UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

# Multiscale simulations of anisotropic particles combining molecular dynamics and Green's function reaction dynamics

Vijaykumar, A.; Ouldridge, T.E.; ten Wolde, P.R.; Bolhuis, P.G.

DOI
10.1063/1.4977515

Publication date
2017

Document Version
Final published version

Published in
Journal of Chemical Physics

License
Article 25fa Dutch Copyright Act

Link to publication

Citation for published version (APA):
Vijaykumar, A., Ouldridge, T. E., ten Wolde, P. R., & Bolhuis, P. G. (2017). Multiscale simulations of anisotropic particles combining molecular dynamics and Green's function reaction dynamics. *Journal of Chemical Physics*, *146*(11), [114106]. https://doi.org/10.1063/1.4977515

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

*UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)*

Download date:09 Mar 2023

# Multiscale simulations of anisotropic particles combining molecular dynamics and Green's function reaction dynamics

Adithya Vijaykumar,[1,2] Thomas E. Ouldridge,[3] Pieter Rein ten Wolde,[1,a)] and Peter G. Bolhuis[2,b)]
[1]*FOM Institute AMOLF, Science Park 104, 1098 XG Amsterdam, The Netherlands*
[2]*Van 't Hoff Institute for Molecular Sciences, University of Amsterdam, P.O. Box 94157, 1090 GD Amsterdam, The Netherlands*
[3]*Department of Bioengineering, Imperial College London, South Kensington Campus, London SW7 2AZ, United Kingdom*

(Received 18 October 2016; accepted 10 February 2017; published online 21 March 2017)

The modeling of complex reaction-diffusion processes in, for instance, cellular biochemical networks or self-assembling soft matter can be tremendously sped up by employing a multiscale algorithm which combines the mesoscopic Green's Function Reaction Dynamics (GFRD) method with explicit stochastic Brownian, Langevin, or deterministic molecular dynamics to treat reactants at the microscopic scale [A. Vijaykumar, P. G. Bolhuis, and P. R. ten Wolde, J. Chem. Phys. **143**, 214102 (2015)]. Here we extend this multiscale MD-GFRD approach to include the orientational dynamics that is crucial to describe the anisotropic interactions often prevalent in biomolecular systems. We present the novel algorithm focusing on Brownian dynamics only, although the methodology is generic. We illustrate the novel algorithm using a simple patchy particle model. After validation of the algorithm, we discuss its performance. The rotational Brownian dynamics MD-GFRD multiscale method will open up the possibility for large scale simulations of protein signalling networks. *Published by AIP Publishing.* [http://dx.doi.org/10.1063/1.4977515]

## I. INTRODUCTION

Complex systems such as biochemical networks in living cells, catalytic reactions in, e.g., a fuel cell, surfactant/water/oil mixtures, or self-assembling soft matter can be modeled efficiently as reaction-diffusion systems. In such reaction-diffusion systems, the spatial distribution of reactants and the stochastic nature of their interactions are crucial for the system's macroscopic behaviour. At sufficiently low concentrations, the time taken for the reactants to diffuse and randomly find each other is much larger than the time required for the reaction. For example, in cellular systems, the concentrations of proteins are often in the nM – $\mu$M range. Experiments indicate that proteins inside the living cell move by normal diffusion[1] with effective diffusion constants in the $1-10\ \mu m^2\ s^{-1}$ range. This means that, with typical protein cross sections of 10 nm, the time it takes for reactants to find each other is on the order of milliseconds to seconds. This is often much longer than the microsecond time scales on which the actual association events occur once the particles have found each other.[2,3] Reaction-diffusion systems thus often exhibit a strong separation of length and time scales, with the diffusive search process happening on length and time scales of microns and milliseconds to seconds, and the reactions occurring on scales of nanometers and sub-milliseconds.[3] Simulating such systems with conventional, brute-force simulation techniques is notoriously difficult. Indeed, simulating cellular biochemical networks with straightforward brute-force Brownian Dynamics (BD)[4–7] often means that most CPU time is spent on propagating the particles towards one another.[8] To overcome the inefficiency of straightforward BD requires special techniques such as Green's Function Reaction Dynamics (GFRD).[9,10]

GFRD is a mesoscopic technique that decomposes the many particle reaction diffusion problem into sets of one- and two-body problems that can be solved analytically. This is achieved by putting single particles and pairs of particles in the so-called *protective* domains that do not overlap with each other. For each of these domains, the reaction-diffusion problem is solved analytically using Green's functions. This yields for each domain a next event *type* which can either be a reaction in the domain or an escape from the domain, as well as a next event *time*, i.e., the time at which this event occurs. These events are put in a scheduler list which is updated chronologically. This makes GFRD an asynchronous, event-driven algorithm. Since stochastic processes in the individual domains are independent of each other, GFRD is an exact algorithm to simulate large reaction-diffusion systems. As the particles make huge leaps in space and time in GFRD, the computational effort in propagating the particles to one another is greatly reduced, making GFRD orders of magnitude faster than brute force BD. However, the particles are assumed to be idealized spheres interacting via an isotropic potential and the reactions to occur according to intrinsic rates in pair domains. Solving the Green's function for reactive events involving the complex anisotropic potentials required for proper modeling of proteins or other molecules is extremely

a)Electronic mail: p.t.wolde@amolf.nl
b)Electronic mail: p.g.bolhuis@uva.nl

cumbersome and in fact most likely will reduce the efficiency of the GFRD approach substantially. In contrast, straightforward BD is able to naturally simulate orientational dynamics of protein particles with complex anisotropic (effective) interactions.

This observation raises the question whether it is possible to combine the computational power of GFRD with the microscopic detail of BD. In previous work, we introduced a novel multi-scale scheme, called Molecular Dynamics-GFRD (MD-GFRD), which combines GFRD with a microscopic simulation technique such as deterministic molecular dynamics (MD), stochastic Langevin Dynamics (LD), or Brownian Dynamics (BD).[11] In this scheme, GFRD handles diffusion of particles at the mesoscopic scale, while MD, LD, or BD treats the particles that are coming close to each other. In previous work and here, we limit ourselves to BD, although the scheme can very easily formulated for MD and LD. The multi-scale algorithm defines the micro- and meso-scopic regions adaptively on the fly and switches seamlessly between the two techniques based on predefined scenarios.

In this work, we extend MD-GFRD to incorporate the orientational dynamics of particles that interact via an anisotropic potential. As in the original MD-GFRD technique,[11] GFRD is used for propagating the particles towards one another when they are far apart. Once the particles are within a predefined threshold distance from each other, the algorithm switches to BD. The complex orientational dynamics, once the particles are close together, is thus simulated with BD. When the particles are bound, MD-GFRD could in principle continue to simulate these particles with BD. However, in many cases, and typically in cellular systems, the particles are bound much longer than the time it takes to diffuse and thermalise within the interaction well, meaning that dissociation is a rare event. MD-GFRD exploits this separation of time scales by treating the dissociation as a first order reaction, with an intrinsic dissociation rate constant that has been pre-determined. After dissociation, the particles can be propagated again with GFRD. Importantly, however, after dissociation, the particles do not immediately loose their orientational memory, which means that they must be propagated with Green's functions that do not only describe the translational dynamics of the particles but also their orientational dynamics. In this paper, we describe in detail how the MD-GFRD scheme switches between MD and GFRD and how this switching depends on the translational and orientational dynamics of the particles. We also present the Green's functions that allow GFRD to simulate the particles' orientational dynamics.

The remainder of the paper is organized as follows: In Sec. II we first give an overview of the MD-GFRD algorithm. Then we describe how the algorithm simulates the diffusion of particles with rotational degrees of freedom, both for particles in BD and GFRD mode. We discuss how MD-GFRD handles the association-dissociation reactions, and we describe how it switches between BD and GFRD propagation. In many systems, including that studied here, dissociation is a rare event. This means that computing the intrinsic dissociation rate constant, as used by MD-GFRD, requires rare event methodologies, like Transition Interface Sampling (TIS)[12] and Forward Flux Sampling (FFS).[13] Here, we briefly describe how we use FFS to pre-compute the dissociation rate constant. We then illustrate the new technique by simulating the association and dissociation of patchy particles. In many cases, globular proteins can be coarse-grained as the so-called *patchy particles*, where the complex binding sites are modeled as patches on a spherical particle. These patchy particles also play an important role in the modeling of soft matter.[14,15] We demonstrate that the algorithm reproduces quantities that can be obtained analytically such as the equilibrium constants, binding probabilities, and the power spectra of the binding reactions. We end with a discussion of the performance of the algorithm.

## II. METHODS

### A. Summary of multiscale approach

The MD-GFRD algorithm is a generic algorithm that enables simulation of any reaction-diffusion system at the particle level. It allows for mono-molecular reactions of the type $A \rightarrow B + C + \cdots$ and bi-molecular reactions of the type $A + B \rightarrow C + D + \cdots$. By combining these two reactions, any complex biochemical network can be simulated. Here, however, we will limit ourselves to simple association-dissociation reactions $A + B \leftrightarrows C$. The MD-GFRD algorithm distinguishes two types of particles as shown in Fig. 1: (1) BD particles that are propagated collectively in a conventional, brute-force manner using small time steps and (2) GFRD particles that are updated asynchronously in an event-driven manner. Single particles that are sufficiently far away from all other particles according to a predefined cut-off distance are put into protective domains. For each of these domains, the algorithm determines, as in the conventional GFRD scheme,[10] the next-event *type*, which is either a mono-molecular decay reaction (such as dissociation) or an exit of the particle from the domain and the corresponding next-event *time*, which is when this next event will happen. The next-event times of the respective GFRD domains are put in a chronologically ordered event list, which

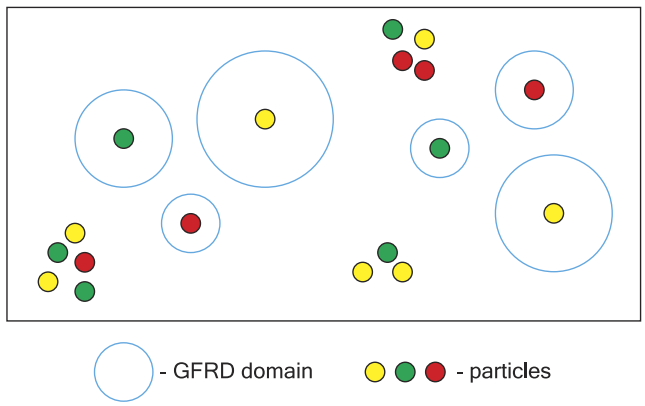


FIG. 1. MD-GFRD scheme: particles that are far away other particles are put into a GFRD domain. For each GFRD domain, the next event *time* and *type* is determined. These next event times are added to a chronologically ordered event list and updated when the simulation time has reached the time of the next event. Particles that are close to other particles are propagated collectively with Brownian dynamics.

is updated only when the simulation time has reached the time of the first next event. The event-driven nature of GFRD allows MD-GFRD to make large jumps in space and time when the domains are large. It is the origin of the high efficiency of the scheme.

The other particles are simulated explicitly with BD. This part of the algorithm takes into account the forces between the particles when they come within the interaction range of the potential from each other. The BD propagation also naturally simulates the association reaction A + B → C: two particles A and B form the bound complex C when they enter the well of the interaction potential. The two monomers A and B in the dimer C could be propagated separately with BD, but it is more efficient to propagate them as a single particle C. The dissociation of C into A and B is then treated as a uni-molecular reaction event, which is added to the event list.

BD propagation is continued until one of the following events occurs: (i) the escape of a particle from a GFRD domain; (ii) the decay of a GFRD particle, e.g., the dissociation of C into A and B; (iii) a BD particle dissociates into its products, e.g., the dissociation of C into A and B; (iv) two BD particles A and B bind each other to form a dimer species C; (v) a BD particle comes too close to a GFRD domain so that the GFRD domain must be burst, which means that a position for the particle in that domain is generated at the current simulation time; and (vi) a BD particle moves sufficiently far away from all other BD particles and GFRD domains, so that it can be put into a GFRD domain. These six possible events are illustrated in Fig. 3. After the event has been executed, the system is updated accordingly; for newly formed GFRD domains, the next-event types and times are determined and inserted into the event list. The propagation of the BD particles is then resumed. The scheme becomes particularly powerful when most particles are in GFRD domains. A key objective is thus to keep the number of BD particles to a minimum.

The multiscale method that we pursue here involves particles interacting via anisotropic potentials. This requires an explicit BD integrator allowing rotational dynamics. Moreover, the GFRD part requires rotational Green's functions. In Secs. II B and II C we provide these ingredients, which constitute the most salient differences of the novel scheme with the previous isotropic MD-GFRD scheme.[11] In Section II E, we discuss in detail how the algorithm switches between GFRD and BD. Secs. II D and II F describe how MD-GFRD handles the dissociation events and how the dissociation rate constant, needed in MD-GFRD, can be computed efficiently. In Sec. II G, we describe the specific interaction potential used to illustrate how orientations can be included in MD-GFRD.

### B. Brownian dynamics of patchy particles

Brownian dynamics is used to simulate the solute particles at the microscopic scales. In this algorithm, the position and the orientation of each solute particle in the BD regime are updated based on the total force and torque acting on the particle. The force and torque contain a deterministic component, which arises from the (solvent-mediated) interaction potential with the other solute particles and the frictional drag from the solvent, and a stochastic component, originating from the stochastic forces exerted by the solvent molecules. Although the interactions between particles are anisotropic, we model the particles as spheres of finite radius. We represent the rigid body orientation of the particles using a four component unit vector known as a quaternion, $q = (q_0, q_1, q_2, q_3)$. The quaternion is an efficient encoding of the rotation matrix, $A$, given by

$$A = \begin{bmatrix} {q_0}^2 + {q_1}^2 - {q_2}^2 - {q_3}^2 & 2(q_1q_2 + q_0q_3) & 2(q_1q_3 - q_0q_2) \\ 2(q_1q_2 - q_0q_3) & {q_0}^2 - {q_1}^2 + {q_2}^2 - {q_3}^2 & 2(q_2q_3 + q_0q_1) \\ 2(q_1q_3 + q_0q_2) & 2(q_2q_3 - q_0q_1) & {q_0}^2 - {q_1}^2 - {q_2}^2 + {q_3}^2 \end{bmatrix},$$

which relates vectors in the stationary lab frame, $\hat{u}_s$, to the vectors in the moving body frame, $\hat{u}_b$ via

$$\hat{u}_s = A^T \hat{u}_b. \quad (1)$$

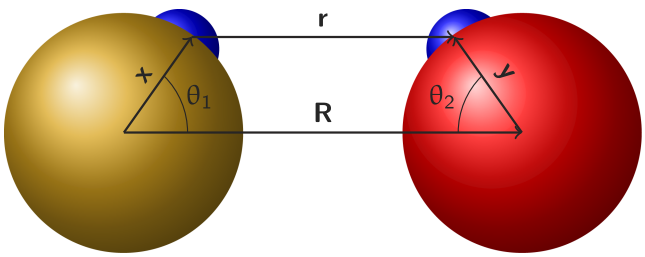


FIG. 2. Each particle may have one or more attractive regions on its surface, called "patches," that facilitate short ranged, highly directional attractive interactions.

For example, the vectors $\hat{u}_b$ might point to the patches on the surface of the particle which are fixed in the body frame.

Each particle has a center of mass and one or more sticky spots on its surface called "patches" (see Fig. 2). The particles interact with each other both via a center of mass isotropic pair potential and via a short ranged isotropic patch-patch interaction. We describe the anisotropic model potential that we employ for illustrative purposes in detail in Sec. II G. We note that the choice of potential is not limited to simple models. In principle any other anisotropic complex potential can be used, even an anisotropic protein-protein interaction derived from all atom MD simulations.

The particles are propagated with the first order Brownian dynamics integrator[16] explained in Algorithm 1:

**Algorithm 1.** The Brownian dynamics integrator[16] used in our multi-scale scheme. We consider $n$ particles in three dimensions with center of mass coordinates $\mathbf{r} = (r^{1\top}, \ldots, r^{n\top})^\top \in \mathbb{R}^{3n}$, $r^j = (r_1^j, r_2^j, r_3^j)^\top \in \mathbb{R}^3$, and rotational coordinates in the quaternion representation $\mathbf{q} = (q^{1\top}, \ldots, q^{n\top})^\top$, $q^j = (q_0^j, q_1^j, q_2^j, q_3^j)^\top \in \mathbb{S}^3$, such that $|q^j| = 1$. Particles are characterized by their mass $m$, the mass moment of inertia $M = \frac{8}{15} m\sigma^2$, and the translational and rotational friction coefficients, $\gamma$ and $\Gamma$, respectively. These parameters can differ among species. Note that $\gamma$ and $\Gamma$ as defined in Ref. 16 contain an implicit dependence on $m$ and $M$, respectively. In the overdamped limit considered here, this dependence is cancelled by factors of $m$ and $M$ that always appear together with $\gamma$ and $\Gamma$. Further more, $\delta t$ is the time step used in the simulations, $\beta = \frac{1}{k_B T}$, $\xi_k$ and $\eta_k^{j,l}$ are the independent and identically distributed (i.i.d.) Gaussian random variables. $\mathbf{f}$ is the total force and $\mathbb{F}$ is obtained from total torque, which follow from the interaction potential. Finally, we define three $4 \times 4$ matrices $S_1 = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix}, S_2 = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, S_3 = \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$.

The particle positions $\mathbf{R}$ and quaternions $\mathbf{Q}$ are updated as follows:

$$\mathbf{R}_0 = \mathbf{r}, \mathbf{Q}_0 = \mathbf{q}, |q^j| = 1, j = 1, \ldots, n,$$
$$\mathbf{R}_{k+1} = \mathbf{R}_k + \frac{\delta t}{\gamma m}\mathbf{f}(\mathbf{R}_k, \mathbf{Q}_k) + \sqrt{\delta t}\sqrt{\frac{2}{\gamma\beta m}}\xi_k,$$
$$Y_k^j = \frac{\delta t}{\Gamma M}\mathbb{F}_j(\mathbf{R}_k, \mathbf{Q}_k) + \sqrt{\delta t}\sqrt{\frac{2}{\Gamma\beta M}}\sum_{l==1}^{3}\eta_k^{j,l} S_l,$$
$$Q_{k+1}^j = \exp(Y_k^j) Q_k^j.$$

### C. Green's functions for rotations

GFRD handles the free diffusion of single particles. A freely moving particle will undergo rotational as well as translational diffusion. Although the interactions between particles are anisotropic, we model the particles as spheres of finite radius for the purpose of modeling diffusion. This assumption allows the decoupling of the rotational and translational diffusion of isolated particles, which is possible since in MD-GFRD the GFRD domains only contain single particles. The Green's functions for translational diffusion are given by the Green's functions for single particles inside single GFRD domains, detailed in previous work.[11] These Green's functions determine (probabilistically) when the particles escape from their respective domains, or what their radial positions inside the domains become when the domains are burst. Although rotational motion does not influence the center-of-mass dynamics of a freely diffusing particle, and hence cannot cause escape from single domains, it is nonetheless important to reproduce the decorrelation of orientations for particles evolving under GFRD. For example, simply drawing orientations at random when a particle leaves a GFRD single domain will lead to unphysically rapid decorrelation of orientations when domains are short-lived and influence properties such as rebinding probability.

More specifically, on bursting or escape from a single domain, a new orientation $\Omega$ is drawn using the Green's function $G(\Omega, \Omega_0, t)$, with $\Omega_0$ being the initial orientation and $t$ the time since domain formation. The Green's functions, expressed in terms of Euler angles $\alpha, \beta, \gamma$, can be found in the literature.[17–19] For particles with spherically symmetric diffusion tensors, the Green's function is

$$G(\alpha, \beta, \gamma, \alpha_0, \beta_0, \gamma_0, t) = \sum_{L=0}^{\infty}\sum_{K,M=-L}^{L}\frac{2L+1}{8\pi^2}D_{K,M}^{(L)*}(\alpha_0, \beta_0, \gamma_0) \times D_{K,M}^{(L)}(\alpha, \beta, \gamma)\exp(-D_r L(L+1)t), \quad (2)$$

where $D_r$ is the threefold degenerate eigenvalue of the diffusion tensor, given by $D_r = k_B T/(8\pi\eta R^3)$ for a particle of radius $R$ in a fluid of viscosity $\eta$. The quantities $D_{K,M}^{(L)}(\alpha, \beta, \gamma)$ and its complex conjugate $D_{K,M}^{(L)*}(\alpha, \beta, \gamma)$ are elements of the Wigner rotation matrices[17–19]

$$D_{K,M}^{(L)}(\alpha, \beta, \gamma) = \exp(-iK\alpha)\, d_{K,M}^{(L)}(\beta)\exp(-iL\gamma), \quad (3)$$

with

$$d_{K,M}^{(L)}(\beta) = ((L+K)!(L-K)!(L+M)!(L-M)!)^{1/2}\sum_{S=\max(0,M-K)}^{\min(L+M,L-K)}\left(\frac{(-1)^{K-M+S}\cdot[\cos(\beta/2)]^{2L+M-K-2S}[\sin(\beta/2)]^{K-M+2S}}{(L+M-S)!S!(K-M+S)!(L-K-S)!}\right). \quad (4)$$

For the purposes of clarity, we emphasize that the Euler angles used here should be understood in the following way. If a body frame $B$ has an orientation $\Omega = (\alpha, \beta, \gamma)$ with respect to some reference frame $F$, then $B$ can be obtained from $F$ by the following:

1. Rotating $F$ around $F_z$ by $\gamma$ to give $F'$.
2. Rotating $F'$ around $F_y$ by $\beta$ to give $F''$.
3. Rotating $F''$ around $F_z$ by $\alpha$ to give $B$.

Moreover, note that the Green's functions are defined without the Jacobian, so that $(\alpha, \beta, \gamma)$ should be drawn from the distribution $\sin(\beta)G(\alpha, \beta, \gamma, \alpha_0, \beta_0, \gamma_0, t)$.

Drawing directly from such a distribution is non-trivial. However, rejection sampling can be used if the maximum of $\sin(\beta)G(\alpha, \beta, \gamma, \alpha_0, \beta_0, \gamma_0, t)$ is known. Physically, the most likely orientation is always aligned with the initial direction, which suggests a rejection scheme in which a trial orientation $(\alpha, \beta, \gamma)$ is drawn uniformly from $([0, 2\pi], [0, \pi], [0, 2\pi])$, and accepted with a probability

$$\frac{\sin(\beta)G(\alpha, \beta, \gamma, \alpha_0, \beta_0, \gamma_0, t)}{\sin(\beta_0)G(\alpha_0, \beta_0, \gamma_0, \alpha_0, \beta_0, \gamma_0, t)},$$

with Euler angles defined with respect to the lab frame. Unfortunately, the angular Jacobian implies that $\sin(\beta)G(\alpha, \beta, \gamma, \alpha_0, \beta_0, \gamma_0, t)$ is not in general maximized by $(\alpha = \alpha_0, \beta = \beta_0, \gamma = \gamma_0)$, violating a requirement of rejection sampling. It is true, however, that $\sin(\beta)G(\alpha, \beta, \gamma, \alpha_0, \beta_0, \gamma_0, t)$

is maximized by $(\alpha = \alpha_0, \beta = \beta_0, \gamma = \gamma_0)$ if $\beta_0 = \pi/2$, $\alpha_0 = 0, \gamma_0 = 0$. We therefore define a new reference frame $F_{\text{temp}}$ for each calculation such that the particle initially has orientation $\Omega_0 = (0, \pi/2, 0)$ with respect to $F_{\text{temp}}$. Using rejection sampling, we can then obtain a new orientation $\Omega = (\alpha, \beta, \gamma)$ with respect to $F_{\text{temp}}$. The particle orientation is updated by first rotating the particle by $-\pi/2$ about the $z$-axis of the original particle frame to obtain a particle aligned with $F_{\text{temp}}$ and then performing rotations $(\alpha, \beta, \gamma)$ about the axes of $F_{\text{temp}}$ as outlined above.

Even with rejection sampling, drawing from the distribution can be computationally challenging due to the costs of evaluating Green's functions. Eq. (2) has an infinite sum that must be truncated; we perform truncation when new contributions are smaller than the current value by a factor of $10^8$. To reduce the cost of the summations, we find it helpful to tabulate factorials. We also note that terms in Eq. (2) can be combined in complex conjugate pairs to eliminate imaginary numbers during the calculation.

Accurate evaluation of the Green's function is most challenging when $D_r t < 1$, when $G(\alpha, \beta, \gamma, \alpha_0, \beta_0, \gamma_0, t)$ is sharply peaked and many terms are needed. For small $D_r t$, we use early rejection, discarding a large fraction of draws of $(\alpha, \beta, \gamma)$ if $(\alpha, \beta - \beta_0, \gamma)$ is large without evaluating $G(\alpha, \beta, \gamma, \alpha_0, \beta_0, \gamma_0, t)$, and compensating for this bias at the acceptance stage. Finally, for values of $D_r t < 0.05$, we use the approximate approach of rotating about a random axis through an angle $\phi = \sqrt{\phi_x^2 + \phi_y^2 + \phi_z^2}$, where $\phi_i$ are i.i.d. random variables drawn from a Gaussian of mean 0 and variance $2D_r t$.[20]

### D. Handling the dissociation/association reaction in MD-GFRD

While particles that are sufficiently far away from each other can be propagated with GFRD, particles that are within a pre-defined cutoff distance from each other will be propagated with MD, or, as we restrict ourselves to here, BD. As described in more detail in Sec. II E, this cut-off distance is beyond the range of the interaction potential, $r_c$. Indeed, the association between two particles, which is driven by their inter-molecular attraction forces, is thus simulated explicitly with BD. Also the dissociation reaction could in principle be simulated with BD: we could explicitly simulate the bound monomers in the dimer A-B, until they dissociate again into A and B. However, the bound state is typically very stable: the time the particles spend inside the potential well is typically much longer than the time it takes for the particles to loose their orientation and thermalise inside the well. Simulating these particles explicitly means that much CPU time would be wasted on propagating them while they simply rattle around each other inside the potential well. In MD-GFRD, we therefore exploit that dissociation is a rare event: when two BD particles meet a predefined criterion signifying that they are deep inside the interaction well, the two "reactants" A and B are replaced by species C. In turn, the dissociation of C into A and B is treated as a first-order reaction $\text{C} \rightarrow \text{A} + \text{B}$ with a dissociation rate constant $k_\text{d}$.

More specifically, when two BD particles come within a distance such that their interaction energy $E$ drops below some predefined threshold $E_{\text{bind}}$, then the particles A and B are replaced by a single particle of species C, with a position that is given by the center-of-mass of the reactants A and B. If space permits, the C particle is directly put into a GFRD domain, which significantly speeds up the simulation. If there is no space to construct a protective domain, the C particle is propagated with BD. The C particle then diffuses, either explicitly with BD or implicitly with GFRD, until it dissociates again into the monomers A and B at a later time $\tau_\text{d}$. Since the interaction well is deep, $\tau_\text{d}$ will be exponentially distributed

$$q_\text{d}(t) = k_\text{d} e^{-k_\text{d} t}. \tag{5}$$

Knowing the dissociation rate constant $k_\text{d}$, the time $\tau_\text{d}$ can thus be sampled from

$$\tau_\text{d} = -k_\text{d} \ln(\mathcal{R}_\text{d}), \tag{6}$$

where $\mathcal{R}_\text{d} \in [0, 1]$ is a uniformly distributed random number.

The intrinsic dissociation rate constant $k_\text{d}$ could in principle be inferred from experiments. However, an approach that is consistent with a brute-force simulation of the same model is to obtain $k_\text{d}$ from a simulation that is performed prior to the MD-GFRD simulation of interest. This pre-simulation can then also be used to generate the distributions of the positions and orientations of A and B at the moment of dissociation. In the MD-GFRD simulation, the positions and orientations of the particles at the moment of dissociation can then be sampled from these distributions, respectively.

In our previous study on isotropic potentials, we determined $k_\text{d}$ by performing a brute force BD simulation of two particles prior to the MD-GFRD simulations.[11] However, the particles in our model interact via an anisotropic interaction potential. This anisotropic interaction is mediated via patches on the surfaces of the particles, see Fig. 2. The range of the patch-mediated interaction must be short, in order to provide a strong anisotropy in the interaction. The short range, however, means that the well of the patch-mediated potential must be deep in order to induce significant binding: the depth of the well, $\sim 25 k_\text{B} T$, is much larger than that of isotropic particles, $\sim 5 k_\text{B} T$. The deep well makes it very hard to obtain good statistics in determining the distribution of dissociation times via brute force simulations. However, it is possible to efficiently compute the dissociation rate with rare event techniques such as Transition Interface Sampling[12] or Forward Flux Sampling (FFS).[13] Here we use the latter technique, which we describe in Section II F.

### E. Coupling BD and GFRD

Now that we have described how MD-GFRD simulates the association and dissociation of two particles A and B, we will discuss how the algorithm switches between BD and GFRD when simulating many particles. At any one point in time, the simulation consists of a set of isolated particles inside GFRD domains that each has a radius of at least $d_{\text{min}}$ and a set of particles that are propagated with BD and interact with each other via a pair potential that has an interaction range $r_\text{c}$. There is also a chronologically ordered next-event list that contains the

times at which the GFRD particles escape from their respective domains, and the times at which the respective particles dissociate, be they in GFRD or BD mode. The particles that are not inside GFRD domains are propagated with brute-force BD until the first next-event happens. This event can be an event from the next-event list, but it can also be the formation of a GFRD domain or the bursting of a GFRD domain when a BD particle comes too close it. After the event has been executed, BD propagation is resumed.

Specifically, before each step of BD propagation, the algorithm checks for the following events, as illustrated in Fig. 3.

#### 1. Escape from a GFRD domain

When the next event in the list is a particle that escapes from a single domain that particle is put at a random center of mass position on the surface of the domain, with an orientation sampled from Eq. (2). The domain is removed and the particle is put in BD mode. This event is shown in Fig. 3 I. Note that at the next BD time step, the algorithm will check whether the particle can be put into a protective GFRD domain again (see Sec. II E 6).

#### 2. Dissociation inside a GFRD domain

When the next event is a particle C inside a GFRD domain decaying into its products at time $t$, the domain is burst and a new radial position $r$ for the reactant is generated according to the normalized translational Green's function $p(r,t,|r_0,t_0)/S(t - t_0|r_0)$, where $r_0$ is the original position of the particle, which is the center of the domain constructed at time $t_0$, and $S(t - t_0|r_0)$ is the survival probability. The reactant is replaced by its products, whose configuration is chosen at random from the ensemble of configurations recorded at the moment of dissociation, obtained in the FFS pre-simulation. This event is shown in Fig. 3 II.

#### 3. Dissociation of a BD particle

When the next event is the dissociation of a BD particle, the particle is replaced by its products, whose configurations are chosen at random from the ensemble of configurations recorded at the moment of dissociation in the FFS pre-simulation. This event is shown in Fig. 3 III.

#### 4. Association of BD particles

When the pair potential energy between two particles becomes smaller than a threshold energy, here taken to be $E_{\text{bind}} = -10k_{\text{B}}T$, the two particles are defined to be in the bound state. The two particles are replaced by a single BD particle at their center of mass. This event is shown in Fig. 3 IV. Note that at the next BD step, the algorithm will check whether the particle can be put into a GFRD domain, as described under Sec. II E 6.

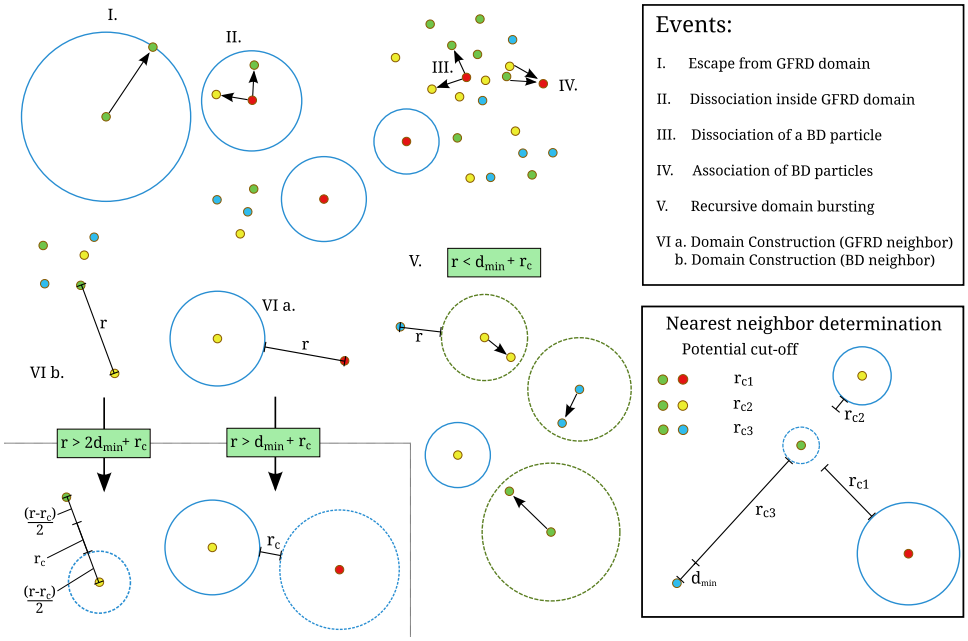


FIG. 3. At each BD time step, the algorithm checks whether BD should be interrupted. The BD propagation is halted when the time of the first next event occurs before the global simulation time at the end of the time step. These next events can be any of the following: I. A particle escapes from a GFRD domain; the position of the particle is updated to a randomly chosen point on the surface of the domain and the domain is removed. II. A particle dissociates inside a GFRD domain; the domain bursts and the particle is updated to a position and orientation sampled using Green's functions and is then replaced by its product particles. III. A BD particle dissociates; it is replaced by its product particles. IV. The binding energy of two BD particles is below the binding threshold; the particles enter the bound state and are replaced by a single product particle. V. The distance from a BD particle to a domain is smaller than $d_{\text{min}} + r_{\text{c}}$; the neighboring domain is burst and the position and orientation of the particle in this domain are updated. This particle may in turn burst another domain and this happens recursively until there is no BD particle within a distance $d_{\text{min}} + r_{\text{c}}$ from any other domain. VI. a. The distance between a BD particle and its nearest neighbor is larger than $d_{\text{min}} + r_{\text{c}}$ in case the nearest neighbor is a GFRD domain; a domain of radius $r - r_{\text{c}}$ is built on the BD particle. b. The distance between a BD particle and its nearest neighbor is larger than $2d_{\text{min}} + r_{\text{c}}$ in case the nearest neighbor is a BD particle; a domain of radius $0.5(r - r_{\text{c}})$ is built on the BD particle of interest. The inset shows the procedure for determining the nearest neighbor, which is the GFRD domain or the BD particle with the closest interaction horizon to the (central green) particle of interest: for BD particles, the relevant distance is the distance minus the sum of the minimum domain radius $d_{\text{min}}$ and the potential interaction range $r_c$, while for a GFRD domain, the relevant distance is the distance to the surface of that domain minus $r_c$. In the example configuration, the blue particle is the nearest neighbor.

#### 5. Recursive domain bursting

When a BD particle comes at time $t$ within a distance of $d_{\min} + r_c$ from the surface of a GFRD domain, the domain is burst and a radial position $r$ of the particle inside that domain is drawn from the normalized translational Green's function $p(r,t|r_0,t_0)/S(t - t_0|r_0)$, where $t_0$ is the time and $r_0$ the position of the center of the domain when it was constructed, and $S(t - t_0|r_0)$ the survival probability. A new orientation of the particles is sampled from Eq. (2). If this particle, after updating its position, comes within a distance of $d_{\min} + r_c$ from another domain that domain is also burst. This may lead to a cascade of domain bursting, which ceases when no BD particle is within a distance of $d_{\min} + r_c$ from any GFRD domain. This event is shown in Fig. 3 V. Note that domains are always at least $r_c$ apart from each other.

#### 6. Domain Construction

For each BD particle, the algorithm determines the nearest neighbor, which is either another BD particle or a GFRD domain. The procedure to determine the nearest-neighbor distance depends on whether the neighbor is a BD particle or a GFRD domain, as shown in the inset of Fig. 3. A BD particle is put into a GFRD domain when the distance $r$ between the particle and its nearest neighbor

(a) is larger than $d_{\min} + r_c$ in case the nearest neighbor is a GFRD domain. A domain of radius $(r - r_c)$ is built around the particle of interest. This event is shown in Fig. 3 VI a.

(b) is larger than $2d_{\min} + r_c$ in case the nearest neighbor is a BD particle. A domain of radius $0.5(r - r_c)$ is built around the particle of interest. This allows enough space to build a domain with a radius of at least $d_{\min}$ around the neighbor, thus preventing the neighbor from prematurely bursting the newly built domain. This event is shown in Fig. 3 VI b.

For the newly constructed domain, the tentative next-event times for the respective tentative event types (e.g., dissociation and escape) are determined, and the event type with the smallest tentative next-event time is added to the event list. To achieve maximum efficiency, the minimal domain size $d_{\min}$ should be as small as is practical.

### F. Computing the dissociation rate with forward flux sampling

The Forward Flux Sampling (FFS) algorithm enables efficient evaluation of rare event kinetics. FFS uses a series of interfaces between the reactant and the product states to construct the transition path ensemble and calculate the corresponding transition rate. Each interface is defined by an order parameter $\lambda$: the reactant state is defined by $\lambda < \lambda_{-1}$ and the product state by $\lambda > \lambda_n$. The remaining interfaces are defined by intermediate values of $\lambda$: $(\lambda_0 \ldots \lambda_{n-1})$. The FFS technique requires that $\lambda_{i+1} > \lambda_i$ for all $i$, and all the trajectories from reactant to product state pass through each interface in succession as shown in Fig. 4. Trajectories starting in the reactant state and reaching product state are rare,

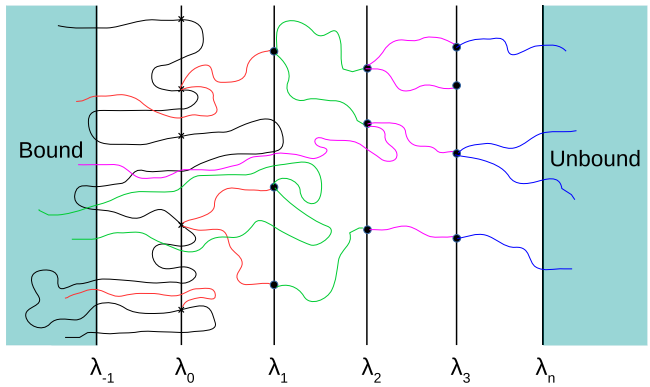


FIG. 4. An illustration of the FFS method. An ensemble of transition paths is generated by starting trial runs from randomly picked configurations on interfaces, which are the end points of previous successful trial runs.

but trajectories starting at an interface and crossing the next interface are more common. This is the central idea used in FFS.[13]

Here we use the "direct" FFS variant, DFFS, to compute the dissociation rate.[21] In this process, the reactant state is the bound A,B dimer, and the product state corresponds to the dissociated dimer. For the purpose of simulating dissociation, we take an order parameter to determine the interfaces based on a combination of the energy of interaction and the inter-particle distance. The reactant bound state interface $\lambda_{-1}$ is defined by a potential energy $E_{\text{bind}}$, while the product state is defined by zero potential energy in addition to an inter-particle distance larger than the cutoff $r_c$.

In the first step of FFS, a brute-force BD simulation is performed to compute the flux $\phi$ of crossing the interface $\lambda_0$ while coming from the bound state. This brute-force simulation generates an ensemble of configurations at $\lambda_0$. In the next step, a trajectory is fired from a randomly chosen configuration from this ensemble; this trajectory is then propagated until it either hits the next interface $\lambda_1$ or returns to the reactant state (i.e., recrosses $\lambda_{-1}$). This procedure is repeated until a sufficiently large number of configurations at the next interface $\lambda_1$ is generated. The fraction of trajectories that makes it from $\lambda_0$ to $\lambda_1$ yields the conditional probability $P(\lambda_1|\lambda_0)$ that a trajectory that comes from the bound state and crosses $\lambda_0$ for the first time will subsequently reach $\lambda_1$ instead of returning to the bound state. This whole procedure is then repeated for all subsequent interfaces until the final interface $\lambda_n$ is reached, signifying the fully dissociated pair. Under the assumption of rare event kinetics, the intrinsic dissociation rate $k_d$ is then given by[12,13]

$$k_d = \phi \prod_{i=0}^{n-1} P(\lambda_{i+1}|\lambda_i). \tag{7}$$

### G. Illustrative anisotropic inter-particle potential

In this section, we describe the interaction potential for the specific patchy-particle system. We reiterate that our multi-scale scheme is independent of the choice of

potential and can in principle be applied with arbitrarily complex potentials.

For convenience, we split our inter-particle potentials into three parts. Every pair of particles experiences a repulsive potential $U_{\text{rep}}(R)$ and an isotropic attractive potential $U_{\text{isoAtt}}(R)$ based on the distance $R$ between the centers of mass. Additionally, each pair of complementary patches interacts through an attractive potential $U_{\text{att}}(r)$ based on the distance $r$ between complementary patches (see Fig. 2). For a pair of particles with a single pair of complementary patches,

$$U(R, r) = U_{\text{rep}}(R) + U_{\text{isoAtt}}(R) + U_{\text{att}}(r). \tag{8}$$

Mediating the attractive interactions through surface-based patches naturally captures short-range contact interactions.

It is common to use 12-6 Lennard-Jones or related potentials in biomolecular modeling. Although the $r^{-6}$ dependence is required for van der Waals interactions between atoms and even between larger entities, in general there is no fundamental reason to choose this functional form in case of complex effective interactions between biomolecules, e.g., hydrophobic interactions. In preliminary simulations, we observed that using Lennard-Jones potentials leads to numerical difficulties, forcing the use of extremely small time steps. The underlying reason is that Lennard-Jones potentials have a large curvature close to the minimum of the bound state, a situation for which the Brownian integrator is poorly suited. This effect is exacerbated by the use of short-ranged anisotropic attractions between particles, which reduces the entropy of the bound state and must be compensated for by stronger attractive potentials, in order to model realistic equilibrium binding constants. Stronger attractive potentials lead to larger second derivatives of the potential. Moreover, requiring potentials to be short-ranged and orientation-specific implies variation over short length and angular scales, again increasing the second derivatives of the potential.

Instead of using a Lennard-Jones type potential, we therefore illustrate our method using piece-wise quadratic potentials similar to those employed elsewhere.[22] These potentials give us more control over the shape, and allow for easier integration with potentials that are short-ranged and highly orientation-specific. We stress that using an alternative potential that is more challenging for the integrator would not remove the advantages of the multi-scale scheme.

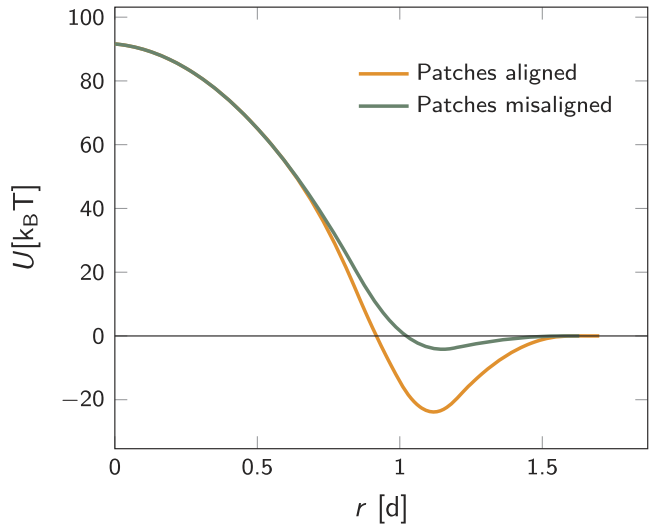


FIG. 5. Inter-particle interactions. Total interaction potential $U_{\text{rep}}(R) + U_{\text{isoAtt}}(R) + U_{\text{att}}(R - 2d_{\text{patch}})$ for two particles with perfectly aligned complementary patches, and the total interaction potential $U_{\text{rep}}(R) + U_{\text{isoAtt}}(R) + U_{\text{att}}(R + 2d_{\text{patch}})$ when the complementary patches are completely misaligned. The existence of patches introduces an attractive bound state with the particles in close contact.

$U_{\text{rep}}(R)$, $U_{\text{isoAtt}}(R)$, and $U_{\text{att}}(r)$ have the form

$$U_i(x) = \begin{cases} \epsilon_i\left(1 - a_i\left(\frac{x}{\sigma}\right)^2\right) & \text{if } x < x_i^\star, \\ \epsilon_i b_i\left(\frac{x_i^c}{\sigma} - \frac{x}{\sigma}\right)^2 & \text{if } x_i^\star < x < x_i^c, \\ 0 & \text{otherwise,} \end{cases} \tag{9}$$

with $i$ = rep, isoAtt, att, respectively. The overall strength $\epsilon_i$, the length scale $\sigma$ (i.e., the particle diameter), the stiffness $a_i$, and the parameter $x_i^\star$, which combined with $a_i$ determines the range of the potential, are free parameters. Cut-offs $x_i^c$ and smoothing parameters $b_i$ are fixed by requiring continuity and differentiability at $x_i^\star$. For our illustrative purposes, we take the following parameters: $\epsilon_{\text{rep}} = 100 k_{\text{BT}}$, $a_{\text{rep}} = 1$, and $R^*_{\text{rep}} = 0.85\sigma$, implying $b_{\text{rep}} = 2.6036$ and $R^c_{\text{rep}} = 1.1764\sigma$; $\epsilon_{\text{att}} = 20 k_{\text{B}}T$, $a_{\text{att}} = 20$, and $r^*_{\text{att}} = 0.1\sigma$, implying $b_{\text{att}} = 5$ and $r^c_{\text{att}} = 0.5\sigma$; and $\epsilon_{\text{isoAtt}} = 10 k_{\text{BT}}$, $a_{\text{isoAtt}} = 1$, and $R^*_{\text{isoAtt}} = 0.85\sigma$, implying $b_{\text{rep}} = 2.6036$ and $R^c_{\text{rep}} = 1.1764\sigma$.

In Fig. 5, we plot the resulting total inter-particle potential as a function of distance $R$ when the two complementary patches are aligned to face each other so that $r = R - \sigma$. A narrow attractive well corresponding to the two particles being in close contact is evident. For comparison, we also show

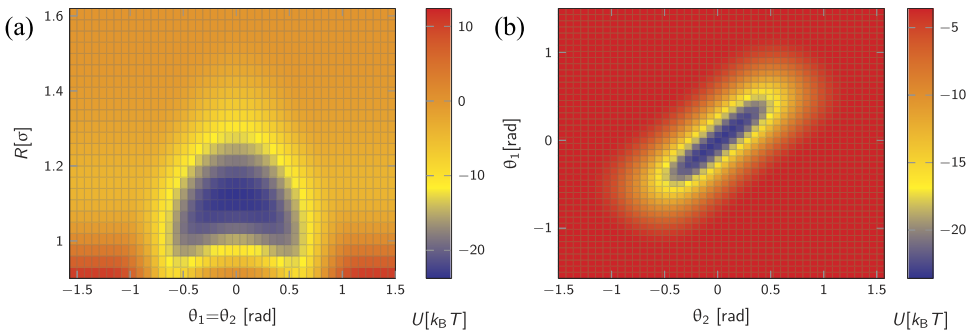


FIG. 6. Total potential $U_{\text{rep}}(R) + U_{\text{isoAtt}}(R) + U_{\text{att}}(r)$ for two particles with complementary patches. (a) Potential as a function of the distance between centres of mass $R$ and alignment of patches with inter-particle vector, $\theta_1$ and $\theta_2$ (see definition in Fig. 2), given $\theta_1 = \theta_2$. (b) Potential as a function of $\theta_1, \theta_2$ given $R = 1.1\sigma$. Note the relatively narrow range of orientations over which strong bonding occurs.

the total inter-particle potential as a function of $R$ when the two complementary patches are misaligned to face opposite each other so that $r = R + \sigma$. In this case, the patches do not contribute to the interaction; the non-specific, isotropic part of the potential, however, still gives rise to a weak attraction. In Fig. 6, we demonstrate the orientational dependence of the attractive potential, showing that the attractive interaction is highly sensitive to misalignment. We note that our choice of potential makes truncation at short distances relatively trivial. This is helpful in allowing rapid switching to GFRD domains once the particles are separated.

For our model potential, the interaction range is set $r_c = 1.6\sigma$, where the pair potential in Eq. (8) has vanished. Moreover, in the MD-GFRD simulations, we set the minimum domain size $d_{\min} = 0.5\sigma$ and the particle diameter to $\sigma = 5$nm.

## III. RESULTS AND DISCUSSION

We test the MD-GFRD simulation using the patchy-particle model described in Section II G. In the simulations, there are three species of particles, A, B, and C, which react according to

$$A + B \leftrightarrows C. \tag{10}$$

The system specific parameters of the simulation are as follows: The particle diameter is $\sigma = 5$ nm, the time step $\delta t = 0.1$ ns, the mass of the particle is $m = 50$ kDa, the mass moment of inertia $M = \frac{8}{15} m\sigma^2$, the translational and rotational diffusion constants are $D_t = 1\ \mu\text{m}^2/\text{s}$ and $D_r = 1.6 \times 10^7 \text{rad}^2/\text{s}$ for all particles, the translational and rotational friction coefficients are $\gamma = \frac{k_B T}{D_t m}$ and $\Gamma = \frac{k_B T}{D_r M}$, respectively, where $k_B = 1.38 \times 10^{-23}\ \text{JK}^{-1}$ is the Boltzmann constant and $T = 300$ K is the temperature of the system. To check if our Green's function and BD correctly describe the rotational motion of isolated particles, we have measured the correlation function of the dot product of the patch vector at a given time $t$ with the initial patch vector. This correlation function decays as $e^{-2D_r t}$, where $D_r = k_B T/\Gamma M$, as expected.

In Sec. III A we first present the results of the FFS-BD pre-simulation used to determine the value of the intrinsic dissociation rate $k_d$. Next, using the value of $k_d$, we perform MD-GFRD simulations in which we compute the probability that A and B are bound, as a function of system size. We compare the results against Monte Carlo simulations and analytical expressions. In Subsection III C, we compute the power spectra for the binding process. Finally, we discuss the performance of the algorithm.

### A. Rate constant determination using FFS-BD pre-simulation

As explained in Sec. II D, it is advantageous to treat the dimer A-B as a single particle C, which then can dissociate again into A and B with an intrinsic rate $k_d$. We used direct FFS to precompute the intrinsic rate constant $k_d$. The interfaces $\lambda_i$ are defined in terms of the interaction energy, as shown in Fig. 7. The bound state interface $\lambda_{-1}$ was defined by $U(R, r) < -10 k_B T$; the dissociated state final interface $\lambda_5$ was set at a distance $R = 1.6\sigma$. Five intermediate interfaces were set

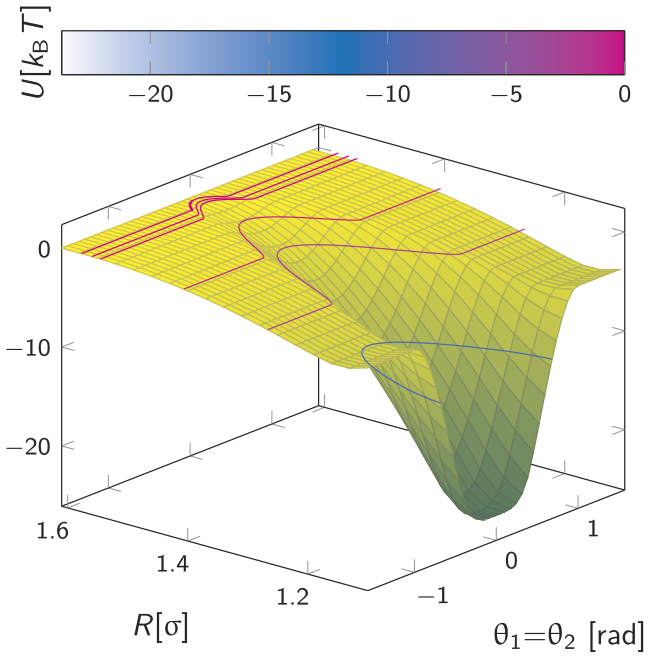


FIG. 7. FFS interfaces were defined by the potential energy: $\lambda_0 = -10 k_B T$, $\lambda_1 = -2.5 k_B T$, $\lambda_2 = -0.75 k_B T$, $\lambda_3 = -0.025 k_B T$, and $\lambda_4 = -0.0075 k_B T$. The final interface $\lambda_5$ was defined by zero energy and a distance $R > 1.6\sigma$. Using these interfaces as starting points for successive trial runs, the particles are driven from the bound to the unbound state.

at, respectively, $\lambda_0 = -10 k_B T$, $\lambda_1 = -2.5 k_B T$, $\lambda_2 = -0.75 k_B T$, $\lambda_3 = -0.025 k_B T$, and $\lambda_4 = -0.0075 k_B T$. A straightforward BD trajectory created 100 000 configurations at the first interface. Subsequently, performing direct FFS yielded 20 000 configurations for each successive interface. Using Eq. (7), we find that for the intrinsic dissociation rate constant $k_d = 4.66\ \text{s}^{-1}$. The configurations at the final interface can be used to draw from when performing the dissociation step in the MD-GFRD, see Sec. II E 2.

### B. Bimolecular reactions

To test the multi-scale scheme, we simulate the bimolecular reaction shown in Eq. (10). In these simulations, we start off with two species of particles A and B, each having one patch on its surface. An A particle can react with a B particle to form a dimer. Also, a C particle can dissociate to form one A and one B, with an intrinsic rate $k_d$ that has been pre-computed using FFS (see Sec. III A). We assume that the mixture is ideal: only species A and B have an attractive interaction $U(R,r)$. All other interaction potentials between pairs A-A, B-B, C-C, C-A, and C-B are repulsive only. We test the scheme for two different scenarios, one starting with a single A and a single B particle and the second starting with two A and two B particles. The simulation results are compared with Monte Carlo simulations of the same model and with analytical expressions.

In the first case, one particle of species A and one particle of species B, each having one patch, are put in a cubic box of volume $V$, with periodic boundary conditions. This means that the number of C particles, $N_C$, is either zero or one. From the computed time average of $N_C$, we calculate the probability $\mathcal{P}_b$ that the A particle is bound to B. We repeat this procedure for different box sizes. In Fig. 8 we compare the value of $\mathcal{P}_b$

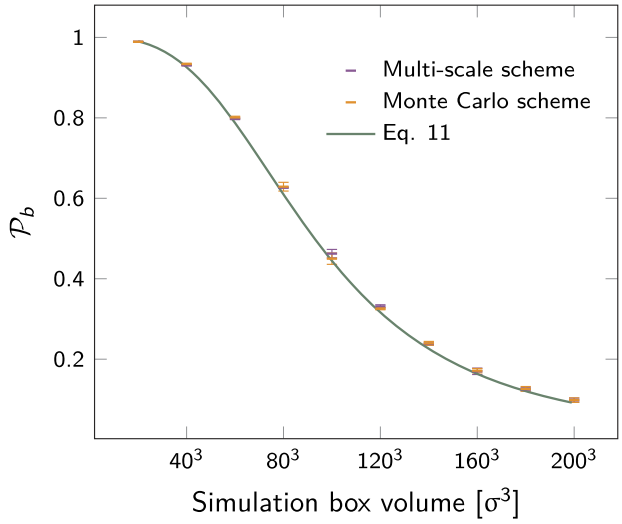


FIG. 8. The probability $\mathcal{P}_b$ that a particle A is bound to a particle B, as a function of the volume of the box. Simulations are performed with one A particle and one B particle in the box. The points with the error bars are the results of the MD-GFRD simulations and the Monte Carlo simulation. These results are validated with the analytical prediction of Eq. (11). It is seen that the agreement is very good. The translational and rotational diffusion constants, which are not important for the value of $\mathcal{P}_b$, are 1 $\mu$m$^2$/s and $1.6\times10^7$ rad$^2$/s.

obtained using the new MD-GFRD algorithm to the results obtained from Monte Carlo simulations of the same system. The figure also shows the analytical result

$$\mathcal{P}_b = \frac{\langle N_C \rangle}{N_A} = \frac{k_{on}}{k_{on} + Vk_{off}} = \frac{\phi(V)}{\phi(V)+1}, \tag{11}$$

where $\langle N_C \rangle$ is the average of $N_C$ and $\phi(V)$ is the ratio of the probability that an A particle is bound versus unbound

$$\phi(V) = \frac{k_{on}}{Vk_{off}} = \frac{K_{eq}}{V}, \tag{12}$$

where $k_{on}$ and $k_{off}$ are the effective association and dissociation rates, respectively, and $K_{eq}$ is the equilibrium constant

$$K_{eq} = \int d\mathbf{R} \int d\hat{\mathbf{u}}_1 \int d\hat{\mathbf{u}}_2 e^{-\beta V(R, r(\mathbf{R}, \hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2))}, \tag{13}$$

where $U(R, r(\mathbf{R}, \hat{\mathbf{u}}_1\hat{\mathbf{u}}_2))$ is the interaction potential given by Eq. (8), with $\mathbf{R}$ the inter-particle vector, $R$ the magnitude of $\mathbf{R}$, and $r$ the distance between the patches of the particles, which depends on $\mathbf{R}$ and the orientation of the two particles denoted by the patch vectors in the stationary lab frame, $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2$, respectively, given by Eq. (1). Solving Eq. (13) analytically is not possible for the complex anisotropic potential used here. However, recently we have shown how in one TIS/FFS simulation both the association rate $k_{on}$ and the dissociation rate $k_{off}$ can be computed,[23] which then allows us to obtain $K_{eq}$ from Eq. (12). Applying this technique to this potential revealed that $k_{on} = 0.135$ $\mu$m$^3$ s$^{-1}$ and $k_{off} = 1.384$ s$^{-1}$. Fig. 8 shows that the results of the MD-GFRD simulations agree very well with both the results of the Monte Carlo simulations and with the analytical predictions.

In the second test, we start with 2 A particles and 2 B particles, which can again interact via the same interaction potential to form species C. We can analytically compute the probability that an A particle is bound to a B particle by carefully summing over all possible configurations[24]

$$\mathcal{P}_b = \frac{\phi(V) + \phi(V)^2}{2(0.25 + \phi(V) + \frac{\phi(V)^2}{2})}, \tag{14}$$

where $\phi(V)$ is given by Eq. (12). The results of the MD-GFRD simulations, the Monte Carlo simulations, and the analytical prediction are shown in Fig. 9. It is seen that the agreement is very good.

### C. Power spectrum

We can use MD-GFRD to compute the power spectrum $P_n(\omega)$ of the time trace of the binding state $n(t)$ of two particles, switching between the bound state with $n(t) = 1$ and the unbound state with $n(t) = 0$. The dotted line in Fig. 10 shows the result. We expect that this power spectrum is given by that of a random telegraph process[25]

$$P(\omega) = \frac{2\mu\mathcal{P}_b(1-\mathcal{P}_b)}{\mu^2 + \omega^2}, \tag{15}$$

where $\omega$ is the frequency, $\mu = k_{on}/V + k_{off}$ is the renormalized/effective decay rate, and $\mathcal{P}_b = k_{on}/(k_{on} + Vk_{off})$ is the binding probability. To predict the power spectrum, we thus need the effective association rate $k_{on}$ and the effective dissociation rate $k_{off}$. As described in Sec. III B, these rates can be computed in a single TIS/FFS simulation.[23] Using the computed values of the rate constants in combination with Eq. (15), we arrive at the analytical prediction of the solid line in Fig. 10. It is seen that the agreement with the MD-GFRD simulation results is excellent. MD-GFRD thus not only reproduces mean quantities but also successfully predicts dynamic quantities.

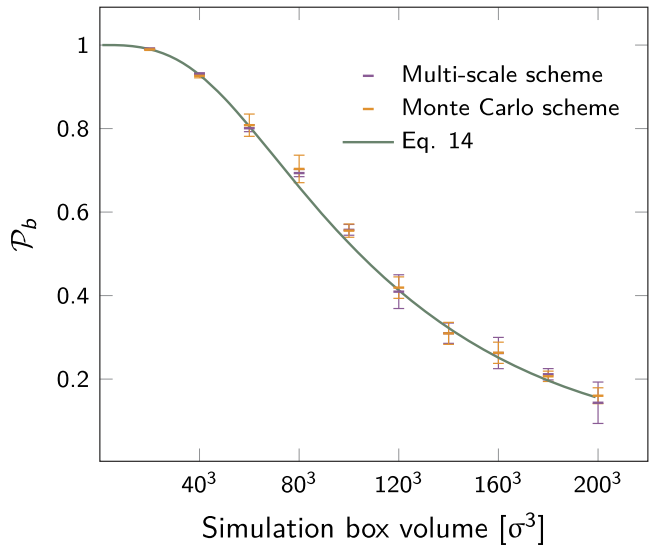


FIG. 9. The probability $\mathcal{P}_b$ that a particle A is bound to a particle B, as a function of the volume of the box. Simulations are performed starting with two A particles and two B particles in the box. The points with the error bars show the results of the new MD-GFRD scheme and the Monte Carlo simulations. These results are validated with the analytical prediction of Eq. (14). It is seen that the agreement is very good. The translational and rotational diffusion constants, which are not important for the value of $\mathcal{P}_b$, are 1 $\mu$m$^2$/s and $1.6\times10^7$ rad$^2$/s.

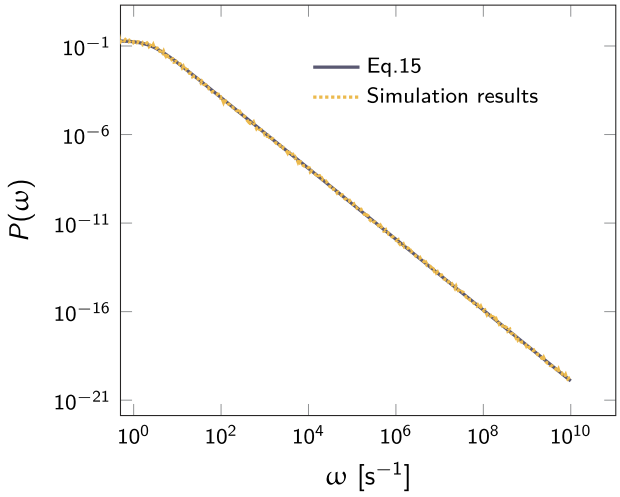


FIG. 10. MD-GFRD successfully predicts the power spectrum $P_n(\omega)$ of the binding state $n(t)$ of two particles switching between the bound state $n(t)$= 1 and the unbound state $n(t)$ = 0. The dotted line shows the results of the MD-GFRD simulations, while the solid line shows the analytical prediction of Eq. 15, where the association rate $k_{on}$ and dissociation rate $k_{off}$ have been computed from a single FFS simulation as described in Ref. 23. Two particles, one of each species A and B, were simulated in a box of side length $100\sigma$.

### D. Performance

The motivation to combine GFRD and MD into a multi-scale scheme is the computational speed up it can provide. Unlike brute force Brownian dynamics which spends a lot of CPU time in propagating the particles toward each other, GFRD makes large jumps in space and time when the particles are far apart from each other and the GFRD domains are large. The computational power of GFRD can thus especially be reaped when the particles are often far apart, which is the case when the concentrations are low. This can be seen in Fig. 11, which shows a comparison of MD-GFRD against brute force BD as a function of concentration. It is seen that MD-GFRD is much more efficient than brute force BD, especially when the concentrations are below a $\mu$M. However, for high concentrations, the performance of MD-GFRD becomes comparable to that of BD. In this regime, the particles are often so close together that no big jumps in time and space can be made. Interestingly, however, the crossover happens only at a mM concentration, which means that for most biologically relevant concentrations, MD-GFRD is much faster than brute-force BD.

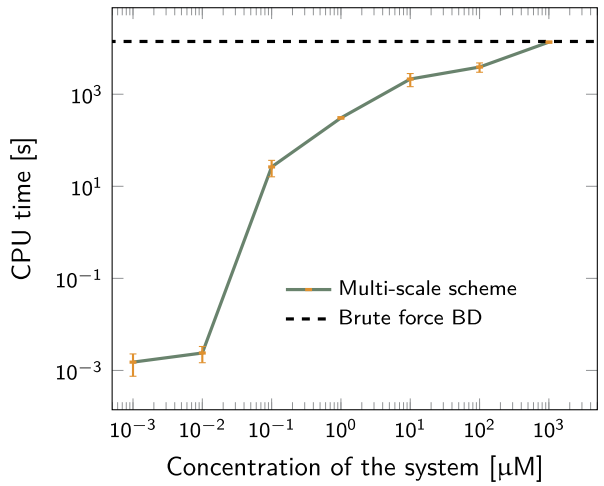


FIG. 11. The CPU time to simulate 1 ms real time as a function of the concentration of A and B, for MD-GFRD (solid line) and BD (dashed line). The concentration is varied by changing the volume of the simulation box, while the number of particles is kept constant at $N_A = N_B = 5$. It is seen that in the biologically relevant concentration range of nanomolar to micromolar, the performance of MD-GFRD is much better than that of brute-force BD, but at higher concentrations, the relative performance of MD-GFRD goes down. This is because at higher concentrations, the particles will be close to each other, and the system cannot capitalize on the potential of MD-GFRD to make large jumps in time and space.

We performed a profiling of the code to establish the overhead associated with the GFRD checks. To determine the percentage of the total time the code spends in building domains and updating positions, the code was first profiled for a low concentration where all the particles are in GFRD. In that case, 50% of the time is spent on nearest neighbour searches and the other 50% is spent on constructing and bursting the domains and updating particles. To determine the overhead associated with GFRD compared to a brute force BD simulation, we profiled the code for a high concentration such that all the particles are in the BD regime. The code spends 60% of the total time integrating the BD particles (neighbour searches, force calculations, torque calculations, and updating position and orientation) and 40% of the time attempting to build domains on these BD particles. The overhead associated with the nearest neighbour searches can be decreased, if the code is optimised such that the nearest neighbour search is performed just once for both the BD and GFRD.

## IV. CONCLUSION

In this work, we extended the MD-GFRD scheme[11] to include the orientational dynamics of the particles, enabling the simulation of reaction and diffusion of particles that interact via anisotropic interaction potentials. This opens up the possibility to treat a whole class of interesting problems. Biomolecules such as proteins and DNA typically interact with each other via anisotropic potentials. In some cases of biological interest, the dynamics at short length and time scales can be integrated out.[3,26–28] For example, a gene regulatory protein that has just dissociated from its promoter on the DNA either rapidly rebinds the DNA or rapidly escapes into the bulk, where it will loose its orientation; conversely, a new protein tends to arrive at the promoter from the bulk in a random orientation. In these cases, we expect that the regulatory proteins can be modeled as isotropic particles that interact with the DNA via effective rate constants, which take into account the anisotropy of the interaction. However, it is now well established that in many systems the dynamics at molecular length and time scales, arising from, e.g., enzyme-substrate rebindings, can qualitatively change the macroscopic behavior of the system at cellular length scales.[10,29] This phenomenon can occur in biochemical networks with multi-site protein modification, which are omnipresent in cellular biology.[10] In such systems, the orientational dynamics cannot be integrated out: the probability that an enzyme which has dissociated from its substrate molecule rebinds to another site on the same substrate molecule to chemically modify it will depend in a non-trivial

manner on the translational and orientational diffusion constants of the particles, their size, and the distance between the patches on the substrate. The MD-GFRD scheme presented here now makes it possible to study the interplay between the microscopic dynamics at the molecular scale and the network dynamics at the cellular scale in this large class of systems.

In addition, the MD-GFRD scheme could more generally be used for soft matter self-assembly where building blocks that are diffusing in the dilute solution come together and bind occasionally to form large complexes and structures.[14]

While the MD-GFRD scheme has been set up for simulating 3D bulk solutions, in principle the method can be extended to include other geometries, where GFRD includes reaction and diffusion in 1D and 2D.[30–32]

Finally, we note that in our multiscale MD-GFRD algorithm, we assume that the interaction potentials are short ranged, which is usually the case in a highly screened environment such as the living cell. Moreover, the algorithm, similar to most algorithms to simulate biochemical networks in time and space, assumes that the particles move by normal diffusion,[4,5,7,33,34] as indeed experiments suggest proteins do inside the living cell.[1] The algorithm cannot straightforwardly treat long-range hydrodynamics interactions, which can be of importance under certain (non-equilibrium) conditions. For future work, it would be of interest to study whether hydrodynamic interactions could be included in MD-GFRD, for instance, via the Oseen tensor approximation.

## ACKNOWLEDGMENTS

This work is part of the Industrial Partnership Programme (IPP) "Computational sciences for energy research" of the Foundation for Fundamental Research on Matter (FOM), which is financially supported by the Netherlands Organization for Scientific Research (NWO). This research programme is co-financed by Shell Global Solutions International B.V.

[1] J. Elf, G.-W. Li, and X. S. Xie, Science **316**, 1191 (2007).
[2] G. Schreiber, G. Haran, and H. X. Zhou, Chem. Rev. **109**, 839 (2009).
[3] P. R. ten Wolde, N. B. Becker, T. E. Ouldridge, and A. Mugler, J. Stat. Phys. **162**, 1395 (2016).
[4] S. S. Andrews and D. Bray, Phys. Biol. **1**, 137 (2004).
[5] J. Lipková, K. C. Zygalakis, S. J. Chapman, and R. Erban, J. Appl. Math. **71**, 714 (2011).
[6] M. B. Flegg, S. J. Chapman, and R. Erban, J. R. Soc., Interface **9**, 859 (2012).
[7] J. Schöneberg and F. Noé, PLoS ONE **8**, e74261 (2013).
[8] J. S. van Zon and P. R. ten Wolde, Phys. Rev. Lett. **94**, 128103 (2005).
[9] J. S. van Zon and P. R. ten Wolde, J. Chem. Phys. **123**, 234910 (2005).
[10] K. Takahashi, S. Tanase-Nicola, and P. R. ten Wolde, Proc. Natl. Acad. Sci. U. S. A. **107**, 2473 (2010).
[11] A. Vijaykumar, P. G. Bolhuis, and P. R. ten Wolde, J. Chem. Phys. **143**, 214102 (2015).
[12] T. S. van Erp, D. Moroni, and P. G. Bolhuis, J. Chem. Phys. **118**, 7762 (2003).
[13] R. J. Allen, P. B. Warren, and P. R. ten Wolde, Phys. Rev. Lett. **94**, 018104 (2005).
[14] S. C. Glotzer and M. J. Solomon, Nat. Mater. **6**, 557 (2007).
[15] A. C. Newton, J. Groenewold, W. K. Kegel, and P. G. Bolhuis, Proc. Natl. Acad. Sci. U. S. A. **112**, 15308 (2015).
[16] R. Davidchack, T. E. Ouldridge, and M. Tretyakov, J. Chem. Phys. **142**, 144114 (2015).
[17] A. Loman, I. Gregor, C. Stutz, M. Mund, and J. Enderlein, Photochem. Photobiol. Sci. **9**, 627 (2010).
[18] H. Versmold, Z. Naturforsch., A **25**, 367 (1970).
[19] L. D. Favro, Phys. Rev. **119**, 53 (1960).
[20] W. H. Furry, Phys. Rev. **107**, 7 (1957).
[21] R. J. Allen, C. Valeriani, and P. Rein ten Wolde, J. Phys.: Condens. Matter **21**, 463102 (2009).
[22] T. E. Ouldridge, A. A. Louis, and J. P. K. Doye, J. Chem. Phys. **134**, 085101 (2011).
[23] A. Vijaykumar, P. G. Bolhuis, and P. R. ten Wolde, Faraday Discuss. **195**, 421 (2016).
[24] T. E. Ouldridge, A. A. Louis, and J. P. K. Doye, J. Phys.: Condens. Matter **22**, 104102 (2010).
[25] K. Kaizu, W. de Ronde, J. Paijmans, K. Takahashi, F. Tostevin, and P. R. ten Wolde, Biophys. J. **106**, 976 (2014).
[26] J. S. van Zon, M. J. Morelli, and P. R. ten Wolde, Biophys. J. **91**, 4350 (2006).
[27] M. J. Morelli, R. J. Allen, and P. R. ten Wolde, Biophys. J. **101**, 2882 (2011).
[28] A. Mugler and P. R. ten Wolde, Adv. Chem. Phys. **153**, 373 (2013).
[29] K. Aokia, M. Yamada, K. Kunida, S. Yasuda, and M. Matsuda, Proc. Natl. Acad. Sci. U. S. A. **108**, 12675 (2011).
[30] J. Paijmans and P. R. ten Wolde, Phys. Rev. E **90**, 032708 (2014).
[31] M. Wehrens, P. R. ten Wolde, and A. Mugler, J. Chem. Phys. **141**, 205102 (2014).
[32] T. R. Sokolowski, *A Computational Study of Robust Formation of Spatial Protein Patterns*, Ph.D. thesis, VU University Amsterdam, 2014.
[33] J. Elf and M. Ehrenberg, Syst. Biol. **1**, 230 (2004).
[34] M. E. Johnson and G. Hummer, Phys. Rev. X **4**, 031037 (2014).