A Diffusion-Based Embedding of the Stochastic Simulation Algorithm in Continuous Space

Marcus Thomas and Russell Schwartz*
Computational Biology Department, Carnegie Mellon University
(Dated: May 21, 2021)

A variety of simulation methodologies have been used for modeling reaction-diffusion dynamics — including approaches based on Differential Equations (DE), the Stochastic Simulation Algorithm (SSA), Brownian Dynamics (BD), Green's Function Reaction Dynamics (GFRD), and variations thereon — each offering tradeoffs with respect to the ranges of phenomena they can model, their computational tractability, and the difficulty of fitting them to experimental measurements. Here, we develop a multiscale approach combining efficient SSA-like sampling suitable for well-mixed systems with aspects of the slower but space-aware GFRD model, assuming as with GFRD that reactions occur in a spatially heterogeneous environment that must be explicitly modeled. Our method extends the SSA approach in two major ways. First, we sample bimolecular association reactions following diffusive motion with a time-dependent reaction propensity. Second, reaction locations are sampled from within overlapping diffusion spheres describing the spatial probability densities of individual reactants. We show the approach to provide efficient simulation of spatially heterogeneous biochemistry in comparison to alternative methods via application to a Michaelis-Menten model.

arXiv:2004.01810v4 [physics.chem-ph] 20 May 2021

I. INTRODUCTION

Simulation methods have become a valuable adjunct to experimental work, facilitating the interpretation of experimental data and inferences about experimentally unobservable aspects of biomolecular dynamics [1], yet accurate simulations remain challenging for many biochemical processes crucial to living systems. The need for improvements in simulation technology is particularly acute for macromolecular assembly systems, which are central to nearly all cellular processes, yet frequently not directly observable experimentally due to their small scale and rapid dynamics [2]. Intractability of experimental approaches is particularly acute for understanding self-assembly *in vivo*, which may operate quite differently from purified *in vitro* models due to such effects as spatial confinement [3–5], macromolecular crowding [6, 7], and influences of extrinsic cellular factors [8]. The challenges of developing simulations that are both accurate and efficient, especially for hard-to-model systems like self-assembly, has led to extensive work on models and algorithms for biochemical simulation seeking to balance computational efficiency with fidelity to the complexity of the underlying biology.

The Gillespie Stochastic Simulation Algorithm (SSA) [9, 10] was particularly influential in establishing a computational framework for efficient sampling of chemical reaction trajectories, especially for small copy-number settings typical of biochemistry in the cell. The SSA has proven a valuable tool for understanding the kinetics of reaction networks, i.e., tracking the evolving populations of interacting reactant species, when older methods based on deterministic differential equation systems are too inaccurate or computationally infeasible [11–13]. Many improvements have been made to efficiency of the basic method either via approximations or for particular spaces of model parameter [14–22]. Yet the SSA is not explicitly spatial and instead treats the reactants as uniformly distributed at all times, aside from transient fluctuations. To better capture spatial heterogeneity, extensions of the SSA have been developed based on the reaction diffusion master equation (RDME), typically partitioning the reaction volume into compartments or voxels for which the usual well-mixed assumption applies in each compartment [23, 24]. In these spatial Gillespie models, reactants can react within a compartment or diffuse to an adjacent compartment. However, there is an inherent conflict between accuracy (smaller compartments imply higher spatial resolution) and the well-mixed assumption (better satisfied with larger compartments and/or diffusion rates). In fact, even in the limit of fast diffusion rates, RDME may not converge to the Chemical Master Equation (CME) underlying the Gillespie algorithm [25].

Brownian dynamics (BD) methods provided an opposite extreme of efficiency/realism tradeoffs for such modeling, allowing detailed, off-lattice spatial dynamics but at much greater computational cost. Coarse-grained BD methods have been widely used in self-assembly modeling, as they can deal well with systems with complicated spatial heterogeneity or geometrically intricate structures [26–33]. However, their need to explicitly model diffusion trajectories of single particles creates high computational demands due to the large gap between timescales of diffusive motion versus those of typical molecular assembly processes. Smoldyn [34] is one particularly prominent example, in which molecules diffuse with ideal Brownian motion and react upon collisions. Smoldyn has been considerably extended and improved since its initial release in 2003, e.g., by the inclusion of rule-based modeling, volume exclusion handling, on-surface diffusion, single

* russells@andrew.cmu.edu

particle tracking, and integration with BioNetGen. However, Smoldyn depends on the use of discrete fixed time steps, creating tradeoffs between accuracy and efficient run time in some problem domains.

Green's function reaction dynamics (GFRD [35, 36]) provided an alternative approach to capture spatial heterogeneity in simulating reaction-diffusion systems while taking advantage of SSA-like efficient discrete event simulation without requiring spatial discretization. Instead of generating sample trajectories from the CME or RDME through MCMC, or numerically solving the many-body Smoluchowski equation as in Brownian Dynamics, GFRD analytically solves the Smoluchowski equation for single molecules and molecular pairs in terms of Green's functions. These Green's functions describe the probability of finding a molecule (pair) at a certain location and time given a known position(s) at an earlier time. A maximum time step is chosen such that, with high probability, at most two molecules come into contact, a requirement for analytical tractability. This single/pairwise interaction assumption becomes more valid with smaller time steps, introducing a tradeoff between accuracy and efficiency. Reactions are incorporated through the boundary conditions, and the method combines into a single step propagation through space and reactions between particles. eGFRD [37] is a more recent exact algorithm which removes the accuracy/efficiency trade-off by including the concept of "protective domains" first developed by Oppelstrup et al. [38]. These domains are geometrically simple mathematical boundaries enclosing single molecules or pairs, each of which requires a distinct Green's function solution yielding next event types (domain escape or reaction) and waiting times. Because the time steps are now domain specific, eGFRD is an asynchronous algorithm allowing increased efficiency in some circumstances, although the additional mathematical complexity comes at significant computational expense.

The Small Voxel Tracking Algorithm (SVTA) [39] offers another strategy for particle-based simulation of reaction-diffusion systems. While SVTA is based on the same underlying physics as eGFRD, its implementation is based on a discrete space model. Instead of protective domains, SVTA constructs one and two particle "corrals," within which single molecules and molecule pairs hop between voxels and potentially interact. More specifically, it is the center of each molecule that hops, since the voxel size is typically smaller than the molecular radius. These small voxel dimensions rule out the use of traditional bimolecular propensity functions that rely on the well-mixed assumption. Because the system state evolves on the time scale of diffusion hops, SVTA does not need to analytically sample locations on the protective domains, an easy task only when the domain is a sphere or other simple shape. It can instead simply keep track of when a diffusion hop places a molecule's center in a voxel identified with the corral. SVTA therefore bypasses the need for a suite of domain specific Green's functions in favor of implementing individual diffusion steps on a lattice, providing a strategy for fast but physically realistic sampling compared to prior off-lattice alternatives.

Similarly, the Microscopic Lattice Method (MLM) [40] simulates lattice-based diffusion with reactions. However, MLM aims at optimizing efficiency by simulating molecules of equal size, and requires that voxel dimensions are larger than molecular radii. Additionally, without corrals or protective domains, MLM relies on periodic boundaries to control the simulation volume and number of molecules. A direct comparison between Chew's MLM and Gillespie's SVTA is unavailable, although each would appear to offer some advantages, the former particularly with respect to efficiency while latter can currently simulate more complex biochemistry.

Despite these advances, the most challenging systems remain out of reach of molecular simulation methods without substantial simplifications [2]. New advances in models and algorithms for efficient but physically realistic simulation remain a pressing concern if the field is to continue to move towards solving the grand challenge of truly comprehensive and predictive models of whole-cell biochemistry.

In the present work, we develop an alternative methodology intended to reduce the computational complexity of eGFRD while maintaining discrete event based system updates. The method makes use of a Green's function representation of possible particle positions as a function of time, as originally proposed in GFRD [35], but with an alternative formulation of the probability function in terms of joint probabilities densities of pairs of interacting particles simultaneously. This reformulation enables a new sampling algorithm for position updates that allows for different tradeoffs of efficiency and precision with current alternatives such as eGFRD, SVTA, and Smoldyn. Our goal is not to present a fully optimized algorithm or simulation tool, but rather to explore and advance an alternative approach for the use of time-dependent reaction propensities as a basis for reaction-diffusion simulation in continuous space, which may offer better trade-offs between realism and efficiency than prior methods in some problem domains.

THEORETICAL FRAMEWORK

In this section, we present some theoretical concepts that will be useful subsequently in explaining our model and its relationship to prior work. Consider the bimolecular association reaction system:

$$A + B \Rightarrow C$$

governed by

$$\frac{d[A](t)}{dt} = \frac{d[B](t)}{dt} = -k(t)[A](t)[B](t) \qquad (1)$$

where A and B are hard-sphere species with radii r_A and r_B and diffusion coefficients D_A and D_B. There are two

traditional treatments of diffusion influenced reactions. The first was introduced by Smoluchowski [41] and later extended by Collins and Kimball (CK) [42]. At time t=0, a single particle of species A is considered fixed at the origin and an initial surrounding concentration gradient is set up for the mobile species B molecules. They showed that

$$k(t) = \Phi(t)/c_0 = (4\pi R^2 D/c_0)(\partial c/\partial r)_{r=R} \qquad (2)$$

where $\Phi(t)$ is the probability flux across a boundary sphere for the A particle at radius R, and c_0 is the initial uniform concentration for species B. The simultaneous diffusion of both species is incorporated by setting D as the sum of their respective diffusion coefficients. In this picture, the concentration gradient for the mobile B species $c(r,t)$, defined as the concentration of the B species at distance r from the origin at time t after the initial condition, is found by solving the diffusion equation

$$\partial c/\partial t = D\nabla^2 c \qquad (3)$$

subject to initial condition $c(r,0) = c_0$ and the radiation boundary condition $D(\partial c/\partial r)_{r=R} = \kappa c(R,t)$ where κ is a specific reaction rate. The solution $c(r,t)$ is a complicated function and obeys the relation

$$k(t)/k_i = \frac{c(R,t)}{c_0} \qquad (4)$$

where k_i is the limiting value $k(t \Rightarrow 0)$. Naqvi et al.[43] (sections III.-IV.) updates this by replacing the diffusion equation with a discrete random walk model from which is obtained in the limit of sufficiently long time and distance scales

$$k(t)/k_0 = \frac{c(R+\Delta,t)}{c_0} \qquad (5)$$

with Δ equal to two thirds the scattering mean free path.

The second treatment is due primarily to Noyes [44] and considers an isolated pair of reactive molecules separating from a nonreactive encounter. They showed that

$$k(t) = k_0\Big[1 - \int_0^t h(t')dt'\Big] \qquad (6)$$

where k_0 is defined as "the rate constant applicable for an equilibrium molecular distribution"[45] and $h(t)dt$ is the "probability two molecules separating from a nonreactive encounter at time zero will react with each other between t and $t+dt$" [44]. This can be recast into the form ([43] Eq.47)

$$k(t)/k_0 = S(t; r_0 = R_0, R) \qquad (7)$$

where R_0 denotes the distance between two molecules separating from a nonreactive encounter at time zero, and the survival probability $S(t; r_0, R)$ is defined as

$$S(t; r_0, R) = 1 - \int_0^t p(t'; r_0, R)dt'. \qquad (8)$$

These two major approaches, based on the diffusion equation and particle-pair standpoint respectively, can be shown to be equivalent under certain assumptions and by a lengthy derivation (see [43], sections IV and V). Our method is most easily identified with the theoretical framework of Noyes, but with a different emphasis on instantiating the physical model so as to enable efficient stochastic off-lattice particle simulations. We describe the novel features of our model in more detail below.
The function $h(t)$ appearing in Noyes' fundamental relation can be inferred as the special case

$$h(t) = p(t; r_0 = R_0, R) \qquad (9)$$

To be clear, the initial separation r_0 is the separation distance immediately after a nonreactive encounter. Naqvi argues that $r_0 \neq R$, the reactive contact distance defined in the boundary condition, but instead $r_0 = R_0 = R + \Delta$. The exact expression for $p(t; r_0, R)$ depends on various assumptions, e.g., that the discrete random walks taken by the particles are accurately described by a continuous diffusion equation. In this case, one needs to make further assumptions about initial conditions and boundary conditions.

In the CK picture, the reaction rate evolves only during the time window beginning with the initial condition and ending with a reaction. The assumption here is that immediately after a reaction, the system returns the concentration surrounding the product molecule to the fixed initial value. As such, the formalism may not be suitable to an event-driven, explicitly spatial simulation. Chew et al. [40] with their microscopic lattice method address this issue by deriving their lattice parameters as analogues to the effective or steady state reaction rates in the continuum CK/Noyes theory. This ensures the model behaves similarly to the theory over suitably long time scales.

While our treatment of diffusion influenced reactions is similar to the particle pair approach in Noyes theory, there are notable differences. Instead of using probabilistic arguments to derive reaction rate functions suitable for a differential equation model, we use them to derive reaction propensities suitable for a discrete event SSA. Our conception is as follows: Given a collection of molecules in an explicit and bounded 3d space, and assuming a maximum diffusion time before which we observe their positions, reaction waiting times can be randomly sampled using pairwise propensity functions. The probability density we focus on is not $h(t) = p(t; r_0 = R_0, R)$, but rather $p(t; r_0, R)$ where the initial separation r_0 is specified for each molecule pair, and R is the center to center distance at which a reaction can occur.

In the remainder of the paper we describe the model and implementation, which we refer to as the Diffusion-Based Embedding of the Stochastic Simulation Algorithm in Continuous Space (DESSA-CS) method, in reference to an earlier space-free method [46] based on an accelerated SSA algorithm [17], and demonstrate its effectiveness in comparison to prior alternatives through application to a Michaelis-Menten model.

II. METHODS

Algorithm 1 summarizes our general procedure for off-lattice spatial stochastic simulation. It makes use of a discrete event structure similar to the stochastic simulation algorithm, with the addition of routines for sampling reaction locations. This sampling is based on diffusion spheres containing n_{sigma} standard deviations of the Gaussian distributions describing each particle, similar to GFRD. The resulting positions (due to reactions and position-only updates) are therefore restricted to be within the diffusion spheres, no matter the choice of n_{sigma} (typically 3-5).

In contrast with existing simulation methods in which the boundaries of the simulation volume are either periodic or reflective, we utilize an alternate approach. The state of each molecule is represented as a probability distribution, therefore we only have access to precise positions immediately following an event, and do not consider velocities at all. For this reason, traditional periodic and reflective boundaries are not well defined. Our approach to reaction location sampling, assuming the waiting time has been accurately sampled previously, is to allow the diffusion spheres of molecules near the boundary to extend some distance beyond the boundary, typically a fraction of the container length. If the sampled location happens to be outside the container, we implement a reflection procedure designed to keep the molecules within the simulation volume while respecting the physics of diffusion.

Algorithm 1 DESSA-CS procedure

1: Initialize Event Queue: For each assembly, consider self events (unimolecular reaction, position-only update) and pair events (bimolecular reaction) and add to the queue the earliest self event and pair event for each assembly.
2: Main Loop:
3: **repeat**
4: Extract the next event on the queue.
5: **if** event is bimolecular and valid **then**
6: sample location for product given waiting time; update data structures; add next self event(s) to the queue; add next potential bimolecular events to the queue.
7: **else if** event is unimolecular event and valid **then**
8: sample locations for both products; update data structures; add next self event(s) to the queue; consider each product and add next potential bimolecular events to the queue.
9: **else if** event is position-only update and valid **then**
10: sample location; update data structures; add next position-only update to the queue; add next potential bimolecular events to the queue.
11: (Apply boundary conditions to product(s) if necessary, before adding new events to the queue.)
12: **until** max allowed simulation time or max allowed number of reactions is reached

SAMPLING BIMOLECULAR REACTION WAITING TIMES

Consider a set of K possible bimolecular reactions, i.e., distinct pairs of individual molecules represented as either point particles or finite spheres, and assume each molecule traverses an explicit 3d space by diffusion. For each molecule pair, k, there exists a reaction propensity $a_k(t;s)dt$ describing the probability of an encounter and subsequent reaction of that pair, within some small time interval $[t,\, t+dt)$ after the most recently executed event at time s. The waiting time, t_{wait}, before the next reaction of reactant pair k can be sampled via the equation [47]

$$\int_0^{t_{wait}} a_k(t \mid s)dt = ln(1/r_k) \tag{10}$$

which determines the time at which the integrated propensity equals an exponentially distributed random variable. r_k is the uniform random number drawn for molecular pair k, for use in sampling an exponential waiting time by the transformation method. Because each of our propensity functions are unique to their associated molecular pair, the reaction channels defined in the original SSA and in Anderson's modified next reaction method [47] at the the species level are now defined at the molecule pair level.

Point Particles

At the moment a given molecule's state is updated, the probability density describing its center of mass is concentrated at a single point, i.e., a Dirac delta function centered on that point. As time progresses, the probability density spreads as a Gaussian. This is the free diffusion Green's function solution of the diffusion equation [35]. The positions of two molecules A and B are therefore described by two independent Gaussian random variables, $x_A(t) \sim\ N[\mu_A, \Sigma_A(t)]$ and $x_B(t) \sim\ N[\mu_B, \Sigma_B(t)]$. In order to evaluate Pr(*encounter* & *reaction* | t), the joint probability of an encounter and a reaction during the interval $[t, t+dt)$, we factor the joint probability as Pr(*encounter* | t) * Pr(*reaction* | *encounter*). The latter factor is expressed using a time-independent *intrinsic* reaction rate constant, c, such that $c\,dt$ is the constant encounter conditioned reaction probability over a small time interval. The constant c is specific to this point particle formalism and not equivalent to the microscopic reaction rates used in Smoluchowski or Collins-Kimball theory.

In evaluating the former factor, Pr(*encounter* | t), we assume the initial positions of A and B are known and ask the following question: given a sampled position $\mathbf{x}_A$ of molecule A taken after time t, what is the probability a sampled position $\mathbf{x}_B$ of molecule B after time t will be close to A? Here "close" means at a distance less than a threshold denoting contact or an encounter.

This question can be answered in the language of distributions of quadratic forms in random variables. We define the quadratic form $Q(t)$ as the squared Euclidean distance between Gaussian random variates $\mathbf{x}_A$ and $\mathbf{x}_B$.

$$X_{B-A}(t) \sim \; N\big(\mu_B - \mu_A, [\Sigma_A(t) + \Sigma_B(t)]\big)$$

$$Q(t) = X_{B-A}(t)^T \; X_{B-A}(t) \tag{11}$$

Thus,

$$Pr(encounter \mid t) = CDF_{Q(t)}(R^2_{enc}) \tag{12}$$
$$= Pr\big(Q(t) < R^2_{enc}\big) \tag{13}$$

where R^2_{enc} is the square of the encounter threshold distance. Theorem 4.2b.1 of Mathai & Provost[48] provides a formula in terms of an infinite power series expansion which we use for evaluation.

$$CDF_{Q(t)}(R^2_{enc}) = \sum_{h=0}^{\infty} (-1)^h z_h(t) \frac{(R^2_{enc})^{(3/2)+h}}{\Gamma\big((3/2)+h+1\big)} \tag{14}$$

The coefficients $z_h(t)$ are defined recursively and depend on $\mu_{AB} = \mu_B - \mu_A$, and $\Sigma_{AB}(t) = (\Sigma_A + \Sigma_B)$. See Appendix A for a full description. Convergence is defined by no change to 5 places after the decimal for 20 successively higher order approximations. For very small t and large initial separation, the approximation can oscillate wildly about zero. In these parameter regions where numerical instability is detected, we set the CDF to zero.

With isotropic diffusion, the reaction propensity given R^2_{enc} and $d = norm(\mu_{AB})$ after time t, and with intrinsic rate c, can be reparameterized as a function of the variance v of $X_{B-A}(t)$ rather than of time directly. This variance is simply the diagonal element of $\Sigma_{AB}(t)$. The reparameterized reaction propensity, denoted $a_k(t)$, is then given by:

$$a_k(t)dt = a_k(v \mid d_k, R^2_{enc,k}, c)dv \tag{15}$$

The time to next reaction can now be determined by evaluating

$$argmin_v \quad \int a_k(v)dv \geq ln(1/r_k) \tag{16}$$

and inferring t_{wait} from the variance value, should it exist. Figure 3 visualizes the wait time sampling procedure. One added complication is that the DESSA-CS algorithm is event driven. After each event, potential new reactions are considered for the product(s) of that most recently executed event. This implies that the position of the product (e.g., reactant A) is known precisely, while its potential partner (e.g., reactant B) has been diffusing for a time t_{offset} and thus has its position represented by a Gaussian random variable. Any integrated propensity up through $v(t_{offset})$ must therefore be discounted when sampling the variance at which a reaction occurs. See Figure 1 for an illustration. The sampling procedure is described in Algorithm 2. For finite sized molecules, the procedure is similar, except the integrated propensities (called *IntF* in Algorithm 2) are expressed directly in terms of times rather than variances. A Matlab implementation of the algorithm is available on GitHub [49].

Our propensity function describing Pr(*encounter* & *reaction* $\mid t$) is equivalent to $p(t; r, R)$ from the Noyes theory under the assumption that the molecules are dimensionless point particles for which there is no minimum separation distance. In this case, there is no need to go beyond the free diffusion Green's function solution to the diffusion equation as there are no boundary conditions enforcing a minimum pairwise separation.

Particles with Finite Size

With molecules of finite size, we will still use integrated reaction propensities to sample reaction waiting times. However, a different mathematical framework is required to construct the propensities. Assume both particles are spherical and R defines the center-to-center distance at contact. In this context, $p(t; r_0, R)$ described in the Theoretical Framework section above is expressed as

$$p(t; r_0, R) = p(r = R, t; r_0, 0) * c \; dt \tag{17}$$

where c now denotes the absorbing/radiation boundary condition parameter, and $p(r, t; r_0, 0)$ is the Green's function solution to the following boundary value problem. Assume $p(r, t; r_0, 0)$ obeys a diffusion equation, and the initial separation between molecules is r_0. This is expressed with the initial condition

$$p(r, 0) = \frac{\delta(r - r_0)}{4\pi r^2}. \tag{18}$$

The two boundary conditions on $p(r, t; r_0, 0)$ ensure that the molecular separation never reaches infinity, and that at contact, the probability of a reaction is accounted for:

$$\lim_{r\to\infty} p(r, t) = 0 \tag{19}$$

$$4\pi R^2 D \left.\frac{\partial p(r, t; r_0, 0)}{\partial r}\right|_{r=R} = c \; p(R, t; r_0, 0) \tag{20}$$

From Chew et al. [40], and Jaeger & Carslaw [50] p. 368), the Green's function solution is

$$p(r, t; r_0, 0) = \frac{1}{8\pi r r_0} \frac{1}{\sqrt{\pi D t}} \Big(exp[-(r - r_0)^2/4Dt] + exp[-(r + r_0 - 2R)^2/4Dt] - 2B\sqrt{\pi D t}\; exp[B^2 Dt + B(r + r_0 - 2R)] * erfc(\frac{(r_0 - R)}{2\sqrt{Dt}} + B\sqrt{Dt}) \Big) \tag{21}$$

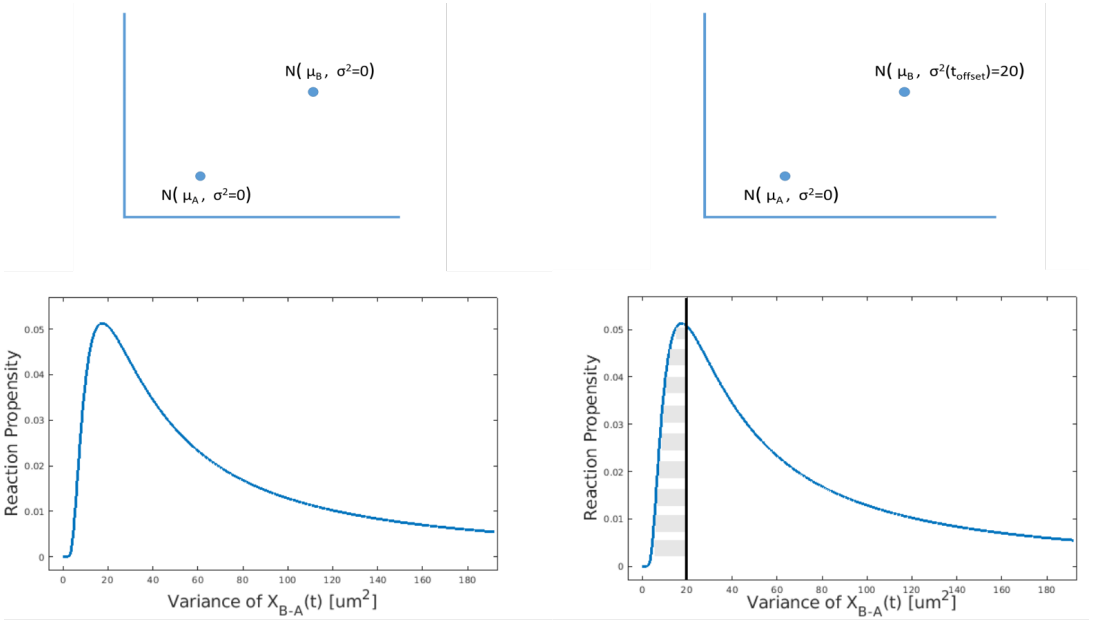

Figure 1. The figures on the left depict two molecules, A and B, described as Gaussians with means separated by $d = 7.235\mu m$. The point-particle reaction propensity grows as the variance increases, reaching a peak just before $20\mu m^2$ and then decreases monotonically. The figures on the right are more typically encountered in the algorithm. The most recent reaction for A was just executed and wait times are being sampled for the $A + B$ reaction. B has already been diffusing for a time t_{offset}, thus, propensity function integration begins not at zero variance, but instead at variance equal to $20\mu m^2$.

where $B = (1+\frac{c}{4\pi RD})/R$. Note that the Green's function also depends on c through B. The propensity function is

$$a(t)dt = p(R, t; r_0, 0)\ c\ dt \tag{22}$$

and the time to next reaction, t_{wait}, can be determined from the integrated propensity by evaluating

$$\frac{B}{4\pi R^2 r_0}\left[erfc[\frac{B(r_0-R)}{2\sqrt{\tau}}]-\left(exp(Br_0-BR+\tau)\ erfc[\frac{Br_0-BR+2\tau}{2\sqrt{\tau}}]\right)-1\right]_0^{\tau-max} - ln(1/r_k) = 0 \tag{23}$$

with $r_k \sim uniform[0, 1]$ and $\tau = tDB^2$. The waiting time is inferred as $t_{wait} = \tau/DB^2$. As in the point particle context, when molecule B of the molecular pair has been diffusing for a time t_{offset} when we are sampling reactions for molecule A, the integrated propensity up through t_{offset} must first be subtracted from the L.H.S. of Eq. 23. Alternately, Eq. 21 can be numerically integrated.

Validation of the Wait Time Sampling Procedure

Noyes theory is formulated in terms of the probability two molecules will re-collide (and potentially rebind) following a nonreactive encounter, therefore comparing the theoretical and simulated rebinding time probability densities is a useful test of our Gillespie-inspired wait time sampling procedure. Following Chew et al. [40], we consider both the activation-limited and diffusion-influenced cases (Figure 2). These are distinguished by $c/\nu_{AB} < 1$ and $c/\nu_{AB} \geq 1$, respectively. See Gillespie et al. [39] for a derivation of this relation. The parameter c is the boundary value parameter appearing in the finite particle propensity function and is related to the collision frequency $\nu_{AB} = 4\pi R(D_A + D_B)$ between A and B molecules in a hypothetical nonreactive system as:

$$c = \frac{P(reaction \mid encounter)\nu_{AB}}{g(r = R)\rho_B},$$

where $g(r)$ is the particle pair correlation function in a liquid phase and ρ_B is the relative density of B molecules. See [51]. Figure 2 shows, for three values of c/ν_{AB}, the theoretical density and the results of our simulations. We computed the integrated reaction propensity at 70,000,000 time points linearly spaced in the range $[10^{-8}$,1]. On the order of 10^9 wait time samples were drawn for each of the three ratios and then aggregated into bins of width $w_{bin} = 5 \times 10^{-9}$s. Simulated probability densities for a subset of bins were computed as

$$pdf(bin) = \frac{N_{bin}}{N_{total}\ w_{bin}},$$

where N_{bin} is the number of samples in the bin and N_{total} is the total number of samples. As the total number of

samples grows, the simulated values approach the theoretical density.

Algorithm 2 Sampling Bimolecular Wait Times - Point Particle Representation

1: (Pre-simulation) Define vector of variance values, $\mathbf{v} = [0, V_{max}]$
2: (Pre-simulation) Define the curve $IntF(\mathbf{v} \mid d, R_{enc}^2, c)$ as the cumulative sum of reaction propensity values along the points $\mathbf{v}$. $\{IntF(\mathbf{v} \mid d, R_{enc}^2, c)\}$ is then the set of integrated propensity curves at increasing d, computed once, before the simulation begins. If desired, further sets of curves can be precomputed for alternate values of R_{enc} and c.
3: (At run time) For reactant pair k = (A,B), select the appropriate curve, $IntF(\mathbf{v} \mid d_k, R_{enc}^2, c)$
4: Evaluate $IntF(v_{t_{offset}})$, the integrated propensity to be discounted, at the variance value corresponding to t_{offset}, i.e., $6D_b t_{offset}$.
5: Set $v^* \leftarrow argmin_v\ IntF(\mathbf{v}) \geq ln(1/r_k) + IntF(v_{t_{offset}})$
6: If v^* exists, t_{wait} is the solution to $v^* = 6D_a t_{wait} + 6D_b(t_{wait} + t_{offset})$
7: Else, no reaction is sampled. Update particle positions.

SAMPLING BIMOLECULAR REACTION LOCATIONS

Again we make use of the labels A and B for the specific molecules undergoing the next association reaction. At this time, the spatial region available for the reaction consists of the intersection of the diffusion spheres bounding their independent Gaussian probability distributions. In order to correctly sample from this region, henceforth called the *overlap volume* (OV), we first introduce the concept of equiprobable rings.

Equiprobability Rings

The line AB connecting the initial known positions of A and B defines an axis of symmetry in the sense that within the OV there exist rings centered on this axis, whose points are equidistant from A and equidistant from B. The rings are therefore sets of equiprobability points from which molecule positions might be sampled. Each ring is uniquely defined by two numbers: the magnitude, r_A, of any vector from the initial position of A to a point on the ring, and the CCW angle, θ_A, between the vector and the line AB. After sampling (r_A, θ_A), we choose the reaction location uniformly at random from on the ring. The joint probability density describing (r_A, θ_A) can be factored as $p(r_A|t)$ and the conditional probability $p(\theta_A|r_A, t)$, which suggests a sequential sampling procedure. First determine r_A and then use it to determine θ_A.

Diffusion Sphere Overlap Volume

While the OV grows continuously due to diffusion, for the purpose of location sampling at a given time we have found it useful to classify it into one of five distinct cases. These cases are not inherently meaningfully different in the theory behind them, but provide a convenient way of describing the evolution of the OV as well as distinguishing the integration regions involved in sampling, r_A and θ_A, for purposes of clearer exposition. For example, if the OV is identical to the diffusion sphere of A (as in cases 3 and 5), θ_A may take on any value in $[0, 2\pi]$, however if the OV has an irregular shape, certain angles may be prohibited. Figure 4 visualizes the two trajectories possible for the OV. The first trajectory applies when $D_B > 4D_A$ and passes through cases 1, 2, 3 and 5. The second trajectory applies when $D_A < D_B < 4D_A$ and passes through cases 1, 2, 4 and 5. Given the current system time t, the waiting time until the next reaction of A and B, t_{wait}, and the system time at which the state B was last updated, we can infer $t_{A-elapsed}$ and $t_{B-elapsed}$, the durations during which each had been diffusing before the reaction, which includes the waiting time to the reaction. Using $t_{A-elapsed}$ and $t_{B-elapsed}$ to define the diffusion spheres at the moment the molecules react, we can infer the OV case.

Case 2 begins when the radius of the faster diffusing particle (here, B) is equal to d, the distance between the Gaussian means of A and B. This radius can be computed as $R_B(t) = n_{sigma}\sqrt{6D_B t}$, where n_{sigma} is the number of standard deviations bounded by the sphere. See Fig. 5 for an illustration of the integration variables in Case 2. The starting time is given by

$$t_{start-2} = \frac{d^2}{6D_B n_{sigma}^2} \tag{24}$$

Starting times for cases 3-5 are calculated as follows:

Path 1	Case 3 Start	Case 5 Start
$D_B > 4D_A$	$d > R_A(t)$ $R_A(t) + d = R_B(t)$ $t_{start-3} = t_\gamma$	$d = R_A(t)$ $R_A(t) + d < R_B(t)$ $t_{start-5} = t_{\sim\gamma}$
Path 2	Case 4 Start	Case 5 Start
$D_A < D_B < 4D_A$	$d = R_A(t)$ $R_A(t) + d > R_B(t)$ $t_{start-4} = t_{\sim\gamma}$	$d < R_A(t)$ $R_A(t) + d = R_B(t)$ $t_{start-5} = t_\gamma$

where

$$t_\gamma = \frac{1}{6n_{sigma}^2(D_A - D_B)^2}(2D_A^2 n_{sigma}^2\gamma + 2D_B^2 n_{sigma}^2\gamma - 4D_A D_B n_{sigma}^2\gamma + D_A d^2 + D_B d^2), \tag{25}$$

$t_{\sim\gamma} = \frac{d^2}{6D_A n_{sigma}^2}$, and $\gamma = \sqrt{\frac{D_A D_B d^4}{n_{sigma}^4(D_A - D_B)^4}}$.

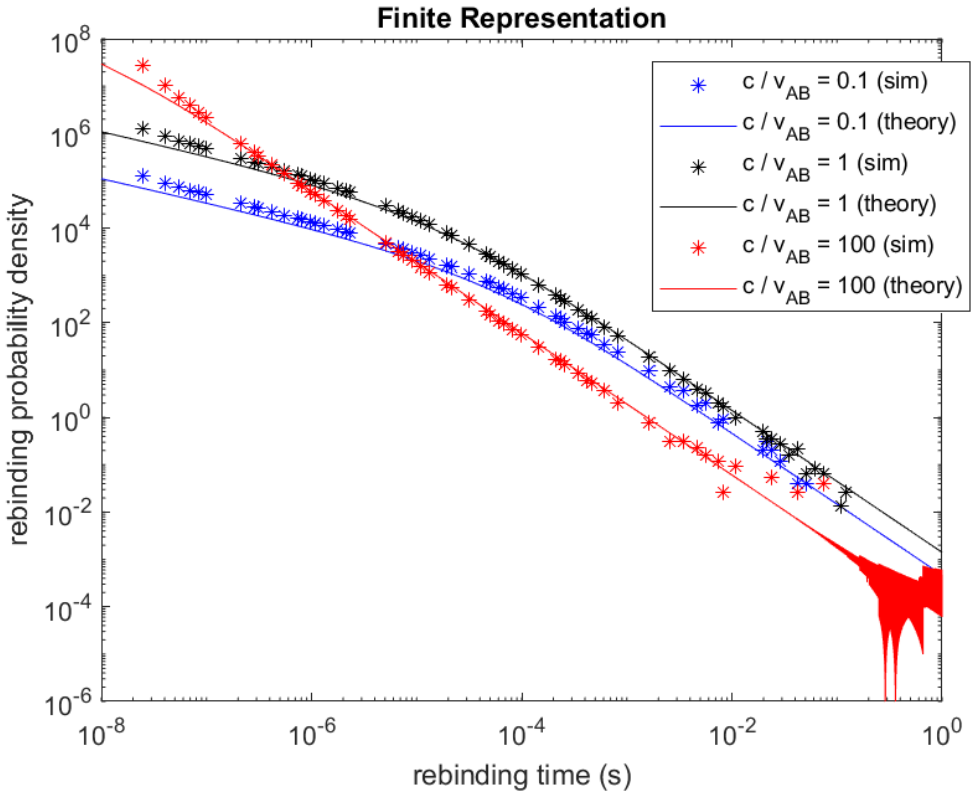

Figure 2. Rebinding time probability density from Noyes theory. We compare the theoretical curves in the finite particle representation with values computed from simulations at $c/\nu_{AB} = 0.1$, 1, and 100. Deviations from the theory at larger rebinding times are explained by the fact that more samples are required to characterize the probability densities at these limits than were drawn in our analysis. Simulation parameters were: $D_A = 1\mu m^2 s^{-1}$, $D_B = 1\mu m^2 s^{-1}$, $r_0 = 0.01001\mu m$. At the largest times, numerical instabilities can appear in the Green's function computations, leading to the noise seen in the $c/\nu_{AB} = 100$ theoretical (theory) data.

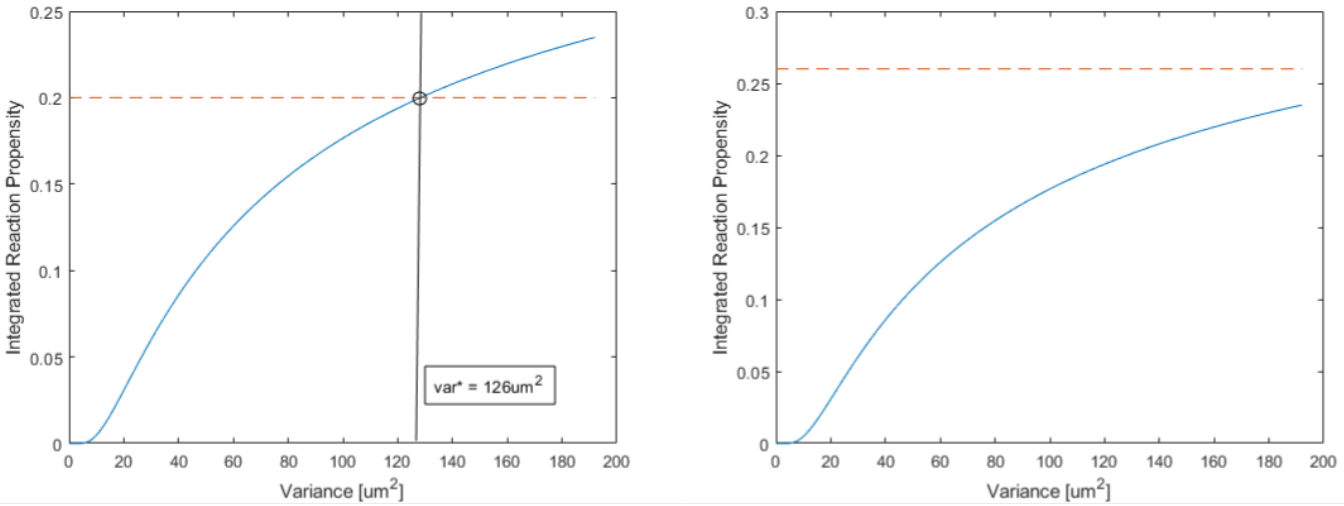

Figure 3. Examples of successful and unsuccessful sampling of a biomolecular reaction in the point particle representation. In both subfigures, the solid curve is the integrated reaction propensity associated with two reactants described by Gaussians with means separated by $8\mu m$. The left subfigure shows the successful sampling of a bimolecular reaction waiting time as there is a variance value (and thus, a waiting time) at which the integrated reaction propensity equals the exponentially distributed random number, 0.2. In the right subfigure, the exponentially distributed random number is 0.26, and so there is not sufficient integrated propensity for a reaction to occur. In the finite particle representation, the x-axis will represent the waiting time directly instead of variance.

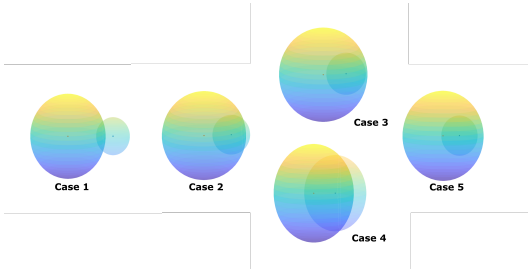

Figure 4. Cases of potential overlap of diffusion spheres in the process of sampling waiting time to a biomolecular reaction. Shown are the diffusion sphere intersections at increasing time points. It is assumed here that $D_B > D_A$. Case 1: The OV contains neither μ_A nor μ_B. Case 2: The OV contains μ_A only, and is not identical to either diffusion sphere. Case 3: The OV contains μ_A only, and is identical to the diffusion sphere of A. Case 4: The OV contains μ_A and μ_B, but is not identical to either diffusion sphere. Case 5: The OV contains μ_A and μ_B, and is identical to the diffusion sphere of A.

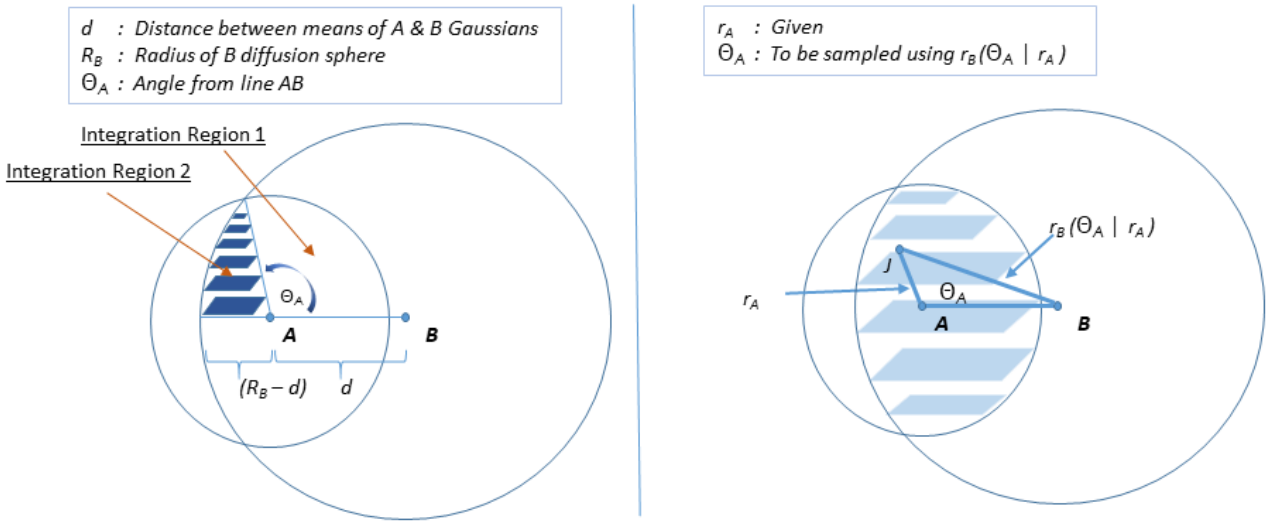

Figure 5. (**Left**) Visualizing the regions of integration for $w(r_A)$ in Case 2. (**Right**) Visualizing θ_A, r_A, and $r_B(\theta_A)$ in Case 2. The equiprobability ring passes through point J, perpendicular to the plane of the page.

Case 1

Sampling r_A

In order to sample r_A correctly, we re-weight the probability density in the OV, i.e., compute a posterior probability. Define $h_{ring}(\theta_A)$ as the radius of the ring whose points are at distance r_A and for which the top most point defines a line with A at angle θ_A. The circumference of this ring is $2\pi h_{ring}(\theta_A)$. Integrating this circumference over the available θ_A range allows us to determine the size of the set of points at distance r_A.

$$p_{reweighted}(r_A, t) = w(r_A) * p(r_A, t) \tag{26}$$

with

$$w(r_A) = \frac{[TotalProbability - at - r_A]}{\int_{OV} dr\Big([TotalProbability - at - r] * p(r,t)\Big)},$$

$$\int p_{reweighted}(r_A, t) dr_A = \int w(r_A) p(r_A, t) = 1, \tag{27}$$

and

$$p(r,t) = \frac{1}{\sqrt{12\pi D_A t}} exp(-r^2/12 D_A t) \tag{28}$$

$$w(r_A) = \frac{\int_0^{\theta_{max}(r_A)} d\theta_A 2\pi h_{ring}(\theta_A)}{\int_{r_{lb}}^{r_{ub}} dr\Big[\big(\int_0^{\theta_{max}(r)} d\theta(r) 2\pi r \sin(\theta)\big) p(r,t)\Big]}$$
$$= \frac{r_A\big(cos(\theta_{max}(r_A)) - cos(0)\big)}{\int_{r_{lb}}^{r_{ub}} dr\Big[r\big(cos(\theta_{max}(r)) - cos(0)\big)p(r,t)\Big]}$$
$$= \frac{r_A\left(\frac{r_A^2+d^2-R_B^2}{2r_A d} - 1\right)}{\Big[term1 + term2\Big]} \quad (29)$$

$$term1 = \frac{1}{4d}(d^2 + 6D_A t - R_B^2)\big[erf\big(r_{ub}/\sqrt{12D_A t}\big) - erf\big(r_{lb}/\sqrt{12D_A t}\big)\big]$$

$$term2 = \frac{1}{\sqrt{12\pi D_A t}} 6D_A t\big[(r_{lb} - 2d)exp(-r_{lb}^2/12D_A t) - (r_{ub} - 2d)exp(-r_{ub}^2/12D_A t)\big]$$

The upper limit of integration, θ_{A-max}, is calculated by considering the triangle defined by the three points: A, B, I. The base (AB) length is d. The side BI has length R_B since I is the point at which (r_A, θ_A) intersects the OV, i.e., a point on the B diffusion sphere. The remaining side length is r_A. From the law of cosines, θ_{A-max} is calculated in terms of the side lengths.

$$\theta_{A-max}(r) = cos^{-1}\left(\frac{r^2 + d^2 - R_B^2}{2rd}\right) \quad (30)$$

The lower and upper bounds, r_{lb} and r_{ub}, on r_A defining the OV are $[(d - R_B), R_A]$.

Sampling $\theta_A|r_A, t$

The tuple (θ_A, r_A) uniquely defines a ring of equiprobability points within the OV from which a single reaction location can be chosen uniformly at random. Thus, the probability with which a given θ_A is sampled should be proportional to the size of the corresponding ring.

Consider the triangle defined by the points A, B, J where J is a point in the OV at (θ_A, r_A). The length of side BJ is $r_B(\theta_A)$ and can be computed with the Law of Cosines. The height of this triangle, h_{ring}, is again the radius of the ring passing through point J.

$$p(\theta_A|r_A, t) = p(r_B(\theta_A)|t) * RingCircumference$$

$$p(\theta_A|r_A, t) = \frac{1}{\sqrt{12\pi D_B t}}\, exp\left(-\frac{r_B(\theta_A)^2}{12D_B t}\right) * 2\pi h_{ring} \quad (31)$$

$$r_B^2(\theta_A) = r_A^2 + d^2 - 2r_A d\, cos(\theta_A) \quad (32)$$

$$h_{ring} = r_A\, sin(\theta_A) \quad (33)$$

$$\theta_A \in [0, \theta_{A-max}]$$

Case 2

Sampling r_A

$$w(r_A) = \int_0^{\theta_{A-max}(r_A)} d\theta_A 2\pi r_A sin(\theta_A) \; * \Big(\int_0^{R_B-d} dr\big[\int_0^{\theta_{max}(r)} d\theta(r) 2\pi r \sin(\theta)\big] * p(r,t) + \int_{R_B-d}^{R_A} dr\big[\int_0^{\theta_{max}(r)} d\theta(r) 2\pi r \sin(\theta)\big] * p(r,t)\Big)^{-1} \quad (34)$$

$$\theta_{A-max}(r) = cos^{-1}\Big(\frac{max[r^2, (R_B - d)^2] + d^2 - R_B^2}{2d\; max[r, (R_B - d)]}\Big) \quad (35)$$

Figure 5 provides a visual description of the relevant Case 2 variables. Variables for the other cases are defined similarly. For any r_A less than or equal to $(R_B - d)$, the full angular range of region 2 is available, i.e., $\theta \in (0, \pi)$. As r_A increases from $(R_B - d)$ to R_A, the available positions within region 2 decrease to 0. We capture this dependence with the angle integration limits, $(0, \theta_{A-max})$, where $\theta_{A-max} = \pi$ for $r_A \leq (R_B - d)$. The logic behind the form of $w(r_A)$ is analogous to case 1, however.

Sampling $\theta_A|r_A, t$

Sampling here is analogous to case 1, with updates to the available angle ranges for a given r_A.

$$p(\theta_A|r_A, t) = \frac{1}{\sqrt{12\pi D_B t}}\, exp\left(-\frac{r_B(\theta_A)^2}{12D_B t}\right) * 2\pi h_{ring} \quad (36)$$

With $r_B^2(\theta_A) = r_A^2 + d^2 - 2r_A d\, cos(\theta_A)$, $h_{ring} = r_A\, sin(\theta_A)$, and $\theta_A \in [0, \theta_{A-max}]$.

Case 3

Sampling r_A

In this case, the full range in r_A ($\in [0, R_A]$) is available. Therefore, no re-weighting of probabilities is needed.

$$p(r_A|t) = \frac{1}{\sqrt{12\pi D_A t}}\, exp\left(-\frac{r_A^2}{12D_A t}\right) \quad (37)$$

Sampling $\theta_A|r_A, t$

Sampling here is analogous to case 1, but with the full range of angles available.

$$p(\theta_A|r_A,t) = \frac{1}{\sqrt{12\pi D_B t}}\ exp\left(-\frac{r_B(\theta_A)^2}{12D_B t}\right) * 2\pi h_{ring} \tag{38}$$

With $r_B^2(\theta_A) = r_A^2 + d^2 - 2r_A d\ cos(\theta_A)$, $h_{ring} = r_A\ sin(\theta_A)$, and $\theta_A \in [0, \pi]$.

Case 4

Sampling r_A

Sampling here is analogous to case 2.

$$w(r_A) = \int_0^{\theta_{A-max}(r_A)} d\theta_A 2\pi r_A sin(\theta_A)\ * \Big(\int_0^{R_B-d} dr\Big[\int_0^{\theta_{max}(r)} d\theta(r) 2\pi r \sin(\theta)\Big] * p(r,t) + \int_{R_B-d}^{R_A} dr\Big[\int_0^{\theta_{max}(r)} d\theta(r) 2\pi r \sin(\theta)\Big] * p(r,t)\Big)^{-1} \tag{39}$$

$$\theta_{A-max}(r) = cos^{-1}\Big(\frac{max[r^2, (R_B-d)^2] + d^2 - R_B^2}{2d\ max[r, (R_B-d)]}\Big) \tag{40}$$

Sampling $\theta_A|r_A, t$

Sampling here is also analgous to case 2.

$$p(\theta_A|r_A,t) = \frac{1}{\sqrt{12\pi D_B t}}\ exp\left(-\frac{r_B(\theta_A)^2}{12D_B t}\right) * 2\pi h_{ring} \tag{41}$$

With $r_B^2(\theta_A) = r_A^2 + d^2 - 2r_A d\ cos(\theta_A)$, $h_{ring} = r_A\ sin(\theta_A)$, and $\theta_A \in [0, \theta_{A-max}]$.

Case 5

Sampling r_A

In this case, the full range in r_A ($\in [0, R_A]$) is available. Therefore, no re-weighting of probabilities is needed.

$$p(r_A|t) = \frac{1}{\sqrt{12\pi D_A t}}\ exp(-\frac{r_A^2}{12D_A t}) \tag{42}$$

Sampling $\theta_A|r_A, t$

Sampling here is analogous to case 1, but with the full range of angles available.

$$p(\theta_A|r_A,t) = \frac{1}{\sqrt{12\pi D_B t}}\ exp\left(-\frac{r_B(\theta_A)^2}{12D_B t}\right) * 2\pi h_{ring} \tag{43}$$

With $r_B^2(\theta_A) = r_A^2 + d^2 - 2r_A d\ cos(\theta_A)$, $h_{ring} = r_A\ sin(\theta_A)$, and $\theta_A \in [0, \pi]$.

Determining Bimolecular Reaction Locations by Rejection Sampling

Because PDFs in each case may be complicated functions, we cannot always sample from them directly. Instead, we first draw a sample of our variable x (i.e. r_A or θ_A) uniformly from its feasible range. In order to determine whether this sample is accepted or rejected, we utilize an envelope function, $Q(x)$ whose probability density at all feasible points is at least as great as that of the PDF from which we want an observation. The sample x is accepted if $q(x)$ drawn uniformly from $[0, Q(x)]$ is less than $p(x)$.

One potential issue is that volume exclusion should prevent sampled locations from leading to particle overlap. We define two molecules to be overlapping if the distance between their Gaussian means is less than R, the minimum allowed separation, *and* neither molecule has been diffusing for longer than $R^2/6D$. In the event overlap is detected, a new location is sampled. This procedure for handling volume exclusion in location sampling is not highly optimized in the demonstration implementation presented here, and can significantly impact the run time as the particle density increases.

Simulation Boundaries

Figure 6 illustrates our method for ensuring all particles remain within the simulation volume. We treat this volume as a cube bounded by planes about which a particle may be reflected if its initially sampled position exceeds the plane. The algorithm samples an unconstrained reaction location and the displacements for both particles are noted. Next, assume the reaction location happens to be outside the simulation volume. Each molecule can be considered to have travelled along a linear path from its initial location to the reaction location, with one piece of the path within the simulation volume and one piece outside. Because the unconstrained spatial probability densities describe radial displacements from either particle's initially known location, application of reflective boundary conditions need only guarantee both particles' piecewise linear paths each sum to the noted displacements, and terminate within the simulation volume.

This procedure is strictly correct only if the wait time sampling, i.e., computing the integrated reaction propensities, is correct. The point (finite) particle reaction propensities described in this paper do not take into account the boundaries of the simulation volume. Error is therefore introduced in wait time sampling for molecules diffusing long enough to encounter a boundary. However, given a wait time t, reaction location sampling depends only on the possible net displacements of either particle after diffusing for t units of time. We can therefore assume free diffusion to sample the location and then use our reflecting procedure if necessary.

III. RESULTS

Application: Michaelis-Menten

We applied DESSA-CS to the well known Michaelis-Menten enzymatic reaction system within a $90\mu m^3$ volume. The original benchmark was developed by Andrews [34] and updated by Chew et al. [40] to account for the extreme run time demands of eGFRD. Figures 7 and 8 display our results for the updated benchmark, displaying the population dynamics for molecular species E, S, ES and P, which obey the binding rules $E + S \Leftrightarrow ES \Rightarrow P$. Figure 9 shows comparable results for eGFRD. Note that the different models approximate and parameterize the physical system in different ways and so it is not possible to conduct perfectly equivalent simulations by each method, but we have chosen parameter values so as to approximate the same physical conditions in each method as closely as they allow in our choice of rate constants. Simulation run times for the point particle and finite particle representations respectively were 20 seconds and 90 seconds, roughly two orders of magnitude faster than eGFRD (see Figure 5 of Chew et al. [40] and Table I below).

In the point particle simulations, the data set (computed before run time) consisted of 3000 linearly spaced distances from R_{enc} to 2^*d_{maxD}, where d_{maxD} corresponds to the mean square displacement due to diffusion at T_{maxD}, the max allowed diffusion time. At each distance, the integrated propensity was computed at 50,000 time points (i.e. variances). There were 40,000 linearly spaced time points from $1*10^{-6}$ to $1*10^{-2}$ where the curvature is often highest, and 10,000 linearly spaced time points from $1 * 10^{-2}$ to T_{maxD}. This 3000 by 50,000 data set was computed in 14.5 minutes in the Go language (golang). The propensity function integration error for a given distance value and time duration depends on the number of integration intervals - here 50,000 for the full duration. We used the trapezoid method, whose error at each time point can be upper bounded by $Err(\Delta t) = \frac{\Delta t^3}{12N^2} * K * c$, where K is the maximum magnitude of the second derivative of the CDF, c is the intrinsic reaction rate, and N is the number of integration intervals over the duration Δt. These are not likely to be tight upper bounds due to the presence of inflection points in the CDF graph. Our golang integration code, including a method to print error bounds and integrals to text files, can be found in the GitHub repository.

In the finite particle simulations, the data set consisted of 3000 integrated propensity curves at the same distance values, each of which was evaluated at 5500 time points. Computation of the integrated propensity data set required $\sim 1m$. Performing these numerical integrations is possible in Matlab but requires symbolic computation to evaluate the integrands. The result was that the same data set takes on the order of days to compute. In golang, the necessary numeric precision was achieved using its big math package which implements arbitrary-precision arithmetic.

In general, both the finite and point particle representations run more efficiently when each molecule is allowed to diffuse farther outside the boundaries before wait times are sampled. We will refer to a cubic simulation volume as having dimension $(L\mu m)^3$ and the fraction of the cube length beyond which a particle may diffuse as l_B. It is then useful to analyze the behavior of simulations as these vary. The Michaelis-Menten Benchmark requires that $90\mu m^3 = L(1 + l_B)$. In Figure 10, we plot simulation trajectories at multiple values of l_B (i.e., $l_B = 0.03$, $l_B = 0.1$, and $l_B = 0.3$) for the finite particle representation. The plot demonstrates that while the kinetics do not change appreciably with changes in l_B, the run time does. The respective run times are 2693s, 777s, and 157s. The results are qualitatively similar in the point particle representation case. As the relative distance allotted to the cube's length increases, so does the run time. This results from the fact that as l_B decreases, so does the maximum diffusion time of molecules near the boundary, leading to a much higher number of position updates compared with the roughly unchanging number of reaction events. Even though the change in kinetics is minimal across the parameter values examined, the effective association rate does show a weak inverse dependence on l_B.

IV. CONCLUSIONS

We have presented a novel event-based method for simulating reaction diffusion systems in continuous space and in the presence of planar or curved boundaries. As in the Gillespie algorithm and related methods, we sample bimolecular reaction waiting times by utilizing propensity functions. However, with the introduction of 3d space, the reaction propensities now depend explicitly on the time reactants diffuse, allowing them to encounter one another. The result is that we integrate the propensity function of each reactant pair in order to determine whether (and when) a reaction is possible in a specified duration. While the method is inspired by ideas from GFRD and eGFRD, our method for sampling reaction locations given the waiting time is, to our knowledge,

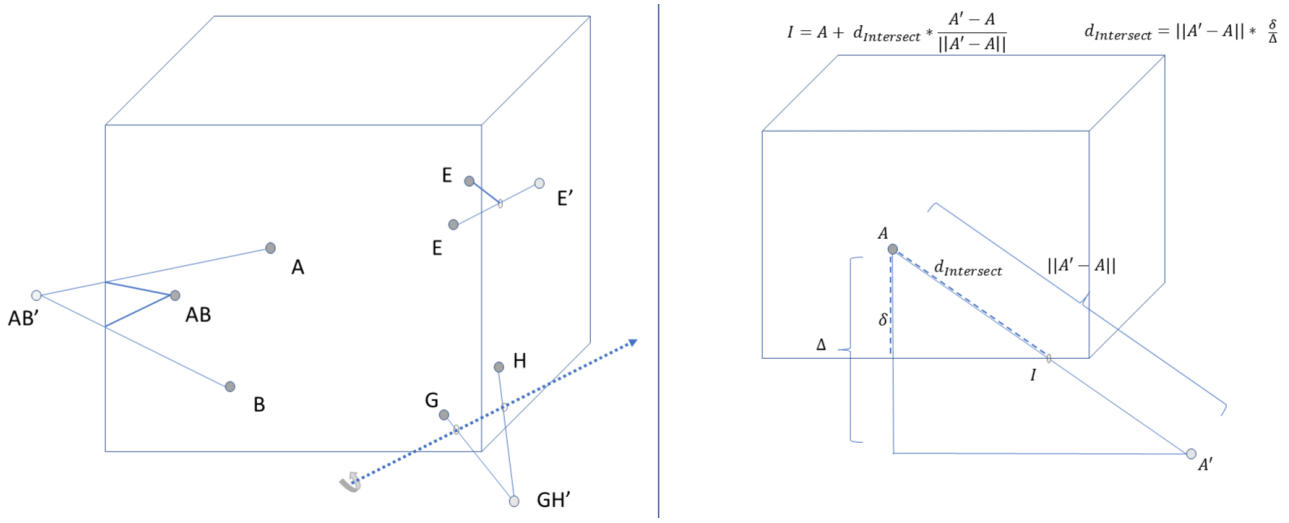

Figure 6. (**Left**) Shown are applications of the reflective boundary condition after a position-only-update event (e.g. E) or after a bimolecular reaction event (e.g. A&B, G&H). (*Single Reflection*) In the bimolecular case, we reflect about an axis defined by the two intersection points of the lines connecting the reactants with the product, and the boundary. This ensures that the distances traveled by both particles remains the same. When these lines exit the simulation box through the same face (e.g. A&B), the reflection axis is parallel to the face. When the lines exit though different faces (e.g. G&H), the axis must be computed and the reflection can be implemented with the Rodrigues rotation formula in the appropriate reference frame. (*Multiple Reflection*) Depending on the location of the reactants and the distances they travel, the post-reflection location may end up outside a different boundary, though to a lesser extent. We simply need to update the reactant positions to be the boundary intersection point(s) and reapply the reflection procedure. In principle, this procedure works for any simulation volume, including those with curved boundaries. (**Right**) For a cubic simulation volume, we determine through which face (and at what point) a reactant (A) first passed if it is found outside the simulation volume. In this case, the pre-reflection location A' exceeds the simulation volume along more than 1 dimension which means it is necessary to compute $d_{Intersect}$ for each 2d plane exceeded by A', and then compute the intersection point, I, for the face with minimum $d_{Intersect}$.

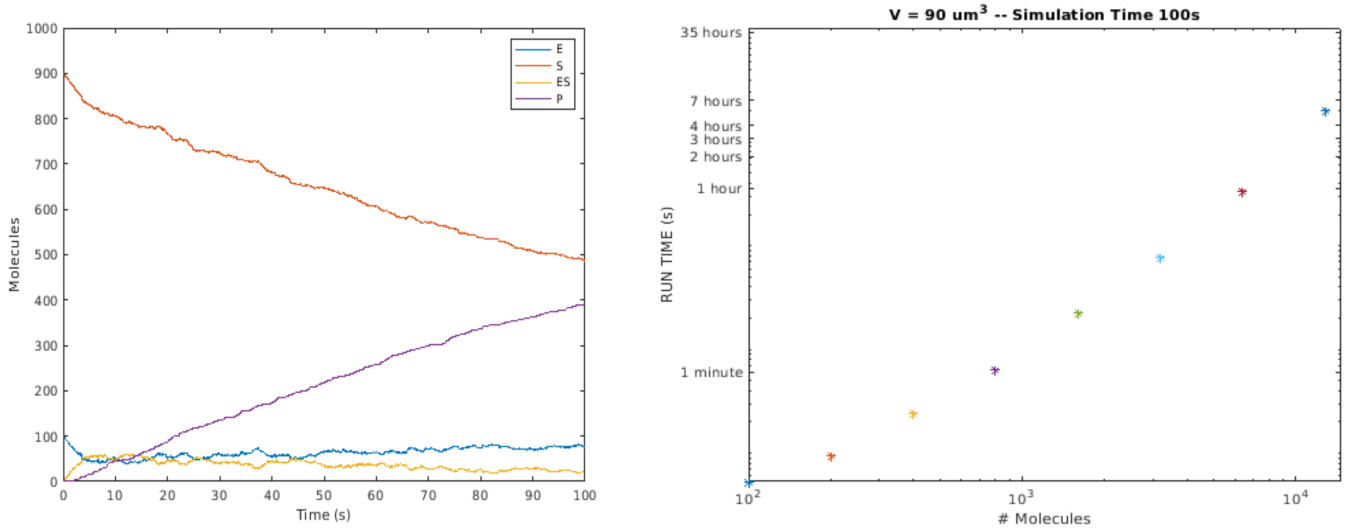

Figure 7. Point particle representation. (**Left**) Time evolution of 1000 molecules in the Michaelis-Menten model with DESSA-CS. Unimolecular rate constant $k_{uni} = 0.1s^{-1}$ (governing $ES \Rightarrow E + S$ and $ES \Rightarrow E + P$) and diffusion coefficient $D = 1\mu m^2 s^{-1}$ are taken from Figure 5 of Chew et al. [40]. In order to reproduce similar dynamics, we chose the point-particle reaction propensity constant $c = 2.5 * 10^7 s^{-1}$ (with $R_{enc}^2 = 0.01^2 \mu m$). (**Right**) The run time for the model increases roughly linearly in log space with the number of molecules - [100,200,400,800,1600,3200,6400,12800] at a fixed volume of 90 μm^3.

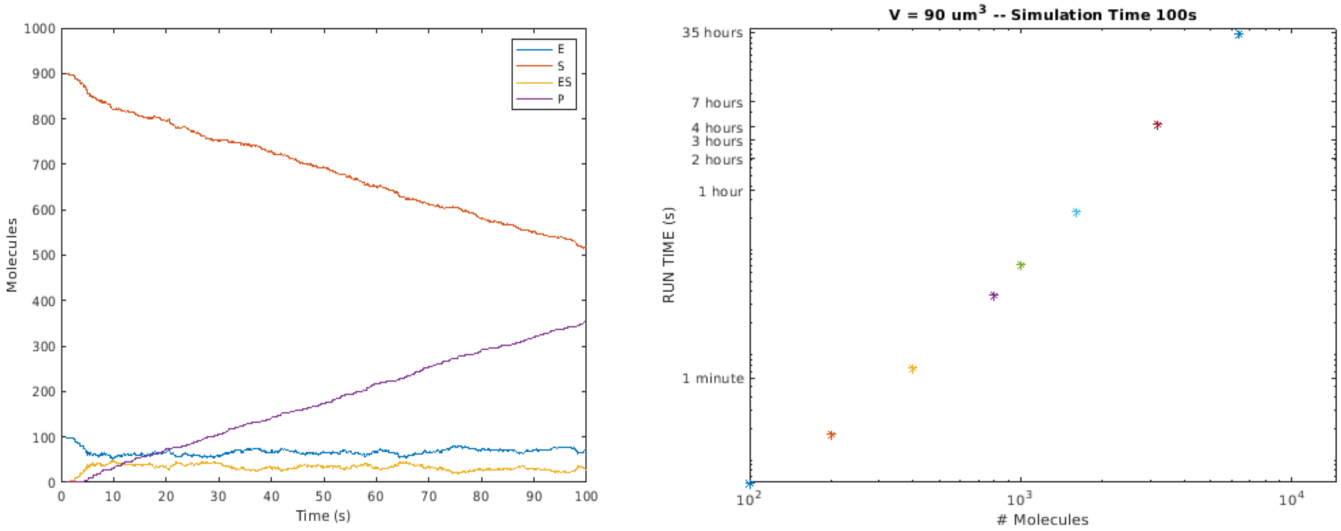

Figure 8. Finite particle representation. (**Left**) Time evolution of 1000 molecules in the Michaelis-Menten model with DESSA-CS. Unimolecular rate constant $k_{uni} = 0.1s^{-1}$ (governing $ES \Rightarrow E + S$ and $ES \Rightarrow E + P$), diffusion coefficient $D = 1\mu m^2 s^{-1}$, boundary condition parameter $c = 1.2 * 10^3 s^{-1}$, and $R_{enc} = 0.01\mu m$. (**Right**) The run time for the model increases roughly linearly in log space with the number of molecules - [100,200,400,800,1600,3200,6400,12000] at a fixed volume of 90 μm^3.

1000 Molecules, 100s Simulation Time					
Local Workstation: Ubuntu 14.04 LTS, 128 GB memory, Intel Xeon E5-2630 2.40GHz					
[40] Workstation: Ubuntu 16.04 LTS, 48 GB memory, Intel Xeon X5680 3.33GHz					
Software	Run Time	Sim Parameters	Boundary Type	Space / Time Steps	Workstation
DESSA-CS	100s	$r = 0nm$, $l_B = 0.1$	reflective	off-lattice / sampled	local
DESSA-CS	728s	$r = 10nm$, $l_B = 0.1$	reflective	off-lattice / sampled	local
eGFRD	10,561s	$r = 10nm$	periodic	off-lattice / variable	local
eGFRD	2,412s	$r = 1nm$	periodic	off-lattice / variable	[40]
eGFRD	3,246s	$r = 10nm$	periodic	off-lattice / variable	[40]
Smoldyn	20s	$\Delta t = 1ms$	periodic	off-lattice / fixed	[40]
Smoldyn	298s	$\Delta t = 67\mu s$	periodic	off-lattice / fixed	[40]
Spaciocyte MLM	13s	$\Delta t = 1ms$, $r = 38.73nm$	periodic	spatial lattice / fixed	[40]
Spaciocyte MLM	276s	$\Delta t = 67\mu s$, $r = 10nm$	periodic	spatial lattice / fixed	[40]

Table I. Method Comparison on Updated Benchmark from Chew et al. [40]. Diffusion coefficients are $1\mu m^2 s^{-1}$. The local eGFRD simulation was run using the open source simulation environment E-Cell version 4 [52]. DESSA-CS parameter l_B describes, as a percentage of the container length, how far a molecule's diffusion sphere may extend beyond the container during location sampling.

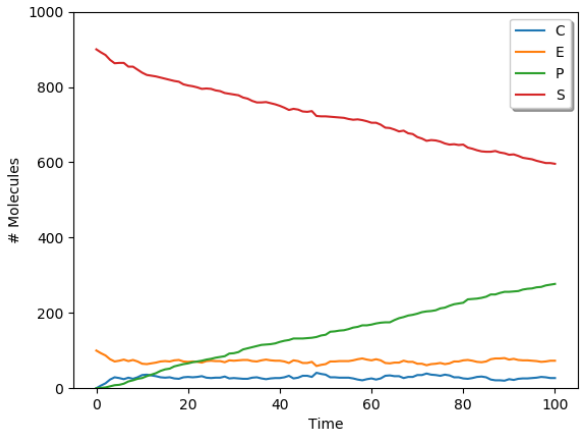

Figure 9. Time evolution of 1000 molecules in the Michaelis-Menten model with eGFRD in the E-Cell v4 environment. Unimolecular rate constant $k_{uni} = 0.1s^{-1}$ (governing $ES \Rightarrow E + S$ and $ES \Rightarrow E + P$), diffusion coefficient $D = 1\mu m^2 s^{-1}$, intrinsic bimolecular rate constant $k_{bimol} = 1 * 10^{-2} s^{-1}$, and particle radius $r = 0.01\mu m$ are taken from Figure 5 of Chew et al. [40]

novel relative to other spatial simulation methods. For point particles, we rely on two assumptions: (1) that reactions must happen in the region both reactants' diffusion spheres overlap and (2) that the probability distributions characterizing the possible distances either reactant has traveled are Gaussian and independent. This implies there are rings of equiprobable points at constant distance from the Gaussian means of the reactants and suggests a method to sample such a ring: first, sample a distance, r_A from one reactant, and then sample the angle w.r.t. the axis connecting the means given r_A. Each ring is uniquely determined by this distance and angle, and the reaction location can then be selected uniformly at random from on the ring. In the case of molecules with finite size, Green's functions governing radial separation must be used for wait time sampling, however the same ring sampling procedure applies to reaction locations. We compare our method with its most relevant competitor, eGFRD, on the modified Michaelis-Menten benchmark model described in Chew et al.[40]. The dynamics displayed by eGFRD, Spatiocyte (implementing a microscopic lattice method), and Smoldyn are quantitatively similar. DESSA-CS shows a substantial improvement over the run time of eGFRD, achieving run times more comparable to the discrete-time alternatives Smoldyn and Spaciocyte MLM. The method as presented leaves several avenues for extension and improvement in future work. DESSA-CS is able to achieve its comparatively high run time efficiency by exploiting the fact that, under certain assumptions, wait time sampling can be described by a deterministic part applicable in many circumstances, and a stochastic part specific to each reactant pair. We can therefore perform much of the expensive deterministic computations once, independently of each simulation run. Those assumptions include isotropic diffusion as the primary method of transport, and that reactions between distinct pairs of molecules are described by time-inhomogenous Poisson processes with mean parameter equal to the integrated propensity (this implies exponentially distributed waiting times). These are reasonable assumptions, yet both may be relaxed in future work. Numerically integrating these reaction propensities when considering new reactions at every step of the simulation can lead to the same computations being performed thousands or millions of times. Only the sampling of the exponentially distributed random numbers must be performed for all potential bimolecular reactions. The overall accuracy of the method is dependent on the resolution of the pre-computed integrated propensity curves. In future work, we will investigate more formally the nature of the accuracy/efficiency trade off associated with our handling of reflective boundaries. We will also consider methods for updating our unconstrained integrated reaction propensities to account for the boundaries of the simulation volume. One other avenue for improvement is the choice of programming language for the simulator. Matlab was chosen for its ease of use, as well as its testing and plotting infrastructure. Moving fully to Golang or C would likely lead to substantially greater run time efficiency.

ACKNOWLEDGMENTS

M.T. was supported in part by US National Institutes of Health award T32EB009403. R.S. was supported in part by U.S. National Institutes of Health awards 1R21CA216452 and 1R01HG010589. We are grateful to Dr. Jim Faeder for helpful advice on this project.

Appendix A: Evaluating $CDF_Q(y)$

Following Mathai and Provost [48], we consider the p dimensional Gaussian distributed random variable $X \sim N(\mu, \Sigma)$, $\sum > 0$ and the quadratic form $Q = X^T AX$, $A^T = A$. Q has the following alternate representation in terms of its eigenvalues:

$$Q = \sum_{j=1}^{p} \lambda_j (U_j + b_j)^2 \tag{A1}$$

Let P be a pxp matrix which diagonalizes $\Sigma^{1/2} A \Sigma^{1/2}$, i.e. $P^T \Sigma^{1/2} A \Sigma^{1/2} P = diag(\lambda_1, ..., \lambda_p)$, and $PP^T = I$. Now, $\mathbf{U^T} = (U_1, ..., U_p)$, $\mathbf{U} = P^T \Sigma^{-1/2} (X - \mu)$, $\mathbf{b^T} = (b_1, ..., b_p) = (P^T \Sigma^{-1/2} \mu)^T$, and the U_j's are mutually independent standard normal variables. We are interested in the distribution function (i.e. CDF of Q) which we here define as $F_p(\lambda, \mathbf{b}; y)$. It can be shown (see sections 4.1-4.2) that

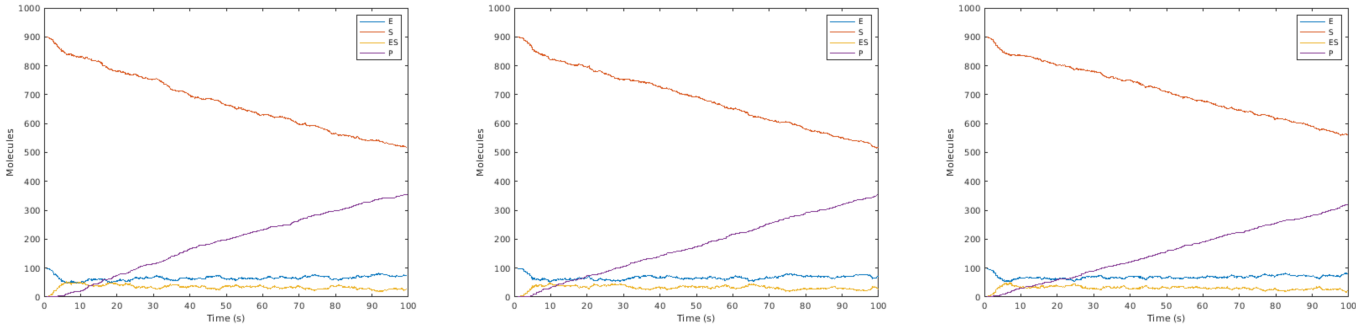

Figure 10. Varying the simulation volume's boundary parameter for the finite particle representation. Time evolution of 1000 molecules in the Michaelis-Menten model with DESSA-CS. Unimolecular rate constant $k_{uni} = 0.1s^{-1}$ (governing $ES \Rightarrow E + S$ and $ES \Rightarrow E + P$), diffusion coefficient $D = 1\mu m^2 s^{-1}$, boundary condition parameter $c = 1.2 * 10^3 s^{-1}$, and $R_{enc} = 0.01\mu m$. From left to right: $l_B = 0.03$ (runtime = 2693s), $l_B = 0.1$ (runtime = 777s), and $l_B = 0.3$ (runtime = 157s).

$$F_p(\lambda, \mathbf{b}; y) = \sum_{k=0}^{\infty} (-1)^k z_k \frac{y^{p/2+k}}{\Gamma(p/2+k+1)}, \tag{A2}$$

$0 < y < \infty$, with the following recursively defined coefficients.

$$z_0 = exp(-\frac{1}{2}\sum_{j=1}^{p} b_j^2) \prod_{j=1}^{p} (2\lambda_j)^{-1/2} \tag{A3}$$

$$z_k = \frac{1}{k}\sum_{r=0}^{k-1} d_{k-r} z_r, \quad k \geq 1 \tag{A4}$$

$$d_k = \frac{1}{2}\sum_{j=1}^{p} (1 - k b_j^2)(2\lambda_j)^{-k}, \quad k \geq 1 \tag{A5}$$

[1] C. Kaya, M. H. Cheng, E. R. Block, T. M. Bartol, T. J. Sejnowski, A. Sorkin, J. R. Faeder, and I. Bahar, Heterogeneities in axonal structure and transporter distribution lower dopamine reuptake efficiency, eNeuro **5** (2018).

[2] M. Thomas and R. Schwartz, Quantitative computational models of molecular self-assembly in systems biology, Physical biology **14**, 035003 (2017).

[3] M. Chevreuil, D. Law-Hine, J. Chen, S. Bressanelli, S. Combet, D. Constantin, J. Degrouard, J. Möller, M. Zeghal, and G. Tresset, Nonequilibrium self-assembly dynamics of icosahedral viral capsids packaging genome or polyelectrolyte, Nature communications **9**, 3071 (2018).

[4] E. Lopez-Fontal, A. Grochmal, T. Foran, L. Milanesi, and S. Tomas, Ship in a bottle: confinement-promoted self-assembly, Chemical science **9**, 1760 (2018).

[5] D. Wang, M. Hermes, R. Kotni, Y. Wu, N. Tasios, Y. Liu, B. De Nijs, E. B. Van Der Wee, C. B. Murray, M. Dijkstra, *et al.*, Interplay between spherical confinement and particle shape on the self-assembly of rounded cubes, Nature communications **9**, 2228 (2018).

[6] N. O. Junker, F. Vaghefikia, A. Albarghash, H. Höfig, D. Kempe, J. Walter, J. Otten, M. Pohl, A. Katranidis, S. Wiegand, *et al.*, The impact of molecular crowding on translational mobility and conformational properties of biological macromolecules, The Journal of Physical Chemistry B (2019).

[7] G. R. Smith, L. Xie, B. Lee, and R. Schwartz, Applying molecular crowding models to simulations of virus capsid assembly in vitro, Biophysical journal **106**, 310 (2014).

[8] B. Van Treeck, D. S. Protter, T. Matheny, A. Khong, C. D. Link, and R. Parker, Rna self-assembly contributes to stress granule formation and defining the stress granule transcriptome, Proceedings of the National Academy of Sciences **115**, 2734 (2018).

[9] D. T. Gillespie, A rigorous derivation of the chemical master equation, Physica A: Statistical Mechanics and its Applications **188**, 404 (1992).

[10] D. T. Gillespie, A diffusional bimolecular propensity function, The Journal of chemical physics **131**, 164109 (2009).

[11] D. T. Gillespie, A. Hellander, and L. R. Petzold, Perspective: Stochastic algorithms for chemical kinetics, The Journal of chemical physics **138**, 05B201_1 (2013).

[12] D. J. Warne, R. E. Baker, and M. J. Simpson, Simulation and inference algorithms for stochastic biochemical reaction networks: from basic concepts to state-of-the-art, Journal of the Royal Society Interface **16**, 20180943 (2019).

[13] A. Nag, M. I. Monine, J. R. Faeder, and B. Goldstein, Aggregation of membrane proteins by cytosolic cross-linkers: theory and simulation of the lat-grb2-sos1 system, Biophysical journal **96**, 2604 (2009).

[14] D. T. Gillespie, Approximate accelerated stochastic simulation of chemically reacting systems, The Journal of Chemical Physics **115**, 1716 (2001).

[15] M. Rathinam, L. R. Petzold, Y. Cao, and D. T. Gillespie, Stiffness in stochastic chemically reacting systems: The implicit tau-leaping method, The Journal of Chemical Physics **119**, 12784 (2003).

[16] Y. Cao, D. T. Gillespie, and L. R. Petzold, The slow-scale stochastic simulation algorithm, The Journal of chemical physics **122**, 014116 (2005).

[17] F. Jamalyaria, R. Rohlfs, and R. Schwartz, Queue-based method for efficient simulation of biological self-assembly systems, Journal of Computational Physics **204**, 100 (2005).

[18] N. Misra and R. Schwartz, Efficient stochastic sampling of first-passage times with applications to self-assembly simulations, The Journal of chemical physics **129**, 204109 (2008).

[19] D. F. Anderson, Incorporating postleap checks in tau-leaping, The Journal of chemical physics **128**, 054103 (2008).

[20] M. W. Sneddon, J. R. Faeder, and T. Emonet, Efficient modeling, simulation and coarse-graining of biological complexity with nfsim, Nature methods **8**, 177 (2011).

[21] R. M. Donovan, A. J. Sedgewick, J. R. Faeder, and D. M. Zuckerman, Efficient stochastic simulation of chemical kinetics networks using a weighted ensemble of trajectories, The Journal of chemical physics **139**, 09B642_1 (2013).

[22] Y. T. Lin, S. Feng, and W. S. Hlavacek, Scaling methods for accelerating kinetic monte carlo simulations of chemical reaction networks, The Journal of Chemical Physics **150**, 244101 (2019).

[23] F. Baras and M. M. Mansour, Reaction-diffusion master equation: A comparison with microscopic simulations, Physical Review E **54**, 6139 (1996).

[24] S. A. Isaacson, The reaction-diffusion master equation as an asymptotic approximation of diffusion to a small target, SIAM Journal on Applied Mathematics **70**, 77 (2009).

[25] S. Smith and R. Grima, Breakdown of the reaction-diffusion master equation with nonelementary rates, Physical Review E **93**, 052135 (2016).

[26] R. Schwartz, P. W. Shor, P. E. Prevelige Jr, and B. Berger, Local rules simulation of the kinetics of virus capsid self-assembly, Biophysical journal **75**, 2626 (1998).

[27] G. K. Bourov and A. Bhattacharya, The role of geometric constraints in amphiphilic self-assembly: A brownian dynamics study, The Journal of chemical physics **119**, 9219 (2003).

[28] M. F. Hagan and D. Chandler, Dynamic pathways for viral capsid assembly, Biophysical journal **91**, 42 (2006).

[29] R. A. Kerr, T. M. Bartol, B. Kaminsky, M. Dittrich, J.-C. J. Chang, S. B. Baden, T. J. Sejnowski, and J. R. Stiles, Fast monte carlo simulation methods for biological reaction-diffusion systems in solution and on surfaces, SIAM journal on scientific computing **30**, 3126 (2008).

[30] B. T. Castle and D. J. Odde, Brownian dynamics of subunit addition-loss kinetics and thermodynamics in linear polymer self-assembly, Biophysical journal **105**, 2528 (2013).

[31] P. Castro-Villarreal, A. Villada-Balbuena, J. M. Méndez-Alcaraz, R. Castañeda-Priego, and S. Estrada-Jiménez, A brownian dynamics algorithm for colloids in curved manifolds, The Journal of chemical physics **140**, 214115 (2014).

[32] S. J. Bachmann, M. Petitzon, and B. M. Mognetti, Bond formation kinetics affects self-assembly directed by ligand–receptor interactions, Soft matter **12**, 9585 (2016).

[33] A. Donev, C.-Y. Yang, and C. Kim, Efficient reactive brownian dynamics, The Journal of chemical physics **148**, 034103 (2018).

[34] S. S. Andrews, N. J. Addy, R. Brent, and A. P. Arkin, Detailed simulations of cell biology with smoldyn 2.1, PLoS computational biology **6**, e1000705 (2010).

[35] J. S. van Zon and P. R. Ten Wolde, Green's-function reaction dynamics: a particle-based approach for simulating biochemical networks in time and space, The Journal of chemical physics **123**, 234910 (2005).

[36] J. S. van Zon and P. R. Ten Wolde, Simulating biochemical networks at the particle level and in time and space: Green's function reaction dynamics, Physical review letters **94**, 128103 (2005).

[37] T. R. Sokolowski, J. Paijmans, L. Bossen, T. Miedema, M. Wehrens, N. B. Becker, K. Kaizu, K. Takahashi, M. Dogterom, and P. R. ten Wolde, egfrd in all dimensions, The Journal of chemical physics **150**, 054108 (2019).

[38] T. Opplestrup, V. V. Bulatov, G. H. Gilmer, M. H. Kalos, and B. Sadigh, First-passage monte carlo algorithm: diffusion without all the hops, Physical review letters **97**, 230602 (2006).

[39] D. T. Gillespie, E. Seitaridou, and C. A. Gillespie, The small-voxel tracking algorithm for simulating chemical reactions among diffusing molecules, The Journal of chemical physics **141**, 12B649_1 (2014).

[40] W.-X. Chew, K. Kaizu, M. Watabe, S. V. Muniandy, K. Takahashi, and S. N. Arjunan, Reaction-diffusion kinetics on lattice at the microscopic scale, Physical Review E **98**, 032418 (2018).

[41] M. Von Smoluchowski, Mathematical theory of the kinetics of the coagulation of colloidal solutions, Z. Phys. Chem **92**, 129 (1917).

[42] F. C. Collins and G. E. Kimball, Diffusion-controlled reaction rates, Journal of colloid science **4**, 425 (1949).

[43] K. R. Naqvi, S. Waldenstrøm, and K. Mork, Kinetics of diffusion-mediated bimolecular reactions. a new theoretical framework, The Journal of Physical Chemistry **86**, 4750 (1982).

[44] R. Noyes, Models relating molecular reactivity and diffusion in liquids, Journal of the American Chemical Society **78**, 5486 (1956).

[45] R. Noyes, Effects of diffusion rates on chemical kinetics, in *Progress in Reaction Kinetics*, Vol. 1, edited by S. Benson (1961) pp. 129–160.

[46] T. Zhang, R. Rohlfs, and R. Schwartz, Implementation of a discrete event simulator for biological self-assembly systems, in *Proceedings of the 37th conference on Winter simulation* (Winter Simulation Conference, 2005) pp. 2223–2231.

[47] D. F. Anderson, A modified next reaction method for simulating chemical systems with time dependent propensities and delays, The Journal of chemical physics **127**, 214107 (2007).

[48] A. M. Mathai and S. B. Provost, *Quadratic forms in random variables: theory and applications* (Dekker, 1992).

[49] M. Thomas, DESSA-CS (2020).

[50] H. S. Carslaw and J. C. Jaeger, Conduction of heat in solids, Oxford: Clarendon Press, 1959, 2nd ed. (1959).

[51] H. Van Beijeren, W. Dong, and L. Bocquet, Diffusion-controlled reactions: A revisit of noyes' theory, The Journal of Chemical Physics **114**, 6265 (2001).

[52] K. Kaizu, K. Nishida, Y. Sakamoto, S. Kato, T. Niina, N. Nishida, M. Koizumi, N. Aota, , and K. Takahashi, E-Cell version 4.