly slower than the exact diffusion process. However, if the angles between the linear segments are small, this error will also be small, and it is therefore important to ensure that the discretization of $\Gamma$ is fine enough to make the angles small enough. Note that the process defined by Eq. (12) can be modified to account for other types of motion on curves, such as active transport. For instance, $d$ could be sampled from another distribution than the normal distribution, or be completely deterministic. Molecules on the same line segment can react according to the PDF obtained by solving Eq. (6) with boundary condition given by Eq. (7).

Many one-dimensional structures in cells are not stationary, but rather move around in

9

space. They can also grow and shrink. We simulate this process with an operator splitting scheme. Choose a time step $\Delta t_{\text{split}}$. First we keep the lines constant in space and propagate the molecules according to the algorithm described above for the full time step $\Delta t_{\text{split}}$. The second step is to move the lines according to some transformation $T(\mathbf{p}, t) : \mathbb{R}^3 \times \mathbb{R}_+ \to \mathbb{R}^3$ in space, for the full time step $\Delta t_{\text{split}}$. $T$ can both move the line in space but also change the length of the line. We apply $T$ to all the points, such that the linear segments at time $t + \Delta t_{\text{split}}$ are defined by the pair of points

$$(T(\mathbf{p}_1^1, \Delta t_{\text{split}}), T(\mathbf{p}_1^2, \Delta t_{\text{split}})), \ldots, (T(\mathbf{p}_N^1, \Delta t_{\text{split}}), T(\mathbf{p}_N^2, \Delta t_{\text{split}})).$$

The position of the molecules are now no longer guaranteed to be on the lines. Each molecule is therefore projected down to the closest point on its respective line. For high accuracy we should therefore choose the time step $\Delta t_{\text{split}}$ such that each molecule is projected only a short distance. After each step it is possible that some molecules on the line that were previously separated by some small distance are now overlapping. If this is the case they are moved apart and placed at a distance equal to their reaction radius. Molecules in space that overlap with the lines are also moved such that they are at a distance from the line equal to their reaction radius. We then repeat this process until the final time $t_f$. If higher accuracy is needed it would be possible to consider higher order splitting schemes, such as Strang splitting[34].

The algorithm outlined above is summarized in Algorithm 1. Note that this is only a summarized version of the GFRD algorithm. There are some special cases that require extra care, such as pairs of molecules that are also close to a boundary. For a more detailed description of how to handle such cases the reader is referred to Refs. 15–17.


## IV. NUMERICAL RESULTS

We show the practical use of the algorithm by studying both examples with straight lines through a three-dimensional domain as well as lines with non-constant curvature. We demonstrate how to use the algorithm to simulate active transport by letting molecules attached to a line move with a deterministic speed in a predetermined direction. Finally, we let the lines move in space as well as grow and shrink.


### A. Polymer with road blocks embedded in a cylinder

We consider a cylinder with radius $R = 10^{-6}$, height $H = 2 \cdot 10^{-6}$ and aligned along the x-axis. Inside the cylinder we consider a line of length $H$, for instance modeling a polymer, going through the center of the cylinder. The polymer has a reaction radius $r_l = 10^{-9}$. Molecules of species $A$ can associate to the polymer to form molecules $A_{\text{cyl}}$. Molecules $A_{\text{cyl}}$ can dissociate from the polymer to become species $A$. Thus, we consider the reaction

$$A \underset{k_d}{\overset{k_r}{\rightleftharpoons}} A_{\text{cyl}}.$$
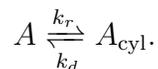
where $A$ is a molecule diffusing in three-dimensional space and $A_{\text{cyl}}$ is a molecule bound to the polymer. The diffusion constant of $A$ is $D_A = 10^{-12}$ and the diffusion constant of $A_{\text{cyl}}$ is $D_{A_{\text{cyl}}} = 10^{-14}$. The intrinsic reaction rates are chosen to be $k_r = 10^{-11}$ and $k_d = 50$. The

---

ALGORITHM 1: GFRD with moving one dimensional structures

---

1: Initialize the system at the time $t = 0$ and choose a final time $t_f$. Choose a time step $\Delta t_{\text{split}}$ for the operator splitting.

2: Divide the set of all molecules into subsets of one or two molecules such that pairs are formed by molecules that are each others nearest neighbor. Set $t_{\text{loc}} = 0$.

3: **while** $t_{\text{loc}} < \Delta t_{\text{split}}$ **do**

4:     Choose a time step $\Delta t \leq \Delta t_{\text{split}} - t_{\text{loc}}$ such that molecules are unlikely to interact with more than one other molecule during that time step. $\Delta t$ should also be chosen such that no molecule is likely to interact with a two-dimensional boundary or a one-dimensional curve and a neighbor during the time step. To fulfill this requirement certain pairs and single molecules may have to be updated in several smaller steps during $\Delta t$.

5:     Pairs of molecules are updated according to the PDF obtained by solving the three-dimensional Smoluchowski equation (1), as described in Sec. II A.

6:     Molecules close to a line are updated according to the PDF obtained by solving the two-dimensional Smoluchowski equation (10), as described in Sec. III A.

7:     Molecules on a line are updated according to the PDF obtained by solving the one-dimensional Smoluchowski equation (6), as described in Sec. II A.

8:     Molecules that dissociate from the cylinder at time $\tau < \Delta t$ are placed at the reaction radius of the molecule and the cylinder. If a molecule in free space dissociates we place the products at a distance equal to the sum of the reaction radii of the two molecules.

9:     Update $t_{\text{loc}} = t_{\text{loc}} + \Delta t$.

10: **end while**

11: Update the linear segments according to the transformation $T$.

12: Update the time $t = t + \Delta t_{\text{split}}$.

13: Repeat (2)-(12) until the final time $t_f$.

---

problem setup is shown in Fig. 1.

Due to the cylindrical symmetry of the problem, the reactions with the polymer can be reduced to a two-dimensional problem. The average time $T_{\text{bind}}$ until a molecule $A$ reacts with the polymer can thus be computed using a formula for association times on disks, derived in Ref. 10. We can therefore compare the microscopic simulation with a corresponding mesoscopic simulation of the reaction

$$A \underset{k_d^{\text{meso}}}{\overset{k_r^{\text{meso}}}{\rightleftharpoons}} A_{\text{cyl}}$$
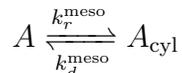
where $k_r^{\text{meso}} = 1/T_{\text{bind}}$ and, due to the detailed balance condition, $k_d^{\text{meso}} = k_d k_r^{\text{meso}}/(A_{\text{disk}} k_r)$. Here $A_{\text{disk}} = \pi R^2$ is the area of the face of the cylinder. In Fig. 2 we show that for this simple problem, the microscopic simulations agree well with spatially homogeneous mesoscopic simulations, performed with the SSA.

Next we consider a more involved example. The geometrical setup is the same as above, but now we let the molecules of species $A_{\text{cyl}}$ react with an operator site at the center of the line, thus considering the line a coarse model of DNA. The operator site is denoted by $B$
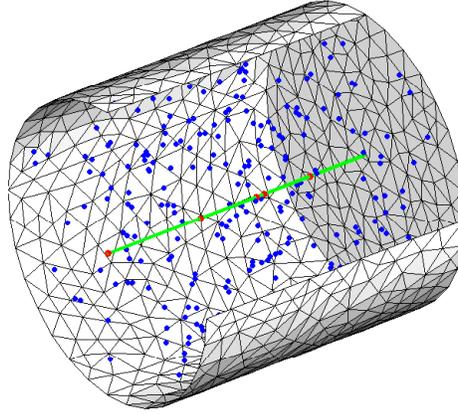
FIG. 1: The problem setup. Molecules of species $A$ (blue dots) diffuse in the cylindrical volume. They react with the line and transforms into the one-dimensional species $A_{\text{cyl}}$ (red dots). The line (green) extends throughout the domain through the center of the cylinder. The outer boundary is approximated by piecewise linear planes and is reflective.
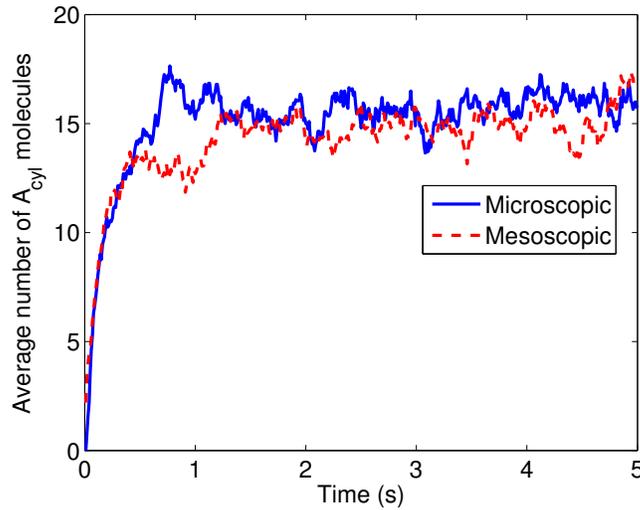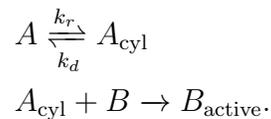


FIG. 2: The average number of $A_{\text{cyl}}$ at the microscopic level and the mesoscopic level. The mesoscopic simulation agree well with the microscopic simulation. The average is taken over 20 trajectories and we start the simulation with 250 $A$-molecules uniformly distributed in the domain.

and we consider the reactions

$$A \underset{k_d}{\overset{k_r}{\rightleftharpoons}} A_{\text{cyl}}$$
$$A_{\text{cyl}} + B \rightarrow B_{\text{active}}.$$

12
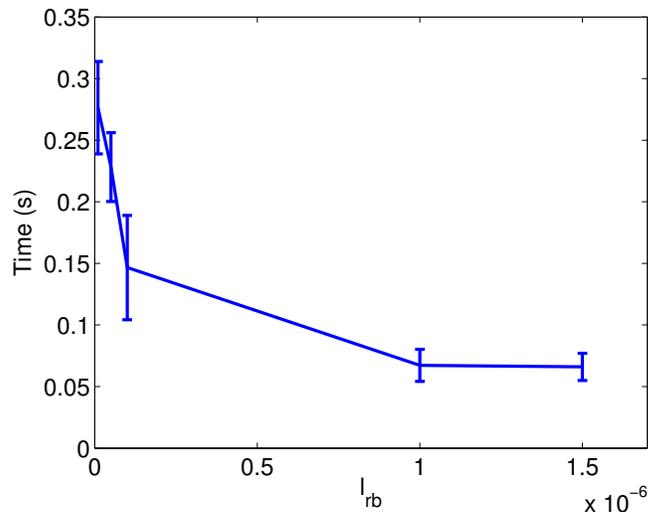
FIG. 3: The average time until an $A$-molecule binds to the operator site decreases with the distance between the road blocks. As the distance becomes greater than the average sliding distance of an $A_{\text{cyl}}$, the decrease levels off. The parameters in the simulations are $k_r = 10^{-11}$, $k_d = 50$ and $D = 10^{-12}$ for all species. The simulation is started with 250 $A$-molecules in free space.

We assume that the operator site is absorbing and of radius $10^{-9}$. On both sides, at the distance $l_{\text{rb}}$ from the operator site, we place a road block. Road blocks are modeled by stationary, non-reactive (but reflective) molecules of radius $10^{-9}$. Thus, molecules binding to the DNA at a distance larger than $l_{\text{rb}}$ from the operator site, will not be able to find it since molecules cannot diffuse through the road blocks. In Ref. 26 a similar problem is simulated with a mesoscopic method. The microscopic parameters cannot easily be transformed to corresponding mesoscopic parameters, but the qualitative behaviour can be compared.

We compute the average time, $T_{\text{active}}$, until an $A_{\text{cyl}}$-molecule binds to the operator site, and this time will depend on the length $l_{\text{rb}}$ between the road blocks. In Fig. 3 we show that this time decreases with increasing length between the road blocks, as was shown in Refs. 26 and 35. The decrease in the binding time levels off as $l_{\text{rb}}$ becomes greater than the average distance that a molecule slides on the DNA before dissociating.

## B. Spirals embedded in a sphere

One-dimensional structures in living cells are often not well approximated by straight lines due to their curvature, and they may also have a complicated spatial distribution relative to each other. It is therefore important to be able to simulate molecules reacting with and on lines of non-constant curvature. To demonstrate the flexibility of the method
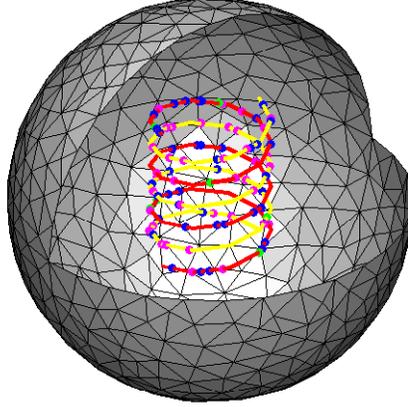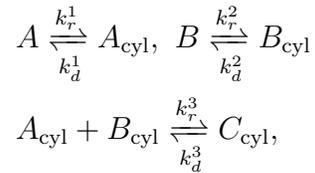
13

FIG. 4: We have simulated association to and dissociation from two spirals (yellow and red lines) embedded in a sphere. Molecules in space diffuse and react with the spirals, and molecules on the spirals diffuse and react with each other. Note that the molecules that are free in space are not plotted in the figure above.

we therefore consider the following system of reactions

$$A \underset{k_d^1}{\overset{k_r^1}{\rightleftharpoons}} A_{\text{cyl}}, \ B \underset{k_d^2}{\overset{k_r^2}{\rightleftharpoons}} B_{\text{cyl}}$$

$$A_{\text{cyl}} + B_{\text{cyl}} \underset{k_d^3}{\overset{k_r^3}{\rightleftharpoons}} C_{\text{cyl}},$$

where $A_{\text{cyl}}$, $B_{\text{cyl}}$ and $C_{\text{cyl}}$ are molecules bound to one of the lines and $A$ and $B$ are molecules diffusing in three-dimensional space that can react with the lines. We will consider two spirals $\Gamma_1$ and $\Gamma_2$ along the y-axis, parameterized by

$$\Gamma_1(a(s), b(s), c(s)) = (-3 \cdot 10^{-7}, 0, 0) + (2.5 \cdot 10^{-8}s, r_c \cos(2\pi s), r_c \sin(2\pi s))$$
$$\Gamma_2(a(s), b(s), c(s)) = (-3 \cdot 10^{-7}, 0, 0) + (2.5 \cdot 10^{-8}s, -r_c \cos(2\pi s), -r_c \sin(2\pi s)),$$

with $0 \leq s \leq 3$ and $r_c = 3 \cdot 10^{-7}$. Each spiral is divided into 30 linear segments. The setup is shown in Fig. 4.

We start with 150 molecules of species $A$ and 150 molecules of species $B$, with initial positions sampled from a uniform distribution. The reactions rates are chosen to be $k_r^1 = k_r^2 = 10^{-11}$, $k_d^1 = k_d^2 = 50$, $k_r^3 = 10^{-6}$ and $k_d^3 = 20$. The diffusion rate for the molecules in space is $10^{-12}$ and the diffusion rate for the molecules on the spirals is $10^{-14}$. We simulate the system for 2 seconds and the evolution of the system is shown in Fig. 5.

## C. Active transport on moving microtubules

Molecules in living cells do not only move by normal diffusion but can also be actively transported. For instance, intracellular transport on the cytoskeleton, a structure in the cell composed of fibers such as actin and microtubules, is an important mechanism for directed transport of cargo in the cell[36]. Cargo attach to the fibers and are then transported by
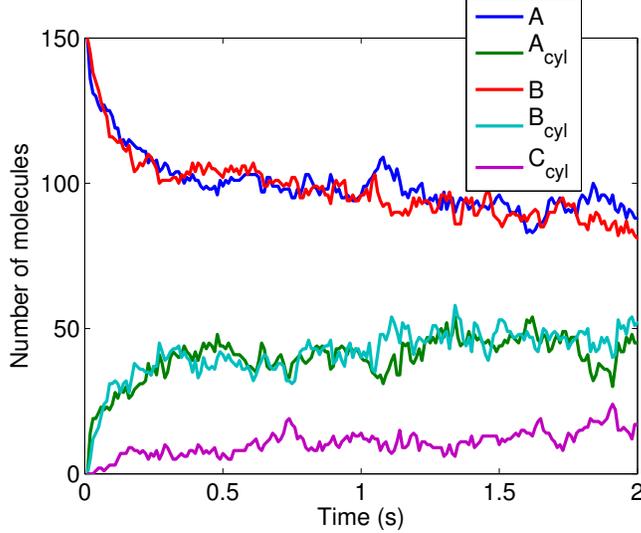
FIG. 5: The time evolution of the total number of molecules of the different species in the system.

molecular motors: myosins along actin and dynein and kinesins along microtubules.

We will consider a much simplified model of active transport from the cytosol towards the center of a cell. Molecules bind to microtubules, modeled by straight lines passing through the origin. We assume that the transport is of constant speed in the direction towards the origin and we do not explicitly model the molecular motors. Other more complex and more biologically accurate models of active transport could be incorporated into the method by modifying Eq. (12) in a suitable way. The model used in this example is not meant to be a very accurate model of such processes, but is rather meant to serve as an example of the flexibility of the algorithm.

We initialize the system by sampling 25 points, $\mathbf{p}_1, \ldots, \mathbf{p}_{25}$, uniformly on a sphere of radius $R = 10^{-6}$. 25 straight line segments through the origin are then defined by the pairs $(\mathbf{p}_i, -\mathbf{p}_i)$, $i = 1, \ldots, 25$. We start with 100 $A$-molecules in space, with initial positions sampled from a uniform distribution. The $A$-molecules can then associate and dissociate to the microtubules through the reaction $A \underset{k_d}{\overset{k_r}{\rightleftharpoons}} A_{\text{cyl}}$. $A$-molecules in space diffuse with the diffusion constant $D = 10^{-12}$, but molecules on the microtubules move deterministically towards the center of the sphere with constant speed $\sqrt{2 \cdot 10^{-14} \Delta t}/40$. The microtubules are rotated $\Delta\theta$ and $\Delta\phi$ radians around the x- and y-axis per time step $\Delta t$, where

$$\begin{aligned} \Delta\theta &= 2\pi\Delta t X_1 \\ \Delta\phi &= 2\pi\Delta t X_2. \end{aligned} \tag{13}$$

$X_1$ and $X_2$ are random numbers sampled from a normal distribution with standard deviation 1. $\Delta t_{\text{split}}$ is chosen to be 0.01. Thus, the microtubule $l_i$ defined by the points $(\mathbf{p}_i, -\mathbf{p}_i)$ at time $t$, will be updated at time $t+\Delta t$ by rotating $p_i$ and $-p_i$ around the x- and y-axis $\Delta\theta$ and $\Delta\phi$ degrees, with $\Delta\theta$ and $\Delta\phi$ given by Eq. (13). The trajectory of one of the microtubules is plotted in Fig. 6 and two snapshots of a simulation, one at $t_1 = 0.05s$ and one at the final time $t_f = 1.0s$ are shown in Fig. 7.

In Fig. 8 the distribution of molecules in the radial direction is plotted. The initial distribution is uniform in space. Due to the transport towards the center, this distribution
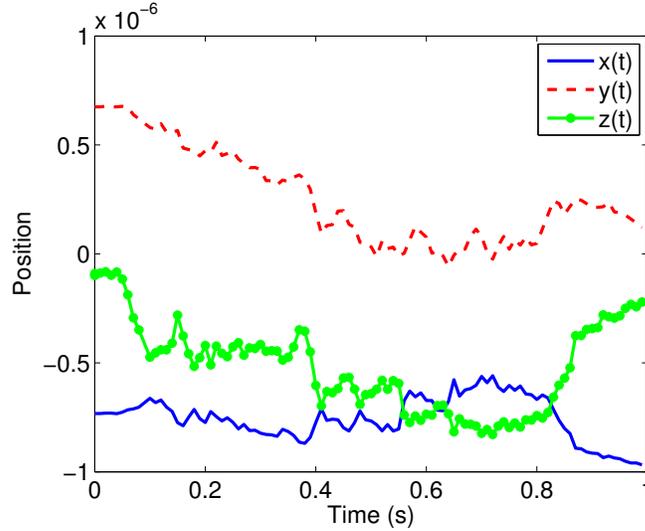
15

FIG. 6: The trajectory of one of the microtubules. The position of the $x$-, $y$- and $z$-coordinates of one of the endpoints of the microtubule is plotted. The microtubule is rotated according to Eq. (13).


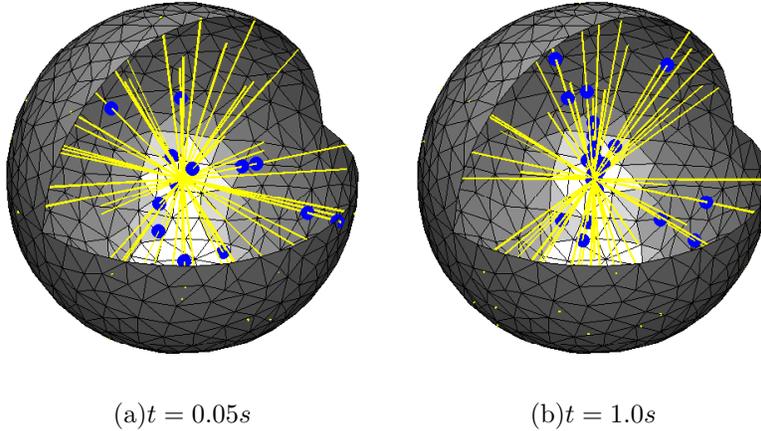
(a)$t = 0.05s$          (b)$t = 1.0s$

FIG. 7: Lines are yellow and the blue dots are molecules bound to the lines. Molecules in free space are not plotted. Molecules on the lines move by active transport towards the center of the sphere and the lines move in space according to Eq. (13).

is shifted and more molecules are close to the center of the sphere at the final time $t_f = 1.0s$. The reaction rates in the simulation are $k_r = 10^{-11}$ and $k_d = 10^3$.

## D. Growing and shrinking microtubules

Microtubules are highly dynamic and can, in addition to moving around in space, both grow and shrink. Therefore, as a final example, we will consider fibers embedded in a sphere and whose lengths can both increase and decrease.

The sphere has a radius $R = 10^{-6}$, with two straight fibers passing through the origin. The fibers are initially of length $l = 2R$. The lines rotate in the same way as in Sec. IV C
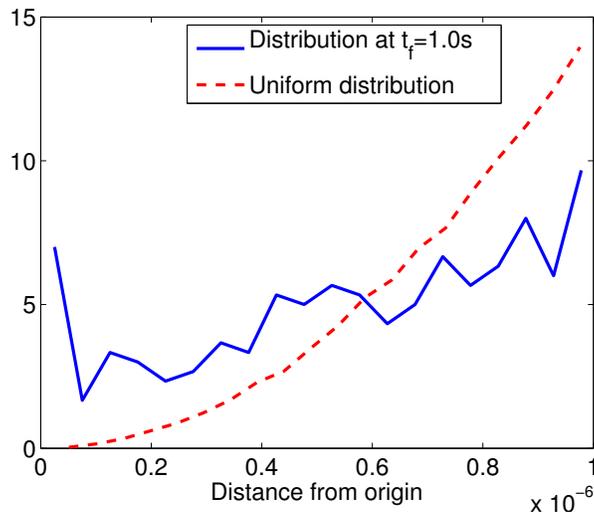
16

FIG. 8: The radial distribution of the positions of the molecules. The molecules have a uniform initial distribution in the sphere (dashed line). The distribution at the final time $t_f = 1.0s$ is shifted towards the center of the sphere. The distribution at the final time, plotted in the figure above, is based on the mean of three trajectories.

but will also grow and shrink. After each time step, the length of the line is updated, and molecules are projected down to the closest point on the line. Note that the closest point could in fact be the endpoint of the line if the molecule was close to the end of the line before the length was updated. One possibility is to let such molecules be transformed to molecules in free space, but in this example they will stay bound to the line. Now, assume that the length of the line at time $t_n$ is given by $l(t_n)$. Then the length at $t_n + \Delta t$ is given by

$$l(t_n + \Delta t) = l(t_n) + X,$$

where X is a random number sampled from a normal distribution with mean 0 and standard deviation $\sqrt{2D_l \Delta t}$, with the restrictions that $l(t_n + \Delta t) < 2R$ and $l(t_n + \Delta t) > 10^{-8}$. Thus, if the line is defined by the points $(\mathbf{p}_n, -\mathbf{p}_n)$ at time $t_n$, it will be defined by the points $l(t_n + \Delta t)/l(t_n) \cdot (\mathbf{p}_n, -\mathbf{p}_n)$ at time $t_{n+1}$.

Intuitively one would expect that when the length of the fibers decrease, the number of bound molecules will also decrease, under the condition that the fibers grow and shrink at a time scale that is longer than or similar to the time scale for the system to reach steady state. Thus, if we choose parameters such that this condition is fulfilled, we would expect to see a correlation between the number of molecules of species $A_{\mathrm{cyl}}$ and the total length of the two lines.

We start the simulation with 600 molecules of species $A$ uniformly distributed in space. The diffusion constants are $D_A = 10^{-12}$ and $D_{A_{\mathrm{cyl}}} = 10^{-14}$. In Fig. 9 we plot the total length of the two lines and the total number of molecules bound to a line. In Fig. 9 a trajectory of the system, with $D_l = 10^{-12}$, is plotted together with the sum of the length of the lines. As expected the two trajectories are clearly correlated.
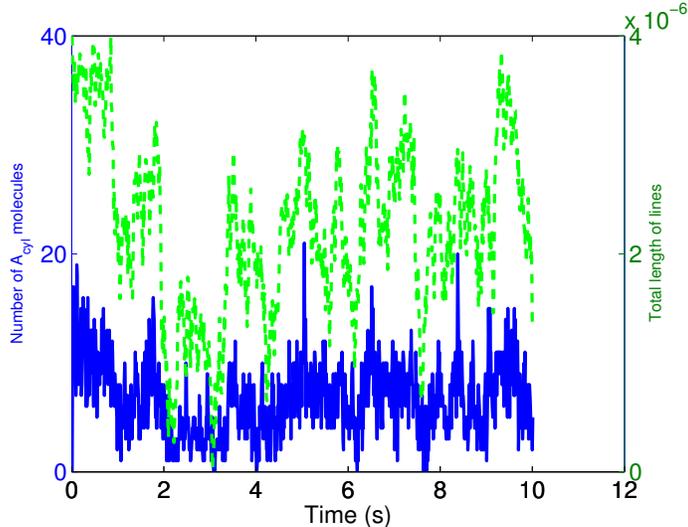
FIG. 9: The number of $A_{\text{cyl}}$-molecules is plotted together with the sum of the length of the two lines. The two time series are correlated with correlation coefficient 0.713 and a 95% confidence interval of the correlation coefficient is given by $[0.669, 0.752]$.

## V. CONCLUSIONS

We have described an algorithm to simulate biochemical systems in complex geometries with embedded and dynamic one-dimensional manifolds. Molecules in space can associate to and dissociate from the embedded curves. Molecules in space and on the lines are simulated using the GFRD algorithm.

The accuracy of the algorithm depends on the discretization of the lines and the time step chosen in the algorithm. The simulation of the molecules and the spatial transformation of the lines are split in time according to a first order operator splitting scheme and the error will therefore depend on the time step chosen for this splitting, as well as the curvature of the lines.

We have demonstrated the flexibility of this method in four numerical examples. First we have shown that the time until a molecule binds to an operator site on DNA depends on the length between two road blocks on each side of the operator site. As the length increases, the time until a molecule binds decreases. The decrease levels off as the distance gets bigger than the average sliding distance of a molecule on the DNA, in agreement with Refs. 26 and 35. In the second example we consider two spirals embedded in a sphere. Molecules in space undergo a reversible reaction with the spirals and molecules on the spirals also react reversibly. In a third example we have considered a simple model of active transport. Molecules in a sphere react reversibly with embedded lines and move deterministically towards the center, while bound. The lines also move in space and we have shown how the radial distribution of the molecules is shifted due to the active transport. In a final example we have simulated lines that are growing and shrinking in space and shown how the dynamics of the lines correlates to the concentration of the molecules in the system.

## VI.   ACKNOWLEDGMENTS

## REFERENCES

[1] H. H. McAdams and A. Arkin, Trends Gen. **15**, 65 (1999).

[2] J. Elf and M. Ehrenberg, Syst. Biol. **1**, 230 (2004).

[3] D. T. Gillespie, J. Comput. Phys. **22**, 403 (1976).

[4] J. Hattne, D. Fange, and J. Elf, Bioinformatics **21**, 2923 (2005).

[5] B. Drawert, S. Engblom, and A. Hellander, BMC Syst. Biol. **76**, 1 (2012).

[6] S. Wils and E. D. Schutter, Front. Neuroinf. **3**, 15 (2009).

[7] S. A. Isaacson, SIAM J. Appl. Math. **70**, 77 (2009).

[8] S. Hellander, A. Hellander, and L. Petzold, Phys. Rev. E. **85** (2012).

[9] R. Erban and J. Chapman, Phys. Biol. **6**, 046001 (2009).

[10] D. Fange, O. G. Berg, P. Sjöberg, and J. Elf, Proc. Natl. Acad. Sci. USA **107**, 19820 (2010).

[11] S. N. V. Arjunan and M. Tomita, Syst. Synth. Biol. **4**, 35 (2010).

[12] M. v. Smoluchowski, Z. phys. Chemie **92**, 129 (1917).

[13] M. Doi, Journal of Physics A: Mathematical and General **9**, 1465 (1976).

[14] M. Doi, Journal of Physics A: Mathematical and General **9**, 1479 (1976).

[15] K. Takahashi, S. Tănase-Nicola, and P. R. ten Wolde, Proc. Natl. Acad. Sci. USA. **107**, 2473 (2010).

[16] J. S. van Zon and P. R. ten Wolde, J. Chem. Phys. **123**, 234910 (2005).

[17] S. Hellander and P. Lötstedt, J. Comput. Phys. **230**, 3948 (2011).

[18] A. Donev, V. V. Bulatov, T. Oppelstrup, G. H. Gilmer, B. Sadigh, and M. H. Kalos, J. Comput. Phys **229**, 3214 (2010).

[19] M. Tomita, K. Hashimoto, K. Takahashi, T. S. Shimizu, Y. Matsuzaki, F. Miyoshi, K. Saito, S. Tanida, K. Yugi, J. C. Venter, and C. A. H. 3rd, Bioinformatics **15**, 72 (1999).

[20] S. S. Andrews and D. Bray, Phys. Biol. **1**, 137 (2004).

[21] R. A. Kerr, T. M. Bartol, B. Kaminsky, M. Dittrich, J.-C. J. Chang, S. B. Baden, T. J. Sejnowski, and J. R. Stiles, SIAM J. Sci. Comput. **30**, 3126 (2008).

[22] A. Hellander, S. Hellander, and P. Lötstedt, Multiscale Model. Simul. **10**, 585 (2012).

[23] M. Klann, A. Ganguly, and H. Koeppl, Bioinformatics **28**, 549 (2012).

[24] M. B. Flegg, S. J. Chapman, and R. Erban, J. R. Soc. Interface **9**, 859 (2011).

[25] A. J. Mauro, J. K. Sigurdsson, J. Shrake, P. J. Atzberger, and S. A. Isaacson, ArXiv e-prints (2013), arXiv:1302.0793 [math.NA].

[26] S. Wang, J. Elf, S. Hellander, and P. Lötstedt, "Stochastic reaction-diffusion processes with embedded lower dimensional structures," Tech. Rep. 2012-034 (Department of Information Technology, Uppsala University, 2012).

[27] A. Hellander and P. Lötstedt, Multiscale Model. Simul. **5**, 1691 (2010).

[28] T. Mitchison and M. Kirschner, Nature **312**, 237 (1984).

[29] J. C. J. H. S. Carslaw, *Conduction of Heat in Solids* (Oxford University Press, 1959).

[30] H. Kim and K. J. Shin, Phys. Rev. Lett. **82**, 1578 (1999).

[31] M. H. Bani-Hashemian, S. Hellander, and P. Lötstedt, Commun. Comput. Phys. **13**, 958 (2013).

[32] T. Carlsson, T. Ekholm, and C. Elvingson, J. Phys. A: Math. Theor. **43**, 505001 (2010).

[33] E. A. J. F. Peters and T. M. A. O. M. Barenbrug, Phys. Rev. E **66**, 056701 (2002).

[34] G. Strang, SIAM J. Numer. Anal. **5**, 506 (1968).

[35] P. Hammar, P. Leroy, A. Mahmutovic, E. G. Marklund, O. G. Berg, and J. Elf, Science **336**, 1595 (2012).

[36] J. Howard, Annu. Rev. Physiol. **58**, 703 (1996).