

Master Thesis
Jan Grzegorzewski

Reaction-Diffusion Dynamics With Fractional Brownian Motion

Winter term: 2016

Contents

1	Fractional Brownian Motion	3
1.1	Brownian Motion	3
1.2	Fractional Brownian Motion	5
1.3	Algorithm	9
2	Particle Based Reaction Diffusion	22
2.1	Diffusion Influenced Bimolecular Reaction	23
2.2	RevReaddy with fBm	28
3	An Enzymatic Reaction With Fractional Brownian Motion	30
3.1	Michaelis-Menten (MM) kinetics	30
3.2	Simulation Model	32
3.3	Results for Normal Diffusion	32
3.3.1	Results for One-way Reactions	36
3.3.2	Results for the MM mechanism	40
3.4	Results for fBm	42
3.5	Conclusion	43
4	Summary	44
5	Appendix	46
5.1	From Central Limit Theorem to Gaussian Distribution	46
5.2	From Gaussian Distribution to Gaussian Transition Probability	47
5.3	Einstein Formula	47
5.4	Erban-Chapman	48
5.5	Michaelis-Menten Kinetics	49

Motivation

Anomalous diffusion can be observed in many different areas of nature, in particular, related to heterogeneous media like porous rocks, gels and crowded biological media with its prominent phenomena macromolecular crowding, confinement and macromolecular adsorption [1]. These environments exhibit remarkable characteristics like anomalous diffusion with its most popular power-law behaviour of the mean-square-displacement ($\delta r^2(t) \propto t^\alpha$) [2], which violates the Einstein formula $\delta r^2(t) = 2dDt$ and thereby the central limit theorem. Anomalous diffusion is described by various theoretical models like the time random walks (CTRW) with an unconventional waiting time-distribution, the Lorentz models with excluded volume and a highly ramified remaining space or fractional Brownian motion (fBm). The latter is modelling a stochastic process with long-range correlations of the increments. fBm was first introduced as a family of Gaussian random functions by Mandelbrot and Van Ness in 1968 and motivated by examples in economics [3]. In contrast to different models of anomalous diffusion, the fBm approach is plainly phenomenologically defined by a power-law increase of the MSD and thus perfectly qualifies as a starting point to study effects arising from a power-law increase of the MSD.

This thesis focuses on the impact of fBm on enzymatic reaction kinetics in three-dimension space. Enzymes are large biological molecules, mostly proteins. Metabolic processes in cells rely vitally on enzymes. They help to break down large nutrient molecules like carbohydrates, fats and proteins during in the process of digestion. Other enzymes contrariwise help to form large and complex molecules from smaller once. Enzymes are involved in storage and release of energy, processes of respiration, vision, muscle contraction and transmission of nerve impulses. In fact, biochemical reactions are controlled mainly by enzymes. In general, they act as catalysts by lowering the activation energy. As stated by the Arrhenius law the activation energy is then increasing the reaction rates. The speed of reactions is typically increased by a factor of $10^6 - 10^{14}$. Pioneer work by Leonor Michaelis and Maud Menten [4] simplified enzymatic reaction kinetics and formulated a model based on the law of mass action. As one of the best known and important models of enzyme kinetics, it is of great interest to study effects from anomalous diffusion on Michaelis-Menten like reactions. For this purpose a particle-based simulation with an fBm integrator of a Michaelis-Menten like reactions was set up. Spatial and temporal effects induced by fBm were studied.

The thesis is organised into four chapters: The first chapter sets up theoretical foundations for fBm. Subsequently, fBm generating algorithms are studied. Chapter 2 deals with models describing reactions in general but focuses on particle-based reaction diffusion. A particle-based reaction diffusion software RevReaDDy, which served as a starting point for the implementation is introduced. Chapter 3 focuses

more specifically on enzyme reactions. Followed by an enzymatic simulation model set up with RevReaDDy. Finally, results of the simulation are discussed and related to existing literature. Chapter 4 summarises the content of the thesis.

1 Fractional Brownian Motion

The Wiener process is a continuous-time stochastic process, which models white noise. It is applied to finance, biology, physics and much more if no or only weak correlations of the underlying processes are present. Brownian motion is the random motion of particles suspended in a fluid, which is modelled by a Wiener process. Velocities of particles, however, are always correlated at least for very short lag times in the physical world. It is often referred to as the super-diffusive regime. A 1μ silica bead in water for examples forgets its initial velocities in about $0.3ns$ and the corresponding displacement of about $1pm$ [5]. For longer times Brownian motion can be assumed.

Fractional Brownian Motion (fBm) is a more general family of stochastic Gaussian processes than standard Brownian motion with long-range correlations as its defining property. The main objective of this chapter is to explore the theoretical foundation for fBm starting with Brownian motion in the following section. Further algorithms generating fBm are introduced and analysed in terms of accuracy and performance.

1.1 Brownian Motion

Standard Brownian Motion (Bm) is a very important and well studied stochastic process. It describes the erratic motion of mesoscopic particles, which first were documented by Jan Ingenhousz in 1785, in particular for coal dust on the surface of alcohol [2]. Later on, in 1827 Robert Brown observed the erratic motion of pollen grains. Brownian motion has a Gaussian propagator, which has its origin in the Central Limit Theorem (CLT) for a sum of independent and identically distributed random variables (i.i.d.). Let's assume a set of N independent variables $\{X_j\}$ with a finite variance $\sigma_j^2 = \langle X_j^2 \rangle$ and the mean $\langle X_j \rangle = 0$. The definition of another random variable Y is given by:

$$Y = \frac{1}{\sqrt{N}} \sum_{j=1}^N X_j \quad (1.1)$$

This scenario in which a random variable is defined by the sum of another can be observed generically in nature. X_j are called increments of Y . The seemingly innocent assumption of independent increments result in a Gaussian distribution of $\rho(y)$ in the limit of large N with $\rho(y)dy = P(y < Y < y + dy)$

$$\rho(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{y^2}{2\sigma^2}} \quad (1.2)$$

A derivation of this very interesting result can be found in the appendix 5.1. Microscopic processes with i.i.d. velocities, thus have a Gaussian distribution for the overall change in position. Indeed, even the increments of Brownian motion are distributed Gaussian as a result of i.i.d. velocities at shorter time scales. A random walk in the limit of very small time steps, therefore, converges to Brownian motion.

With Bayes' theorem and an initial delta-distribution one can show that transition probability $T_t(y|0) = \rho_t(y)$ is equivalent to the particle density distribution. One can find the calculation in the appendix section 5.2. The above-mentioned elaborations motivate the reason for a process with a Gaussian transition probability. A process with Gaussian distributed transition probability and independent increments is called Wiener process (Standard Brownian motion).

Brownian motion described by the Wiener process is a stochastic process $\{W_t\} : \Omega \rightarrow \mathbb{R}^d$ with $W_t(\omega)$ being the position of a particle with $\omega \in \Omega$ at time $t \in T$ in the observation time $T = [0, \infty)$. It has a fixed $x \in \mathbb{R}^d$ as its origin. The transition probabilities are [6]:

$$T_t(y|x) := (2\pi t)^{-\frac{d}{2}} \exp\left(-\frac{\|x-y\|^2}{2t}\right) \text{ for } y \in \mathbb{R}^d, t > 0 \quad (1.3)$$

$$T_0(y|x) = \delta(x-y)$$

The Wiener process is a Gaussian process with mean $\langle W_t \rangle_y = x$ and particle position $W_0 = x$ at $t = 0$. Its variance is $\langle W_t^2 \rangle_y = t$. It has stationary, independent, Gaussian increments. Following from the definition, Brownian motion has the property called Brownian scaling:

$$\{\hat{W}_t := \frac{1}{c} W_{c^2 t}\}_{t \geq 0} \quad (1.4)$$

A Wiener process has self-similar and fractal paths as a result from Brownian scaling.

However, it is a purely mathematical model with a missing connection to the strength of diffusion of a particle. The process, in fact, is a normalised diffusion process with the diffusion constant $D = \frac{1}{2}$. A mathematical description of diffusion was first realised by Fick in 1855 [7]. He applied partial differential equations for heat transfer first solved by Fourier in 1822 [8]. Fick's second law of diffusion describes how diffusion causes the concentration ($c(\mathbf{r}, t)$) to change over time:

$$\frac{\partial}{\partial t} c(\mathbf{r}, t) = -\nabla J(\mathbf{r}, t) = D \Delta c(\mathbf{r}, t) \quad \text{with} \quad \Delta = \nabla^2 \quad (1.5)$$

With Fick's first law $J(\mathbf{r}, t) = -\nabla c(\mathbf{r}, t)$ and D being the flux of particles and the diffusion coefficient, respectively. The notation of bold letters is chosen to describe vectors (e.g. \mathbf{r}). Fick's first law is a result from the linear response theory. Fick's second law can be derived from mass conservation and Fick's first law. The most general form of mass conversation with a production term σ is described by the

continuity equation:

$$\frac{\partial}{\partial t} c(\mathbf{r}, t) = -\nabla J(\mathbf{r}, t) + \sigma \quad (1.6)$$

Albert Einstein in 1905 proposed a similar equation to Fick's second law of diffusion to solve one-dimensional Brownian motion [9]. This was the first derivation and application of probabilistic stochastic theory. Derived from the kinetic gas theory and a statistical description collisions in 1906 Marian von Smoluchowski could give similar results to Einstein [10]. The concentration $c(\mathbf{r}, t)$ is proportional to the transition probability $P(\mathbf{r}, t) \propto c(\mathbf{r}, t)$ and $\int d\mathbf{r} P(\mathbf{r}, t) = 1$. The transition probability of the Wiener process builds the foundation for a more physical description of particle motion by a propagator:

$$P(r, t) = (4\pi Dt)^{-d/2} \exp\left(-\frac{r^2}{4Dt}\right) \quad (1.7)$$

The propagator $P(r, t)$ can be interpreted as the probability to move the distance r during the time t . d is the dimension of the space of \mathbf{r} . One can show that this propagator is, in fact, solving the diffusion equation eq. (1.5) with the meaningful initial condition of vanishing concentration at boundaries:

$$c(\pm\infty, t) = 0 \quad (1.8)$$

The calculation can be found in the appendix 5.3. The most important property of Brownian motion in the context of this thesis is the mean square displacement ($\delta r^2(t)$), which grows linearly with the number of steps:

$$\delta r^2(t) = \langle [\mathbf{R}(t) - \mathbf{R}(0)]^2 \rangle = 2dDt \quad (1.9)$$

Here $\mathbf{R}(t)$ is the position of the particle at time t . Brownian scaling was mentioned as a property of a Wiener process. It obviously also applies as a property to the propagator of Brownian motion. A scale-free form of the Gaussian propagator exists as a more intuitive consequence of Brownian scaling:

$$P(r, t) = r^{-d} \mathcal{P}_{gauss}(\hat{r}) \quad (1.10)$$

$$\text{with } \hat{r} = \frac{r}{\sqrt{2Dt}} \quad \text{and} \quad \mathcal{P}_{gauss}(\hat{r}) = (2\pi)^{-d/2} \hat{r}^d \exp\left(-\frac{\hat{r}^2}{2}\right) \quad (1.11)$$

The scale free form of Bm can later on be compared to the scale free form of fractional Brownian motion, which will be introduced next. This property results in a neat possibility of analysis of Bm and fBm generating algorithms later on.

1.2 Fractional Brownian Motion

In the previous section, the MSD has been shown to be linear with time as a result of the central limit theorem. In normal liquids, this behaviour can be seen already at

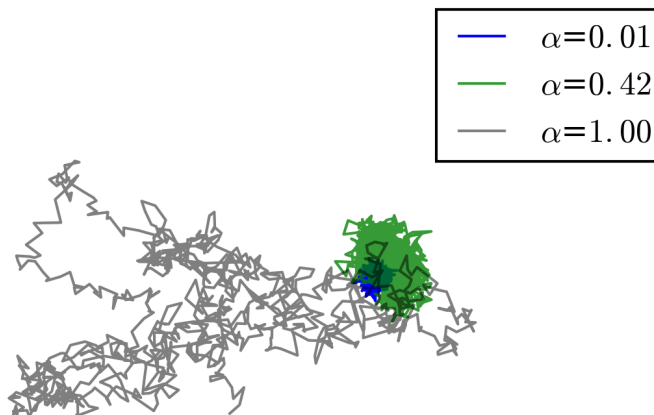


Figure 1.1: A three dimensional fBm trajectories for the anomalous coefficient $0.01 \leq \alpha \leq 1$ generated with the Lowen algorithm are shown. The general diffusion constant, the time step and trajectory length were chosen as follows: $K_\alpha = 1$, $\Delta t = 1$, $M = 1000$.

time scales higher than picoseconds [2]. Nevertheless, many experiments show that the MSD has a power law behaviour ($\delta r^2(t) \propto t^\alpha$ for $0 < \alpha < 1$) over certain time windows from microseconds to even seconds. Thus, the central limit theorem does not hold, at least in these time scales. It can be shown that persistent correlations of the increments are present. In soft matter, like polymers, a sub-diffusive behaviour is typically present in a time window but finally, the linear MSD takes over. Fractional Brownian motion idealises this situation by imposing that the central limit theorem is violated even for all time scales. The basic feature of fBm's is that the "span of interdependence" between their increments can be said to be infinite [3].

Just like the Wiener process, continuous-time Fractional Brownian motion (ctfBm) is a continuous-time Gaussian process $\{B_t^\alpha\} : \Omega \rightarrow \mathbb{R}^d$ at time $t \in T$ in the observation time $T = [0, \infty)$. It is fully specified by its mean $\langle B_t \rangle = 0$ and its covariance function:

$$\text{Cov}[B_t^\alpha, B_s^\alpha] = \frac{\sigma^2}{2} [t^\alpha - 2(s-t)^\alpha + s^\alpha] \quad \text{for} \quad t < s \quad (1.12)$$

With the mean and the covariance defined, the mean square displacement for fBm can be derived as:

$$\langle (B_t^\alpha - B_s^\alpha)^2 \rangle = (s-t)^\alpha \sigma^2 \quad (1.13)$$

Fractional Brownian motion exhibits a power law behaviour of the mean square displacement and therefore indeed motivates to act as a model for anomalous diffusion.

Fractional Brownian motion can be alternatively expressed in terms of its incremental sequence. The incremental sequence is a related stochastic process called continuous-time fractional Gaussian noise (ctfGn). CtfGn is a continuous-time

Gaussian process $\{X_t^\alpha\} : \Omega \rightarrow \mathbb{R}^d$ at time $t \in T$ in the observation time $T = [0, \infty)$. The mean is zero $\langle X_t^\alpha \rangle = 0$ and variance $\text{Var}[X_t^\alpha] = \langle (X_t^\alpha)^2 \rangle = \sigma^2$:

$$B_t^\alpha = \int_s^t X_{t'}^\alpha dt' + B_s^\alpha \quad \text{for} \quad t < s \quad (1.14)$$

The covariance function for ctFGn can be derived from the covariance function of fractional Brownian motion:

$$\text{Cov}[X_t^\alpha, X_s^\alpha] = \frac{d^2}{dt ds} \text{Cov}[B_t^\alpha, B_s^\alpha] = \sigma^2 \frac{d^2 |t - s|^\alpha}{dt ds} \quad (1.15)$$

$$= \alpha(\alpha - 1)\sigma^2 |t - s|^{\alpha-2} + \alpha\sigma^2 |t - s|^{\alpha-1} \delta(t - s) \quad (1.16)$$

The first term of this covariance function is describing the correlation of the increments. For Wiener white noise ($\alpha = 1$) the first terms disappears:

$$\text{Cov}[X_t, X_s] = \sigma^2 \delta(t - s) \quad (1.17)$$

The description of fractional Brownian motion in terms of fractional Gaussian noise is often convenient because of fGn's stationarity. It is possible to apply the Wiener-Khinchin theorem on a stationary random process. Thus, a spectral decomposition exists. The power spectral density can be written as the Fourier transform of the auto-covariance function:

$$S(\omega) = \int_{-\infty}^{\infty} \text{Cov}[X_t, X_0] \exp[-2\pi i \omega t] dt \quad (1.18)$$

The power spectral density for fGn for $\omega \rightarrow 0$ can be written as [2]:

$$S(\omega)_{fGn} = \sigma^2 \Gamma(1 + \alpha) (-i\omega)^{1-\alpha} \quad (1.19)$$

For the Wiener white noise ($\alpha = 1$), the result simplifies to a constant:

$$S(\omega)_{Gn} = \sigma^2 \quad (1.20)$$

For fBm generating algorithms, the discrete time version of fractional Brownian motion (fBm) and Gaussian noise (fGn) are interesting as well. The relation between them can be written as:

$$B_t^\alpha = \sum_{i=0}^k X_{t_i}^\alpha \quad \text{with} \quad t_i = t_0 + i\Delta t \quad (1.21)$$

The covariance function of discrete time fractional Gaussian noise is:

$$\text{Cov}[X_n^\alpha, X_m^\alpha] = \langle (B_{n+1}^\alpha - B_n^\alpha)(B_{m+1}^\alpha - B_m^\alpha) \rangle \quad (1.22)$$

$$= \frac{\Delta t^\alpha \sigma^2}{2} [(n-m-1)^\alpha - 2(n-m)^\alpha + (n-m+1)^\alpha] \quad (1.23)$$

$$\text{for } n \geq m \quad (1.24)$$

Due to stationarity one can write the auto-covariance function as:

$$\text{Cov}[X_0^\alpha, X_n^\alpha] = \frac{\Delta t^\alpha \sigma^2}{2} [(n-1)^\alpha - 2n^\alpha + (n+1)^\alpha] \quad (1.25)$$

A more detailed study of fBm and fGm can be found in [11].

Similar to the line of argument for Brownian motion, also fBm needs a connection to the strength of diffusion. This connection is introduced by the variance of fGm as:

$$\sigma^2 = 2dK_\alpha \quad (1.26)$$

The mean square displacement for fBm follows from eq. (1.13) as:

$$\delta r^2(t) = 2dK_\alpha t^\alpha \quad (1.27)$$

Here $K_\alpha > 0$ is the generalised diffusion coefficient of the physical dimension $\text{cm}^2/\text{sec}^\alpha$. The more physical description in terms of a propagator is the correlation function of the single particle density $P(\mathbf{r} - \mathbf{r}', t - t') = V \langle \rho(\mathbf{r}, t) \rho(\mathbf{r}', t') \rangle$. Here, the single particle density $\rho(\mathbf{r}, t) = \langle \delta(\mathbf{r} - \mathbf{R}(t)) \rangle$ describes the density of a particle, which is localised at position $\mathbf{R}(t)$. V is the volume. Just like Brownian motion, fractional Brownian motions increments $\Delta \mathbf{R}$ are Gaussian distributed with zero mean. The single particle density results in:

$$P(\mathbf{r}, t) = [4\pi K_\alpha t^\alpha]^{-d/2} \exp \left[\frac{-|\mathbf{r}|^2}{4K_\alpha t^\alpha} \right] \quad (1.28)$$

The propagator of fBm can be transformed into a scale-free form. A scale-free form will be relevant in the analysis of the algorithm. It is related to the scale-free form of standard Brownian motion eq. (1.11):

$$P(\mathbf{r}, t) = \mathbf{r}^{-d} \mathcal{P}_{\text{gauss}}(\hat{\mathbf{r}}) \quad (1.29)$$

$$\text{with } \hat{\mathbf{r}} = \frac{\mathbf{r}}{\sqrt{2K_\alpha t^\alpha}} \quad \text{and} \quad \mathcal{P}_{\text{gauss}}(\hat{\mathbf{r}}) = (2\pi)^{-d/2} \hat{\mathbf{r}}^d \exp \left(-\frac{|\hat{\mathbf{r}}|^2}{2} \right) \quad (1.30)$$

Velocities ($\xi(t)$) in the physical world can be described by fractional Brownian noise. The velocity auto-correlation function (VACF) ($Z(t)$) can be expressed from the

auto-covariance function of fGn as defined in eq. (1.15):

$$Z(|t - t'|) = \frac{1}{d} \langle \boldsymbol{\xi}(t) \bullet \boldsymbol{\xi}(t') \rangle = \frac{1}{2d} \frac{d^2}{dt^2} \delta r^2(t - t') \quad (1.31)$$

The frequency representation of the velocity auto-correlation function of fractional Brownian motion is defined analogously to eq. (1.19) as:

$$\tilde{Z}(\omega) = K_\alpha \Gamma(1 + \alpha) (i\omega)^{1-\alpha} \quad (1.32)$$

For the analysis of fBm generating algorithms an indicator, if the created increments indeed are Gaussian, will be of interesting. One can show that for Gaussian propagators with zero-mean all but the second cumulants vanish. For non-Gaussian transport also further cumulants are non-zero. The cumulants can be obtained by expanding the characteristic function of $\Delta R(t)$ for small wave numbers $k \rightarrow 0$. Its logarithm returns the cumulants. From the cumulants, the so-called non-Gaussian parameter can be defined [2]:

$$\alpha_2 = \frac{d\delta r^4(t)}{(d+2)[\delta r^2(t)]^2} - 1 \quad (1.33)$$

The non-Gaussian parameter shows non-Gaussian transport if $\alpha_2 > 0$. This, in turn, does not lead to sure Gaussian transport for $\alpha_2 > 0$. Cumulants of higher order still can be non-zero.

1.3 Algorithm

Effects of anomalous diffusion increase their relevance for long times. Thus, a fractional Brownian motion generating algorithm for preferably long trajectories is going to be searched for. Exactness and performance of the algorithm matter. In literature, several algorithms are described. Promising exact methods are: Cholesky [12], Hosking [13], Davis-Harte [12] and Lowen [14] method. Cholesky method claims to perform $\mathcal{O}(M^3)$ with the length of trajectory M and $\mathcal{O}(M^2)$ for every next trajectory. Davis-Harte and Lowen method claim to be exact and fast $\mathcal{O}(M \log(M))$ [15][14]. There are further algorithms with even faster performance (e.g. the $RMD_{3,3}$ -method [12] with $\mathcal{O}(M)$). However, they are approximations. In order to choose an algorithm, four algorithms were analysed in terms of accuracy and speed. 1. Cholesky 2. Davis-Harte 3. Lowen and 4. an own naive algorithm. All the methods are based on one or the other correlation functions described in the previous section.

Cholesky Method

The Cholesky method [12] is based on the auto-covariance function of fGn as described by eq. (1.12). The algorithm is shown in table 1.1. This approach is simple but as claimed by Ton [12] unfortunately not efficient with an algorithmic scaling of

1. Calculate the covariance matrix of fGn:

$$\Gamma = (\text{Cov}[\mathbf{x}_i, \mathbf{x}_j]) \quad , \quad (1.34)$$

$$\text{Cov}[\mathbf{x}_i, \mathbf{x}_j] = \Delta t^\alpha \sigma^2 (i^\alpha + j^\alpha - |i - j|^\alpha) / 2 \quad (1.35)$$

2. Calculate the eigenvalues λ and eigenvectors \mathbf{v} of Γ
3. Compute Σ the square root of Γ by the eigenvalues.
4. Draw for every entrance of the vector $\boldsymbol{\eta}$ a standard Gaussian distributed random number. Length of the vector is defined by the length M of the trajectory:

$$\eta_k = \mathcal{N}(0, 1) \quad , \quad k = 0, 1, 2, \dots, M \quad (1.36)$$

5. Calculate the fGn by $B_k^\alpha = \Sigma \boldsymbol{\eta}_k$.

Table 1.1: The table shows Choleskys fBm generating method.

$\mathcal{O}(M^3)$ with the trajectory length (M). Performance and accuracy of all algorithms are discussed in detail in the last two sections of the chapter.

A Naive Algorithms

In this section, a naive algorithm based on the VACF in the frequency domain is introduced. This was the first attempt of a fast fBm algorithm, without any insight from the existing literature. The idea is to compute the increments in the frequency domain and eventually transform them back to the time domain. The algorithm is described in table 1.2. In the first step of the algorithm, four times the number of increments in comparison to the trajectory length are generated. It counteracts a problem arising from the boundary problem of the FFT. The problem is visualised in fig. 1.2. The figure shows an ensemble-averaged mean square displacement for $\alpha = 0.5$ for two versions of the algorithm. The blue curve is the curve without any modification and the green curve for the algorithm with the modification introduced in step 4. The blue curve is for now of interest. It can be seen that on average all the particles travel back to their initial position. It is not shown in the plot, however, this is even true for every trajectory. The FFT has a downside compared to the continuous Fourier transform, which causes this effect. The VACF is zero at zero-frequency $\tilde{Z}_{l=0} = 0$. From eq. (1.39) also the first increment in the frequency

1. Generate $4M$ independent normally distributed random increments:

$$\eta_k \sim \mathcal{N}(0, \sqrt{\Delta t}) \quad , \quad k = 0, 1, 2, \dots, 4M \quad (1.37)$$

2. Fourier transform the increments via FFT into the frequency domain:

$$\tilde{\eta}_l = \sum_{k=0}^{4M-1} \eta_k \exp \left[\frac{-i2\pi lk}{4M} \right] \Delta t \quad , \quad l = 0, 1, 2, \dots, 4M \quad (1.38)$$

3. Generate fractional increments in the frequency domain with the VACF:

$$\tilde{\xi}_l = \tilde{\eta}_l \sqrt{2\text{Re}(\tilde{Z}_l)} \quad (1.39)$$

Here, \tilde{Z}_l is the discrete VACF in the frequency domain:

$$\tilde{Z}_l = K_\alpha \Gamma(1 + \alpha) (i2\pi l \Delta \omega)^{1-\alpha} = K_\alpha \Gamma(1 + \alpha) \left(\frac{il2\pi}{4M\Delta t} \right)^{1-\alpha} \quad (1.40)$$

4. Replace the zero-increment in the frequency domain with:

$$\tilde{\xi}_{l=0} = \mathcal{N}(0, \sqrt{2K_\alpha(4M\Delta t)^\alpha}) \quad (1.41)$$

5. Reverse Fourier transform the increments into the time domain:

$$\xi_k = \frac{1}{4M} \sum_{l=0}^{4M-1} \tilde{\xi}_l \exp \left[\frac{2\pi ilk}{4M} \right] \Delta \omega \quad (1.42)$$

6. Build the cumulative sum only for the first quarter of the total number of generated increments:

$$B_n^\alpha = \sum_{k=0}^{n-1} \xi_k, \quad n = (0, 1, \dots, M) \quad (1.43)$$

Table 1.2: The table shows a naive fBm generating algorithm.

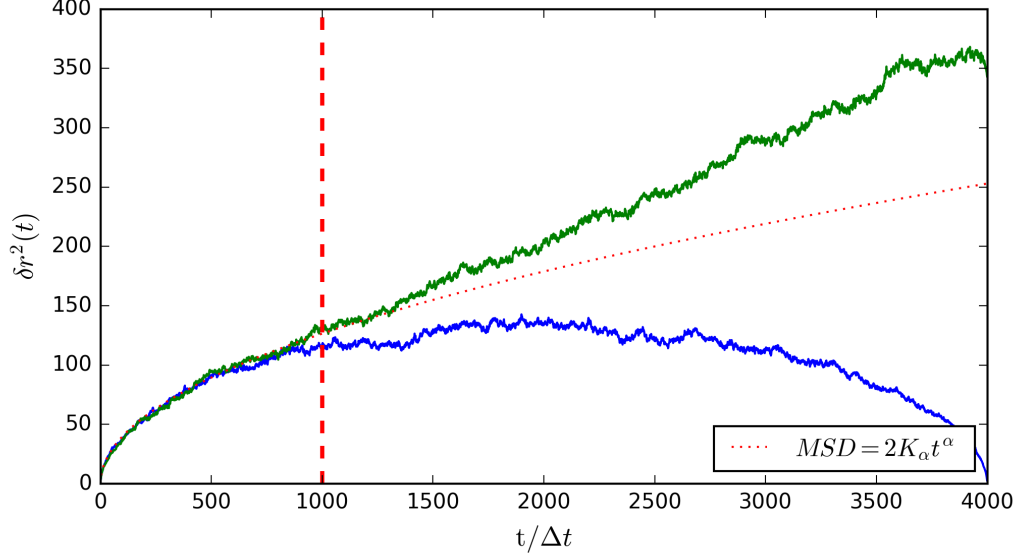


Figure 1.2: An ensemble average mean square displacement of two versions of the naive algorithm. The green curve is the MSD of the algorithm with the correction introduced in step 4 and the blue curve without.

domain is zero $\tilde{\xi}_{l=0} = 0$. Due to eq. (1.38) also the following relation holds:

$$\tilde{\xi}_{l=0} = \sum_{k=0}^{M-1} \xi_k e^0 \Delta t = \Delta R_M \quad (1.44)$$

ΔR is the distance between the starting point and the position of the particle. Therefore, the particle would travel after M steps back to its initial position. The naive idea is to generate a trajectory four times longer than the desired length and only take the first forth into consideration. This can be thought of as a finite-time correction. The cut is displayed as a vertical red dashed line. The plot suggests not a perfect but at least a better agreement of the generated data to the theoretical mean square displacement displayed as a dotted line for short times. Hereafter the normally distributed random numbers are transformed into the frequency domain to apply Wiener-Khinchin's theorem by multiplying with the square root of two times VACF. Besides taking only one-quarter of the trajectory a modification to the algorithm was introduced in step 4. Its impact in the MSD is shown in fig. 1.2 as the green line. Instead of accepting eq. (1.44), the zero-frequency increment is calculated as follows:

$$\tilde{\xi}_{l=0} = \mathcal{N}(0, \sqrt{2K_\alpha (M\Delta t)^\alpha}) \quad (1.45)$$

This equation would be correct if we assumed fractional Brownian motion to be a Markovian process. This is certainly is not the cause. This approximation seems

to improve the accuracy only slightly. Finally in step 4 the increments are back transformed into the time domain and in step 5 the cumulative sum of fGn is calculated, which yields fBm. The described algorithm can be performed independently for every Cartesian component of the three dimensional fractional Brownian motion. Different Cartesian components are independent. Without any extensive analysis of the naive algorithm one can already conclude draw backs in terms of convergence close to the overall simulation time.

Lowen Algorithm

1. Compute the auto-covariance function R_n^ξ of a periodic stochastic process ξ_n :

$$R_n^\xi = \begin{cases} \frac{1}{2} [1 - (\frac{n}{N})^\alpha] & \text{for } 0 \leq n \leq M \\ R_{2M-n}^\xi & \text{for } M \leq n \leq 2M \end{cases} \quad (1.46)$$

2. Transform the auto-covariance function via FFT. The result is called the spectral density of the stochastic process ξ_n :

$$S_k^\xi = \text{FFT}(R_n^\xi; k) \quad (1.47)$$

3. Calculate $\tilde{\xi}_k$ the Fourier transform of the stochastic process ξ_n :

$$\tilde{\xi}_k = \begin{cases} 0 & \text{for } k = 0 \\ \exp(i\theta_k) \eta_k \sqrt{S_k^\xi} & \text{for } 0 \leq k \leq M \\ \eta \sqrt{S_k^\xi} & \text{for } k = M \\ \tilde{\xi}_{2M-k}^* & \text{for } M \leq k \leq 2M \end{cases} \quad (1.48)$$

$\{*\}$ denotes the complex conjugate, θ_k is a random phase uniformly distributed in $(0, 2\pi]$. η_k is a random Gaussian variable with zero mean and variance 1 [$\eta_k \sim \mathcal{N}(0, 1)$].

4. Perform the inverse Fourier transform on $\tilde{\xi}_k$ and use the first half of the resulting stochastic process ξ_n multiplied by a factor:

$$\xi_n = \frac{\text{FFT}^{-1}(\tilde{\xi}_k; n)}{2M} \quad (1.49)$$

$$B_n^\alpha = \sqrt{2K_\alpha M^\alpha \Delta t^\alpha} (\xi_n - \xi_0) \quad \text{for } 0 \leq n \leq M \quad (1.50)$$

Table 1.3: The table shows Lowen's fBm generating algorithm [14].

An algorithm by Lowen [14] was implemented. Lowen's algorithm claims to be both fast ($\mathcal{O}(N \log N)$) and exact. However, our implementation shows not to be exact, which is shown in the analysis part. It starts with a auto-covariance function R_n^ξ . Its second derivative yields the auto-covariance of discrete time fGn described by eq. (1.25). R_n^ξ can be thought of as an auto-covariance function of a stationary periodic version of fBm. Lowen's algorithm is described in table 1.3. Let's check the resulting auto-covariance function of B_n^α resulting from the algorithm:

$$\text{Cov}[B_n^\alpha, B_m^\alpha] = 2K_\alpha \Delta t^\alpha N^\alpha \langle (\xi_n - \xi_0)(\xi_m - \xi_0) \rangle \quad (1.51)$$

$$= 2K_\alpha \Delta t^\alpha N^\alpha \langle \xi_n \xi_m \rangle - \langle \xi_0 \xi_m \rangle - \langle \xi_0 \xi_n \rangle + \langle \xi_0^2 \rangle \quad (1.52)$$

$$= K_\alpha \Delta t^\alpha [n^\alpha - 2(m-n)^\alpha + m^\alpha] \quad \text{for } n < m \quad (1.53)$$

The auto-covariance function indeed satisfies the condition for fBm with variance similar to eq. (1.26). In step 3 the phase and amplitude of the Fourier transform ξ_m were chosen to be random as suggested in [16]. The second derivative of $R_\xi(n)$ is positive. $R_\xi(n)$ results in a periodic function with non-negative curvature. Its Fourier transform S_k^ξ is, real, symmetric and non-negative for all k. S_k^ξ is then a valid power spectral density function of the discrete-time periodic process ξ_n with period $2n$ and R_n^ξ its valid auto-covariance function [14]. This algorithm is actually using the property of Brownian scaling eq. (1.4) to generate a realisation of fBm with non-negative auto-covariance function.

Davis-Harte Algorithm

Davis-Harte algorithm [15] is related to Lowen's Algorithm. It is also based on a FFT. The author claims the algorithm to be exact and its performance to scale with the length of the trajectory as ($\mathcal{O}(N \log N)$). However, in contrast to Lowen's algorithm, it is based on the covariance function of fGn described by equation eq. (1.25). Thus, it actually generates the increments of fBm. The algorithm is shown in table 1.4.

Accuracy Analysis

The algorithms were implemented in Python and C++. For the C++ implementation, a wrapper to Python was added. All algorithms were analysed by an analysis class. The C++ implementation is using the FFTW library for the Fast Fourier Transform [17] and Mersenne-Twister (gsl_rng_mt19937) [18] as the random number generator. Python implementations were developed beforehand and only serve as a reference. Python uses the numpy.fft library for the Fast Fourier Transform and also the Mersenne-Twister random number generator.

Insights of stochastic algorithms can be obtained by analysing observables. Observables of a stochastic process are generally averaged values. The motivation for fBm was its power law behaviour of the MSD. Therefore it is reasonable to analyse first the MSD. It can be calculated as the ensemble- or time average. A difference in time and ensemble average would show a violation of ergodicity. A comparison of all

1. Compute the auto-covariance function R_n^ξ of a periodic stochastic process ξ_n :

$$R_n^\xi = \begin{cases} K_\alpha \Delta t^\alpha [(n-1)^\alpha - 2n^\alpha + (n+1)^\alpha] & \text{for } 0 \leq n \leq M \\ R_{2M-n}^\xi & \text{for } M \leq n \leq 2M \end{cases} \quad (1.54)$$

2. Transform the auto-covariance function via FFT. The result is called the spectral density of the stochastic process ξ_n :

$$S_k^\xi = \text{FFT}(R_n^\xi; k) \quad (1.55)$$

3. Calculate $\tilde{\xi}_k$ the Fourier transform of the stochastic process ξ_n :

$$\tilde{\xi}_k = \begin{cases} \sqrt{2S_k^\xi M}(\eta_k) & \text{for } k = 0 \\ \sqrt{S_k^\xi M}(\eta_k + i\eta_k) & \text{for } 0 \leq k \leq M \\ \sqrt{2S_k^\xi M}(\eta_k) & \text{for } k = M \\ \tilde{\xi}_{2M-k}^* & \text{for } M \leq k \leq 2M \end{cases} \quad (1.56)$$

$\{*\}$ denotes the complex conjugate, η_k is a random Gaussian variable with zero mean and variance 1 [$\eta_k \sim \mathcal{N}(0, 1)$].

4. Perform the inverse Fourier transform on $\tilde{\xi}_k$ and use the first half of the resulting stochastic process ξ_n :

$$\xi_n = \frac{1}{\sqrt{2M}} \text{FFT}^{-1}(\tilde{\xi}_k; n) \quad \text{for } 0 \leq n \leq N \quad (1.57)$$

5. perform the cumulative sum on the increments:

$$B_n^\alpha = \sum_{j=0}^n \xi_j \quad \text{for } 0 \leq n \leq M \quad (1.58)$$

Table 1.4: The table shows Davis-Harte fBm generating algorithm [15].

four implemented algorithms can be seen in fig. 1.3. Here $\Gamma = \sqrt{2K_\alpha \Delta t}$ is going to be the scale defining quantity in space. The MSD is displayed in Γ^2 . MSDs for the algorithm are scaled with $k = 0.1, 1, 10, 100$ for Davis-Harte, Cholesky, Naive, Lowen, respectively for a better visualisation. The theoretical scaled MSD $\delta r^2(t) = 2K_\alpha kt$ are displayed as solid black lines. All algorithm show good accuracy between ensemble

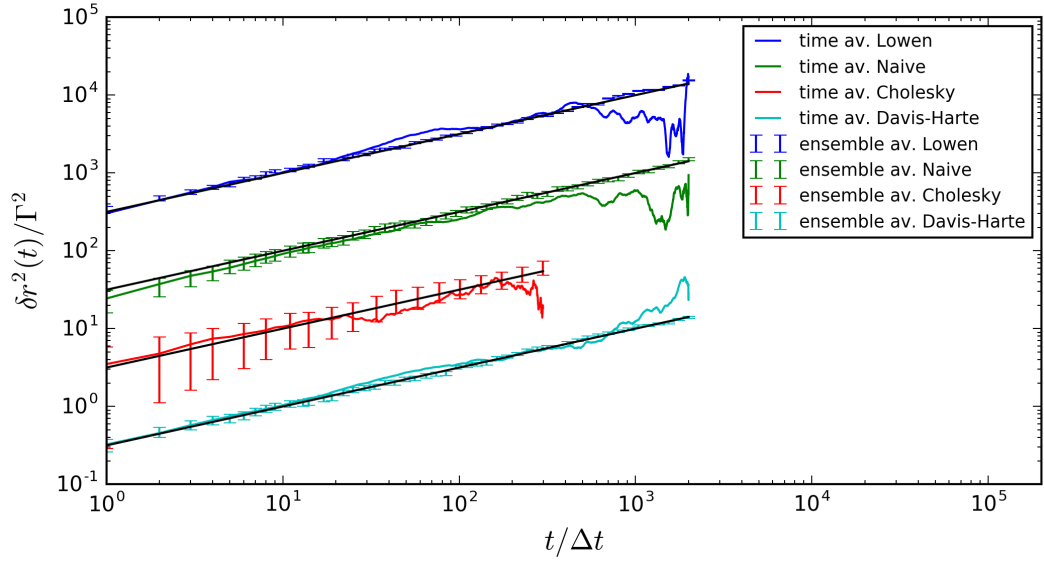


Figure 1.3: Comparison of MSD between time-average, ensemble-average over $N = 1000$ trajectories with trajectory length $M = 2000$ (Cholesky: $M = 300$) and $\alpha = 0.5$ for four algorithms. Cyan coloured line: Davis-Harte with $k = 0.1$, red coloured: Cholesky $k = 1$, green coloured: Naive $k = 10$, blue colour Lowen $k = 100$.

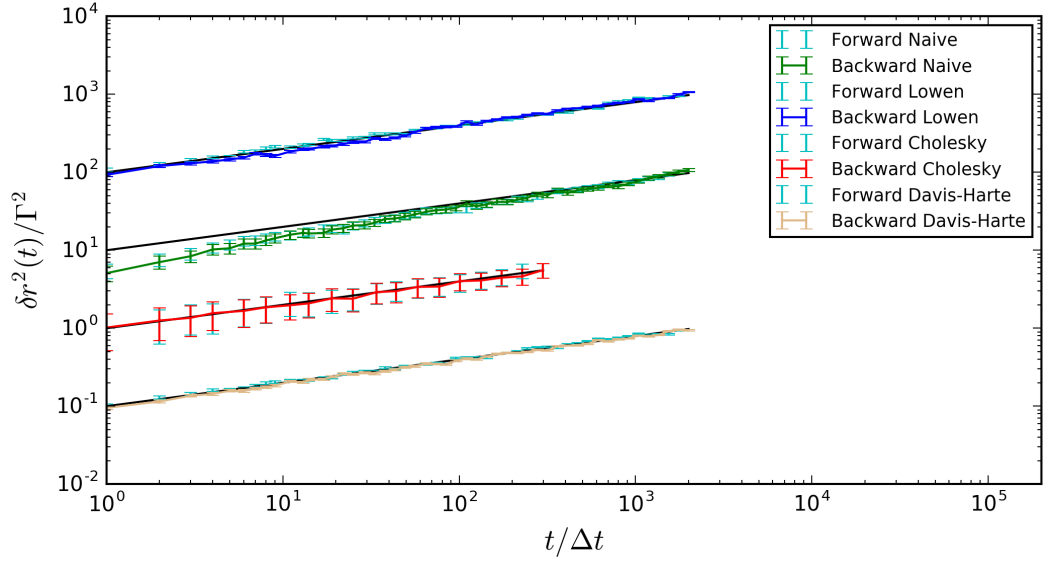


Figure 1.4: The plot show ensemble averaged MSDs for forward time and backward time over $N = 2000$ (Cholesky: $M = 300$) trajectories with trajectory length $M = 1000$ $\alpha = 0.3$, brown coloured line: Davis-Harte with $k = 0.1$, red coloured: Cholesky $k = 1$, green coloured: Naive $k = 10$, blue colour Lowen $k = 100$. All backward trajectories are cyan coloured

ble and time averaged MSD, besides for long lag times for the time averaged MSD due to decreasing statistics. The naive algorithm tends to result in too small MSDs for small lag times particularly for small α , which is going to be shown later on. This is caused by the finite number of samples in the discrete Fourier transform. In the following figures all MSD plots are ensemble averages.

Secondly, time-reversibility had been checked. FBm is an equilibrium process and thus should be time reversible [19]. A time reversed trajectory can be defined as:

$$\overleftarrow{B}_n^\alpha = B_M^\alpha - \sum_{j=0}^M \xi_{M-j} \quad (1.59)$$

In fig. 1.4 MSDs of normal fBm and reversed time fBm described in eq. (1.59) are plotted with the previously introduced scaling parameter k for better visualisation. No difference can be seen. Thus all algorithms generate an equilibrium process. α had been chosen to be 0.3. Thereby the naive algorithm has an even stronger deviation from the desired power law because of the problem of the finite number of samples in the discrete Fourier transform. A strong influence of α on the accuracy of MSD for the naive algorithm is suspected. Hence, the influence of α on the MSD accuracy was studied next. The influence of α on the MSD for all four algorithms is shown in fig. 1.5. No deviations for Lowen, Cholesky and Davis-Harte can be seen. Cholesky and Davis-Harte algorithm work also in the range of $1 < \alpha < 2$. In the limit of Brownian motion $\alpha = 1$, the artefacts of the naive algorithm vanish. Cholesky, Lowen and Davis-Harte show no deviation to the expected MSD for all α .

FBm's property of self-similarity can be used to plot a scale-free version of the density distribution ($\mathcal{P}_{gauss}(\hat{r})$). The scale-free version was introduced in eq. (1.30). The scaled density distribution is not dependent on time. Time rescaling leads to a "data collapse". In fig. 1.6 histograms over $N = 10000$ single trajectories at various times ($100 \leq t \leq 1000$) were calculated and rescaled according to eq. (1.30). The dashed line shows the expected distribution. For Cholesky and our algorithm no deviations from the theoretical values are obvious. For Lowen's algorithm, systematically small changes to the theoretical value can be observed. However, the distribution of fBm should stay a Gaussian for all times. As introduced in eq. (1.33) the non-Gaussian parameter ($NGP(t)$) should be zero. The Non-Gaussian parameter has been plotted for $0 \leq t \leq 1000$ (≤ 356 for Cholesky) $\alpha = 0.5$ in fig. 1.7 a). Our and Cholesky algorithm perform as expected. Lowen however, shows deviation from a Gaussian distribution increasingly towards the end of the trajectory. In fig. 1.7 b) the influence of α on the non-Gaussian parameter for the Lowen algorithm was tested. For small α the distribution is closer to a Gaussian. In our implementation, Lowen's algorithm is not exact as claimed in [14]. A modification to the algorithm was introduced in point 3 in table 1.3. The modification introduces the correlation the Davis-Harte style, as shown on table 1.4, with random complex numbers generated in the rectangular representation. The third point of Lowen algorithm was exchanged with the third point of Davis-Harte algorithm. The resulting accuracy of MSD and a comparison of the non-Gaussian parameter are shown in fig. 1.8.

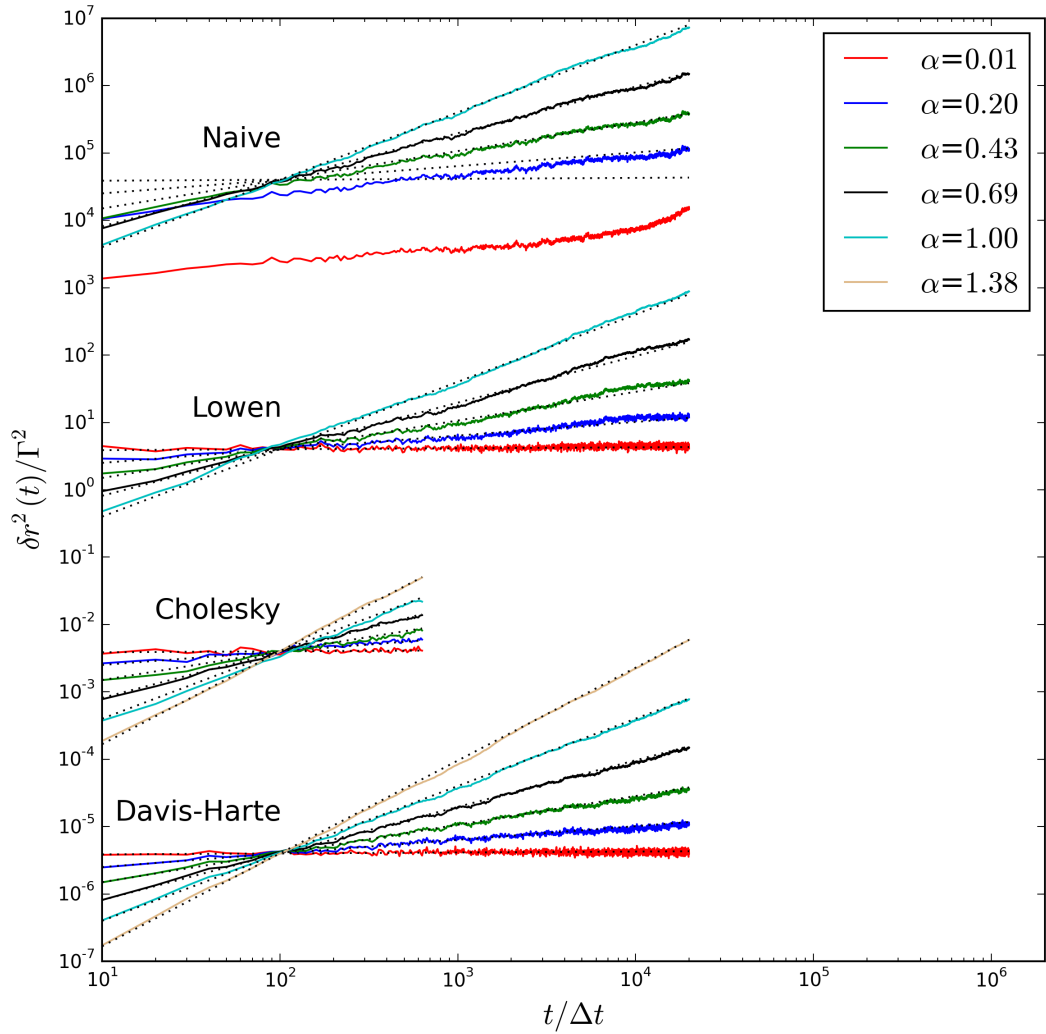


Figure 1.5: The figure shows ensemble averaged MSDs for all four algorithms over $N = 400$ trajectories with the trajectory length $M = 2000$ (Cholesky: $M = 64$) and $0.01 \leq \alpha \leq 1.0$ (Cholesky and Davis-Harte $\alpha \leq 1.38$).

Performance Analysis

The performance with sample size (trajectory length) and number of samples (amount of trajectories) was analysed in fig. 1.9 b). Thereby the most promising, in terms of exactness, Cholesky algorithm scales with the trajectory length M with $\mathcal{O}(M^3)$. Our application of fBm desires long trajectories. It is not reasonable to use such a slow algorithm. Even if every next trajectory with the same α and length scales with $\mathcal{O}(M^2)$ [12]. In fig. 1.9 a) one can see Cholesky to scale slow with the amount of trajectories N . Although one should keep in mind that the trajectory length had to be chosen $M = 128$ compared to $M = 1000$ as for the naive and Lowen's algorithm to get comparably small values. Lowen's and the naive algorithm perform

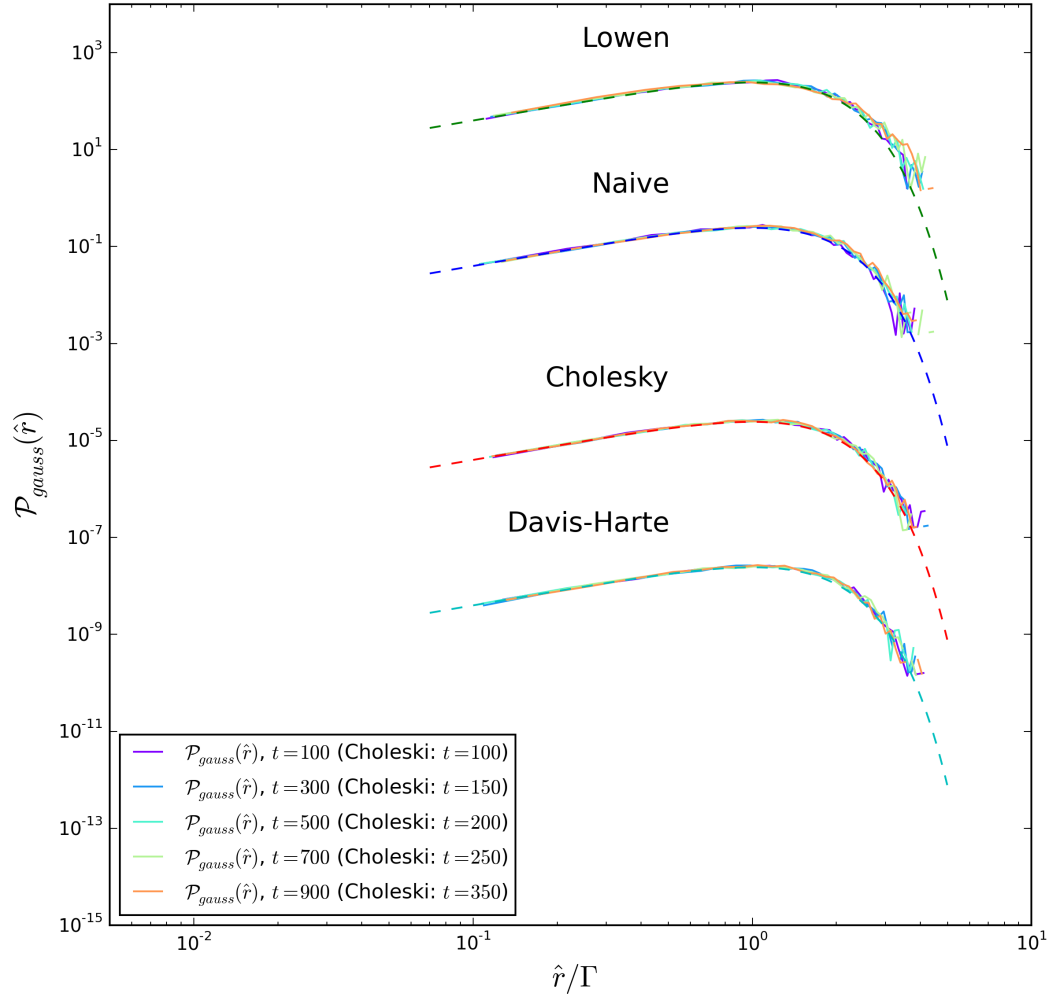


Figure 1.6: All four plots show the scale free form of the propagator $\mathcal{P}_{gauss}(\hat{r})$ at different times as introduced in eq. (1.30) as an histogram over $N = 10000$ trajectories for $\alpha = 0.5$ at different times $100 < t < 900$.

both $\mathcal{O}(M \log(M))$ in respect to the trajectory length as a result of the FFT and $\mathcal{O}(N)$ in respect to the amount of trajectories.

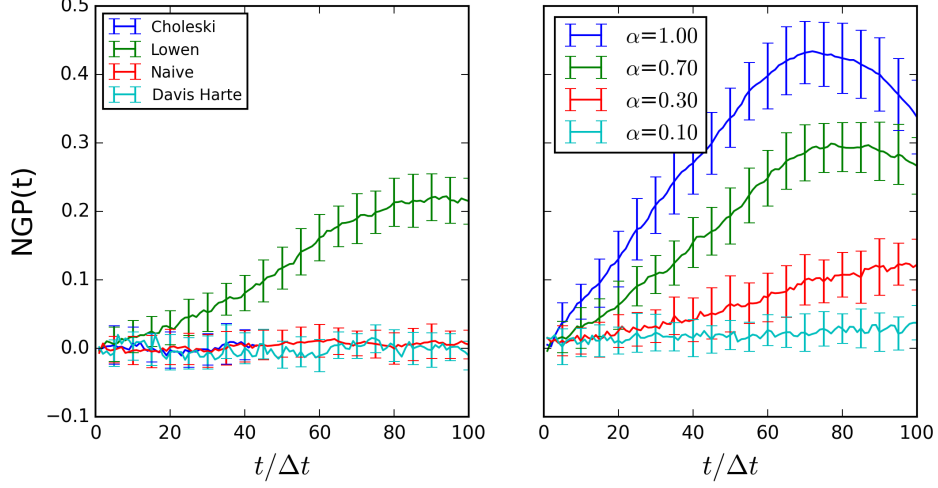


Figure 1.7: The left figure shows the Non-Gaussian-Parameter as introduced in eq. (1.33) for all four algorithms with $N = 5000$, $M = 100$ and $\alpha = 0.5$ averaged over 30 non-Gaussian-Parameter with its variance displayed as an error bar. The right figure shows the Non-Gaussian-Parameter for Lowen's algorithm with $N = 5000$ and $M = 100$ for various $0.1 \leq \alpha \leq 1.0$ averaged over 30 non-Gaussian-Parameter with its variance displayed as an error bar.

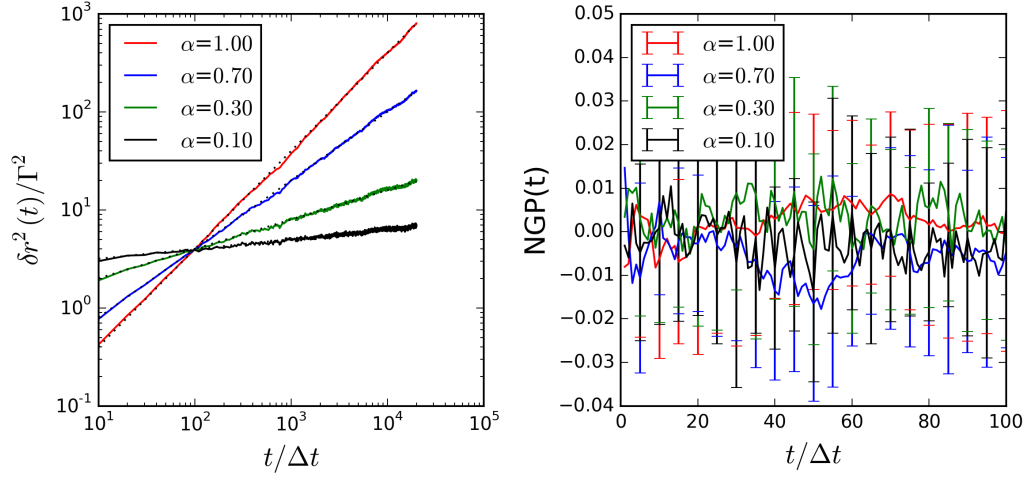


Figure 1.8: The left figure shows the MSD for various $0.1 \leq \alpha \leq 1.0$ generated with the modified Lowen algorithm with $M, N = 2000$. The right figure shows the Non-Gaussian-Parameter for the modified Lowen algorithms with $N = 5000$, $M = 100$, $\Delta t = 0.1$ for various $0.1 \leq \alpha \leq 1.0$ averaged over 30 non-Gaussian-Parameter with its variance displayed as an error bar.

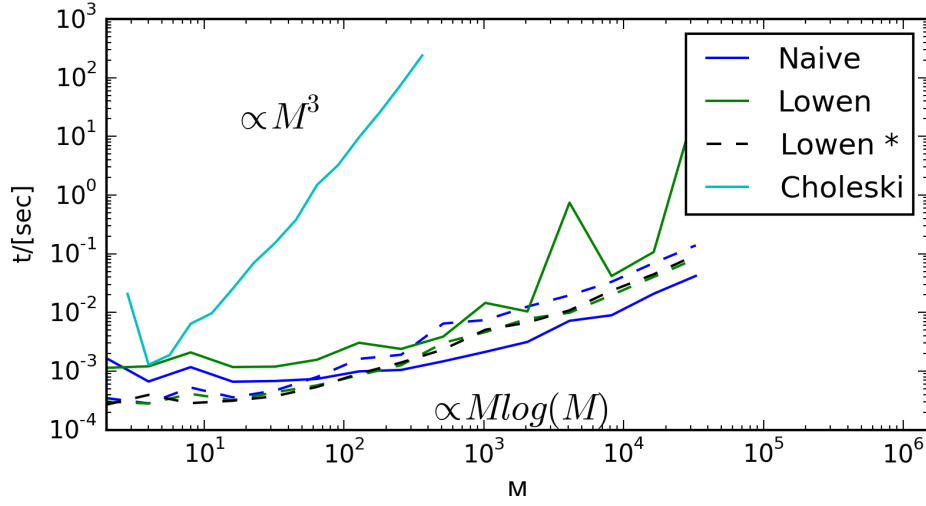


Figure 1.9: Algorithmic scaling of computational time with the trajectory length M for a single trajectory $N = 1$. Dashed lines are algorithms implemented in C++. Solid Lines are algorithms implemented in Python.

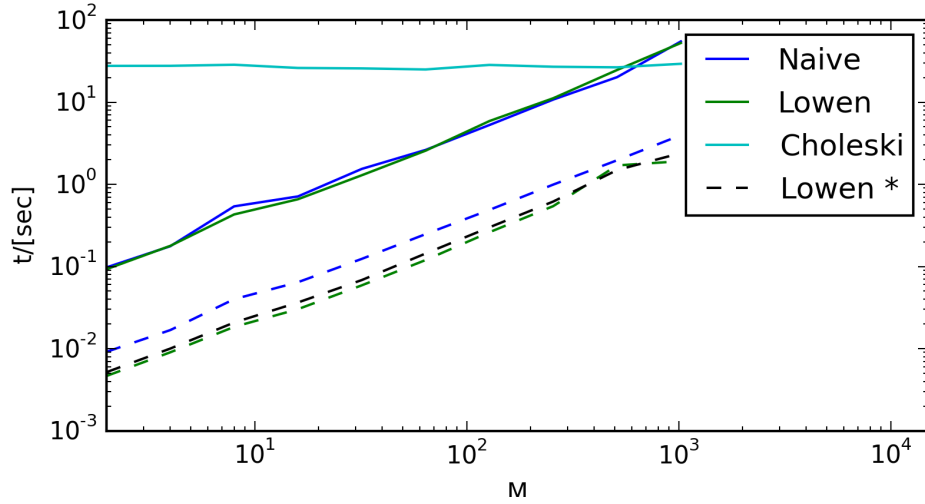


Figure 1.10: Algorithmic scaling of computational time with the number N of trajectories and trajectory length $M = 1000$ (Cholesky $M = 128$). Dashed lines are algorithms implemented in C++. Solid Lines are algorithms implemented in Python.

2 Particle Based Reaction Diffusion

Chemical reactions are ubiquitous in nature. They describe transformations of one or several chemical substances. Transformations of chemical substances are described by formation and breaking of chemical bonds. Physically correct these bonds should be described quantum mechanically. Quantum mechanics is not needed for a description at cellular scales. Not to mention that for such large systems these calculation cannot be performed. Hence, coarse grained models for reactions at cellular scale were developed. But reactions in a broader sense can also be a transformations of one state to another one (e.g. protein changing its conformation). Biochemical reaction give rise to the complexity of life and are of key importance for processes in living cells. Enzymatic reactions as an example are involved in storage and release of energy, processes of respiration, vision, muscle contraction, transmission of nerve impulses. Indeed Biochemistry is a whole area of research.

A Early pioneer general work on reactions by Guldberg and Waage already in 1864 [20] resulted in the law of mass action, which states the proportionality of the reaction rate to the product of concentrations for elementary molecular reactions. Thereby spatial resolution is neglected. The system is described in terms of a concentration (well-mixed assumption). The mass action law leads to a deterministic model described by ordinary differential equations (ODEs). Quantities like reaction kinetics and amount of reactants in equilibrium can be obtained. Deterministic models however are not resolving stochastic phenomena. Stochastic models provide a more detailed understanding of the reaction-diffusion processes. Such a description is often necessary for the modelling of biological systems, where small molecular abundances of some chemical species make deterministic models inaccurate or even inapplicable [21]. Stochastic models based on the mass action law are characterised by the chemical master equation (CME). CME describes the temporal evolution of the probabilities for the system to occupy each different state. The interesting topic of chemical kinetics described by CME will not be part of the elaborations. Nevertheless, for the sake of completeness, it was to be mentioned that stochastic simulations based on CME with the Gillispie algorithm [22] are performing very efficiently. Even spatial resolutions can be achieved with simulations based on the reaction-diffusion master equation (RDME [23]). Still the well-mixed assumption has to hold at least for the compartments of the RDME model. A more detailed model is given by particle-based reaction-diffusion dynamics (PBRD)[24, 25]. Thereby all particles are modelled as points or spheres and are resolved in space and time. However, the size particularly of common biological systems surpass computational feasibility. So further coarse graining gets necessary. Difficulties in simulation of biological systems arise on the one hand from very large systems and on the other from small molecular abundances of some chemical species. The well-mixed assumption does not hold.

A less prominent approach is a PBRD model with fractional Brownian dynamics. Fractional Brownian motion approximates the motion of a particle in crowded biological medium. The motion of the media itself has not to be modelled explicitly. Therefore PBRD simulations can be performed just for the relevant reacting particles. Stochastic particle based models conventionally rely on Brownian or Langevin dynamics for the description of diffusion. Brownian dynamics was introduced in chapter 1. Langevin dynamics will not be examined explicitly. In short, it is further includes also the relaxation of particles momentum and reaches Brownian dynamics in the limit of strong damping, or at long time scales.

This chapter deals with some theoretical considerations on the impact of diffusion on reaction rates. The first section examines Smoluchowski pioneer work [26] on diffusion limited reaction kinetics of a bi-molecular reaction and Erban/Chapman extensions [27] for diffusion influenced bi-molecular reactions. Thereby a dependence of macroscopic reaction rates and the diffusion of the particles can be seen. These results shall motivate investigations of reaction diffusion dynamics with a different diffusion model. Along the way the law of mass action is shown to follow under certain conditions from the diffusion eq. (1.5) motivated and introduced in the first chapter. These derivations further motivate both, concentration and PBRD models. RevReaDDy, a realisation of the latter, is then introduced and explained in the second section of this chapter.

2.1 Diffusion Influenced Bimolecular Reaction

Uni- and bi-molecular elementary reactions are most relevant in nature. According to collision theory tri-molecular elementary reactions are negligible. Diffusion is not relevant for uni-molecular reactions. It is reasonable to start our considerations on how diffusion influences chemical reactions with the Smoluchowski problem for kinetics of a bi-molecular chemical reaction in solution. One particular area of research on which the Smoluchowski relation had significant impact in the last few decades is biology. This is unsurprising as a vast number of biomolecular systems involve dilute and minute diffusing molecular populations undergoing continuous reaction. Understanding how these biological systems operate is complicated and is in itself a whole field of research; systems biology. The Smoluchowski result has provided a very powerful tool for theoretical investigation of microscopic biochemical reaction-diffusion processes [28]. It was extended by P. Debye in 1942 to add intermolecular forces. However, especially in biological systems Brownian motion do not apply for all time-scales. These environments exhibit very often anomalous diffusion, which can be addressed by fBm.

The following derivation shall give the kinetics of a simple model for a bi-molecular reaction and motivate a particle based simulation scheme. The reaction scheme for a bi-molecular reaction with reactants A and B associate into a complex AB with a rate k_d^+ and decay into A and B with a rate k_d^- . The chemical reaction can be

written as:



Free diffusion of particle A and B with the diffusion constants D_A and D_B , respectively, are assumed. Further no interactions between them are present. In fact this problem can be described by a modified version of the diffusion eq. (1.5) motivated in the first chapter. The joint concentration field $\rho_t(\mathbf{r}_A, \mathbf{r}_B)$ for a bi-molecular system in a solution obeys:

$$\frac{\partial \rho_t(\mathbf{r}_A, \mathbf{r}_B)}{\partial t} = (D_A \nabla_A^2 + D_B \nabla_B^2) \rho_t(\mathbf{r}_A, \mathbf{r}_B) \quad (2.2)$$

The complexity of the problem can be reduced by substituting the positions of the particles A and B with their relative distance $\mathbf{r} = \mathbf{r}_A - \mathbf{r}_B$. It is convenient to introduce even further substitutions:

$$D = D_A + D_B \quad \mathbf{R} = \frac{D_B \mathbf{r}_A + D_A \mathbf{r}_B}{D_A + D_B} \quad (2.3)$$

the Laplace operator in terms of new coordinates result in:

$$\nabla_A^2 = \left(\nabla_r + \frac{D_B}{D} \nabla_R \right)^2 \quad (2.4)$$

$$\nabla_B^2 = \left(\nabla_r + \frac{D_A}{D} \nabla_R \right)^2 \quad (2.5)$$

By inserting eq. (2.4) and (2.5) in eq. (2.2) the following result can be obtained:

$$\frac{\partial \tilde{\rho}_t(\mathbf{r}, \mathbf{R})}{\partial t} = \left(D \nabla_r^2 + \frac{D_B D_A}{D_A + D_B} \nabla_R^2 \right) \tilde{\rho}_t(\mathbf{r}, \mathbf{R}) \quad (2.6)$$

This equation describes two independent diffusion processes, one in the coordinate \mathbf{r} and one in the coordinate \mathbf{R} . The solution can be obtained by the product ansatz $\tilde{\rho}_t(\mathbf{r}, \mathbf{R}) = \rho_t(\mathbf{r}) q_t(\mathbf{R})$. Integration over \mathbf{R} results in:

$$\frac{\partial \tilde{\rho}_t(\mathbf{r}, \mathbf{R})}{\partial t} = D \nabla_r^2 \rho_t(\mathbf{r}) + \frac{D_B D_A}{D_A + D_B} \int_{\partial V} \nabla_R q_t(\mathbf{R}) d\mathbf{a} \quad (2.7)$$

In the previous equation the stokes theorem was applied. The integral in the second term is zero due to conservation of probability. The problem is isotropic, hence $r = |\mathbf{r}|$ and $\nabla_r^2 = \left(\partial_r + \frac{2}{r} \right) \partial_r$. The equation reduce to one dimension in the form of the continuity eq. (1.6) without any production term:

$$\frac{\partial \rho_t(r)}{\partial t} = - \left(\frac{\partial}{\partial r} + \frac{2}{r} \right) j_t(r) \quad j_t(r) = D \frac{\partial \rho_t(r)}{\partial r} \quad (2.8)$$

The stationary distribution satisfies:

$$-\left(\frac{\partial}{\partial r} + \frac{2}{r}\right)j^s(r) = 0 \quad (2.9)$$

$$\frac{dj^s(r)}{j^s} = -\frac{2}{r}dr \quad (2.10)$$

$$j^s(r) = Ar^{-2} \quad (2.11)$$

$$\rho^s(r) = \rho^s(\sigma) - \frac{\int_{\sigma}^r j^s(r')dr'}{D} = \rho^s(\sigma) + \frac{A}{D} \left(\frac{1}{r} - \frac{1}{\sigma}\right) \quad (2.12)$$

Here σ is defining a distance between particle A and B. It is going to be the scale defining quantity in space. For particles A and B closer than σ instantaneous reactions occur and the joint concentration equals zero ($r \leq \sigma \Rightarrow \rho_t(r \leq \sigma) = 0$). Now let's assume only a single molecule B being at the position $r = 0$ and a constant concentration of particles A at infinity ($\rho_t(r \rightarrow \infty) = c_A$). The frame of reference is the B particle. This approximation, where the distribution of particle A around a single particle B is studied without any B particles, besides the one in the origin, should also be true for a low concentrations of B particles relative to A particles ($c_B \gg c_A$). Two particles B in proximity would disturb the concentration profile. The solutions for $\rho^s(r)$ and $j_t^s(r)$ for these boundary conditions are:

$$\rho^s(r) = c_A \left(1 - \frac{\sigma}{r}\right) \quad j^s(r) = -D\sigma c_A r^{-2} \quad (2.13)$$

With eq. (2.1) in mind it should be clear that the flux of particles at σ results in a production of c_{AB} . The production of the concentration of c_{AB} is then:

$$\frac{dc_{AB}}{dt} = 4\pi\sigma D c_A c_B \quad (2.14)$$

Along the way several assumptions were made:

- The system can be described by a diffusion constant which is independent of time and position.
- No interaction between particles are present (ideal gas).
- The system has a constant background concentration of A particles $\rho_t(r \rightarrow \infty) = c_A$.
- The dynamics are stationary.
- The reaction is limited by diffusion. Since reactions occur instantaneously at $r = \sigma$.

Still the solution gives the kinetics for a simple bi-molecular system. It is consistent with collision theory, which states that the collision frequency is proportional to the concentrations of the colliding particles. Pioneer work on reaction kinetics by Guldberg and Waage already in 1864 [20] resulted in the law of mass action which

states for elementary uni- and bi-molecular reactions, the proportionality of the reaction rate (r_f) to the concentration.

$$r_f = k_f c_A c_B \quad (2.15)$$

This equation provides the basis for a description of reaction kinetics in terms of ordinary differential equations (ODEs) like eq. (2.14).

Smoluchowski's problem can be in fact generalized for reactions to occur with a finite microscopic rate λ_+ within a radius σ , derived by Erban and Chapman [27]. However, as a small remark, the microscopic rate has to be distinguished from the macroscopic rate. As an example microscopic rates are related to the Kramers problem, describing the time dependence of a particle passing a potential barrier e. g. commonly in biology protein binding / unbinding. The generalized problem can be described similarly to Smoluchowski's problem eq. (2.8) with an production term continuity eq. (1.6) :

$$\frac{\partial}{\partial t} c(\mathbf{r}, t) = -\nabla J(\mathbf{r}, t) - \lambda_+ c(\mathbf{r}, t) H(\sigma) \quad (2.16)$$

Here $H(\sigma)$ is the Heaviside step function. The derivation of a stationary solution of the second-order ODEs analogously to the Smoluchowski problem can be found in the appendix 5.4. The solution shows again the possibility of describing the production with a macroscopic reaction rate $k_+(\kappa, \sigma)$:

$$k_+ = 4\pi D \sigma \left(1 - \frac{1}{\sigma \kappa} \tanh \sigma \kappa \right) \quad (2.17)$$

Here $\kappa = \sqrt{\lambda_+/D}$ and $1/\kappa$ can be intuitively interpreted as the penetration length of A particles into the shell with distance σ around B particles. The scenario in two dimensions is visualised in fig. 2.2. The law of mass eq. (2.15) still holds for a bi-molecular reaction not only in the diffusion limit but also for reactions occurring with a certain microscopic rate and under the well mixed assumption and normal diffusion. ODEs for the system can be formulated. Their solutions give the kinetics and concentrations in equilibrium, but no information on the spatial distribution are contained.

For the generalised problem with a microscopic rate λ_+ a diffusion and microscopic reaction limit can be shown. For $\lambda_+ \rightarrow \infty$ reaction occurs immediately and the diffusion limit is reached. If $\lambda_+ \ll \frac{D}{\sigma^2}$ a Taylor expansion of the Tangens Hyperbolicus in eq. (2.17) is given by:

$$\tanh x \rightarrow 0 \approx x - \frac{x^3}{3} \quad (2.18)$$

and eq. (2.17) can be written approximated as:

$$k_+ = \frac{4\pi\sigma^3\lambda_+}{3} \quad (2.19)$$

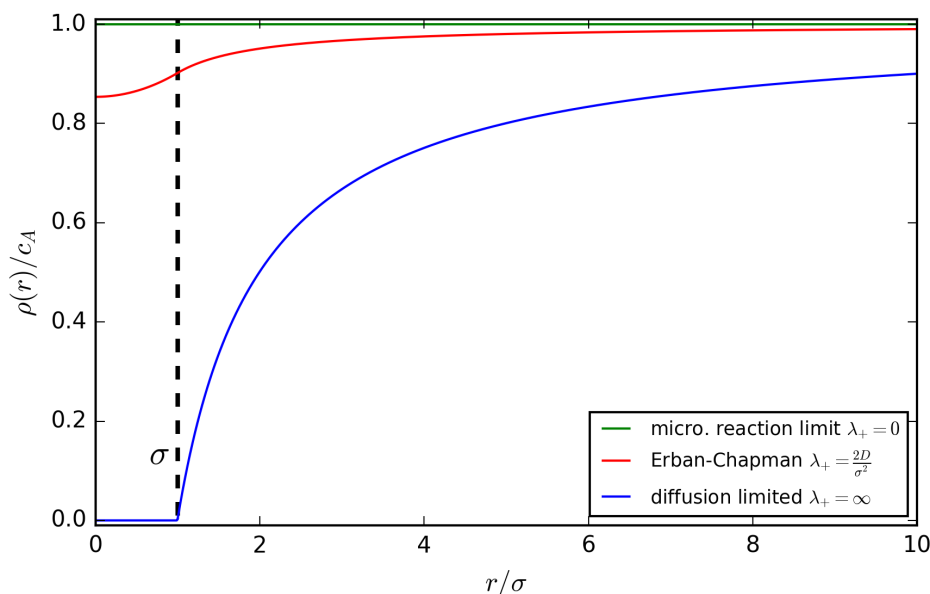


Figure 2.1: Density distribution from eqs. (5.19) and (5.20) for a bi-molecular reaction from the Erban-Chapman derivation for three scenarios:

1. The microscopic reaction limit for $\lambda_+ = 0.0$ displayed as a green solid line
2. An intermediate scenario with $\lambda_+ = 2d/\sigma^2$ displayed as a blue solid line.
3. Smoluchowski's result for diffusion limited reactions $\lambda_+ = \infty$.

The macroscopic reaction rate no longer depends on the diffusion but solely on the microscopic reaction rate. The stationary distributions for both limits and an intermediate scenario are plotted in fig. 2.1.

The previous elaborations show that macroscopic reaction rates dependent on diffusion for bi-molecular reactions. For simple reaction diffusion systems the evolution and the stationary distribution of the density distribution can be obtained by solving Smoluchowski's equation (PDE). For the bi-molecular reaction the law of mass action could be seen to follow for spatial homogeneity. Then reaction kinetics can be described for easy systems in terms of ODEs. Nevertheless complex systems cannot be solved. Further, for small molecular abundances of some chemical species deterministic models get inaccurate or even inapplicable. Stochastic models are necessary. Via chemical master equations (CMEs) and reaction diffusion master equations (RDMEs) analytical stochastic results can be obtained. Deduced from the Master equation also stochastic simulations based on Monte Carlo methods, like the prominent Gillespie algorithm [22], can be performed. This chapter however aims to introduce PBRD. The implementation of a PBRD simulation introduced in the next section follows the theoretical description of reactions with microscopic reaction

rates within a reaction radius in this section.

2.2 RevReaddy with fBm

RevReaDDy is a PBRD tool developed by Christoph Fröhner within his PHD project under supervision of Frank Noé. It is not published. Many modification to the tool had to be done in the process of this thesis. ReReaDDy is based on the simulation package ReaDDy by Johannes Schöneberg and Frank Noé [29]. Reaction-diffusion dynamics of all particles can be resolved in space and time. Particles are modelled as spheres with possible particle interaction potentials. The tool aims to bridge the gap between molecular simulation and CME/RDME simulations. It is written in C++. However, the user can setup up and simulate the desired reaction scheme in Python via a Python binding. Particle types with a specific radius and diffusion constant can be defined. Particles of a predefined particle type can be placed anywhere into a simulation box with variable box size and periodic boundary conditions. A desired simulation length and width of the time step have to be set. Conventionally particles propagate by Brownian Dynamics. In the process of the thesis also a fractional Brownian motion integrator with the circulant method by Davis-Harte [12] and Lowen [14] was implemented. For Brownian Dynamics, potentials with two different potential types, Lennard Jones and soft repulsion, can be introduced to act in-between particles and between fixed geometries in the simulation box and particles. Thereby a temperature has to be defined. Particles can be e.g. limited to diffuse on a surface by a potential. It can be chosen from various reaction types to affect the particle types. For uni-molecular reactions like conversion $A \xrightleftharpoons[k']{k} B$, annihilation $A \xrightarrow{k} \emptyset$, decay $A \xrightarrow{k} B + C$ the probability that one particle undergoes a reaction within the Δt is simply approximated as:

$$P(n) = 1 - \exp[-\lambda \Delta t n] \approx \lambda \Delta t n \quad \text{for} \quad P(1) \ll 1 \quad (2.20)$$

Here $n = t/\Delta t$. A more sophisticated algorithm, which draws the reaction time from the correct distribution is the Gillispie algorithm [22]. It could be implemented into RevReaDDy as an improvement. Microscopic reaction rates λ for such reactions are independent of the position.

For the decay reaction a weighting factor and reaction distance shall be defined. They determine how far each of the particles are placed from the previously existing complex. With weighting factor equals zero particle A is placed at the position of the pre-existing complex and particle B is placed with a uniform distribution in the sphere with reaction radius σ . For bi-molecular reactions like the enzymatic reaction with complex building $A + B \xrightleftharpoons[k']{k} A + C$ or fusion $A + B \xrightarrow{k} C$ relative positions of the reactants are important for the reaction to take place. Analogous to the setup up of Erban and Chapman in section 2.1 a microscopic reaction rate λ within a reaction distance σ have to be defined. If the reactants come closer than σ a reaction takes place with probability $P = \lambda \Delta t$. The problem is visualised in fig. 2.2. Thereby the time step cannot be chosen arbitrarily large. The mean square

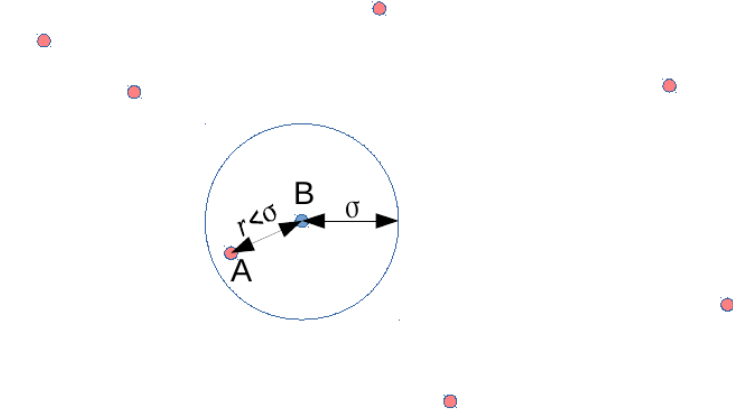
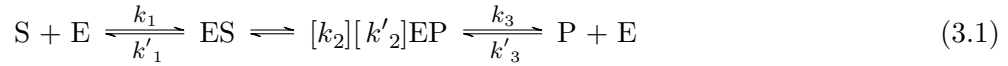


Figure 2.2: The Figure shows a visualisation of a bi-molecular reaction between particle A and B in RevReaDDy. Particle B are coloured blue and particle A are coloured red.

displacement for Brownian motion in three dimensions is given as $\delta r^2(\Delta t) = 6D\Delta t$. The mean displacement during one time step has to be a lot smaller than the squared penetration length $\delta r^2(\Delta t) \ll 1/\kappa^2$. Otherwise differences between the time of particles staying within a radius σ to another particle would differ significantly in comparison to $\Delta t \rightarrow 0$. Thus, also the macroscopic reaction rates would become less accurate. For fractional Brownian motion with $\alpha < 1$ and the effect increases. A variety of observables are provided in RevReaddy, which are written into HDF5 files [30] during the simulation. Possible observables are: Acceptance rate, energy, mean square displacement, particle number, radial distribution functions, amount of reactions, increments and positions. Observables can mostly be defined to consider any particle type over any period in time.

3 An Enzymatic Reaction With Fractional Brownian Motion

So reactions with enzymes involved take place a lot faster than without. Enzymes do not change the equilibrium concentrations. They have a defined amino acid sequence and a three-dimensional structure. The specific three-dimensional structure contains an active site for a substrate molecule to bind and orient and a catalytic site to reduce the chemical activation energy. Usually, an enzyme molecule has only one active site, and the active site fits with one specific type of substrate. So enzymes mostly have a high specificity for a particular type of reaction. Catalysis follows the binding of the substrate to the binding site. The catalytic site lowers the activation energy for the reaction. After reaction takes place the product is released. The enzyme is free to start the reaction circle with a new substrate. From a single enzyme and single substrate to a single enzyme and single product many reversible intermediate states can be present in the reaction mechanism. A simple representation of the mechanism with two complex intermediate states can be written as:



k are the forward rates and k' the backward rates. S and P are substrates and products, respectively. ES and EP are intermediate complexes. By the law of mass action ODEs can be formulated but not solved.

The focus of this chapter is to study the kinetics and spatial distribution of the Michaelis-Menten mechanism, an even simpler mechanism, without and with fractional Brownian motion. Some related theoretical work on the fractional reaction diffusion equation of the Michaelis-Menten mechanism had been done by Abdullah [31]. Monte Carlo simulations with the anomalous diffusion modeled by a random walk and immobile obstacles were performed by Berry and Schnell [32, 33]. An overview on stochastic modeling of in vivo reactions was given by Turner [34]. Recently first passage time simulations with fBm were performed by Jae-Hyung Jeon, A. V. Chechkin and Ralf Metzler [35]. Nevertheless, a Michaelis-Menten mechanism with fractional Brownian motion was not yet studied.

3.1 Michaelis-Menten (MM) kinetics

Leonor Michaelis and Maud Menten proposed in 1913 a mathematical model for enzyme kinetics. Their most important contribution was to assume an enzyme-substrate to be present [4]. In their honor it is also called Michaelis-Menten complex.

A simpler version of eq. (3.1) was proposed as the Michaelis-Menten mechanism:



The rate of the backward reaction from products and enzymes to complexes was assumed to be negligible. They postulated the relation between the reaction velocity, formation of products, and the concentration of substrates. This relation is captured by the so called Michaelis-Menten equation:

$$v = \frac{v_{max}c_s}{K_M + c_s} \quad (3.3)$$

v_{max} is the maximum reaction velocity. The Michaelis constant K_M is the substrate concentration with 50% of maximum velocity and c_s is the concentration of the substrate. A derivation of the Michaelis-Menten equation was done by Briggs and Haldane in 1925 assuming the mass action law, mass conservation and a quasi-steady state approximation (QSSA). It states that after a negligible initial transition period ($t > t_c$) the concentration of the complex vary slowly as if in a quasi-steady state $dC_{ES}/dt = 0$. The approximation was shown to hold for $c_{E_0} \ll K_M$ or $K_M \ll c_{S_0}$ [36]. A derivation of eq. (3.3) can be found in the Appendix 5.5. In 1997 a closed form for the kinetics of the Michaelis-Menten mechanism with the quasi-steady state approximation was derived by Schnell and C.Mendoza [37]:

$$c_S(t) = K_M W \left(\frac{c_{S_0}}{K_M} \exp \left(\frac{-v_{max}t + c_{S_0}}{K_M} \right) \right) \quad (3.4)$$

$$c_{ES}(t) = \frac{c_{E_0}c_S(t)}{K_M + c_S(t)} (1 - \exp(-(K_M + c_S(t))k_+)) \quad (3.5)$$

$$c_E(t) = c_{E_0}(t) - c_{ES}(t) \quad (3.6)$$

$$c_P(t) = c_{S_0} - c_S(t) + c_{ES}(t) \quad (3.7)$$

$v_{max} = k_2c_{E_0}$ is the maximum velocity, $W(x)$ is the omega function and the Michaelis-Menten constant is defined as:

$$K_M = \frac{k_- + k_c}{k_+} \quad (3.8)$$

The quasi-steady state approximation was applied in the derivation of $c_S(t)$. A time independent c_{ES} is conventionally solved for the transitions period $t < t_c$ by approximating $c_S(t < t_c) \approx c_{S_0}$. For eq. (3.5) the result for the time independent c_{ES} during the transitions period ($t < t_c$) was extended to the total period of time ($0 < t < \infty$) by applying time depended $c_S(t)$ from eq. (3.4). For normal diffusion theses equation will be used to fit the concentrations resulting from the simulation. Still for a small number of particles deterministic approach may differ from the averages of the stochastic simulation [34].

The quasi equilibrium approximation is the second very common approximation typically applied on setups with high enzyme concentrations. It is assumed that the

binding step quickly reaches an equilibrium state ($c_S c_E / c_{ES} = K_S$). $K_S = k_- / k_+$ is the dissociation constant and the rate of product formation can be written as:

$$v = \frac{V_{max} c_S}{K_S + c_S} \quad (3.9)$$

The assumption was shown to be true for $c_{E_0} \gg K_M$ or $k_- \gg k_c$. [36].

A description of the Michaelis Menten mechanism with fractional Brownian motion by fractional fractional diffusion reaction equation [38] was done by Abdullah [31]. An earlier description of reaction kinetics with fractional diffusion was done by Kopelman [39] for a percolation cluster with a scaling theory and an upgraded time depended reaction rates. This concepts were applied by Berry [32] and Schnell [33] for the Michaelis Menten mechanism.

3.2 Simulation Model

The simulation of the Michaelis-Menten mechanism eq. (3.2) and fractional Brownian motion with the circulant method by Davis-Harte [12] was setup up with RevReaDDy. Four particle types were defined: substrate, product complex and enzyme. Substrates and products were set to diffuse with the general diffusion constant $K_\alpha = 1/6$. Enzymes and complexes are modeled to be immobile with $K_\alpha = 0$. Enzymes are often even 100 times heavier than substrates or products and therefore diffuse a lot slower. A single enzyme was set in the origin $(x, y) = (0, 0)$ and 20 substrates were placed randomly in the simulation box with the box size l and periodic boundary conditions. Two association/decay reaction types were defined: $S + E \xrightleftharpoons[k_-]{k_+} ES$ and $P + E \xleftarrow{k_c} ES$ with intrinsic reaction rates λ_+, λ_- and λ_c for complex formation, decay into substrate and enzyme and decay into product and enzyme, respectively. The reaction radius σ was set for both reaction types to be 1. A desired simulation length $M = 2^{14}$ and length of the time step $\Delta t = 0.05$ was set. $1/\kappa = \sqrt{\lambda_+}/D$ can be defined as the penetration length into the the reaction square. Reaction distance should be larger than the mean displacement traveled with one step. Otherwise the time of one particle staying within the reaction radius and thus the reaction rate would have large errors due to finite time step errors. In this simulation setup with $\sigma = 1$ and $\Delta t = 0.05$ and regardless of α the particles on average travel for $n = 20$ steps $\sqrt{6K_\alpha n \Delta t^\alpha} \sigma$ to overcome the distance σ . Further the reaction probability as approximated in eq. (2.20) should be a lot smaller than $P(\Delta t) = \lambda \Delta t \ll 1$. Otherwise that approximations fails to describe the correct reaction probability $P(t) = 1 - \exp(-\lambda t)$. Simulations for various $\lambda_+, \lambda_-, \lambda_c$, box size l and α were performed. The results are discussed in the next two sections.

3.3 Results for Normal Diffusion

First simulations with normal diffusion were set up to have a connection to the results from the Smoluchowski and Erban-Chapman problem. The difference in

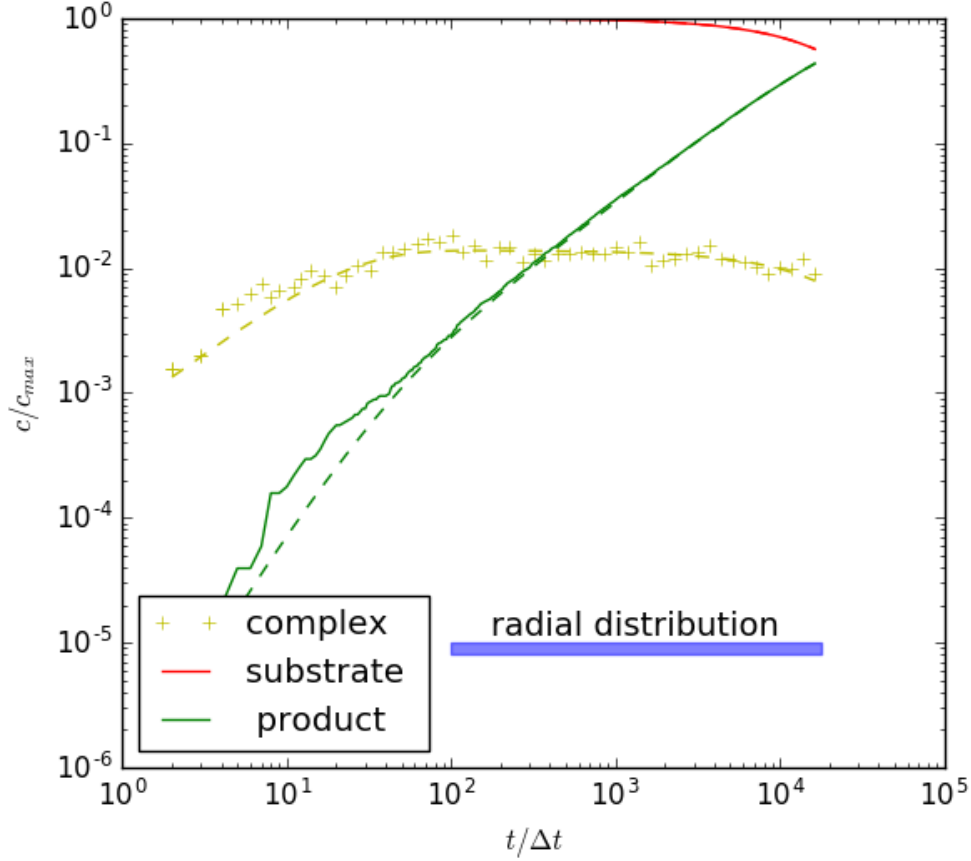


Figure 3.1: Normalised concentrations of substrate, product and complex and their fit with the quasi-steady state approximation for the substrate concentration are shown in the double logarithm plot. The blue bar in the plot indicates the area where the radial distribution was observed, which is plotted in fig. 3.2. Parameters were set to: $\lambda_+ = 0.1$, $\lambda_- = 0$, $\lambda_c = 1.0$ and the box size $l^3 = 8^3$.

kinetics of the Michaelis Menten mechanism to the Erban-Chapman problem is a possible blocking of the enzyme. Thus slowing down the reaction speed and a higher local concentration of the substrates in close proximity to the enzyme because of the back reaction. At first a reaction similar to the Erban-Chapman problem was aimed to simulate. Therefore $\lambda_- = 0$ was chosen to neglect the effect of a higher local concentration of substrates and $\lambda_+ > \lambda_c$ to neglect enzyme blocking. Thereafter the normal MM mechanism was studied.

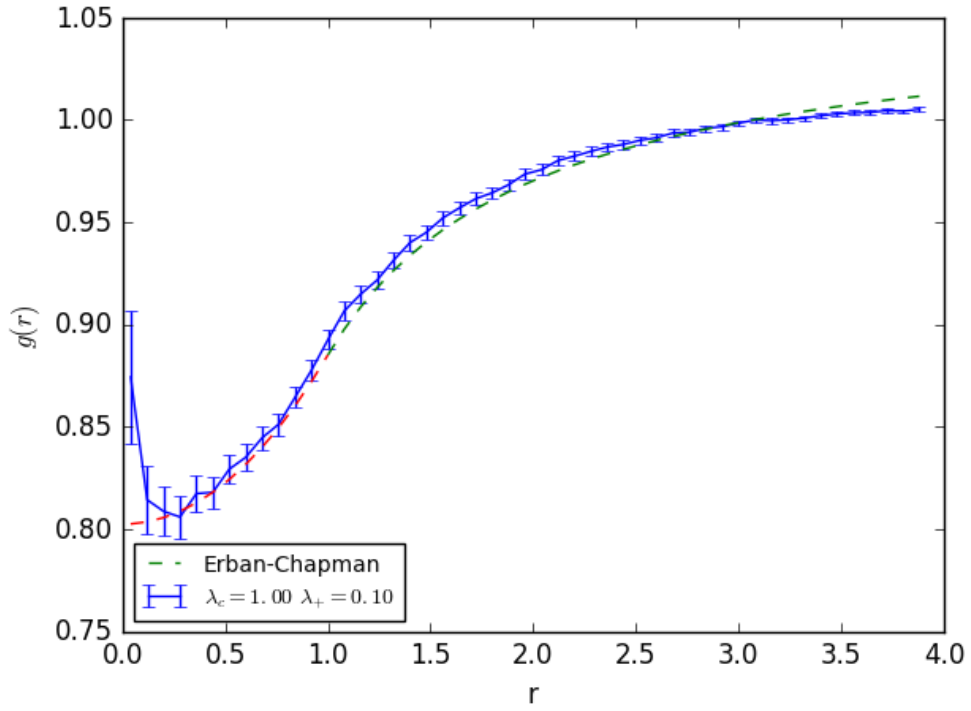


Figure 3.2: The radial distribution function of the substrates around the enzyme / complex was observed for $5 < t < 900$ and plotted as a blue line with an error-bar of the standard deviation. The Erban-Chapman result was plotted as a green dashed line for bigger than the reaction radius σ and a red dashed line for smaller than σ . Parameters were set to: $\lambda_+ = 0.1$, $\lambda_- = 0$, $\lambda_c = 1.0$ and the box size $l^3 = 8^3$.

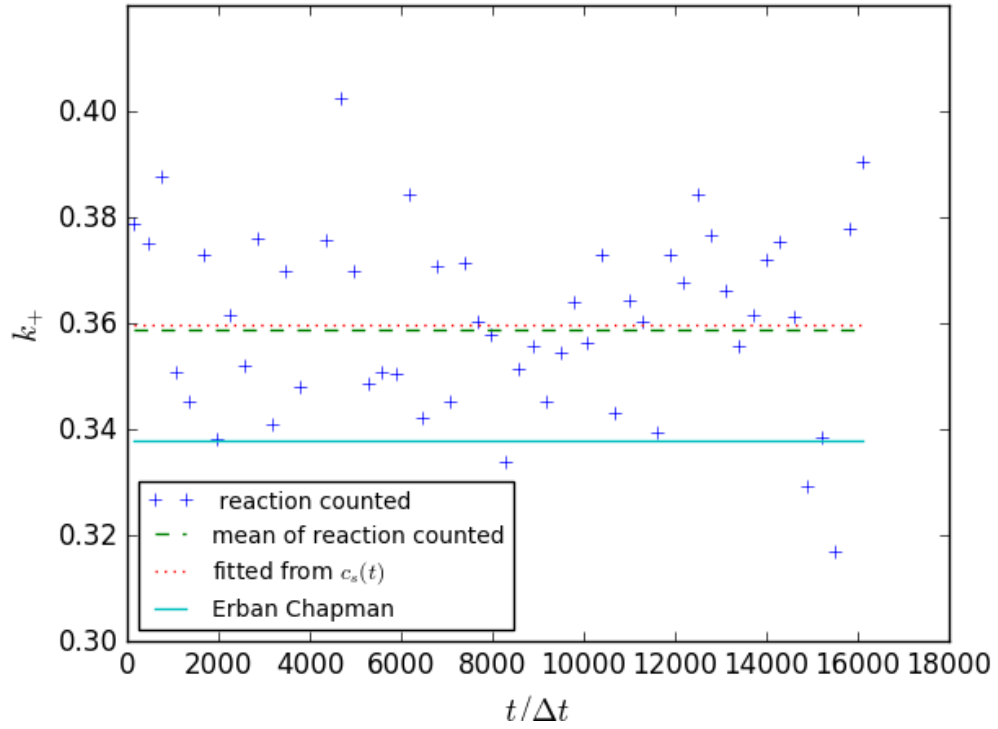


Figure 3.3: A comparison of k_+ from three different calculations. k_+ was calculated from the number of reactions with eq. (3.11) and marked as blue crosses. The Mean of this k_+ is plotted as a green dashed line. The result for k_+ from the fit of the substrate concentration eq. (3.4) with the QSSA is plotted as a red dotted line and the result from the Erban-Chapman eq. (2.17) problem is plotted as a solid cyan colored line. Parameters were set to: $\lambda_+ = 0.1$, $\lambda_- = 0$, $\lambda_+ = 1.0$ and the box size $l^3 = 8^3$.

3.3.1 Results for One-way Reactions

The first simulations was performed under normal diffusion $\alpha = 1$ and with $\lambda_- = 0$ in a simulation box of size $l^3 = 8^3$. The reaction mechanism results in:



$\lambda_+ = 0.1$ was set to be the rate limiting step in the reaction chain and $\lambda_c = 1.0$ was set to reproduce approximately the Erban-Chapman problem. The average normalized number of substrates $c_S(t)/c_{S_0}$, the normalized number of products $c_P(t)/c_{S_0}$ and the averaged number of complexes $c_{ES}(t)$ (over 2544 runs) were observed and visualized in a loglog plot in fig. 3.1. Every run had a new random (uniformly distributed) initial position of the substrate molecules. The simulated concentrations are shown as solid lines. The curves were fitted to obtain k_+ with the quasi-steady state approximation from eq. (3.4). The microscopic reaction rates and macroscopic reaction rate for a uni-molecular reaction should be similar $\lambda_c = k_c$. With all macroscopic rates available, the MM kinetics with the quasi-steady state approximation for the substrate were plotted from eqs. (3.4) to (3.7) as dashed lines in fig. 3.1. A blue bar in the plot indicates the time during which the radial distribution was observed. It was assumed that after 20 steps a stationary distribution was reached. The assumption was based on an approximated constant average number of complexes from there on. The radial distribution of the substrates around the enzyme/complex from 0 to the boundary of the simulation box $0 < r < 4.0$ was plotted in fig. 3.2. It was obtained by averaging over a normalized histogram of substrate particles. The Erban-Chapman solution eqs. (5.19) and (5.20) was plotted with dashed lines. The Erban-Chapman radial distribution is reproducing with good accuracy the observed radial distribution obtained by averaging over the histograms. For a better agreement substrate particles should be placed at the boundaries during the simulation to get a stationary concentration. The macroscopic reaction rate k_+ was calculated from the number of reactions $N_{reaction}$ of the reaction $S + E \xrightleftharpoons{k_+} ES$. k_+ can be calculated from the reaction number change as:

$$k_+ = \frac{dN_{reaction}}{dt} \frac{1}{V c_{S_0} c_{E_0}} \quad (3.11)$$

V is the volume of the simulation box. Numerically the difference of the average reaction number was calculate. k_+ was very noise. It could be smoothed by averaging over 300 time steps. The result was plotted in fig. 3.3 as blue crosses. The average of these blue crosses was plotted as a dashed green line. The result for k_+ from the fit of the substrate concentration eq. (3.4) with the QSSA was plotted as a red dotted line and the result form the Erban-Chapman eq. (2.17) problem was plotted as a solid cyan colored line.

On the other hand a setup where reaction kinetics are limited by diffusion was tried to setup. Therefore $\lambda_c = 1.0$, $\lambda_- = 0$ and the box size $l^3 = 8^3$ were set fixed and $1 \leq \lambda_+ \leq 3.0$ scanned until diffusion becomes the rate limiting step. The macroscopic rates were obtained by fitting the substrate concentration with

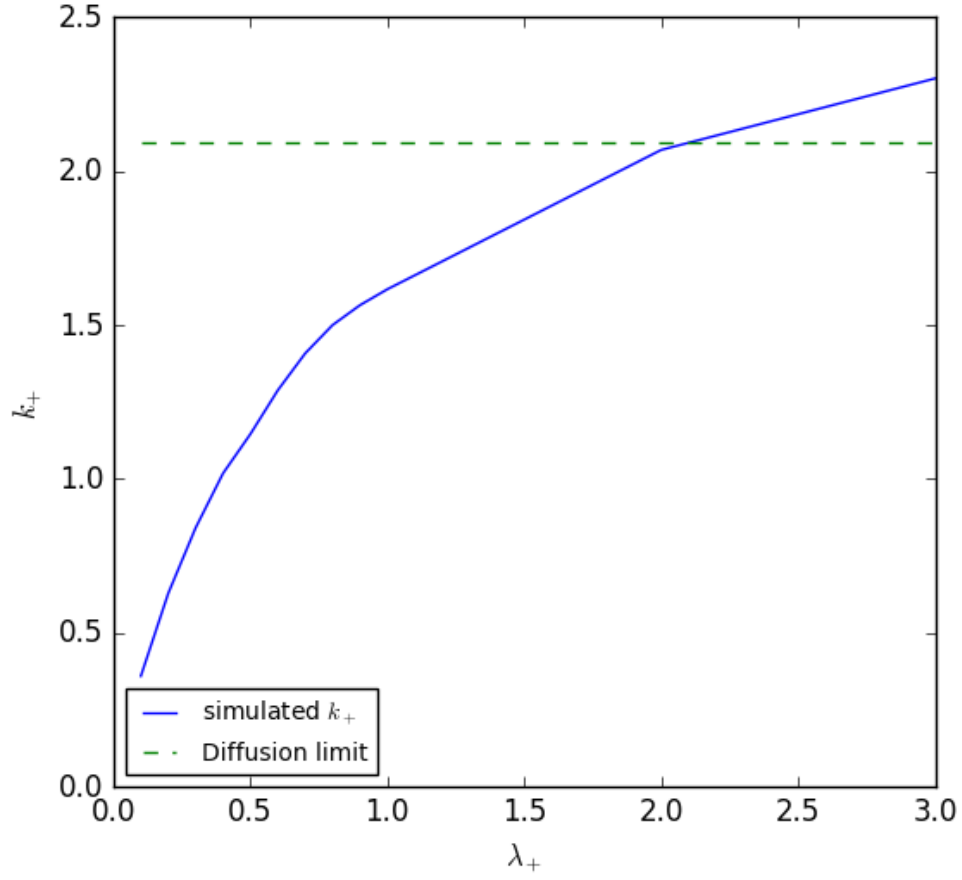


Figure 3.4: Parameters were set to: $\lambda_c = 1.0$, $\lambda_- = 0$ and the box size $l^3 = 8^3$. The blue solid line shows macroscopic rates k_+ in respect to microscopic rate $1 \leq \lambda_+ \leq 3.0$. The macroscopic rates are obtained with a fit of the substrate concentration eq. (3.4) with the QSSA. The diffusion limit eq. (2.14) is plotted as the dashed blue line.

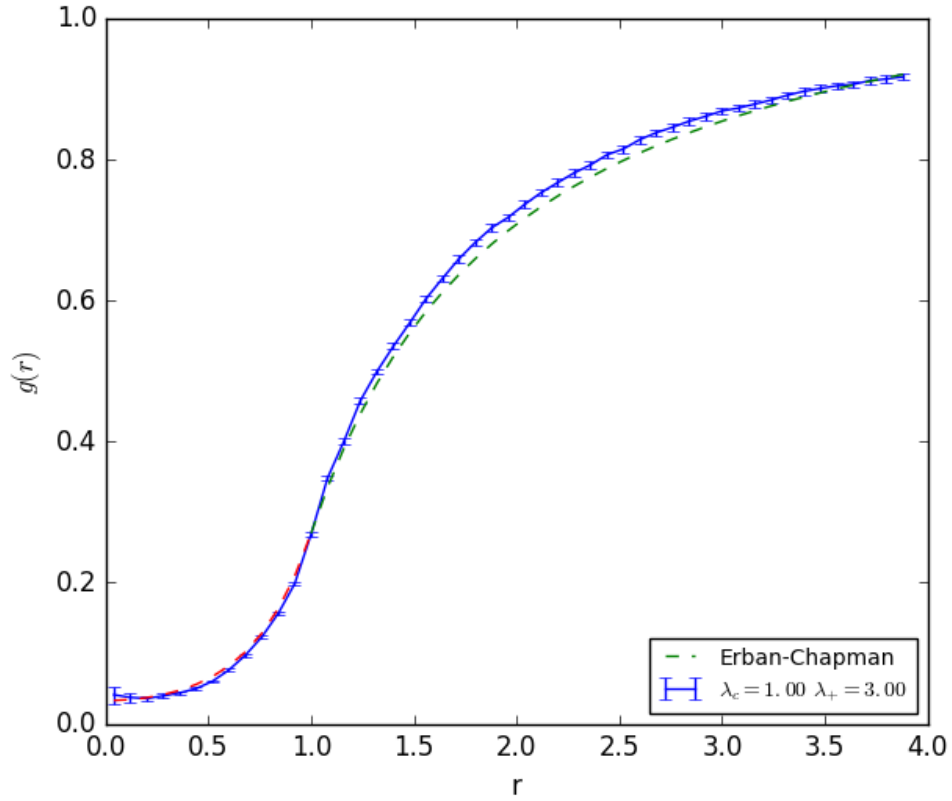


Figure 3.5: The radial distribution function of the substrates around the enzyme / complex was observed for $5 < t < 900$ and plotted as a blue line with an error-bar of the standard deviation. The Erban-Chapman result was plotted as a green dashed line for bigger than the reaction radius σ and a red dashed line for smaller than σ . Parameters were set to: $\lambda_+ = 3.0$, $\lambda_- = 0$, $\lambda_c = 3.0$ and the box size $l^3 = 8^3$.

eq. (3.4). The fastest possible reaction should be limited by eq. (2.14) to be $k_+ = 2.094$. Even higher reaction rates were observed for $\lambda_+ = 2.0$ and 3.0 possibly because no extra particles were placed at the boundaries to indeed have a quasi stationary distribution. The radial distribution function was again obtained by a histogram observed for $5 < t < 900$ and compared with good agreement to the Erban-Chapman result in fig. 3.5. Even for the closest scenario to reactions being limited by diffusion the quasi steady state approximation still should hold because of $0.002 \approx c_{E_0} \ll K_M \approx 0.04$ as stated by [36].

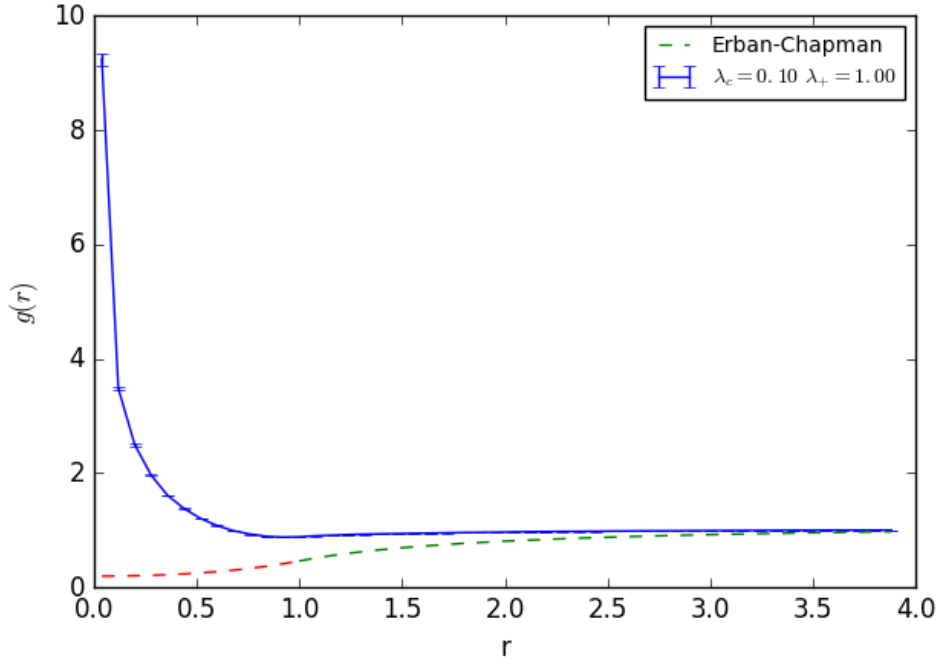


Figure 3.6: The radial distribution function of the substrates around the enzyme / complex was observed for $5 < t < 900$ and plotted as a blue line with an error-bar of the standard deviation. The Erban-Chapman result was plotted as a green dashed line for bigger than the reaction radius σ and a red dashed line for smaller than σ . Parameters were set to: $\lambda_+ = 1.0$, $\lambda_- = 1.0$, $\lambda_c = 0.1$ and the box size $l^3 = 8^3$.

3.3.2 Results for the MM mechanism

The first scenario of the MM mechanism was $\lambda_- > \lambda_c$ and $\lambda_+ > \lambda_c$ to have a local higher concentration of substrates around the enzyme. The radial distribution function was plotted in fig. 3.6 with the Erban-Chapman result as a reference. Determining $k_{+fit} = 4.34$ from the fit of eq. (3.4) results in a lot higher reaction rate than the Erban Chapman result $k_{+erban} = 1.25$. It is the result of the high local concentration of substrates around the enzyme. eq. (3.4) however, is based on the mass action law with homogeneously distributed substrates. The fits from the quasi steady state approximation with k_{+fit} and the average and not local concentration are still in well agreement with the data from the simulation. This can be seen in x-log plot in fig. 3.7. Also the method of counting reactions eq. (3.11) is based on an average concentration and results in similar macroscopic reaction rate $k_{+count} = 4.39$.

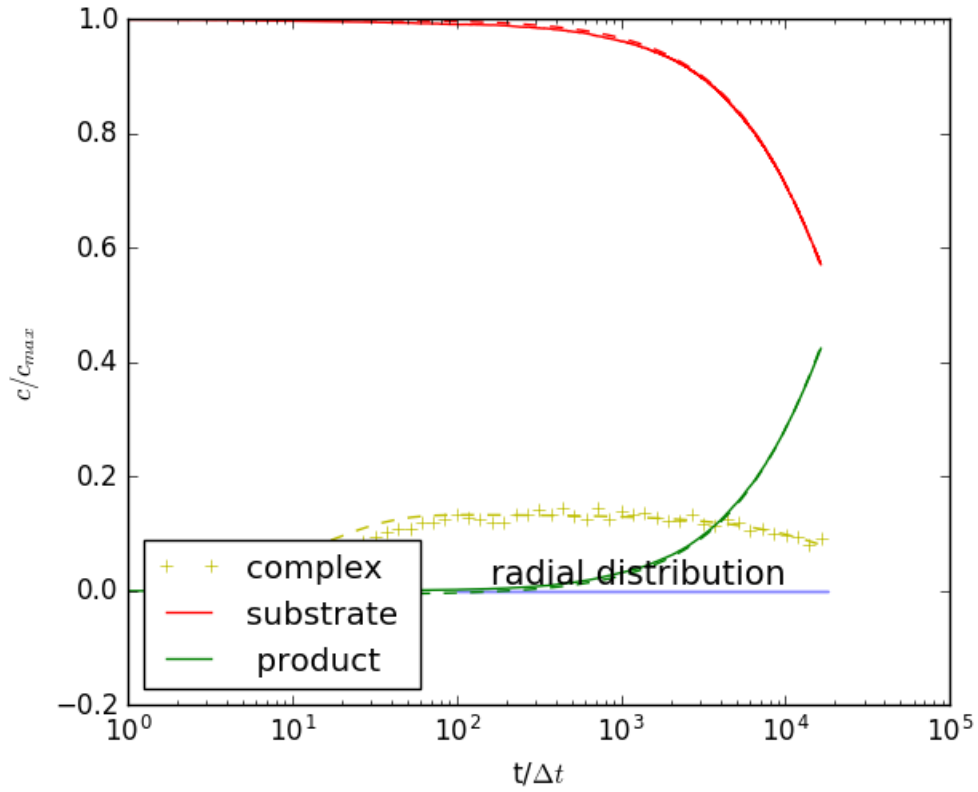


Figure 3.7: Normalized concentrations of substrate, product and complex and their fit with the quasi-steady state approximation for the substrate concentration are shown. The blue bar in the plot indicates the area where the radial distribution was observed, which is plotted in fig. 3.2. The remaining Parameters were set to: $\lambda_+ = 1$, $\lambda_- = 1$, $\lambda_c = 0.1$ and the box size $l^3 = 8^3$.

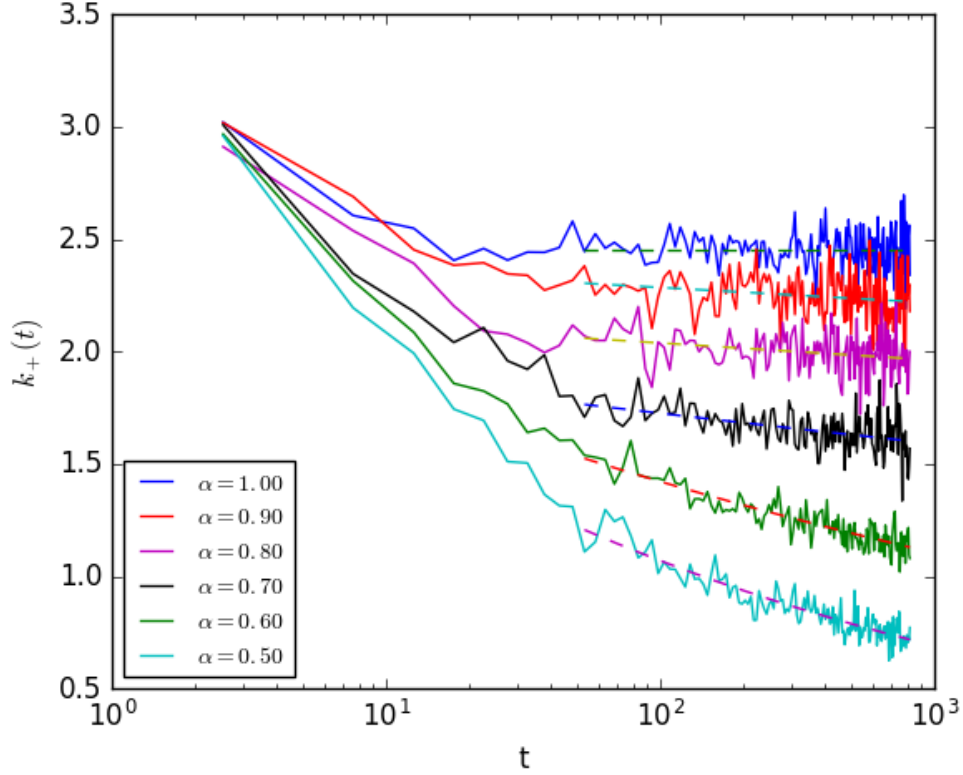


Figure 3.8: Form the number of reactions $k_{+count}(t)$ was calculated via eq. (3.11) and displayed as solid lines for $0.5 \leq \alpha \leq 1.0$. Witheq. (3.12) the fractal like kinetics exponent h was fitted and displayed as dashed lines in the plot. The remaining parameters were set to: $\lambda_+ = 1.0$, $\lambda_- = 1$, $\lambda_c = 1.0$ and the box size $l^3 = 8^3$.

3.4 Results for fBm

The mass action law does not hold anymore for reactions with fractional Brownian motion. Kopelman suggested for percolation cluster, in which particles diffuse anomalously, to have time depended macroscopic reaction rates for bi-molecular reaction of the form [39]:

$$k(t) = k_0 t^{-h} \quad \text{for} \quad 0 \leq h \leq 1 \quad \text{and} \quad t \geq 1 \quad (3.12)$$

to do: check how anomalous diffusion is related to fractal like kinetics factor h fig. 3.8. Can be shown that $k(t)$ indeed describes the kinetics of MM with fractional Brownian motion.

3.5 Conclusion

The simulation model is capable of reproducing Smoluchwoskis and Erban-Chapman theoretical results for a Bi-molecular reaction. The simulation model can be described by Michaelis Menten kinetics for the quasi steady state approximation for normal diffusion but not for anomalous diffusion. Michaelis Menten reactions in a diffusion controlled regime with fractional Brownian motion are of fractal type. The modeling of biological reaction-diffusion dynamics with fBm can be interesting in a large system with a high number of particles not involved in the reaction mechanism but a small number of reactants, which diffuse anomalously. The large number of particles not have to be simulated explicitly. There influence on diffusion of the reactants is incorporated in fBm. The missing particle-particle interactions might get negligible with a low density of particles.

4 Summary

In cells, reaction diffusion simulations are particularly challenging. On one hand, large numbers of macromolecules densely packed causes classical particle-based simulations to reach computational limits very early. On the other hand, small molecular abundances of particular species cause simulations based on a "well-mixed assumption" to be inaccurate, although computationally favourable. For a specific reaction diffusion system, only specific reactions are of interest. Most of the molecules thereby are not involved in reactions but induce anomalous diffusion. The aim of this thesis is to tackle the impact of anomalous diffusion on reaction diffusion systems.

Diffusion was chosen to be modelled by fractional Brownian motion. This approach approximates the influence of the crowded medium by long-range correlations of the trajectory increments. First, theoretical foundations for fBm were set. Most importantly auto-covariance functions for fractional Brownian noise and motion have been introduced. Subsequently, fBm generating algorithms have been studied in terms of accuracy and performance. They rely on Wiener-Khintchin's theorem and theses auto-covariance functions. A naive, Cholesky, Davis-Harte and Lowen's fBm generating algorithm were implemented and analysed in terms of accuracy and performance. Our implementation of Lowen's algorithm turned out not to be exact as claimed by the author[14]. A slightly modified version of Lowen's algorithm was introduced. The modified version turned out to overcome the problem of a non-Gaussian distribution. Davis-Harte and modified Lowen algorithms seem to be both, exact and fast $\mathcal{O}(M \log(M))$ with the trajectory length (M).

Theoretical foundations for Bi-molecular reactions have been introduced. Theses elaborations motivated and verified the underlying particle based simulation model for Bi-molecular reactions in RevReaDDy. However, the theory is not applicable for fBm. No prognoses,

The algorithms analysis suggest Davis-Harte and Lowen algorithm to be well suited for an implementation in RevReaDDy. Both algorithms generate all increments beforehand, it caused the major modification to RevReaDDy. In the process of implementation further modifications to RevReaDDy were necessary to address.

Besides the implementation of fBm many modifications to RevReaDDy were done to deal with our set of questions. A new reaction type and new observables were implemented. Some old observables were modified. A simulation setting with one enzyme and 20 randomly distributed substrates for the Michaelis Menten mechanism was set up and studied under normal Brownian motion and under fractional Brownian motion. Theoretical results from the Smoluchowski problem of a diffusion controlled bi-molecular reaction and theoretical results for a bi-molecular reaction with microscopic reactions within close proximity to the enzyme could be reproduced by the simulation. It could be shown that the quasi steady-state assumption is in-

deed reproducing the kinetics of the Michaelis Menten mechanism for our setting, although this approach is not considering a non uniform spatial distribution.

For fractional Brownian motion with $\alpha < 1$ and $t < 1$ a general slowing down of reactions compared to normal diffusion was shown. Reaction rates are time depended for MM with fBm if reaction speed is controlled by diffusion, as Kopelman suggested for particles performing reactions in a percolation cluster ($k(t) = k_0 t^{-h}$). The kinetics are of fractal type.

As an outlook of the thesis a mathematical model for fractal Brownian motion and Michaelis menten mechanism for 3D could be developed. FBm with forces could be studied. Although a algorithm with possible potentials included will probably be very slow.

5 Appendix

5.1 From Central Limit Theorem to Gaussian Distribution

In the following the Central Limit Theorem will be applied to calculate the distribution of Y , $\rho(y)dy = P(y < Y < y + dy)$ in the limit of large N , with Y being defined as the sum of a random variable:

$$Y = \frac{1}{\sqrt{N}} \sum_{j=1}^N X_j \quad (5.1)$$

The Generating function for a random variable Y is:

$$G_Y(k) = \langle e^{ikY} \rangle = \int e^{ikY} \rho(y) dy \quad (5.2)$$

eq. (5.1) can be inserted into the generating function, which results in:

$$\begin{aligned} G_Y(k) &= \langle e^{\frac{ik}{\sqrt{N}} \sum_{j=1}^N X_j} \rangle \\ G_Y(k) &= \langle \prod_{j=1}^N e^{\frac{ik}{\sqrt{N}} X_j} \rangle \end{aligned}$$

If all X_j are independent, then:

$$\begin{aligned} G_Y(k) &= \prod_{j=1}^N \langle e^{\frac{ik}{\sqrt{N}} X_j} \rangle = e^{\sum_{j=1}^N A_j(\frac{k}{\sqrt{N}})} \\ \text{with } A_j(\frac{k}{\sqrt{N}}) &= \ln \langle e^{\frac{ik}{\sqrt{N}} X_j} \rangle \end{aligned} \quad (5.3)$$

For large N behavior, we assume $\frac{k}{\sqrt{N}} \ll 1$ and expand

$$A_j(\frac{k}{\sqrt{N}}) = \ln(1 + \langle X_j \rangle \frac{ik}{\sqrt{N}} - \langle X_j^2 \rangle \frac{k^2}{2N} + \mathcal{O}(N^{-\frac{3}{2}})) \quad (5.4)$$

with a finite variance $\sigma_i^2 = \langle X_i^2 \rangle$ and the mean $\langle X_i \rangle = 0$

$$A_j(\frac{k}{\sqrt{N}}) = -\sigma_j^2 \frac{k^2}{2N} + \mathcal{O}(N^{-\frac{3}{2}}) \quad (5.5)$$

Thus, the generating function for large N is:

$$G_Y(k) = e^{-\frac{\sigma^2 k^2}{2}} \quad (5.6)$$

with $\sigma = \frac{1}{N} \sum_{j=1}^N \sigma_j^2$

The distribution of Y can be calculated via the inverse Fourier Transform:

$$\rho(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\frac{\sigma^2 k^2}{2}} e^{iky} dk \quad (5.7)$$

$$= \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{y^2}{2\sigma^2}} \quad (5.8)$$

$\rho(y)$ results in a Gaussian distribution.

5.2 From Gaussian Distribution to Gaussian Transition Probability

The conditional distribution function to be in x at time t if visited position y at time s can be written due to Bayes' theorem as a transition probability from y to x in time $t - s$ multiplied with the probability to be in y at time s :

$$\rho_{t,s}(x, y) = T_{t-s}(x|y)\rho_s(y) \quad (5.9)$$

Further due to particle conservation another relation holds:

$$\rho_t(x) = \int \rho_{t,s}(x, y) dy \quad (5.10)$$

Having an initial condition $\rho_s(x) = \delta(x - y)$:

$$\rho_{t,s}(x|y) = \int \rho_{t,s}(x, y) dy = \int T_{t-s}(x|y)\rho_s(y) dy \quad (5.11)$$

$$= \int T_{t-s}(x|y)\delta(x - y) dy = T_{t-s}(x|y) \quad (5.12)$$

5.3 Einstein Formula

The derivative of the mean variance of the Gaussian distribution in respect to time is defined as:

$$\frac{d}{dt} \delta \mathbf{r}^2(t) = \frac{d}{dt} \langle \Delta \mathbf{R}^2(t) \rangle = \frac{d}{dt} \int d\mathbf{r} \mathbf{r}^2 c(\mathbf{r}, t) = \int d\mathbf{r} \mathbf{r}^2 \frac{\partial}{\partial t} c(\mathbf{r}, t) \quad (5.13)$$

Fick's second law can be applied:

$$= D \int_{-\infty}^{\infty} d\mathbf{r} \mathbf{r}^2 \Delta c(\mathbf{r}, t) \quad (5.14)$$

Assuming a reasonable assumption $c(\pm\infty, t) = 0$ and two times partial integration one can derive:

$$= -2D \int_{-\infty}^{\infty} d\mathbf{r} \mathbf{r} \nabla c(\mathbf{r}, t) \quad (5.15)$$

$$= 2Dd \int_{-\infty}^{\infty} d\mathbf{r} c(\mathbf{r}, t) = 2dD \quad (5.16)$$

For the initial condition $\mathbf{r}(0) = 0$, one gets the Einstein Formula: $\langle (\mathbf{r}(t) - \mathbf{r}(0))^2 \rangle = 2dDt$

5.4 Erban-Chapman

The calculation can be started from eq. (2.8). This derivation can be found in [27]. The stationary distribution can be written as:

$$\left(\frac{\partial}{\partial r} + \frac{2}{r} \right) D \frac{\partial \rho_t(r)}{\partial r} = 0 \quad \text{for} \quad r \geq \sigma \quad (5.17)$$

$$\left(\frac{\partial}{\partial r} + \frac{2}{r} \right) D \frac{\partial \rho_t(r)}{\partial r} = \lambda_+ \rho_t(r) \quad \text{for} \quad r \leq \sigma \quad (5.18)$$

For $r \leq \sigma$ a production term had been added to the continuity equation. The general solution can be written as:

$$\rho_t(r) = a_1 + \frac{a_2}{r} \quad \text{for} \quad r \geq \sigma \quad (5.19)$$

$$\rho_t(r) = \frac{a_3}{r} \exp \left[r \sqrt{\frac{\lambda_+}{D}} \right] + \frac{a_4}{r} \exp \left[-r \sqrt{\frac{\lambda_+}{D}} \right] \quad \text{for} \quad r \leq \sigma \quad (5.20)$$

with a_1, a_2, a_3, a_4 real constants. Just like in the calculation in diffusion limit the distribution at $r \rightarrow \infty$ is defined as the concentration of particle A $\rho_t(r \rightarrow \infty) = C_A$. Thus $a_1 = C_A$ and due to the continuity of the density distribution, $a_3 = -a_4$. Further, due to the continuity of the flux $j_s(r)$ at $r = \sigma$ also a_2 and a_3 can be determined:

$$a_2 = C_A \left(\sqrt{\frac{D}{\lambda_+}} \tanh \left(\sigma \sqrt{\frac{\lambda_+}{D}} \right) - \sigma \right) \quad (5.21)$$

$$a_3 = \frac{C_A \sqrt{\frac{D}{\lambda_+}}}{2 \cosh \left(\sigma \sqrt{\frac{\lambda_+}{D}} \right)} \quad (5.22)$$

The flux at $r = \sigma$ can be written as:

$$j^s(\sigma) = D \frac{\partial \rho_t(\sigma)}{\partial \sigma} = -D \frac{a_2}{\sigma^2} \quad (5.23)$$

Gauss's theorem states that the negative flux of a quantity through a closed surface is equal to the production of that quantity inside the volume. Further The change of complex AB is proportional to the change of A inside the volume of the sphere with $r = \sigma$. The total flux and thus the change of AB can be written as:

$$\frac{dC_{AB}}{dt} = - \int_{\partial V} j^s(\sigma) da = \int_{\partial V} D \frac{a_2}{\sigma^2} da = -4\pi D a_2 C_B \quad (5.24)$$

$$= 4\pi D C_B C_A \left(\sigma - \sqrt{\frac{D}{\lambda_+}} \tanh \left(\sigma \sqrt{\frac{\lambda_+}{D}} \right) \right) \quad (5.25)$$

5.5 Michaelis-Menten Kinetics

Michaelis-Menten kinetics are describing the following reaction:

$S + E \xrightleftharpoons[k_-]{k_+} ES \xrightarrow{k_c} P + E$. A set of differential equations can be formulated as a result of mass conservation and the mass action law:

$$\frac{dc_P(t)}{dt} = k_c c_{ES}(t) \quad (5.26)$$

$$\frac{dc_E(t)}{dt} = k_c c_{ES}(t) + k_- c_{ES}(t) - k_+ c_E(t) c_S(t) \quad (5.27)$$

$$\frac{dc_S(t)}{dt} = k_- c_{ES}(t) - k_+ c_E(t) c_S(t) \quad (5.28)$$

$$\frac{dc_{ES}(t)}{dt} = -k_c c_{ES}(t) - k_- c_{ES}(t) + k_+ c_E(t) c_S(t) \quad (5.29)$$

With a quasi-steady-state approximation: $\frac{dc_{ES}(t)}{dt} = 0$ one can reduce labelquasis-teadymichaelis to: $k_+ c_E(t) c_S(t) = k_- c_{ES} + k_c c_{ES}(t)$. A Rearrangement of this equation results in the Michaelis-Menten constant:

$$K_M = \frac{k_- + k_c}{k_+} = \frac{c_E(t) c_S(t)}{c_{ES}}. \quad (5.30)$$

From mass conservation law one gets: $c_E(t) = c_{E_0} - c_{ES}(t)$ After inserting this equation into the quasi-steady-state approximation of eq. (5.29), one gets:

$$c_{ES}(t) = \frac{c_{E_0} c_S(t)}{K_M + c_S(t)} \quad (5.31)$$

The combination of this eq. with the first differential eq. (5.27) defines the velocity

of the reaction:

$$v(c_S(t)) = \frac{dc_P(t)}{dt} = k_c \frac{c_{E_0} c_S(t)}{K_M + c_S(t)} = v_{max} \frac{c_S(t)}{K_M + c_S(t)} \quad (5.32)$$

Bibliography

- [1] A. P. Minton: *How can biochemical reactions within cells differ from those in test tubes?*, Journal of cell science **119**, 2863 (2006).
- [2] F. Höfling and T. Franosch: *Anomalous transport in the crowded world of biological cells*, Reports on Progress in Physics **76**, 046602 (2013).
- [3] B. B. Mandelbrot and J. W. Van Ness: *Fractional Brownian Motions, Fractional Noises and Applications*, SIAM Review **10**, 422 (1968).
- [4] L. Michaelis and M. L. Menten: *Die kinetik der invertinwirkung*, Biochem. z **49**, 352 (1913).
- [5] R. Huang, I. Chavez, K. M. Taute, B. Lukić, S. Jeney, M. G. Raizen, and E.-L. Florin: *Direct observation of the full transition from ballistic to diffusive brownian motion in a liquid*, Nature Physics **7**, 576 (2011).
- [6] F. Höfling: *Stochastic processes and correlation functions*, University Lecture (2016).
- [7] A. Fick: *On liquid diffusion*, Poggendorffs Annalen (1855).
- [8] J. Fourier: *theorie analytique de la* (Firmin Didot, 1822).
- [9] A. Einstein: *Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen*, Annalen der Physik **322**, 549 (1905).
- [10] M. von Smoluchowski: *Zur kinetischen theorie der brownschen molekularbewegung und der suspensionen*, Annalen der Physik **326**, 756 (1906).
- [11] H. Qian: *Fractional brownian motion and fractional gaussian noise*, in *Processes with Long-Range Correlations* (Springer, 2003), pp. 22–33.
- [12] D. Ton: *Simulation of fractional Brownian motion*, Master’s thesis (2004).
- [13] J. R. M. Hosking: *Modeling persistence in hydrological time series using fractional differencing*, Water Resources Research **20**, 1898 (1984).
- [14] S. B. Lowen: *Efficient generation of fractional brownian motion for simulation of infrared focal-plane array calibration drift*, Methodology And Computing In Applied Probability **1**, 445 (1999).

- [15] R. B. DAVIES and D. S. HARTE: *Tests for hurst effect*, Biometrika **74**, 95 (1987).
- [16] M. Timmer, J.;Koenig: *On generating power law noise*, Astronomy and Astrophysics **300**, 707 (1995).
- [17] *Fftw*, <http://www.fftw.org/> (????), accessed: 20.11.2016.
- [18] M. Matsumoto and T. Nishimura: *Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator*, ACM Transactions on Modeling and Computer Simulation **8**, 3 (1998).
- [19] P. Horvai, T. Komorowski, and J. Wehr: *Finite time approach to equilibrium in a fractional brownian velocity field*, Journal of Statistical Physics **127**, 553 (2007).
- [20] P. Waage and C. M. Gulberg: *Studies concerning affinity*, Journal of Chemical Education **63**, 1044 (1986).
- [21] R. Erban, J. Chapman, and P. Maini: *A practical guide to stochastic simulations of reaction-diffusion processes* (2007).
- [22] D. T. Gillespie: *Exact stochastic simulation of coupled chemical reactions*, The Journal of Physical Chemistry **81**, 2340 (1977).
- [23] S. Winkelmann and C. Schütte: *The spatiotemporal master equation: Approximation of reaction-diffusion dynamics via markov state modeling*, The Journal of Chemical Physics **145**, 214107 (2016).
- [24] J. Schöneberg, A. Ullrich, and F. Noé: *Simulation tools for particle-based reaction-diffusion dynamics in continuous space*, BMC Biophysics **7** (2014).
- [25] G. Gruenert, B. Ibrahim, T. Lenser, M. Lohel, T. Hinze, and P. Dittrich: *Rule-based spatial modeling with diffusing, geometrically constrained molecules*, BMC Bioinformatics **11**, 307 (2010).
- [26] M. v. Smoluchowski: *Versuch einer mathematischen Theorie der Koagulationskinetik kolloider Lösungen* (????).
- [27] R. Erban and S. J. Chapman: *Stochastic modelling of reaction–diffusion processes: algorithms for bimolecular reactions*, Physical Biology **6**, 046001 (2009).
- [28] M. B. Flegg: *Smoluchowski reaction kinetics for reactions of any order* pp. 1–30 (????).
- [29] J. Schöneberg and F. Noé: *Readdy - a software for particle-based reaction-diffusion dynamics in crowded cellular environments*, PLOS ONE **8**, 1 (2013).

- [30] The HDF Group: *Hierarchical Data Format, version 5* (1997-NNNN), /HDF5/.
- [31] F. A. Abdullah: *Using fractional differential equations to model the michaelis-menten reaction in a 2-d region containing obstacles*, ScienceAsia **37**, 75 (2011).
- [32] H. Berry: *Monte carlo simulations of enzyme reactions in two dimensions: fractal kinetics and spatial segregation.*, Biophysical journal **83**, 1891 (2002).
- [33] S. Schnell and T. E. Turner: *Reaction kinetics in intracellular environments with macromolecular crowding: Simulations and rate laws*, Progress in Biophysics and Molecular Biology **85**, 235 (2004).
- [34] T. E. Turner, S. Schnell, and K. Burrage: *Stochastic approaches for modelling in vivo reactions*, Computational Biology and Chemistry **28**, 165 (2004).
- [35] J.-H. Jeon, A. V. Chechkin, and R. Metzler: *First passage behavior of multi-dimensional fractional brownian motion and application to reaction phenomena*, in *First-Passage Phenomena and Their Applications* (World Scientific Pub Co Pte Lt, 2014), pp. 175–202.
- [36] B. O. Palsson: *On the dynamics of the irreversible michaelis-menten reaction mechanism*, Chemical Engineering Science **42**, 447 (1987).
- [37] S. Schnell and C. Mendoza: *Closed form solution for time-dependent enzyme kinetics*, Journal of Theoretical Biology **187**, 207 (1997).
- [38] S. Yuste and K. Lindenberg: *Subdiffusion-limited reactions*, Chemical Physics **284**, 169 (2002).
- [39] R. Kopelman: *Fractal reaction kinetics*, Science **241**, 1620 (1988).