Reversible integrators for particle based reaction-diffusion simulations

MSc thesis

Christoph Fröhner

November 6, 2015

Contents

1 Introduction

Biological life relies on the interplay of proteins, membranes and other molecules, of which thousands of copies exist in a typical cell. To understand the function of every molecule in a given process, one approach is to model the system in a computational environment. With the help of such simulations one has the opportunity to resolve e.g. conformational changes of a protein or the transport of an ion through a membrane. In classical Molecular Dynamics (MD) simulations this is achieved by an all-atom model where time is integrated in discrete timesteps which are on the order of femtoseconds to resolve the fast dynamics of water molecules. This is computationally expensive to do, since often the rare events happen on large timescales that are beyond the scope of MD simulations. To be able to resolve e.g. whole reaction chains or the transport of many molecules through a cell in a spatio-temporal manner one can use reaction-diffusion methods. One of these are deterministic reaction-diffusion equations. These govern the time evolution of a set of concentrations of species $\mathbf{u}(\mathbf{x}, t)$ that generally depend on time and space. The time evolution can then be found by solving the following set of equations

$$\frac{\partial \mathbf{u}(\mathbf{x}, t)}{\partial t} = \mathbf{D}\Delta\mathbf{u}(\mathbf{x}, t) + f(\mathbf{u}), \tag{1.1}$$

where $\partial/\partial t$ is the time derivative, $\mathbf{D}$ is the diffusion tensor, Δ is the Laplacian and $f(\mathbf{u})$ is an arbitrarily complicated function that describes the behavior of the system. The problem with such a description is that it relies on concentrations of particles. In systems where the difference of having one particle or zero particles of a certain species is crucial, this description cannot be used to study such a system.[1]

Another method to study cellular systems are stochastic *particle-based reaction-diffusion simulations* (which are the focus of this thesis). In such simulations objects of interest such as proteins are considered as particles. Dynamics are realized by random motion of particles in discrete timesteps. The time evolution of such a method could be written as

$$\boldsymbol{\psi}(t + 1) = \mathcal{R}\,\mathcal{D}\,\boldsymbol{\psi}(t), \tag{1.2}$$

where $\boldsymbol{\psi}(t)$ is now a vector containing all particles, their positions and properties at timestep t, $\mathcal{D}$ propagates the particles in space due to diffusion, $\mathcal{R}$ performs reactions, i.e. particles change their species or form other species by means of bimolecular reactions.[2]

[1] Another method which is able to resolve single particles combined with reactions is the chemical master equation, but it is not further mentioned here because we focus on methods that treat space continuously.

[2] We will call $\mathcal{D}$ and $\mathcal{R}$ *propagators*, not to be confused with the propagators of Quantum Mechanics which reprensent probabilities. For us they will represent operators acting on the state of a system.

Often water molecules are not modeled explicitly in a reaction-diffusion simulation to reduce computational cost. Using such a method enables to explore longer times as compared to Molecular Dynamics simulations and at the same time being able to consider systems with only a few particles of a species as opposed to deterministic reaction-diffusion equations. There are already software packages using different algorithms and levels of detail to adress this topic [1].

When modeling real world systems, another important concept is the *detailed balance* condition. It connects the probability of any elementary process with its reverse process. Consider any system undergoing a transition from state a to state b, the detailed balance condition reads

$$p(a \to b)\pi(a) = p(b \to a)\pi(b). \tag{1.3}$$

Here $p(a \to b)$ denotes the transition probability from state a to b and $\pi(a)$ is the equilibrium probability of being in state a which is often related to the energy of state a. This relation is a sufficient condition for the system to be in equilibrium. For systems in equilibrium the entropy is maximal and approaching this maximum is a behavior which is expected from real world systems. This is why *detailed balance* is often demanded of computational models that represent real world applications. One can immediately see that the relation is invariant upon exchanging states a and b, which clarifies why the term *microscopic reversibility* is often used as a synonym, if the transition $a \to b$ occurs during a time interval. Having introduced the most important concepts mentioned in this thesis, we can pose the actual question.

The question for this particular project is: How does an integrator, that enforces (local) detailed balance on dynamics $\mathcal{D}$ and/or reactions $\mathcal{R}$, influence the results of particle based reaction-diffusion simulations?

Answering these questions will improve the method of reaction-diffusion itself. A method which can be used to understand processes such as the synaptic vesicle exocytosis, where clustering of Syntaxin appears to be a key factor [1, 2]. Also it will help to understand the effects of crowding on reactions in living cells [3], which is experimentally hard to observe. Schöneberg et al. [1] provide the following fields of application for highly detailed reaction-diffusion simulations:

- Compartmentalization and confinement
- Macromolecular crowding
- Particle-particle potentials
- Oligomerization and self assembly
- Molecular shape

Most cellular environments feature these aspects, making reaction-diffusion simulations an appropriate tool to investigate these systems. An application example is depicted in figure 1.1.

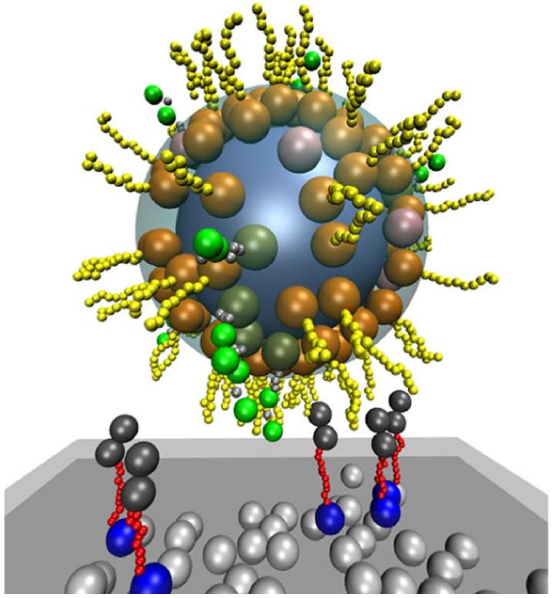


Figure 1.1: A high level of detail in reaction-diffusion simulations enables shaping of elongated proteins such as syntaxin shown as red particles with Habc domains in dark grey or Synaptobrevin in orange/yellow. Vesicle exocytosis can be modeled with the help of these. Figure by Schöneberg et al. [4].

2 Literature review

To this day there are several software packages implementing reaction- diffusion using different methods. Some examples are Cell++ [5], Smoldyn [6], eGFRD [7] and ReaDDy [4]. Cell++ treats particles as points and continuous space is divided into cubic subvolumes of different types, which defines diffusion properties of molecules in that region. It also features molecule densities (called *small molecules*) in addition to discrete particles (called *large molecules*). The large molecules are propagated using Brownian Dynamics in continuous space, while the small molecules' densities are constant within one cubic subvolume and change depending on their neighbouring cube's density according to the diffusion equation. Smoldyn similarly propagates particles in space and time with a fixed timestep and lets second order reactions happen when two particles are within a certain range of each other.

eGFRD [7] uses a different approach compared to the first two. Instead of incrementing the time in equally spaced steps van Zon et al. developed an event based method called Green's Functions Reaction Dynamics. This method uses analytical solutions of the Smoluchowski equation to draw new positions for isolated particles instead of propagating them step by step. Also if two particles are within a certain range a solution of the Smoluchowski equation for a two particle system is used to determine the probability of a second order reaction. This leads to a very efficient algorithm since there is little computation needed to propagate isolated particles for large timesteps. In these approaches particle repulsion is governed by an overlap rejection.

ReaDDy [4] utilizes potentials to deal with particle-particle-interaction or particle-geometry-interaction and propagates spherical particles with a fixed timestep using Brownian Dynamics. Reactions occur with a given microscopical reaction rate when particles come closer than a certain reactive distance depending on the particle species. The highly detailed computational models of ReaDDy have proved useful in investigating mesoscopic biological systems thereby supporting experimental results and their interpretation. In the work of Gunkel et al. [8] different reaction-diffusion models of the arrangement of rhodopsin in retinal rod-cells were compared. One model supports the idea of single photon responses by spatially confining transducin molecules. The feature of interaction potentials in ReaDDy has helped Ullrich et al. [9] to understand the clustering of Synatxin proteins on the neuronal membrane, which is involved in the fusion of synaptical vesicles with the neuronal membrane. The cluster size and their distribution are in agreement with experimental observations from super-resolution microscopy. From the computational model insight on the interaction between Syntaxin proteins is gained. The model and methods used throughout the present work are based on the approach of ReaDDy as will be further explained in section 4.

To combine the idea of Monte Carlo methods with Brownian Dynamics is long known

and turns out to be faster converging than standard Metropolis Monte Carlo [10, 11]. In such a *Smart* Monte Carlo method the moves are not proposed by uniform displacement of particles but rather a Brownian Dynamics step (also considering forces). Also to combine Grand Canonical Monte Carlo (GCMC) [12] with Brownian Dynamics has been applied. Im et al. [13] used this method to simulate an ion channel, where the Grand Canonical character (meaning the fluctuating particle numbers) was deployed in the surrounding region of the ion channel but not within the ion channel itself.

Morelli and Ten Wolde [14] proposed a method to implement a reaction-diffusion algorithm that obeys detailed balance with respect to a bimolecular reaction. In their approach particles are displaced subsequently according to Brownian Dynamics and volume exclusion is accomplished by rejection of such a step. If a step would lead to an overlap of possible reaction partners, e.g. A and B that can react to C, this reaction step is accepted with the correct probability that fulfills the detailed balance condition. The backward reaction / dissociation event is performed vice versa, i.e. it occurs with the correct probability and places the educts such that reversibility is guaranteed. This method reproduces the particle concentrations of an equivalent system of ordinary differential equations or chemical master equations.

Using a Metropolis-Hastings [15] method to enforce detailed balance together with a highly detailed particle based reaction diffusion simulation as it is proposed here has not been realized. The motivation for that approach shall be given in the following.

3 Motivation

It was stated before that the approach used here is based on Schöneberg et al.'s ReaDDy software [4], where spherical particles are propagated in time using Brownian Dynamics and reactions occur with pre-specified microscopic rates. In their paper is an example system which they use to test binary reactions in a crowded environment. The motivation for improving their method is best explained by means of this example system.

The given system consists of spherical particles confined to a cubic volume. There are three particle species in this system called A, B and C. All species have a specific radius and diffusion constant. The particle radii are chosen such that the sum of volume of particles A and B is equal to the volume of particle C. A snapshot of this system is given in figure 3.1.

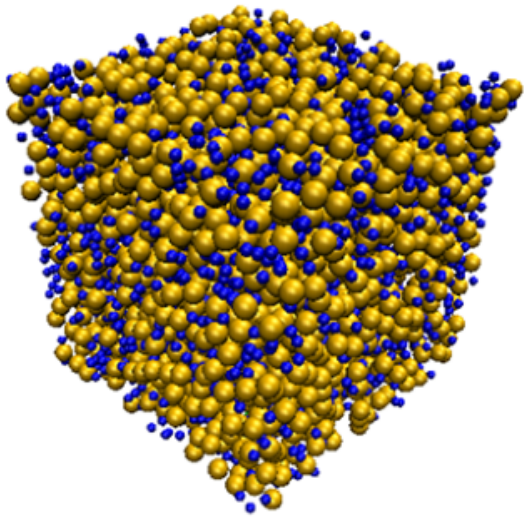


Figure 3.1: The system shown here is used to test bimolecular reactions in which two blue (small) particles react to one yellow (large) particle and vice versa. The volume occupation by particles is 30%. Figure taken from Schöneberg et al. [4].

There is one reaction that can occur forward and backward, namely

$$A + B \underset{k_{\text{off}}}{\overset{k_{\text{on,micro}}}{\rightleftharpoons}} C \tag{3.1}$$

where k_{off} is the reaction rate that describes how frequent a particle of type C dissociates into two particles A and B in units of s^{-1}. The *on* reaction is characterized by $k_{\text{on,micro}}$,

which is the microscopic rate of two particles A and B reacting if they have formed an encounter complex before by diffusing close to each other up to a certain distance. So the units of this microscopic rate $k_{\text{on,micro}}$ is also s^{-1}. If one is to observe the overall rate called $k_{\text{on,macro}}$ with which the *on*-reaction takes place, one would expect that it is lower than the microscopic rate since particles first have to encounter. For the approximation that the system is well stirred or in other words, that particles can move freely without volume exclusion effects, an analytic solution is at hand. This rate provided by Erban et al. [16] reads

$$\begin{aligned} k_{\text{on,macro}} &= 4\pi D_{AB}\left[R_{AB} - \sqrt{\frac{D_{AB}}{k_{\text{on,micro}}}}\tanh\left(R_{AB}\sqrt{\frac{k_{\text{on,micro}}}{D_{AB}}}\right)\right] \\ D_{AB} &= D_A + D_B \\ R_{AB} &= R_A + R_B, \end{aligned} \tag{3.2}$$

which relates the microscopic $k_{\text{on,micro}}$ with the macroscopic reaction rate $k_{\text{on,macro}}$. Here D_A and D_B are the diffusion constants, R_A and R_B are the radii of particles A and B respectively. Note that the units of the macroscopic rate are now $\text{m}^3\,\text{s}^{-1}$. With the help this one can obtain a solution of the concentrations of particle species by means of ordinary differential equations (ODE), which represents the well stirred approximation. This ODE solution can be compared to the results of the ReaDDy software. Exactly this is depicted in figure 3.2. One can clearly see that in the case of no particle repulsion the well stirred case is exactly represented, due to the fact that the dynamic timestep is four magnitudes faster than the reaction rates. The more interesting part here is the case when particle-particle potentials are switched on (solid lines in the plot). In ReaDDy harmonic repulsion potentials are used, which in effect reduces the reaction volume available to particles A and B in the *on*-reaction. Hence the effective macroscopic *on*-rate observed is lower than before. This results in slower reaction kinetics, meaning that an equilibrium state is reached much slower. Furthermore with particle repulsion the equilibrium ratios of concentrations differ from the well stirred case.

The actual problem about these observations is that we cannot assure that reactions are done carefully enough, because in ReaDDy an *on*-reaction takes place only when particles A and B meet. That means that the effective *on*-rate depends strongly on the particle densities. Whereas the *off*-reaction always takes place with the same effective rate, which is purely described by the microscopic rate k_{off}. So the *off*-rate is completely independent of the environment. This fact hints towards an imbalance of the bimolecular reaction being present.

To illustrate the problem further assume an *off*-reaction taking place in a very crowded system. For example one particle C dissociates into particles A and B and new positions for those particles are drawn randomly in the vicinity of the reaction taking place. This case is shown in figure 3.3. It could happen that particles overlap strongly after the reaction because one of the newly created particles was placed inside another particle. Such a state would mean a drastic increase in potential energy and be physically unreasonable. To avoid such states one must carefully place new particles and/or be able to reject such

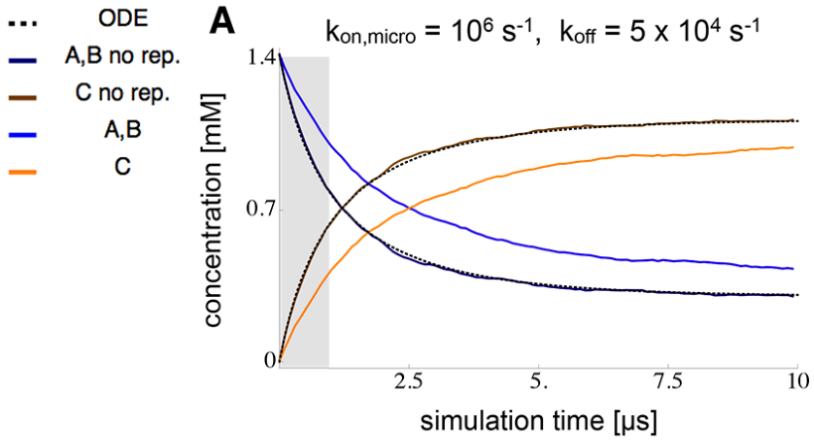


Figure 3.2: The concentrations of particle species A, B and C as a function of time starting with only A and B particles in the example system of ReaDDy. The given rates correspond to the bimolecular reaction given in equation 3.1. The dotted line gives the solution from ordinary differential equations, described by the macroscopic rates k_{off} and $k_{\text{on,macro}}$ (see equation 3.2). The solid lines give the concentrations computed from ReaDDy simulation, once with and without volume exclusion / particle repulsion. The timestep used for Brownian Dynamics was 10^{-10} s. Figure taken from Schöneberg et al. [4].

steps. The overall goal is to impose detailed balance on the simulation, meaning that every timestep forward will be executed considering the step backward. The exact scheme will be given in section 5.

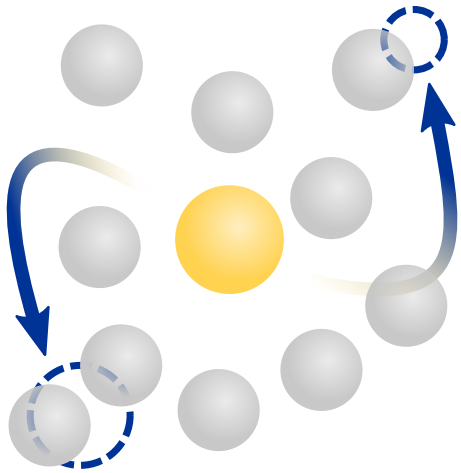


Figure 3.3: A particle C in the middle dissociates into two particles A and B (dashed circles). The system is so crowded that at the new position of one particle there was already a particle, leading to a large overlap of particles after the reaction and an increase in potential energy.

4 Model

In this chapter we will explain the model used to realize the reaction- diffusion simulations. Afterwards the new features of the integration scheme that implement detailed balance are stated in section 5.

To achieve simulations of biological system with the help of reaction- diffusion, we will use a similar model as ReaDDy [4]. This means that every object of interest is modeled by spherical particles that have a certain radius. All of these particles live in a continuous three- dimensional space at a certain fixed time. In contrast to space, the time is discretized in our model and the timespan between two points in time is usually referred to as *timestep*, often denoted as τ. Given a system of particles at a certain time t, the next state of the system at time $t + \tau$ is calculated using Brownian Dynamics. Also there will be interaction potentials involved, which are used to realize particle repulsion or cellular geometry. These aspects describe the dynamics part of our propagator $\mathcal{D}$ mentioned in equation (1.2). Finally particles will be able to react, meaning that their type can change with a fixed rate (unimolecular reaction) or that they react together with other particles to form yet another particle (bimolecular reaction). This would describe the reactions part of our propagator $\mathcal{R}$. With the explanations given in this section we will be able to propagate a given system from state $\boldsymbol{\psi}(t)$ to $\boldsymbol{\psi}(t + \tau)$. In the style of equation (1.2) we will denote this as

$$\boldsymbol{\psi}(t + \tau) = \mathcal{R}_0 \, \mathcal{D}_0 \, \boldsymbol{\psi}(t) \tag{4.1}$$

where $\mathcal{D}_0$ is the Brownian Dynamics propagator and $\mathcal{R}_0$ our default reaction propagator, and $\boldsymbol{\psi}(t)$ is a short-hand notation for the state of the whole system. The details of $\mathcal{D}_0$ and $\mathcal{R}_0$ are now explained in detail.

4.1 Brownian Dynamics

We start by stating the Langevin [17] equation for a single particle

$$m\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = -\frac{1}{\mu}\frac{\mathrm{d}x}{\mathrm{d}t} - \nabla U(x) + \eta(t), \tag{4.2}$$

which is an equation of motion similar to Newton's equation of motion for particles involving a force due to a potential energy U and a non- deterministic noise term η. Further m is the mass of the particle, x its position and μ its mobility. Such an equation will result in a partly random motion of particles. This is why such an equation of motion is called *stochastic*. The exact form of η is yet unspecified, but we know its first and

second moment

$$\begin{aligned} \langle \eta(t) \rangle &= 0 \\ \langle \eta(t)\eta(t') \rangle &= 2\,\delta(t-t')\,k_B T\,/\mu, \end{aligned} \tag{4.3}$$

where $\delta(t-t')$ is the Dirac delta function, k_B is the Boltzmann constant and T is the temperature. The delta function states also that the random forces acting on the particle are uncorrelated.

From the equation (4.2) we finally want to arrive at an expression giving us the new position at a timestep τ later given the old position x. Therefore we make an important assumption by setting the left hand part of equation (4.2) to zero. Assuming that there is no net acceleration *for large times*, which means that we will make an error on short times, but the long term behavior is mostly determined by the friction, noise and the potential energy. We further multiply the equation with μ and arrive at the following expression

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -\frac{D}{k_B T}\nabla U(x) + \xi(t) \tag{4.4}$$

where we defined the random velocity $\xi(t) := \mu\,\eta(t)$. Making use of the Einstein-Smoluchowski relation $D = \mu\,k_B T$ which relates the diffusion constant D with the mobility and the temperature, we can rewrite the moments (4.3) of the random velocity

$$\begin{aligned} \langle \xi(t) \rangle &= 0 \\ \langle \xi(t)\xi(t') \rangle &= 2D\,\delta(t-t'). \end{aligned} \tag{4.5}$$

We now discretize the time so that we can integrate it successively. Therefore an Euler-Maruyama method is used. This transition from continuous to discrete time means the following changes

$$\begin{aligned} \frac{\mathrm{d}x}{\mathrm{d}t} &\to \frac{x_{t+\tau} - x_t}{\tau} \\ \delta(t-t') &\to \frac{\delta_{tt'}}{\tau}, \end{aligned} \tag{4.6}$$

where x_t is the position before the timestep and $x_{t+\tau}$ is the position after the timestep. The Dirac delta function changes to a Kronecker delta divided by the timestep to get the correct normalization. After defining the random position change from the random velocity $R_t = \tau\,\xi(t)$, we can finally formulate the equation that propagates particles as it is done in ReaDDy and in this work as well (at first)

$$x_{t+\tau} = x_t - \tau\frac{D}{k_B T}\nabla U(x_t) + R_t. \tag{4.7}$$

This integration scheme is often called *Brownian Dynamics*[1]. We can again write the moments of the random position change, but we will from now on associate the drawn random numbers with a *normal* distribution or *Gaussian* distribution. Up to now we have given the values in one spatial dimension, but the whole equation is easily generalized

[1] In the literature the random component R_t is ofttimes called the differential of a Wiener process dW_t and the integration scheme written in rescaled units as $dX_t = -\nabla U(X_t)dt + \sqrt{2\beta^{-1}}dW_t$ with $\beta = 1/k_B T$

for three dimensions if spatial variables become vectors (which will be bold print in this document)

$$\begin{aligned} \langle R_t \rangle &= 0 \\ \langle R_t R_{t'} \rangle &= 2D\tau\, \delta_{tt'} \quad \text{(1D)} \\ \langle \mathbf{R}_t \mathbf{R}_{t'} \rangle &= 6D\tau\, \delta_{tt'} \quad \text{(3D)}. \end{aligned} \tag{4.8}$$

With this we have the tools to describe the dynamics of any given system consisting of particles and potentials U, which leads to the next part of our model.

4.2 Interaction potentials

We have mentioned how particles are propagated in time if the potentials U and the resulting forces $-\nabla U$ are known. In this work we divide potentials in two classes. First order potentials and second order potentials. First order potentials are only dependent on the coordinates of one particle and are used to realize cellular geometries, e.g. membranes or other confinements. Second order potentials depend on the relative coordinate of two particles and are used to realize particle interactions, e.g. particle repulsion or attraction.

As a first order potential example, a harmonic repulsive infinite sized wall is useful for confining particles to a volume. The potential energy U and the force $\mathbf{F}$ of such a wall is defined as

$$U = \begin{cases} \kappa r^2 & \text{if } r \leq 0 \\ 0 & \text{if } r > 0 \end{cases} \quad \mathbf{F} = \begin{cases} -2\kappa r\, \mathbf{n} & \text{if } r \leq 0 \\ 0 & \text{if } r > 0, \end{cases} \tag{4.9}$$

where κ is a parameter specifying the strength of repulsion, $\mathbf{n}$ is the normal vector of the plane and r is the closest distance of the particle to the wall defined as

$$r = (\mathbf{r} - \mathbf{x}_0) \cdot \mathbf{n}.$$

Here $\mathbf{x}_0$ is an arbitrary point of the plane and $\mathbf{r}$ is the position of the particle.

For second order potentials, we will mostly use harmonic repulsion of particles and sometimes Lennard-Jones interaction. The harmonic repulsion of two particles i and j at positions $\mathbf{r}_i$ and $\mathbf{r}_j$ is defined as

$$\begin{aligned} U &= \begin{cases} \kappa \left(r_{ij} - \rho_{ij}\right)^2 & \text{if } r_{ij} \leq \rho_{ij} \\ 0 & \text{if } r_{ij} > \rho_{ij} \end{cases} \\ \mathbf{F}_i &= \begin{cases} 2\kappa \left(r_{ij} - \rho_{ij}\right) \mathbf{r}_{ij}/r_{ij} & \text{if } r_{ij} \leq \rho_{ij} \\ 0 & \text{if } r_{ij} > \rho_{ij}, \end{cases} \end{aligned} \tag{4.10}$$

where $\mathbf{r}_{ij} = \mathbf{r}_j - \mathbf{r}_i$ and r_{ij} is the euclidean norm of $\mathbf{r}_{ij}$. Further $\rho_{ij} = \rho_i + \rho_j$ is the sum of the radii of the two particles. In essence this means, if the particles' centers are closer than the sum of their radii, the particles will repulse each other. The strength of the repulsion is determined by the parameter κ and repulsion itself is rather soft and will allow particles to overlap partly. This soft potential is especially sensible when we consider complete

proteins as a sphere. It also lets us choose a larger timestep as compared to hard-sphere potentials.

The Lennard-Jones interaction potential is well studied and is used here mostly to test if Brownian Dynamics works as expected. Also it is an alternative representing hard spheres as this potential is rather stiff and requires smaller timesteps to minimize discretization errors and keep the simulation stable. To represent a biological system with reaction-diffusion however it is not recommended. The potential U and the force $\mathbf{F}$ between particles i and j with positions $\mathbf{r}_i$ and $\mathbf{r}_j$ read

$$\begin{aligned} U &= 4\epsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^{6} \right] \\ \mathbf{F}_i &= \frac{24\epsilon}{r_{ij}^2} \left[2 \left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^{6} \right] \mathbf{r}_{ij} \end{aligned} \tag{4.11}$$

where ϵ is an energy parameter and σ is a parameter describing the range of the potential. Often σ is chosen to represent the sum of the radii of the particles. We will however in accordance with the soft potential set σ such that when the distance of particles is exactly the sum of the radii, the potential is at its minimum, which is the case at $2^{1/6}\sigma$. So if $\rho_{ij} = \rho$ is the sum of the particles' radii, then $\sigma = 2^{-\frac{1}{6}} \rho$. Further we will use a cutoff-distance of 2.5ρ.

With the integrator and the potentials we have everything at hand to simulate a system with fixed number of particles, which defines our propagator $\mathcal{D}_0$ stated in equation (4.1). In most biological applications particles change their state or react with one another, which has to be covered as well.

4.3 Reactions

In ReaDDy [4] and also in this work the most important reactions are unimolecular reactions and bimolecular reactions. These can for example represent the state change of

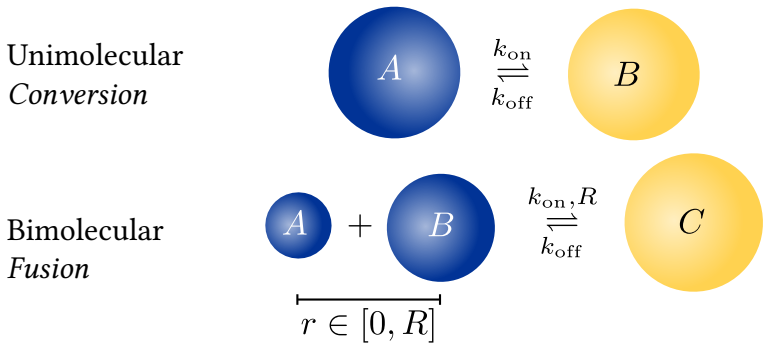


Figure 4.1: Reversible reactions that are considered. Note that all rates are microscopic by construction and come in units of s^{-1}.

a protein from an inactive to an active state or the association of a protein and ligand or the activation by a messenger particle.

Unimolecular reactions (*Conversion*) can occur at all times with a probability

$$p = 1 - e^{-k_{\text{on/off}}\,\tau}, \tag{4.12}$$

which is the Poisson probability of a reaction event occuring during the timespan τ.[2] For small timesteps compared to the characteristic reaction time $k_{\text{on/off}}^{-1}$, also the approximation

$$p \approx k_{\text{on/off}}\,\tau \tag{4.13}$$

holds. When a unimolecular reaction occurs, e.g. in the *on* direction, then particle A would be destroyed and a new particle B will be placed at the same position.

For bimolecular *on* reactions (*Fusion*) the mechanism is different. In addition to the two reaction rates given in table 4.1 we also need a reaction distance R. Only when the two educts A and B are within a distance of R a reaction to form a particle C can occur with a probability as given in equation (4.13) where the rate now is the *microscopical* association rate k_{on}.[3] When such a reaction occurs the two particles A and B are deleted and a new particle C is placed in between them such that the center of mass is conserved. To illustrate this consider figure 4.2. The blue dashed lines represent particles A and B at positions $\mathbf{r}_A$ and $\mathbf{r}_B$. They are seperated by a vector $\mathbf{r}_{AB} = \mathbf{r}_B - \mathbf{r}_A$, whose length r is

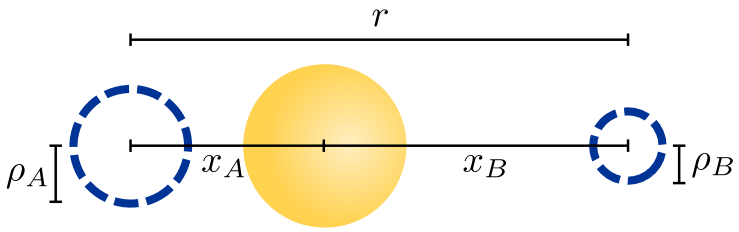


Figure 4.2: Scheme to conserve the center of mass during bimolecular reaction in both directions. r is the distance of particles A and B. The ρ_i correspond to the collision radii of particles and x_i are the weighted distances of the particles from the center of mass.

smaller than the reaction distance $r < R$. The particles' collision radii are ρ_A and ρ_B. We assume that the particles have a mass which depends cubic on their collision radius. If we want to conserve the center of mass when particles A and B are deleted, we have to place the new particle C in between A and B with the exact position weighted with the cubic radii of particles. The new position $\mathbf{r}_C$ of particle C is then

$$\mathbf{r}_C = \mathbf{r}_A + x_A\,\mathbf{r}_{AB} = \mathbf{r}_B - x_B\,\mathbf{r}_{AB} \tag{4.14}$$

[2] This timespan τ is not necessarily the same timestep mentioned in the dynamics/diffusion part. It might be useful to choose the reaction timestep as an integral multiple of the diffusion timestep, as long as the reaction timestep is still small enough compared to the fastest reaction rate

[3] From here on we will call the microscopic association rate k_{on} and the macroscopic $k_{\text{on,macro}}$

Here the weights x_A and x_B represent the mass of the particles normalized according to $x_A + x_B = 1$ and read

$$x_A = \frac{\rho_A^3}{\rho_A^3 + \rho_B^3} \quad x_B = \frac{\rho_B^3}{\rho_A^3 + \rho_B^3} \tag{4.15}$$

In the case of this bimolecular *on* reaction the effective reaction rate $k_{\text{on,macro}}$ then depends on the system. For non- interacting particles in three dimensional space the effective rate was already given in equation (3.2), where R_{AB} then corresponds to the distance R at which the enounter complex is formed.

For the bimolecular *off* reaction (*Fission*), the case is similar to unimolecular reactions and a reaction event can always occur with probability (4.13). When this happens new positions for particles A and B must be found. At first a random orientation $\mathbf{n}$ is created. This is a random unit vector, that is uniform with respect to spherical coordinates. Then the distance of particles r is determined. It is drawn from a distribution $f(r)$, which will be discussed in chapter 5. The new positions $\mathbf{r}_A$ and $\mathbf{r}_B$ are then calculated such that the center of mass is again conserved

$$\begin{aligned} \mathbf{r}_A &= \mathbf{r}_C - r\, x_A\, \mathbf{n} \\ \mathbf{r}_B &= \mathbf{r}_C + r\, x_B\, \mathbf{n} \end{aligned} \tag{4.16}$$

With the execution of unimolecular and bimolecular reactions we have defined our default reactions propagator $\mathcal{R}_0$ stated in equation (4.1). In the following we will extend these by introducing the *reversibility* / *detailed balance* feature, which is the main aspect of this work.

5 Achieving reversibility

In chapter 3 a system is given where forward and backward rates of a bimolecular reaction are not balanced in the sense that in the backwards reaction two new particles are positioned without considering the rest of the system. That could lead to physically unreasonable states, that have a large increase in system energy. To correct this, we will enforce detailed balance. Another motivation to do so is that we want to ensure that the studied systems are sampled correctly, to avoid high energetic states to be visited too frequently as this cannot be assumed for pure Brownian Dynamics [18, 19]. Before doing so we have to clarify if it is sensible to enforce detailed balance on reaction-diffusion simulations that often represent models of biological systems.

Assume you were to model the ATP production of a Mitochondrion by means of stochastic reaction-diffusion simulations. When one treats the Mitochondrion or just the ATP-syntase complex in a closed simulation box, one can never observe the process of chemi-osmosis in its fullness, since the real world process relies on the constant in-flux and out-flux of several components/molecules involved. In other words, the process can only work with an ongoing flow of energy in- and out of the system. If a living system is treated in a closed environment it will simply stop functioning. Qian et al. [20] puts it like "*In terms of physical chemistry, a closed system has no life*".

Another important insight is that a global equilibrium can only exist in a closed system. Most systems that we are interested in live far from global equilibrium. In such systems however one can often observe so called non-equilibrium steady states. The problem that occurs when we want to model systems *realistically* is that the detailed balance condition (1.3) is constructed for systems that have a probability distribution $\pi(a)$ that corresponds to global equilibrium. For non-equilibrium systems we can still formulate a so called *local detailed balance* [21] or *generalized detailed balance* [22], which is a weaker constraint than *global detailed balance*. A non-equilibrium system may violate *global* detailed balance but still fulfill *local* detailed balance [23]. The latter will be implemented in this work and both concepts should converge to the same method when dealing with closed systems. We denote both concepts as

$$\text{global detailed balance:} \qquad \frac{p(a \to b)}{p(b \to a)} = \frac{\pi(b)}{\pi(a)} \tag{1.3}$$

$$\text{local detailed balance:} \qquad \frac{p(a \to b)}{p(b \to a)} = \exp\left(\frac{\Delta S}{k_B}\right) \tag{5.1}$$

By $p(a \to b)$ we denote the transition probability of going from state a to b and $\pi(a)$ is the equilibrium probability of being in state a. The main difference between these two definitions is that instead of the equilibrium probabilities we just need to know the

change of entropy ΔS for the local detailed balance. Since we are dealing with systems in contact with a heat bath and variable particle numbers, we will calculate the entropy based on the grand-canonical ensemble. This means that the energy function of interest is the grand potential Φ [24]. It can be viewed as the free Helmholtz energy of a system with varying particle numbers which creates additional energy terms for adding/removing particles to/from the system. It is defined as

$$\Phi = E - TS - \sum_s \mu_s N_s \tag{5.2}$$

where E describes the internal energy terms, T is the temperature, μ_s is the chemical energy for particle species s, that describes the energy change when the number of s-particles increases from N_s to $N_s + 1$. When we calculate the change of entropy using equation (5.2) we get

$$\Delta S = -\frac{1}{T}\left(\Delta\Phi + \sum_s \mu_s \Delta N_s\right). \tag{5.3}$$

Here we have assumed that the internal energy remains constant $\Delta E = 0$. If any internal configurations of particles should change, we will account for that with a unimolecular reaction that incorporates the appropriate change of energy via the chemical energy. We have further assumed that the chemical energy μ_s of species s does not depend on the state of the system. This leads to our final expression for local detailed balance

$$\frac{p(a \to b)}{p(b \to a)} = \exp\left(-\beta\left[\Delta\Phi + \sum_s \mu_s \Delta N_s\right]\right) \tag{5.4}$$

with $\beta = 1/k_B T$. The term $\Delta\Phi$ then accounts for all potential energy changes that correspond to the interactions that were introduced in chapter 4. Enforcing local detailed balance will still guarantee that entropy increases as the system evolves even without global equilibrium, which allows us to study non-equilibrium systems.

To achieve local detailed balance, we will modify our propagators already mentioned in equation (4.1). An acceptance step in the fashion of Metropolis-Hastings [15] is performed after the system has been propagated for one timestep. To be more precise, one acceptance step is done after the Brownian Dynamics step $\mathcal{D}$ and one after the reaction step $\mathcal{R}$. One could write the new procedure similar to equation (4.1) as

$$\boldsymbol{\psi}(t + \tau) = \mathcal{R}_{\text{acc}}\, \mathcal{R}_0\, \mathcal{D}_{\text{acc}}\, \mathcal{D}_0\, \boldsymbol{\psi}(t) \tag{5.5}$$

In other words this means that we will propagate the particles according to Brownian Dynamics $\mathcal{D}_0$. In the acceptance step $\mathcal{D}_{\text{acc}}$ we will either accept this new state or return to the old state. The same applies for reactions $\mathcal{R}$. In any case the clock will then be advanced from t to $t + \tau$. If we perform the simulations without the acceptance steps we will call it *irreversible*. If we use the acceptance steps, then the simulation will be called *reversible*. We will now derive the expressions for the proposal and acceptance probabilities.

The condition we want to fulfill is given in equation (5.4) when going from state a to state b. Local detailed balance states that this condition must hold for all elementary processes. It also means that the elementary processes, here displacements and reactions, must be reversible. To find the probability of accepting a step or not we will split the conditional probabilities in (1.3) as follows

$$p(a \to b) = q(a \to b)\alpha(a \to b) \tag{5.6}$$

where we have splitted the total probability of going from a to b into the proposal probability $q(a \to b)$ and the acceptance probability $\alpha(a \to b)$. It can be shown that local detailed balance (5.1) is fulfilled by the acceptance probability

$$\alpha(a \to b) = \min\left\{1, \frac{q(b \to a)}{q(a \to b)} e^{-\beta[\Delta\Phi + \sum_s \mu_s \Delta N_s]}\right\}. \tag{5.7}$$

We will now derive the needed expressions of acceptance probabilities in the case of pure particle displacements and then of pure reactions.

5.1 In dynamics

In the case of pure dynamics $\mathcal{D}$ without reactions and thereby a constant number of particles, we can formulate the acceptance probability (5.7) by knowing that our Brownian Dynamics integrator proposes positions that have a mean dependent on the deterministic displacement due to forces (second term on right-hand side of eq. (4.7)) and a variance of $2D\tau$ in every spatial component. In three dimensions the proposal probability $q(a \to b)$ reads

$$q(a \to b) = (4\pi D\tau)^{-3/2} \exp\left[-\frac{(\Delta\mathbf{x} - \beta D\tau \mathbf{F}_a)^2}{4D\tau}\right] \tag{5.8}$$

where $\Delta\mathbf{x} = \mathbf{x}_b - \mathbf{x}_a$ are the displacements of particles, $\mathbf{F}_a$ are the forces acting on particles in state a and $\beta^{-1} = k_B T$ is the inverse thermal energy. The backward probability is defined analogously. The entropy change is represented by the Boltzmann distributions in the constant particle number case

$$e^{-\beta[\Delta\Phi + \sum_s \mu_s \Delta N_s]} = e^{-\beta\Delta U} \tag{5.9}$$

with the potential energy difference $\Delta U = U_b - U_a$. After plugging these expressions into (5.7) and rearranging we get[1]

$$\alpha(a \to b) = \min\left\{1,\ e^{-\beta\left[\frac{1}{2}\Delta\mathbf{x}\cdot(\mathbf{F}_b + \mathbf{F}_a) + \frac{\beta D\tau}{4}\left(\mathbf{F}_b^2 - \mathbf{F}_a^2\right) + \Delta U\right]}\right\} \tag{5.10}$$

Summing up we will propose the new state with probability $q(a \to b)$ during the $\mathcal{D}_0$ step and we will accept it with the probability $\alpha(a \to b)$ during the $\mathcal{D}_{\text{acc}}$ step.

[1]Note that the term $\Delta\mathbf{x} \cdot (\mathbf{F}_b + \mathbf{F}_a)$ is meant to be the dot product or euclidean inner product.

5.2 In reactions

When considering the reaction step $\mathcal{R}_0$ and $\mathcal{R}_{\text{acc}}$, we will multiply the acceptances of every individual reaction occuring. This is necessary because we are considering different types of reactions as already stated in section 4.3. These are unimolecular and bimolecular. Assuming that within a timestep, there are M reaction events proposed. Then the total acceptance probability is

$$\alpha(a \to b) = \min\left\{1, \prod_i^M \alpha_i\right\} \tag{5.11}$$

where the individual α_i have a familiar form

$$\alpha_i = \frac{q_i(b \to a)}{q_i(a \to b)} e^{-\beta[\Delta\Phi + \Delta\mu_i]}. \tag{5.12}$$

Here the $q_i(b \to a)$ and $q_i(a \to b)$ are now the proposal probabilities of a single reaction occuring. We have now denoted $\Delta\mu_i$ as the change of chemical energy during the reaction i. Furthermore we will assume that this change of chemical energy corresponds to the ratio of the microscopic reaction rates

$$\frac{k_{\text{i,on}}}{k_{\text{i,off}}} = e^{-\beta\Delta\mu_i}. \tag{5.13}$$

In the modeling picture this means, that a reaction that consumes a large amount of energy is less frequent than the reverse reaction where the overall system energy is decreased. To illustrate such a situation consider figure 5.1. Assume you know of the interaction of two particles A and A and the according potential energy function looks like in figure 5.1. If you would observe such a system of two A particles at a certain temperature T you would see them often trapped within a distance x_0 of each other but you would also observe them sometimes seperated by a distance larger than x_0. If we want to build a reaction model out of such a system we would treat the two particles individually when they are far apart and treat them as a single particle C when they come closer than the distance x_0. This problem is equivalent to the Kramers problem [25]. Then the rates k_{on} and k_{off} correspond to the change in chemical energy $\Delta\mu$ as given in equation (5.13)

In addition to local detailed balance, we want another constraint to be fulfilled when considering reactions.

> *constraint: in very dilute systems or in the absence of other particles, reactions should occur with an acceptance probability of 1. Only in crowded environments should the acceptance decrease.* (5.14)

This will ensure that the macroscopic reaction rates correspond to those, that can be calculated by means of ordinary differential equations.

We will now derive the expressions for the acceptance probabilities for the two possible reactions based on their proposal probabilities and the according change of entropy.

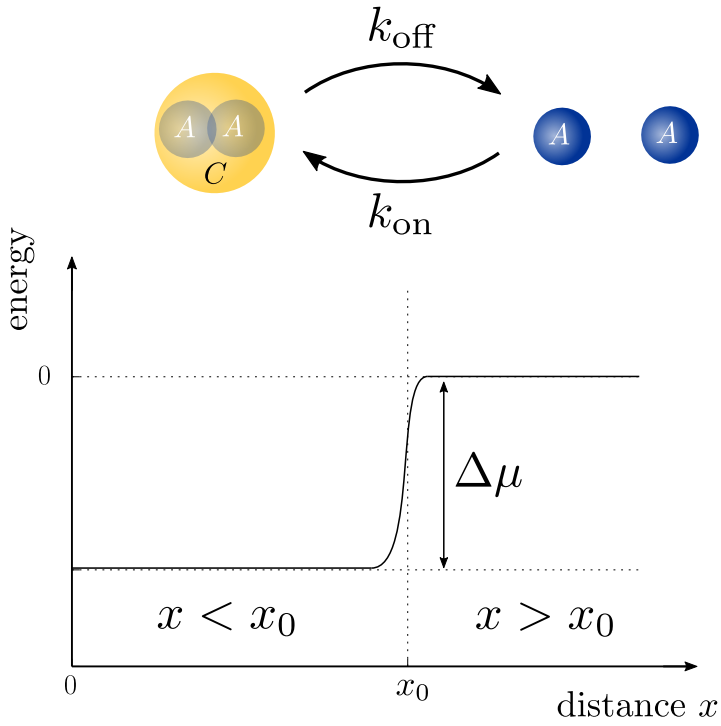


Figure 5.1: Example what a reaction model could describe and how the rates depend on the energy change.

5.2.1 Unimolecular

In the unimolecular reaction

$$A \underset{k_{\text{off}}}{\overset{k_{\text{on}}}{\rightleftharpoons}} B \tag{5.15}$$

we already stated in equation (4.13) the probability of proposing a reaction of this type. If state a is associated with the particle type A and state b with the type B, then we can write

$$q_{\text{on}} = k_{\text{on}}\tau \qquad q_{\text{off}} = k_{\text{off}}\tau\,. \tag{5.16}$$

Applying equation (5.13) we get

$$\frac{k_{\text{off}}}{k_{\text{on}}} = e^{+\beta(\mu_B - \mu_A)} = e^{+\beta\Delta\mu} \tag{5.17}$$

with the chemical energy μ of the particle species A and B. The entropy change in this case reads

$$e^{\Delta S/k_B} = e^{-\beta[\Delta\Phi + \sum_s \mu_s \Delta N_s]} = e^{-\beta\Delta U} e^{-\beta\Delta\mu}. \tag{5.18}$$

Gathering all equations we can state the acceptance probability for both directions of the unimolecular reaction (5.15)

$$\text{Unimolecular reaction:}\quad \alpha_i = e^{-\beta\Delta U}. \tag{5.19}$$

This means that we only have to consider the potential energy changes like particle repulsion etc. when performing a unimolecular reaction.

5.2.2 Bimolecular

Consider the bimolecular (*Fusion*) reaction

$$A + B \underset{k_{\text{off}}}{\overset{k_{\text{on}},R}{\rightleftharpoons}} C \tag{5.20}$$

where R is the reaction distance of this reaction. We will distinguish between the case of the forward/on reaction and the backward/off reaction and calculate the proposal probabilities first.

The forward reaction can occur only when particles A and B are within distance R, then the proposal probability is

$$q_{\text{on}} = k_{\text{on}}\tau \tag{5.21}$$

and particle C will be placed such that the center of mass is conserved as described in chapter 4. The backward reaction $C \to A + B$ can always occur with probability $k_{\text{off}}\tau$ and the particles A and B are put at a distance r away from each other. This distance r is drawn from a probability distribution $f(r)$ so that the probability of putting them at r is $f(r)\mathrm{d}r$. The form of f depends on the interaction potential between particles A and B, called U_{AB}. We denote it as

$$f(r) = Z^{-1} e^{-\beta U_{AB}(r)} \quad \text{with } r \in [0, R] \tag{5.22}$$

and the normalisation constant Z, defined as

$$Z = \int_0^R e^{-\beta U_{AB}(r)} \mathrm{d}r \tag{5.23}$$

In general the definition of f means that distances r where the interaction is weak are favored over distances where the interaction is strong. We can write down the proposal probability of the backward reaction as

$$q_{\text{off}} = k_{\text{off}}\tau\, f(r)\mathrm{d}r \tag{5.24}$$

From equations (5.21) and (5.24) we can formulate the acceptance probability for the forward reaction using equation (5.12) and the fact that the reaction rates cancel out the change of chemical energy

$$\alpha_{\text{on}} = f(r)\, \mathrm{d}r\, e^{-\beta\Delta\Phi}$$

The change of grand potential consists only of potential energy changes ΔU with the environment (other particles or geometrical constraints). Plugging in f we get

$$\alpha_{\text{on}} = Z^{-1}\, e^{-\beta U_{AB}(r)} e^{-\beta\Delta U}$$

This can be easy calculated during the simulation but if we want to fulfill the constraint (5.14) we need to make sure that

$$Z = \int_0^R e^{-\beta U_{AB}(r)} \mathrm{d}r = 1 \tag{5.25}$$

because in the dilute case the term ΔU is exactly $-U_{AB}(r)$ and the acceptance probability becomes 1. The same calculation and arguments hold for the reverse process and we can state the acceptance probabilities for both forward and backward reaction

$$\text{Bimolecular reaction:} \quad \begin{aligned} \alpha_{\text{on}} &= e^{-\beta U_{AB}(r)} e^{-\beta \Delta U} \\ \alpha_{\text{off}} &= e^{+\beta U_{AB}(r)} e^{-\beta \Delta U} \end{aligned} \tag{5.26}$$

Still we have to figure out how to accomplish the correct normalization $Z = 1$. The answer lies in the reaction distance R, that we have not made any assumptions about. In the model of ReaDDy [4] the reaction distance was some arbitrary parameter that had to be chosen by the modeler even without any knowledge about the distance at which reactions become important. We will chose the reaction distance R such that

$$\int_0^R e^{-\beta U_{AB}(r)} \mathrm{d}r = 1 \tag{5.27}$$

We will illustrate how this turns out if you assume that the interaction between particle A and B is a harmonic repulsion as described in section 4.2. We vary the collision distance ρ of particles and solve the root finding problem

$$\int_0^R e^{-\beta U_{AB}(r)} \mathrm{d}r - 1 = 0$$

This is easily solvable with standard numeric methods such as bi-section, since the integrand is never negative and there is always exactly one solution R that solves the problem. The result can be seen in figure 5.2. One can see that in the case of a repulsive interaction the reaction distance found is always larger than the collision distance of particles $R > \rho$. It is also evident that the reaction distance converges to some value as the interaction vanishes

$$R \to \text{const as } \rho \to 0.$$

The question is now what this constant value is. Fortunately there is already an answer to this question. It was already given in chapter 3. The formula provided by Erban [16] relates the microscopic reaction rate k_{on} with the reaction distance R and the macroscopic $k_{\text{on,macro}}$ the case of no interaction. We state it here again

$$k_{\text{on,macro}} = 4\pi D \left[R - \sqrt{\frac{D}{k_{\text{on}}}} \tanh\left(R \sqrt{\frac{k_{\text{on}}}{D}} \right) \right] \tag{3.2}$$

The R in this equation corresponds to the R in figure 5.2 at $\rho = 0$ (no interaction), thus connecting the reaction distance to the other parameters one might already know, e.g. from experiments or other simulations.

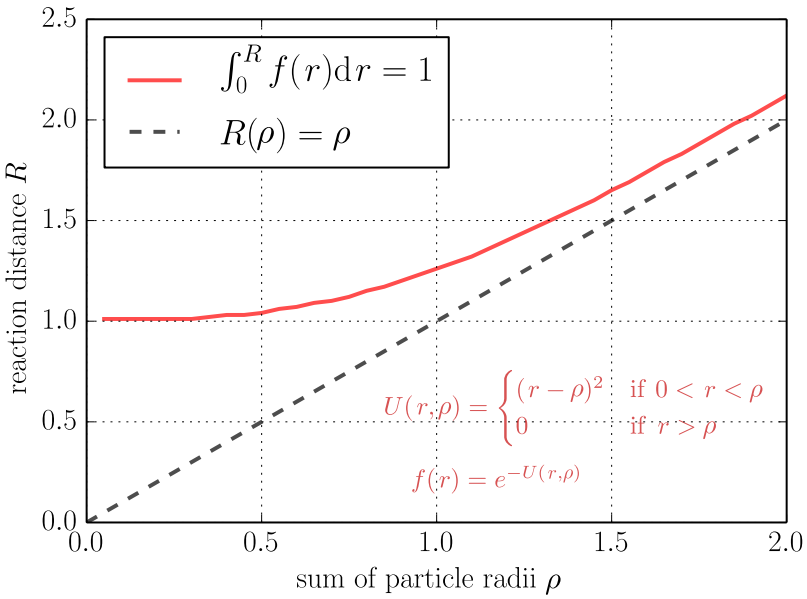


Figure 5.2: The reaction distance R as a function of the collision distance ρ of particles with harmonic repulsion. The dashed line shows "naive" choosing of the reaction distance. The red solid line shows the reaction distance when normalized as given in the text. Inset are the parameters of the used potential, given in rescaled units.

5.3 Summary

We will now give a summary of the new methods in the style of equation (1.2). See table 5.1 for an overview of the different used integration variants. This notation will become useful when we simulate systems differently. For example the algorithm that ReaDDy implements can be described by the notation

$$\boldsymbol{\psi}(t+\tau) = \mathcal{R}_0\mathcal{D}_0\boldsymbol{\psi}(t).$$

In principle one could also choose different timesteps for the dynamics and the reactions, e.g. assuming that the reaction timestep is 100 times the dynamics timestep we could write

$$\boldsymbol{\psi}(t+100\tau) = \mathcal{R}_0(100\tau)\mathcal{D}_0(\tau)^{100}\boldsymbol{\psi}(t)$$

where the argument of $\mathcal{D}$ and $\mathcal{R}$ denote how "far" the system will be propagated.

In the following chapters we will often simulate either the dynamics or the reactions reversibly, i.e. either

$$\boldsymbol{\psi}(t+\tau) = \mathcal{R}_0\mathcal{D}_{\text{acc}}\mathcal{D}_0\boldsymbol{\psi}(t)$$

or

$$\boldsymbol{\psi}(t+\tau) = \mathcal{R}_{\text{acc}}\mathcal{R}_{\text{M}}\mathcal{D}_0\boldsymbol{\psi}(t).$$

Object	Explanation
$\boldsymbol{\psi}(t)$	state of the system at time t
$\boldsymbol{\psi}(t+\tau)$	state of the system one timestep τ later
$\mathcal{D}_0$	Brownian Dynamics propagator, displaces particles $\Delta\mathbf{x} = \beta D\tau\mathbf{F}_t + \sqrt{2D\tau}\,\mathcal{N}_{3\mathrm{D}}$
$\mathcal{D}_{\mathrm{acc}}$	accept the new state with probability $\alpha = \min\left\{1,\ e^{-\beta\left[\frac{1}{2}\Delta\mathbf{x}\cdot(\mathbf{F}_{t+\tau}+\mathbf{F}_t)+\frac{\beta D\tau}{4}\left(\mathbf{F}_{t+\tau}^2-\mathbf{F}_t^2\right)+\Delta U\right]}\right\}$
$\mathcal{R}_0$	Default reaction propagator Unimolecular $A \underset{k_{\mathrm{off}}}{\overset{k_{\mathrm{on}}}{\rightleftharpoons}} B$ Bimolecular $A + B \underset{k_{\mathrm{off}}}{\overset{k_{\mathrm{on}},R}{\rightleftharpoons}} C$ (with arbitrary R)
$\mathcal{R}_{\mathrm{M}}$	Modified reaction propagator Unimolecular $A \underset{k_{\mathrm{off}}}{\overset{k_{\mathrm{on}}}{\rightleftharpoons}} B$ Bimolecular $A + B \underset{k_{\mathrm{off}}}{\overset{k_{\mathrm{on}},R}{\rightleftharpoons}} C$ (with R chosen such that $1 = \int_0^R e^{-\beta U_{AB}(r)}\mathrm{d}r$)
$\mathcal{R}_{\mathrm{acc}}$	accept the new state with probability $\alpha = \min\{1, \prod_i \alpha_i\}$ Unimolecular $\alpha_{\mathrm{on}} = e^{-\beta\Delta U}$ $\alpha_{\mathrm{off}} = e^{-\beta\Delta U}$ Bimolecular $\alpha_{\mathrm{on}} = Z^{-1}e^{-\beta(\Delta U+U_{AB})}$ $\alpha_{\mathrm{off}} = Ze^{-\beta(\Delta U-U_{AB})}$

Table 5.1: Overview of the different methods and their propagators that were introduced.

6 Software *revreaddy*

In the past chapter we have derived all methods in principle. However we are dealing with stochastic methods that require numerical implementation. Therefore a small software-package was written, called *revreaddy*, which is shorthand notation for *reversible ReaDDy*. It is available via github (see [26]). To investigate the local detailed balance and its effects on the dynamics and the reactions of a simulated system, *revreaddy* is mainly written in C++ (with C++11 standard) as a fast object oriented programming language. This C++ code is written as generic as possible to make a simulation kernel that can simulate many different systems only depending on the input by the user (which is mainly the author itself).

This input from the user is in the form of Python scripts. The Python programming/scripting language enables to prepare even complex systems with very little effort. These python scripts have access to the C++ software via a C-extension, which is generated with the help of Cython. This scheme is depicted in figure 6.1. The functions and variables accessible to the user are only to configure a system, i.e. determine its particle types and their propoerties, add interaction potentials to act between particles, add reactions between specified particle types that occur with specific reaction rates. Also the system parameters like the timestep, the temperature and the boxsize - periodic or not - are determined. Furthermore the user has access to start the simulation, after which the C++ program runs for the specified number of timesteps and only returns to the Python layer when it is finished.

How the C++ code works in principle is shown in figure 6.2. The main loop over the range of the maximum number of timesteps is depicted there in pseudocode. This loop is called when the run() function executes. In the beginning the current state is saved to be called again later if the step was rejected. The state consists of the particle positions, the forces acting on them and the total system energy. Then the position of particles are displaced using Brownian Dynamics. To compare the new state to the old state the forces have to be calculated again. The function acceptanceDynamics() determines the acceptance probability of the dynamics-step. According to this probability, the step gets either accepted (nothing happens) or rejected (the old state is invoked again). The same procedure is performed for the reactions. After these two steps the observables for time t are recorded and the next iteration starts.

Observables

The output can also be configured before executing the simulation with the help of the Python scripts. There are several observables that are evaluated on the run and

saved to disk in configurable intervals and formats. These observables are: the mean squared displacement, the radial distribution function, the probability density along one coordinate, the number of particles of a certain type, the total energy, the acceptance probabilities and finally the trajectory. The observables are saved to binary HDF5 files [27, 28]. We will now explain two observables a bit more detailed.

The mean squared displacement (MSD) is a quantity that can give insight on how far particles travel on average within a certain timespan t. This quantity is often written as

$$\text{MSD} = \langle (\mathbf{r}(t) - \mathbf{r}_0)^2 \rangle_{N(t)}$$

where $\mathbf{r}(t)$ is the position of a particle at time t and $\mathbf{r}_0$ was its position at time $t = 0$. The mean $\langle . \rangle_N$ is taken over the ensemble consisting of N particles.[1] In revreaddy one has to specify a particle type at first, whose MSD has to be recorded. When taking the first sample, say at time $t = 0$, the positions $\mathbf{r}_0$ are saved together with the particles unique ids. When another sample is recorded at a later time t then the quantity $\langle (\mathbf{r}(t) - \mathbf{r}_0)^2 \rangle_{N(t)}$ is calculated for the particles whose unique ids were saved at time $t = 0$. This way the ensemble of observed particles $N(t)$ might decrease over time because some of the particles that existed at $t = 0$ do not exist anymore, but at least the expression for the MSD is still exact.

We will derive the theoretical expectation for the MSD in the case of free diffusion. Let $c(x, t)$ be the probability distribution of a particle that shall be positioned at $x = 0$ when $t = 0$ and that diffuses according to a Wiener process. The diffusion equation holds (cf. equation 1.1)

$$\frac{\partial c(x,t)}{\partial t} = D \frac{\partial^2 c(x,t)}{\partial x^2} \tag{6.1}$$

Consider the time evolution of the mean squared displacement of c's x variable

$$\frac{\partial \langle x^2 \rangle}{\partial t} = \frac{\partial}{\partial t} \int x^2 c(x,t) \mathrm{d}x = \int x^2 \frac{\partial c(x,t)}{\partial t} \mathrm{d}x = \int x^2 D \frac{\partial^2 c(x,t)}{\partial x^2} \mathrm{d}x$$

where we have used the exchangeability of the derivative and the integral and equation (6.1). Integration by parts can be applied twice where the boundary terms vanish since $c(x, t)$ is zero for $x \to \pm\infty$. Hence

$$\frac{\partial \langle x^2 \rangle}{\partial t} = 2D \int c(x,t) \mathrm{d}x = 2D$$

since $c(x, t)$ is normalized to 1. We can now write the MSD as a function of the time

$$\begin{aligned} \langle x^2 \rangle &= 2Dt \quad \text{1-dimensional} \\ \langle \mathbf{r}^2 \rangle &= 6Dt \quad \text{3-dimensional} \end{aligned} \tag{6.2}$$

The difference to our measured MSDs is that we average over ensemble of particles instead of the space x of some probability distribution.

[1] In principle the number $N(t)$ can depend on time as well as in reaction-diffusion simulations the number of particles changes.

Another important observable is the radial distribution function, RDF or $g(r)$. It describes how the distances r of particles with respect to each other are distributed. We will use this often to assure that a system is equilibrated, although not shown here for every application example. It is defined as

$$g(r) = \frac{\text{number of particles counted at } r}{\text{number of particles in an ideal gas at } r} = \frac{n(r)}{n_{\text{ideal}}(r)}.$$

In an ideal gas all particles are distributed uniformly in space. If we choose one particle and count all particles that have a distance in the interval between r and $r + \mathrm{d}r$ we will find that

$$n_{\text{ideal}}(r) = \frac{N}{V} 4\pi r^2 \mathrm{d}r$$

where N/V is the number density of the ideal gas and $4\pi r^2 \mathrm{d}r$ is the volume of the spherical shell with width $\mathrm{d}r$. With this the RDF becomes

$$g(r) = \frac{n(r)\, V}{N 4\pi r^2 \mathrm{d}r} \tag{6.3}$$

In revreaddy $\mathrm{d}r$ is the bin size of the histogram which is used to count the number $n(r)$. We will often average this quantity over the whole ensemble of particles N as well as over the timespan T. This is denoted by $\langle g(r) \rangle_{N,T}$.

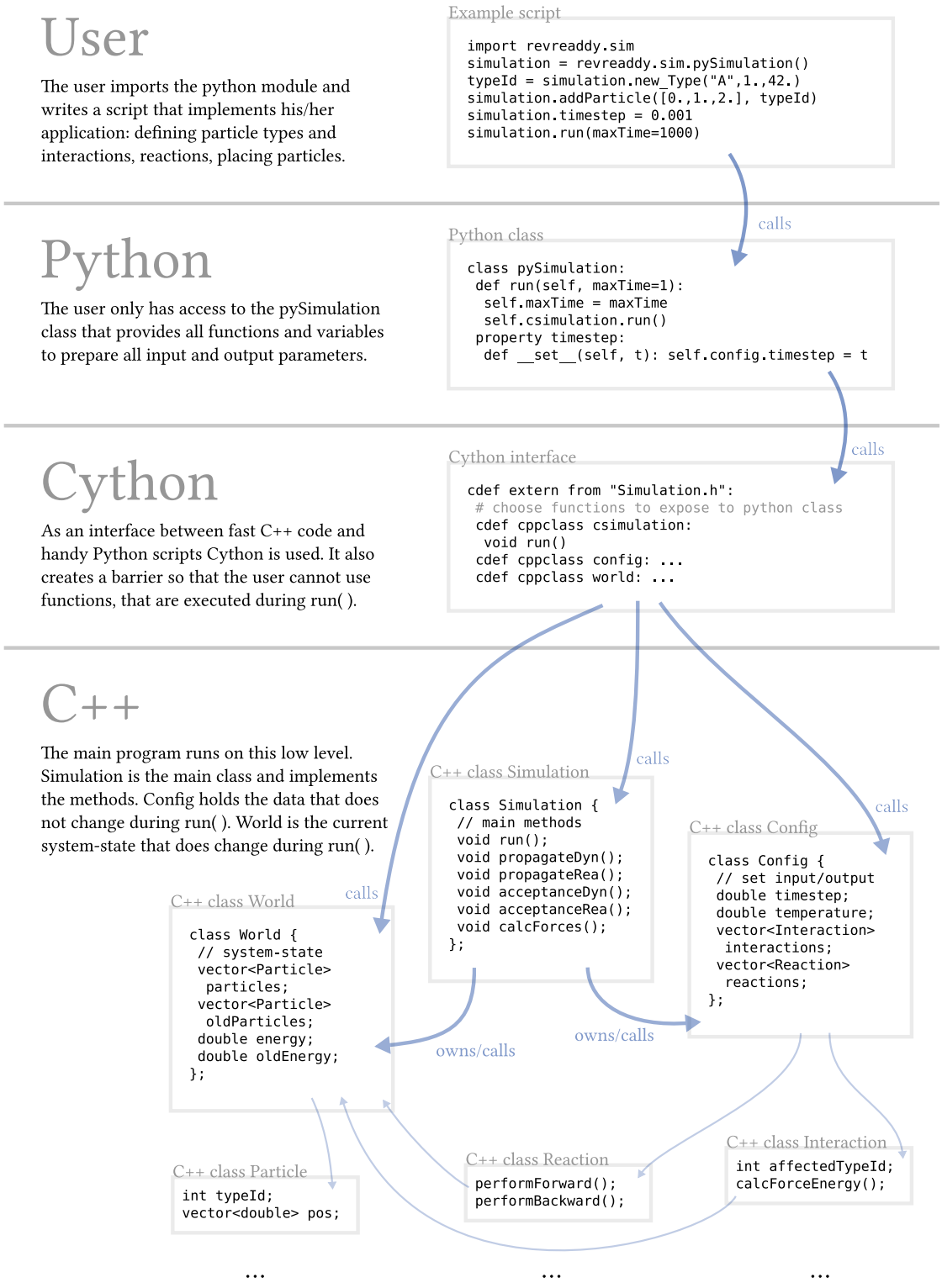


Figure 6.1: Scheme of the revreaddy-software. On the low level C++ side, not all components are shown.

```
state = (particles, forces, energy)                                    ψ(t)
calculateForces(state) # particle-particle, particle-geometry interactions
for t in range(maxTime):
    # dynamics
    oldState = state
    propagateDynamics(state) # Brownian Dynamics according to forces   D_0
    calculateForces(state)
    a = acceptanceDynamics(state, oldState)
    if (a < 1):
        u = random()
        if (u < a):                                                    D_acc
            pass # accept / do nothing
        else:
            state = oldState

    # reactions
    oldState = state
    propagateReactions(state) # according to microscopic rates         R_0/R_M
    calculateForces(state)
    a = acceptanceReactions(state)
    if (a < 1):
        u = random()
        if (u < a):                                                    R_acc
            pass # accept / do nothing
        else:
            state = oldState
```

Figure 6.2: Pseudocode of the revreaddy main loop, given in python-style syntax. The symbols on the right correspond to those given in section 5.3 and represent the propagators of the simulation.

7 Effect on dynamics

We will in this section see how the *reversibility* feature introduced in section 5 will affect the dynamics of well studied systems in four scenarios *Overview*, *Efficiency*, *Deceleration* and *Sampling*. Note that these scenarios are given in rescaled units (see Appendix), which is why there are no units given in the figures and tables.

7.1 Overview

The scenario introduced here will give an overview of how the integrator performs with respect to the acceptance rate and how this can change over time. The system is a typical textbook example, where particles are confined to a cubic volume and then are released into a volume of doubled size. There are 1000 particles and no reactions. All particles are soft particles with the harmonic repulsion given in equations (4.10). In the beginning all particles are in the *left* container, which is surrounded by six wall potentials (see equation 4.9)) that confine the cubic volume. The *right* container has the same size as the other one but they are still seperated by one of the wall potentials. At time t_0 this seperating wall potential is removed so that particles are free to diffuse into the empty second box. Four snapshots at times t_0 to t_3 are given in figure 7.1. These show how the particles fill

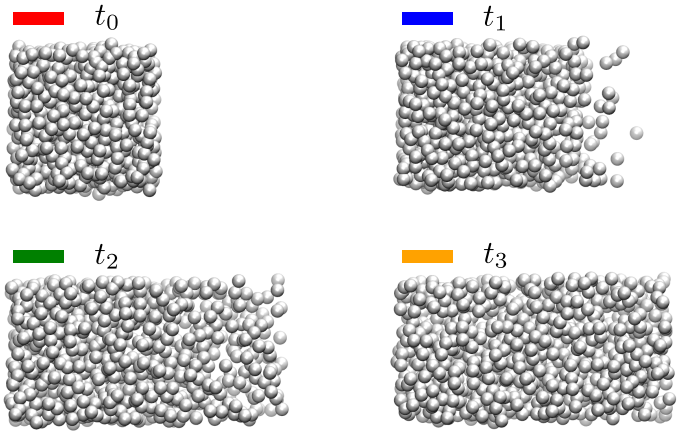


Figure 7.1: Four snapshots of the simulation *Overview* at times t_0 top-left, t_1 top-right, t_2 bottom-left, t_3 bottom-right. Shown is the projection on the x-z-plane, meaning that the x direction is towards the right and the z direction is upwards.

out the available volume. All parameters used are collected in table 7.1. The integration

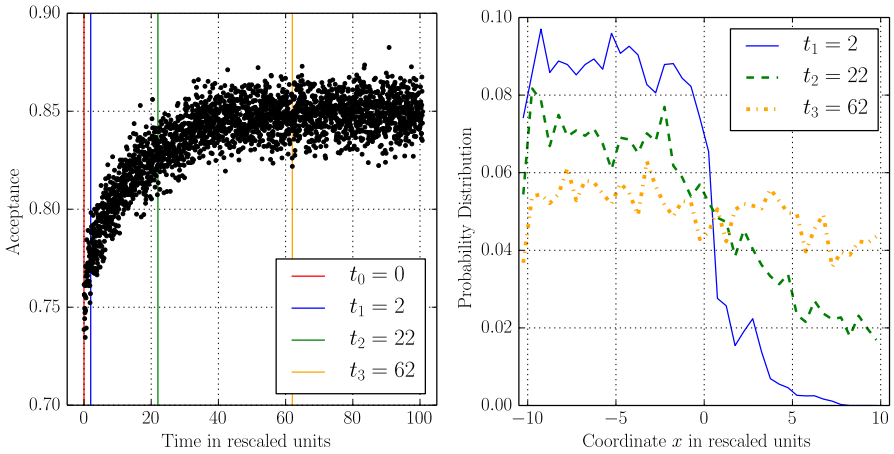


Figure 7.2: The plots for the scenario *Overview*. Left: The acceptance of the dynamic integration step as a function of time shown here with a lagtime of 2000τ. The acceptance is directly obtained from the calculation of eq. (5.10). Right: The probability distribution of finding a particle projected on the x coordinate at different times of the simulation. The times correspond to those given in table 7.1 and figure 7.1.

scheme here can be denoted as (cf. table 5.1)

$$\boldsymbol{\psi}(t+\tau) = \mathcal{D}_{\text{acc}}\mathcal{D}_0\boldsymbol{\psi}(t),$$

which means that no reactions are considered. We only observe the so called dynamical acceptance probability.

parameter	value	parameter	value
number of particles	1000	timestep τ	2×10^{-5}
particle radius	0.5	time t_0	0
particle repulsion coeff. k	10	time t_1	2
wall repulsion coeff. k	10	time t_2	22
volume before t_0	1000	time t_3	62
volume after t_0	2000	periodic boundaries	NO

Table 7.1: Parameters used in the simulation *Overview* given in rescaled units.

The interesting part is now how the acceptance probability defined in equation (5.10) changes as a function of time. This behavior is shown in figure 7.2. It is obvious that the mean acceptance increases as the density of the system decreases. This is mostly due to less particle overlaps and by that smaller differences in energy as the local density of particles decreases. This effect also depends on the timestep. As the timestep increases the acceptance decreases, which is quantified in the next scenario.

7.2 Efficiency

At this point we have seen that the acceptance can change as a function of time. We will show here that it depends also on the size of the timestep τ and the choice of particle potentials.

First we will define the efficiency η, that tells us how far we effectively propagate our system in time when we use a timestep τ

$$\eta = \tau \langle \alpha(t \to t + \tau) \rangle_T \tag{7.1}$$

where $\alpha(t \to t + \tau)$ is the acceptance probability for one timestep of the dynamics of the whole system which was given in equation (5.10). This quantity is averaged over the complete timespan T.

The system used to characterize η is a cubic volume with periodic boundary conditions. It contains 256 particles. The simulation is then run for different timesteps and two different potentials (harmonic repulsion and Lennard-Jones), while the average efficiency is determined. The parameters are given in table 7.2. As in the scenario before we propagate the system according to

$$\boldsymbol{\psi}(t + \tau) = \mathcal{D}_{\text{acc}} \mathcal{D}_0 \boldsymbol{\psi}(t).$$

The results of the measurements of the acceptance and the efficiency are given in figure

parameter	value	parameter	value
number of particles	256	Lennard-Jones coeff. ϵ	1
boxlength	16	timestep τ	10^{-6} - 10^{1}
particle radius	1	averaging timespan T	10000τ
harmonic repulsion coeff. κ	2	periodic boundaries	YES

Table 7.2: Parameters used in scenario *Efficiency* given in rescaled units.

7.3. As expected the acceptance decreases as the timestep increases. This is because the integration error by Brownian Dynamics increases with a more coarse discretization. The Metropolis-Hastings correction then has to make up for that. In general we can also see that a stiffer potential means also a lower acceptance. This becomes clear for example when looking at $\tau = 5 \times 10^{-2}$ in the efficiency plot of figure 7.3. At this point the soft potential is several orders of magnitude more efficient. From the same plot it becomes clear, that the efficiency has a maximum which depends on the system. For both potentials this maximum is in the region where the acceptance is about 50%, which might not be a desirable working point as will be seen in the next section.

7.3 Deceleration

We have seen how the acceptance probability depends on system parameters in the previous two scenarios. We will now see that the new Metropolis-Hastings correction

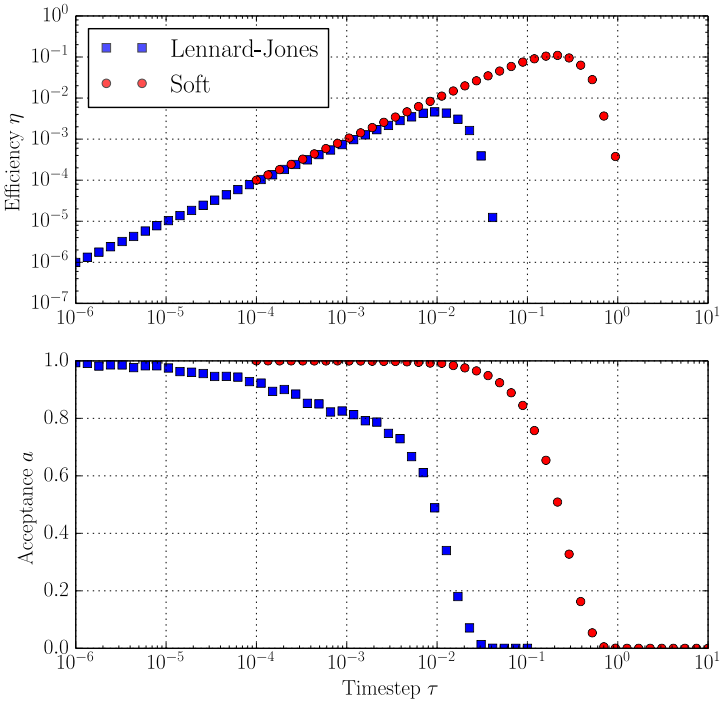
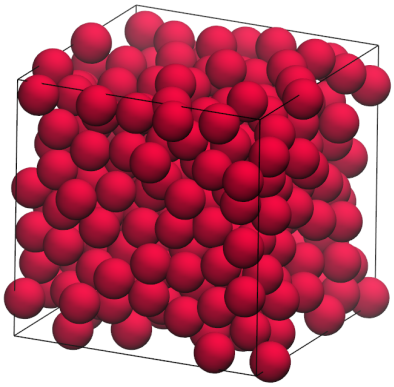


Figure 7.3: Bottom: The acceptance probability averaged over 10000 timesteps as a function of the timestep τ for two different particle interaction potentials. Top: The efficiency η as a function of the timestep τ for two different potentials.

slows the systems dynamics down by a factor of the average acceptance ratio. Therefore we use a system very similar to the *Efficiency* scenario. This time we have 300 particles in a slightly smaller periodic box. To increase the effect of acceptance reduction, the Lennard-Jones interaction is used for particles. Figure 7.4 shows a snapshot of the system and the given parameters. In the given system we measure the mean squared displacement of particles as a function time. This is done with the reversibilty feature (called *reversible*) and without it (called *irreversible*)

$$\begin{aligned} \text{irreversible} \quad & \boldsymbol{\psi}(t+\tau) = \mathcal{D}_0 \boldsymbol{\psi}(t) \\ \text{reversible} \quad & \boldsymbol{\psi}(t+\tau) = \mathcal{D}_{\text{acc}} \mathcal{D}_0 \boldsymbol{\psi}(t). \end{aligned}$$

Each one is repeated 100 times to average over the values and reduce statistical errors. Note that the mean squared displacement is calculated on the fly such that a particle, that leaves the box on one side and enters on the other is recognized by the program. In this way it is possible to track the real travelled distance of particles without having boundary effects, where the mean squared displacement approaches a final value. The results of these measurements are given in figure 7.5. The *irreversible* case shows exact agreement to the free diffusion equation solution for small timescales (ballistic diffusion). In both cases *reversible* and *irreversible* the curves show qualitatively the same behavior but the *reversible* system is slower. Meaning that if you were to rescale the time in the *reversible*



parameter	value
number of particles	300
boxlength	13
particle radius	1
Lennard-Jones coeff. ϵ	1
timestep τ	10^{-5}
periodic boundaries	YES

Figure 7.4: Snapshot of the *Deceleration* scenario (left) and the used parameters given in rescaled units (right).

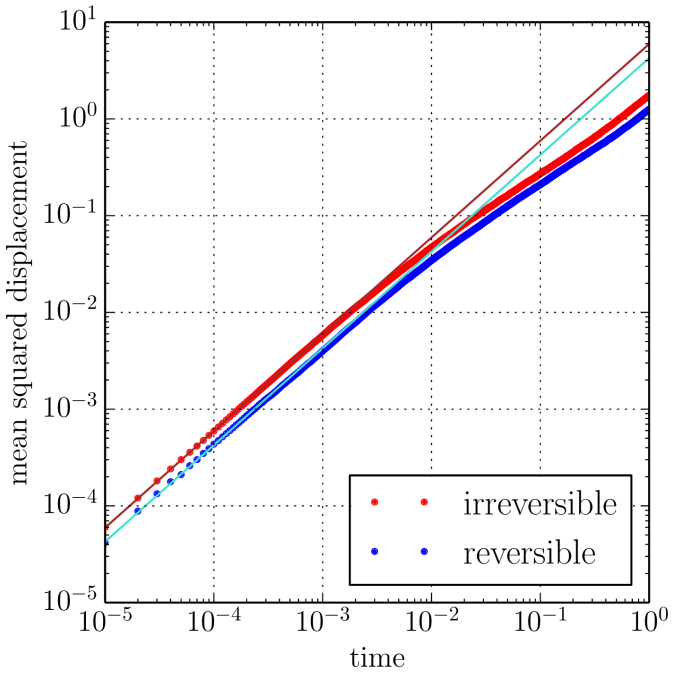


Figure 7.5: The mean squared displacement as a function of time in the *Deceleration* scenario. Errorbars are left out since they are smaller than the dot markers itself. The solid lines are fitted to the ballistic regime $t < 10^{-3}$. The red line exactly represents the $6D$ slope one would expect from the diffusion equation. The blue line has the same slope multiplied by the average acceptance probability. At around $t = 10^{-2}$ subdiffusive behavior starts and towards the end it becomes linear again.

case by a factor of the average acceptance probability you would recover the *irreversible* case. This might work in this example where the average acceptance probability is constant during time. But in general this not the case, see scenario *Overview*, where the acceptance probability depends strongly on time.

One method to make up for that would be to rescale every individual timestep by its calculated acceptance probability. But it turned out that this method leads to wrong sampling of the equilibrium distribution.

7.4 Sampling

We have seen in the previous sections that the acceptance step can slow down the dynamics of any system, depending on how small the acceptance probability is. On the other hand we gain the correct sampling of energetic states, which will be demonstrated here by the *Sampling*. To accomplish this we will perform two single-particle simulations and observe the sampling of the probability density of the particle. A single particle will be diffusing in three dimensions and is subject to a potential energy landscape. This potential only depends on one spatial coordinate z. In the other directions, periodic boundary conditions are realised. In both cases a relatively large timestep is chosen to enhance the effect of discretisation errors in the Brownian Dynamics scheme. As in the scenarios before we will compare the two different integration schemes

$$\begin{aligned} \text{irreversible} \quad & \boldsymbol{\psi}(t+\tau) = \mathcal{D}_0 \boldsymbol{\psi}(t) \\ \text{reversible} \quad & \boldsymbol{\psi}(t+\tau) = \mathcal{D}_{\text{acc}} \mathcal{D}_0 \boldsymbol{\psi}(t). \end{aligned}$$

The two potentials represent well studied toy-models. One of them is the harmonic potential, which is given in rescaled units as

$$\text{harmonic} \quad U(z) = 2z^2. \tag{7.2}$$

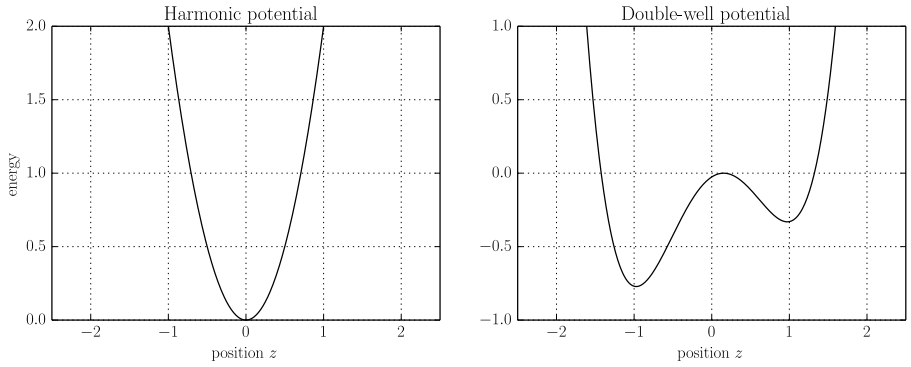
The other potential is a double-well potential that has some extra parameters to determine the distance of the two minima. We denote this potential as

$$\text{double-well} \quad U(z) = 2\left[\left(\frac{z}{L} - \frac{3a}{8}\right)^4 - \left(\frac{z}{L} - \frac{3a}{8}\right)^2 + a\left(\frac{z}{L} - \frac{3a}{8}\right)^3\right] \tag{7.3}$$

where a is a parameter that is here set to $a = 0.3$. Then the L parameter connects to the distance of the two minima like $L = \text{distance of minima}/1.47267$ and we set the distance of minima to 2. The two potentials are plotted in figure 7.6 together with all the parameters used. The results of both potential landscapes can be seen in figure 7.7. The mean acceptance probabilities for the reversible cases are in both simulations

$$\begin{aligned} \text{harmonic potential} \quad & \langle \alpha \rangle_T = 84\% \\ \text{double-well potential} \quad & \langle \alpha \rangle_T = 95\% \end{aligned}$$

The shown exact solution to the problem is a Boltzmann distribution $\rho(z)$



parameter	value
number of particles	1
diffusion coefficient D	0.1
periodic boundary conditions	YES (in x and y)
averaging timespan T	100000τ
timestep harmonic potential τ	2
timestep double-well potential τ	0.1

Figure 7.6: Top: The two potentials used to investigate the sampling of a single Brownian particle. Bottom: Parameters used in scenario *Sampling* given in rescaled units.

$$\rho(z) = Z^{-1} e^{-\beta U(z)} \quad Z = \int_{-\infty}^{+\infty} e^{-\beta U(z)} \mathrm{d}z. \tag{7.4}$$

One can clearly see that the deviation from the exact sampling is much smaller for the reversible case in favor of the metropolised method, which is not suprising since our local detailed balance method becomes the detailed balance method in equilibrium systems. This method is designed such that that the system converges to the correct distribution (in our case $\rho(z)$).

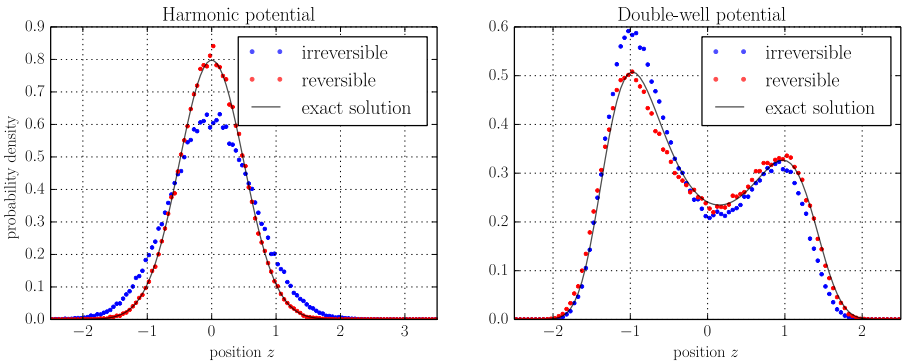


Figure 7.7: A histogram of the particles' visited z positions. The bin width is 0.01 in both cases. The irreversible case coresponds to standard Brownian Dynamics. The reversible case contains an acceptance step (see text).

8 Effect on reactions

Up to now all application examples were performed without reactions, i.e. with constant particle numbers. We will now study two scenarios that include reactions. The first will be the *ReaDDy benchmark* system, which was already mentioned in chapter 3 that focusses on bimolecular reactions. The second example will inlcude both unimolecular and bimolecular reactions that do not conserve total volume of all particles, which is why we call it *volume unconserved*.

8.1 ReaDDy benchmark

The scenario investigated here was mentioned in the ReaDDy paper by Schöneberg et al. [4] and used to demonstrate a bimolecular reaction in a spatiotemporal manner. The reaction of interest reads

$$A + B \underset{k_{\text{off}}}{\overset{k_{\text{on}},R}{\rightleftharpoons}} C$$

where in the forward/on reaction two particles A and B have to come within a distance of R to react with the microscopic rate k_{on}. The backward/off reaction can always occur with the microscopic rate k_{off} . The system starts with only A and B particles of equal amount and eventually relaxes to an equilibrium state, which depends on all system parameters. This procedure is done with and without particle repulsion, which is harmonic repulsion. Particles start to repulse when their distance is smaller than the sum of their radii. The radii ρ of particles A, B and C are chosen such that total volume is conserved during a reaction, i.e. $v_A + v_B = v_C$. The radii of particles then also connect to their diffusion coefficient D and the temperature $k_B T = \beta^{-1}$ via the Einstein-Stokes relation

$$D = \frac{k_B T}{6\pi\eta\rho}$$

where η is the viscosity. The parameters are chosen to represent conditions similar to cellular cytosol at room temperature. Considered is a simulation box which has six confining surfaces that are modeled as repulsive wall potentials (cf. equation (4.9)). The box dimensions are $100 \times 100 \times 100\,\text{nm}^3$ and nm are the natural units used in this setup. The volume occupation by particles in this box is 30%, which means that there are 2357 A and 2357 B particles in the beginning. The timestep is $\tau = 10^{-10}$ s and in the ReaDDy paper the maximum time is $T = 10\,\mu\text{s} = 100,000 \times \tau$.

The main difference between our revreaddy setup and the ReaDDy setup is the integration scheme. In the style of equation (1.2) we formulate the ReaDDy integration scheme as

$$\boldsymbol{\psi}(t+\tau) = \mathcal{R}_0 \mathcal{D}_0 \boldsymbol{\psi}(t)$$

(for explanation of the symbols see table 5.1) and the revreaddy scheme reads

$$\boldsymbol{\psi}(t+\tau) = \mathcal{R}_{\text{acc}}\mathcal{R}_{\text{M}}\mathcal{D}_0\boldsymbol{\psi}(t)$$

where after every reaction step, an acceptance step is included. Note that we also modify the reaction step itself according to the constraint (5.14). In essence this modification means that we increase the reaction distance R to make up for the decreasing reaction volume when introducing particle repulsion. Thus this modification is only applied in the *with-repulsion* case. More information on determining R see section 5.2.2.

parameter	value
number of particles	2357- 4714
particle radius ρ_A	1.5 nm
particle radius ρ_B	3.0 nm
particle radius ρ_C	3.12 nm
diffusion coefficient D_A	143.1×10^6 nm^{2}s^{-1}
diffusion coefficient D_B	71.6×10^6 nm^{2}s^{-1}
diffusion coefficient D_C	68.82×10^6 nm^{2}s^{-1}
reaction distance (ReaDDy) R	4.5 nm
reaction distance (revreaddy w/ rep.) R	4.891 nm
reaction rate k_{on}	10^6 s^{-1}
reaction rate k_{off}	5×10^4 s^{-1}
macroscopic reaction rate $k_{\text{on,macro}}$	3.678×10^8 nm^3 s^{-1}
timestep τ	10^{-10} s
simulation length T	100,000 τ or 200,000 τ
periodic boundaries	NO
box volume	$100 \times 100 \times 100$nm^3
temperature k_BT	2.437 kJ mol^{-1} (20 °C)

Table 8.1: Parameters used in the simulation *Readdy benchmark*.

The main observable in this scenario is the number of particles of a certain species at a given time t. We will also compare these to the solution of a set of ordinary differential equations (ODE). The approach to these is given in the appendix. We will denote the concentrations of particles as $A(t)$, $B(t)$ and $C(t)$. Then the ODEs for this system read

$$\begin{aligned} \frac{\mathrm{d}B(t)}{\mathrm{d}t} = \frac{\mathrm{d}A(t)}{\mathrm{d}t} &= -k_{\text{on,macro}}A(t)^2 + k_{\text{off}}C(t) \\ \frac{\mathrm{d}C(t)}{\mathrm{d}t} &= +k_{\text{on,macro}}A(t)^2 - k_{\text{off}}C(t) \end{aligned} \tag{8.1}$$

Here we have already used the constraints that $A(t) = B(t)\ \forall t$. The macroscopic rate $k_{\text{on,macro}}$ is determined with equation (3.2) from the reaction distances of particles A and B and their diffusion coefficients and the microscopic reaction rate k_{on} [16]. The set

of ODEs (8.1) is naively integrated with an Euler-Maruyama method with the initial conditions

$$A(0) = B(0) = 3.9\,\mathrm{mM} \quad C(0) = 0.$$

The results of the simulations are seen in figure 8.1.[1] The ODE are described well by both methods ReaDDy and revreaddy in the case of no repulsion, although we can see that the revreaddy concentrations vary slightly from the ODE solution which is due to rejections of reaction steps at the boundary of the system, where energy terms arise. More interesting is the case when repulsion is switched on. In the ReaDDy case, this resulted in slower reaction kinetics and a different equilibrium ratio of particle concentrations. In the revreaddy method we can also observe slower reaction kinetics but the concentrations of particles eventually converge to the equilibrium concentrations that are also predicted by the ODE method.

Concluding, our new method seems to assure convergence to the ODE solution. If this is a wanted effect however depends on the model and the system under consideration.

8.2 Volume unconserved

We have seen our new method acting on a system with one bimolecular reaction. We will now look at the behavior, when we consider a system of two coupled reactions I and II

$$A + A \underset{k_{\mathrm{off,I}}}{\overset{k_{\mathrm{on,I},R}}{\rightleftharpoons}} B \underset{k_{\mathrm{off,II}}}{\overset{k_{\mathrm{on,II}}}{\rightleftharpoons}} C. \tag{8.2}$$

All rates are microscopic. We design the system such that in the end, all species have the same concentrations in an ODE description.

$$A(t \to \infty) = B(t \to \infty) = C(t \to \infty) \tag{8.3}$$

This is accomplished by matching the effective, macroscopic rates, i.e.

$$k_{\mathrm{on,macro,I}} \times A(t \to \infty) = k_{\mathrm{off,I}} = k_{\mathrm{on,II}} = k_{\mathrm{off,II}} \tag{8.4}$$

where the macroscopic reaction rate $k_{\mathrm{on,macro,I}}$ can be determined by equation (3.2) that describes the case of no interaction in between A particles. We will design the parameters of particles such that the total volume of particles during the reaction I is not conserved in the following sense

$$2v_A = v_C \quad v_B > v_C \tag{8.5}$$

Note that the ODE approach cannot capture this detail and we will expect deviation. We will also perform this simulation with volume conservation. We will investigate the three cases

[1] Note the deviation of the initial concentrations. It occurs that the scaling of the concentration was miscalculated in the ReaDDy paper (or in this work). However the author could not resolve this issue, but it is only a scaling factor. The qualitative results are the same.

- no interaction (w/o rep.)
- repulsion - volume conserved (w/ rep. vol. cons.)
- repulsion - volume not conserved (w/ rep. vol. uncons.)

Let us find out how to fulfill the condition (8.4). The problem reduces to finding the relation between the on-rate $k_{\text{on,macro,I}}$, the off-rate $k_{\text{off,I}}$ and the equilibrium concentration of $A(t)$. The system of ODEs reads

$$\frac{\mathrm{d}A(t)}{\mathrm{d}t} = -k_{\text{on,macro,I}}\, A(t)^2 + k_{\text{off,I}}\, B(t) \tag{8.6}$$

$$\frac{\mathrm{d}B(t)}{\mathrm{d}t} = +k_{\text{on,macro,I}}\, A(t)^2 - k_{\text{off,I}}\, B(t) - k_{\text{on,II}}\, B(t) + k_{\text{off,II}}\, C(t) \tag{8.7}$$

$$\frac{\mathrm{d}C(t)}{\mathrm{d}t} = +k_{\text{on,II}}\, B(t) - k_{\text{off,II}}\, C(t) \tag{8.8}$$

At equilibrium or for $t \to \infty$ we have a stationary situation

$$\left.\frac{\mathrm{d}A(t)}{\mathrm{d}t}\right|_{t\to\infty} = \left.\frac{\mathrm{d}B(t)}{\mathrm{d}t}\right|_{t\to\infty} = \left.\frac{\mathrm{d}C(t)}{\mathrm{d}t}\right|_{t\to\infty} = 0$$

Utilizing our constraint (8.3), the rate equation (8.6) becomes

$$0 = -k_{\text{on,macro,I}}\, A(t\to\infty)^2 + k_{\text{off,I}}\, A(t\to\infty).$$

We make use of the conservation of particles when starting only with A particles

$$A(t) + B(t) + C(t) = A(0) \forall t$$
$$\Rightarrow A(t\to\infty) = \frac{A(0)}{3}$$

We can find the rate $k_{\text{off,I}}$ as

$$k_{\text{off,I}} = \frac{A(0)}{3} k_{\text{on,macro,I}} \tag{8.9}$$

The procedure of setting the rates is then the following:

- Set particle radii, reaction distance and diffusion coefficients in the no-interaction case
- Set the microscopic on-rate of reaction I
- Find the macroscopic on-rate of reaction I from equation (3.2)
- Use equation (8.9) to get the backward rate
- Set the remaining rates equal to that, see equation (8.4)

The ODEs are then again solved with an Euler-Maruyama integration, with the initial condition $A(0) = 0.00896$ (in rescaled units).

All values in the revreaddy simulation are in rescaled units that correspond to similar conditions as the scenario before (*ReaDDy benchmark*). These conditions are similar to cellular cytosol with 30 % volume occupation when only A and C particles are present. Rescaled units means that we measure lengths in nm, time in ns and energies in k_BT at room temperature, i.e. $2.437\,\mathrm{kJ\,mol^{-1}}$. All parameters in rescaled units are given in table 8.2.

parameter	value
lengthscale	10^{-9} m
timescale	10^{-9} s
energyscale	$k_BT = 2.437\,\mathrm{kJ\,mol^{-1}}$
particle radii ρ_A	2
particle radii ρ_B	2.7501 (vol. uncons.)
particle radii ρ_B	2.5198 (vol. cons.)
particle radii ρ_C	2.5198
diffusion coefficient D_A	0.10734
diffusion coefficient D_B	0.07806
diffusion coefficient D_C	0.08520
reaction distance without repulsion R	4
reaction distance with repulsion R	4.391
reaction rate $k_{\mathrm{on,I}}$	0.001
reaction rate $k_{\mathrm{on,macro,I}}$	0.26032
reaction rate $k_{\mathrm{off,I}}$	7.774×10^{-4}
reaction rates	$k_{\mathrm{off,I}} = k_{\mathrm{on,II}} = k_{\mathrm{off,II}}$
periodic box dimensions	$V = 50 \times 50 \times 50$
initial A particles	1120
initial density	0.00896

Table 8.2: Parameters used in the simulation *volume unconserved*.

The revreaddy simulation procedure is as follows. We start initially with 1120 A particles in the simulation box without any reactions. According to which case is simulated (w/o rep., w/ rep. vol. uncons. or w/ rep. vol. cons.), the system is equilibrated via pure Brownian Dynamics. If the system is equilibrated is checked with the radial distribution function $g(r)$ of the particles. Also the mean squared displacement is observed to check whether the systems behave as expected. These observables are given in figure 8.2. We can see that in the case of no repulsion the particles behave like free particles, as is expected without any interactions. The radial distribution shows that particle distances are uncorrelated and the mean squared displacement reproduces the exact solution of the driftless diffusion equation that was given in section 6. The case with repulsion shows different behavior. From the radial distribution function it is clear that particles are rarely overlapping. It is also evident that the probability of finding a particle at the point where

particles start to overlap is largest (see figure 8.2, dashed red line in the RDF plot). Also the mean squared displacement shows some differences in the case of interaction. We can see ballistic diffusion in the first few timesteps and then a transition region (probably power law behavior), while in the end the exponent seems to be quadratic again. This is in agreement with Höfling et al. [29], who studied diffusion in crowded environments.

After this initial equilibration phase the reactions are switched on and the systems propagate according to

$$\boldsymbol{\psi}(t+\tau) = \mathcal{R}_{\mathrm{acc}}\mathcal{R}_{\mathrm{M}}\mathcal{D}_0\boldsymbol{\psi}(t)$$

(see table 5.1 for reference). During this phase the number of particles are recorded as a function of time. The results of this procedure are given in figure 8.3. In the case without particle repulsion, the concentration as a function of time of the revreaddy simulation exactly represent the ODE solution. This is expected because no interaction also means that no energy terms arise that might decrease the acceptance probability during the reaction step. The center graph in figure 8.3 shows the case of particle repulsion where radii of particles A, B and C are chosen so that total volume is always conserved. The reaction kinetics are slowed slightly. This can be explained by the obstruction that two potential reaction partners face in a crowded environment. Also the equilibrium values of concentrations are shifted. Most prominent for the A particles, the concentration deviates about 20% from its ODE value. This effect might be explained by the dissociation step $B \to A + A$ being rejected more often than the association step $A + A \to B$ because the dissociation certainly leads to more overlap of particles and thus more energetic penalties than the association step, which most often is accompanied by a decrease in energy which is favorable. For B and C the concentrations vary only little. In general this case still reproduces the overall ODE behavior while incorporating realistic crowding effects as well. As in the scenario before this behavior might be wanted or not (further discussion in the conclusion). In the third case we have particle repulsion and the particle B is larger than C, making the state B an improbable one by construction. This is reflected in the results as well. Due to rejection of steps, the whole process is slowed even further. As expected the state B is less populated than the other two. In this case, also C is more favorable than A.

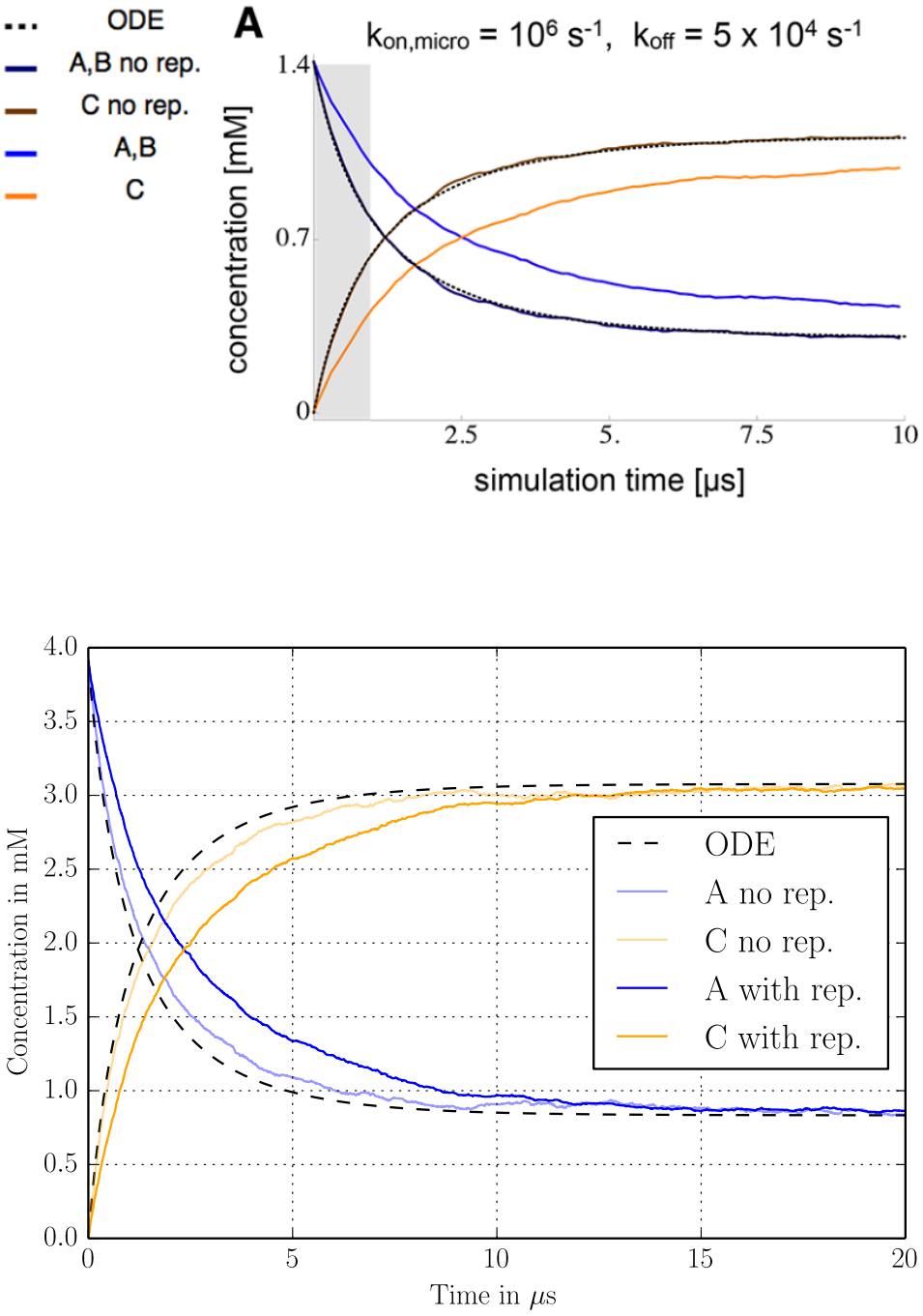


Figure 8.1: Results of the simulation of the bimolecular reaction $A + B \underset{k_{\text{off}}}{\overset{k_{\text{on}},R}{\rightleftharpoons}} C$, namely the concentrations of particles. Top: ReaDDy. Bottom: revreaddy, with a modified reaction scheme. Note the different simulation lengths.

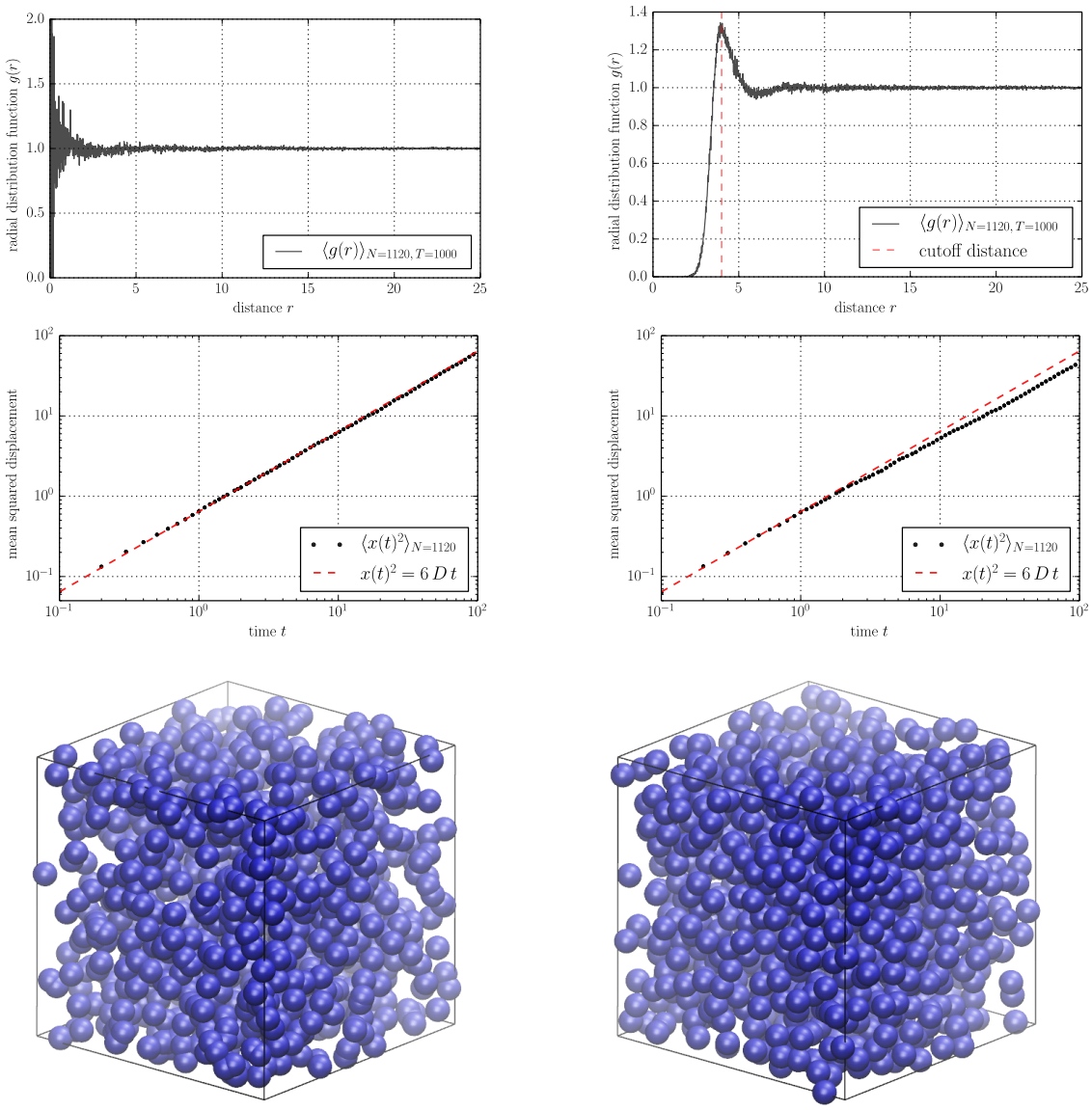


Figure 8.2: Preparation of the scenario *volume unconserved*, with only *A* particles present. Shown are: The radial distribution function averaged over all 1120 particles and 100 timesteps. The mean squared displacement as a function of time averaged over all 1120 particles. A snapshot of the system after equilibration. Left: without particle repulsion. Right: with soft/harmonic repulsion.

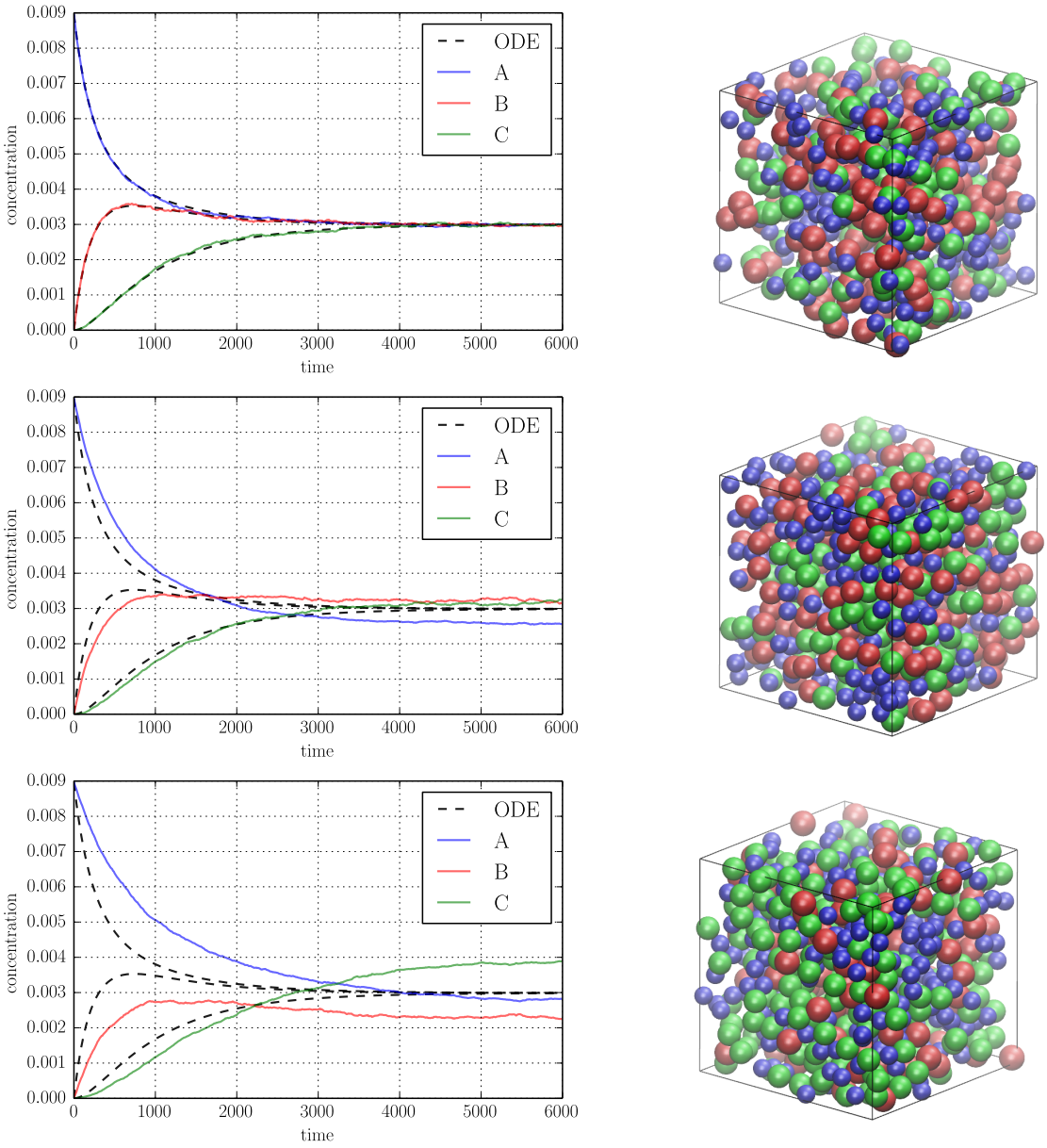


Figure 8.3: Results of the scenario *volume unconserved*. Reactions were performed with the acceptance/rejection scheme. Dynamics are purely Brownian Dynamics. The whole simulation was performed ten times to average the concentrations over a larger ensemble. Shown are particle concentrations as a function of time. Top: The case without particle repulsion. Mid: With repulsion and conserved volume. Bottom: With repulsion but unconserved volume.

9 Conclusion & Outlook

This thesis deals with the method of stochastic particle based reaction-diffusion simulations. More precisely, we applied an alternative scheme to spatiotemporal simulations like the framework of ReaDDy. The main motivation was to enforce detailed balance or local detailed balance to the integration scheme and thus make the simulation more realistic. Especially when considering systems of few particle numbers or in the absence of global equilibrium, care must be taken not to bias the system by "wrong" elementary operations (diffusion and reactions in our case). The method to enforce local detailed balance was to introduce an acceptance step in the style of Metropolis and Hastings. This model was implemented[1] in a software called *revreaddy* and compared to the ReaDDy results. We investigated effects of our method on pure dynamics without reactions and then investigated the effects on reactions with standard Brownian Dynamics.[2]

The effect of our method on dynamics showed that we get slowed down systems but correct sampling of the energetic landscape. In terms of local detailed balance this means that a state of maximum entropy is approached during the course of the simulation. For reactions the results are similar. We get slower reaction kinetics but the equilibrium concentrations of particles show reasonable behavior in comparison to solutions of ordinary differential equations that neglect space. If the behavior of approaching an ODE solution is desirable is left to the modeler. It might prove useful when one knows about the macroscopic kinetics and wants to incorporate spatial effects such as crowding or obstructed diffusion in the model.

The correct "speed" of a system AND the correct sampling can only be satisfied in the limit of the acceptance $\alpha \to 1$. This can always be approached by making the timestep small enough, but this becomes infeasible, mostly for one reason.

The acceptance probability α is an extensive variable of the underlying system. This is iilustrated in figure 9.1 on the left. Assume an equilibrated system (blue box) that can be well described by a set of quantities such as the volume V, number of particles N, entropy S, energy Φ, temperature T, pressure P and chemical potential μ. Now consider a system which is 8 times "bigger". This means that the extensive variables N, V, S, Φ also take new values which are 8 times larger. On the other hand the intensive variables T, P, μ remain the same as before. The acceptance probability is as we have seen in section 7.2 largely dependent on the timestep τ, which is a model parameter, but it is also dependent on the number of particles N and the volume V. This is what makes it hard to observe

[1] The implementation of the method was a large part throughout the research phase.

[2] A combined application was not performed as it lead to dynamically frustrated systems just when the reactions were becoming important with respect to system size. This problem is further discussed in the text.

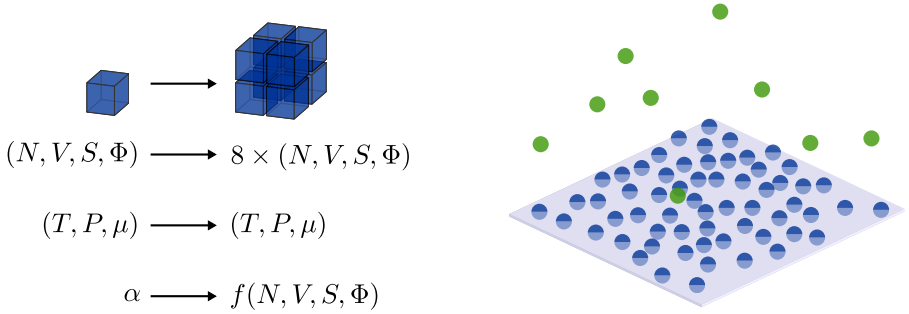


Figure 9.1: Left: The acceptance α is an extensive variable. Right: Two subsystems, membrane bound and free particles, together in one simulation.

large systems with such a method. The exact form of how α scales with the extensive variables is denoted by f. An interesting future task would be to find out what this f is.

Another issue can be illustrated with the help of figure 9.1 on the right. Assume a system with two types of particles. One bound to a membrane and very crowded. The other freely diffusing in space. When one calculates the acceptance α in such a system and propagates particles as we have done it, the "speed" with which the free particles diffuse will highly depend on the acceptance, which is dominated by the effects of the dense membrane system. But in the model that we have described so far the free particles do not even interact with the membrane system. So why should the free particles be influenced by the membrane particles? An integration scheme that resolves this paradoxon and still satisfies local detailed balance is needed. This would also be an interesting question to pose.

References

[1] Johannes Schöneberg, Alexander Ullrich, and Frank Noé. "Simulation tools for particle-based reaction-diffusion dynamics in continuous space". In: *BMC Biophysics* 7 (2014), pp. 1–10.

[2] M K Knowles et al. "Single secretory granules of live cells recruit syntaxin-1 and synaptosomal associated protein 25 (SNAP-25) in large copy numbers." In: *Proceedings of the National Academy of Sciences of the United States of America* 107 (2010), pp. 20810–20815.

[3] Isabel Pastor et al. "Effect of crowding by Dextrans in enzymatic reactions". In: *Biophysical Chemistry* 185 (2014), pp. 8–13.

[4] Johannes Schöneberg and Frank Noé. "ReaDDy–a software for particle-based reaction-diffusion dynamics in crowded cellular environments." In: *PloS one* 8.9 (2013), e74261.

[5] Chris Sanford et al. "Cell++ - Simulating biochemical pathways". In: *Bioinformatics* 22.23 (2006), pp. 2918–2925.

[6] Steven S. Andrews et al. "Detailed simulations of cell biology with Smoldyn 2.1". In: *PLoS Computational Biology* 6.3 (2010).

[7] Jeroen S. van Zon and Pieter Rein ten Wolde. "Green's-function reaction dynamics: a particle-based approach for simulating biochemical networks in time and space." In: *The Journal of chemical physics* 123.2005 (2005), p. 234910.

[8] Monika Gunkel et al. "Higher-Order Architecture of Rhodopsin in Intact Photoreceptors and Its Implication for Phototransduction Kinetics". In: *Structure* 23.4 (Apr. 2015), pp. 628–638.

[9] Alexander Ullrich et al. "Dynamical Organization of Syntaxin-1A at the Presynaptic Active Zone". In: *PLOS Computational Biology* 11.9 (Sept. 2015). Ed. by Kim T. Blackwell, e1004407.

[10] P. J. Rossky, J. D. Doll, and H. L. Friedman. "Brownian dynamics as smart Monte Carlo simulation". In: *The Journal of Chemical Physics* 69.1978 (1978), p. 4628.

[11] Nicholas Metropolis et al. "Equation of State Calculations by Fast Computing Machines". In: *The Journal of Chemical Physics* 21.1953 (1953), pp. 1087–1092. arXiv: 5744249209.

[12] M.P. Allen and D.J. Tildesley. *Computer Simulation of Liquids*. New York: Oxford University Press, 1987.

[13] W Im, S Seefeld, and B Roux. "A Grand Canonical Monte Carlo-Brownian dynamics algorithm for simulating ion channels." In: *Biophysical journal* 79.August (2000), pp. 788–801.

[14] Marco J. Morelli and Pieter Rein ten Wolde. "Reaction Brownian dynamics and the effect of spatial fluctuations on the gain of a push-pull network". In: *The Journal of Chemical Physics* 129.5 (2008), p. 054112.

[15] W. K. Hastings. "Monte carlo sampling methods using Markov chains and their applications". In: *Biometrika* 57.1 (1970), pp. 97–109.

[16] Radek Erban and S Jonathan Chapman. "Stochastic modelling of reaction-diffusion processes: algorithms for bimolecular reactions." In: *Physical biology* 6.4 (2009), p. 046001. arXiv: 0903.1298.

[17] Don S Lemons. "Paul Langevin's 1908 paper "On the Theory of Brownian Motion" ["Sur la théorie du mouvement brownien," C. R. Acad. Sci. (Paris) 146, 530–533 (1908)]". In: *American Journal of Physics* 65.11 (1997), p. 1079.

[18] Nawaf Bou-Rabee and Eric Vanden-Eijnden. "Pathwise accuracy and ergodicity of Metropolized integrators for SDEs". In: *Communications on Pure and Applied Mathematics* 63 (2010), pp. 655–696. arXiv: 0905.4218.

[19] Juan C. Latorre et al. "A structure-preserving numerical discretization of reversible diffusions". In: *Communications in Mathematical Sciences* 9 (2011), pp. 1051–1072.

[20] Hong Qian. "Phosphorylation Energy Hypothesis: Open Chemical Systems and Their Biological Functions". In: *Annual Review of Physical Chemistry* 58.1 (May 2007), pp. 113–142.

[21] Michel Bauer and Françoise Cornu. "Local detailed balance: a microscopic derivation". In: *Journal of Physics A: Mathematical and Theoretical* 48.1 (Jan. 2015), p. 015008. arXiv: 1412.8179.

[22] Bernard Derrida. "Non-equilibrium steady states: fluctuations and large deviations of the density and of the current". In: *Journal of Statistical Mechanics: Theory and Experiment* 2007.07 (July 2007), P07023–P07023.

[23] Sheldon Katz, Joel L. Lebowitz, and H. Spohn. "Phase transitions in stationary nonequilibrium states of model lattice systems". In: *Physical Review B* 28.3 (Aug. 1983), pp. 1655–1658.

[24] Franz Schwabl. *Statistische Mechanik*. Springer-Lehrbuch. Berlin/Heidelberg: Springer-Verlag, 2006.

[25] H.A. Kramers. "Brownian motion in a field of force and the diffusion model of chemical reactions". In: *Physica* 7.4 (Apr. 1940), pp. 284–304.

[26] https://github.com/chrisfroe/revreaddy. [Online; accessed 2015-11-03].

[27] https://www.hdfgroup.org/HDF5/. [Online; accessed 2015-11-03].

[28] http://www.h5py.org/. [Online; accessed 2015-11-03].

[29] Felix Höfling and Thomas Franosch. “Anomalous transport in the crowded world of biological cells”. In: *arXiv* (2013), pp. 1–55. arXiv: 1301.6990.

Appendix

Rescaled units

For generality and simplicity most of the simulation examples are given in rescaled (dimensionless) units. The results can be applied to different situations recovering the real units with some constraint, that will be given here.

The goal is to reduce the integrator (4.7) to a unitless, normalized form to fully utilize machine precision. We also want to keep D as a per-particle property and not eliminate via the Einstein-Smoluchowski relation. We state it here again

$$x_{t+\tau} = x_t - \tau \frac{D}{k_B T} \nabla U(x_t) + R_t. \tag{4.7}$$

Also the variance of the noise term

$$\langle R_t R_{t'} \rangle = 2D\tau \, \delta_{tt'} \tag{4.8}$$

will be affected. We start by defining the new dimensionless quantities

$$\tilde{U} = U/k_B T \quad \tilde{x} = x/h \quad \tilde{t} = t/\lambda \tag{9.1}$$

This also means

$$\tilde{\nabla} = \nabla h \qquad \tilde{\tau} = \tau/\lambda \qquad \tilde{D} = D\frac{\lambda}{h^2} \tag{9.2}$$

Putting this all in (4.7) we get

$$\tilde{x}_{t+\tau} = \tilde{x}_t - \tilde{\tau}\tilde{D}\tilde{\nabla}\tilde{U}(x_t) + \frac{R_t}{h} \tag{9.3}$$

We further define the new rescaled random displacement

$$\tilde{R}_t = R_t/h, \tag{9.4}$$

which alters the second moment

$$\langle \tilde{R}_t \tilde{R}_{t'} \rangle = 2\tilde{\tau}\tilde{D}\delta_{tt'} \tag{9.5}$$

Summarizing we get, when leaving out the tildes, our rescaled integration equation

$$x_{t+\tau} = x_t - \tau D \nabla U(x_t) + R_t \quad \text{with} \quad \langle R_t R_{t'} \rangle = 2\tau D \, \delta_{tt'} \tag{9.6}$$

Its form has not changed a lot except that we do not need to compute $k_B T$ anymore. To recover the real world units, the formulas (9.1) and (9.2) must be used.

ODE description

In some applications we compare the results of the spatio-temporal *revreaddy* algorithm with a description by ordinary differential equations (ODE). This is based on equation (1.1) which reads

$$\frac{\partial \mathbf{u}(\mathbf{x},t)}{\partial t} = \mathbf{D}\Delta\mathbf{u}(\mathbf{x},t) + f(\mathbf{u}), \tag{1.1}$$

For the simplest and most well behaved reaction systems and for non-interacting species we can neglect the diffusion term and are left with

$$\frac{\mathrm{d}\mathbf{u}(t)}{\mathrm{d}t} = f(\mathbf{u}) \tag{9.7}$$

If $\mathbf{u}(t) = [u_A(t), u_B(t), u_C(t), \dots]$ are the individual concentrations of the species A, B and C, then $f(\mathbf{u})$ contains the change of concentration that corresponds to a set of reactions. For example, consider the chemical reaction system

$$\varnothing \xrightarrow{k_+} A \qquad\qquad A \xrightarrow{k_-} \varnothing$$

with reaction rates k_+ and k_-. In this case equation (9.7) becomes

$$\frac{\mathrm{d}u_A(t)}{\mathrm{d}t} = k_+ - k_- u_A(t)$$

with the solution

$$u_A(t) = u_A(0)\, e^{-k_- t} + \frac{k_+}{k_-}\left(1 - e^{-k_- t}\right)$$

Note the different units of the reaction rates k_+ and k_-.

In the present thesis we investigate systems which include unimolecular and bimolecular reactions. In these cases the set of equations can become more complicated. To obtain the solution however we integrate equation (9.7) by means of an Euler-Maruyama scheme.

$$\begin{aligned} \frac{\mathrm{d}\mathbf{u}(t)}{\mathrm{d}t} &\approx \frac{\mathbf{u}(t+1) - \mathbf{u}(t)}{\Delta t} \\ \Rightarrow\ \mathbf{u}(t+1) &= \mathbf{u}(t) + \Delta t \cdot f(\mathbf{u}(t)) \end{aligned} \tag{9.8}$$