Non-Markovian Modeling of Molecular Many-Body Dynamics

Dissertation

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

am Fachbereich Physik

der Freien Universität Berlin

vorgelegt von

Cihan Ayaz

Berlin 2022

Erstgutachter: Prof. Dr. Roland R. Netz

Zweitgutachter: Prof. Dr. Frank Noé

Tag der Disputation: 05.07.2023

Selbstständigkeitserklärung

Name: Ayaz
Vorname: Cihan

Ich erkläre gegenüber der Freien Universität Berlin, dass ich die vorliegende Dissertation selbstständig und ohne Benutzung anderer als der angegebenen Quellen und Hilfsmittel angefertigt habe. Die vorliegende Arbeit ist frei von Plagiaten. Alle Ausführungen, die wörtlich oder inhaltlich aus anderen Schriften entnommen sind, habe ich als solche kenntlich gemacht. Diese Dissertation wurde in gleicher oder ähnlicher Form noch in keinem früheren Promotionsverfahren eingereicht.

Mit einer Prüfung meiner Arbeit durch ein Plagiatsprüfungsprogramm erkläre ich mich einverstanden

Berlin, 15. Dezember 2022

_____

Danksagung

Zunächst möchte ich mich bei meinem Doktorvater Prof. Dr. Roland Netz vom Fachbereich Physik der Freien Universität Berlin bedanken. Die Betreuung durch Prof. Dr. Netz war stets erstklassig. Ich werde die interessanten Diskussionen mit ihm sehr vermissen.

Besonderer Dank gilt meinen Koautoren für die tolle Zusammenarbeit, insbesondere Lucas Tepper, Dr. Laura Scalfi, Dr. Ben A. Dalton, Dr. Philip Loche, Dr. Jan O. Daldrop, Florian N. Brünig, Dr. Julian Kappler. Auch meinen Kollegen Sina Zendehroud, Laura Lavacchi, Henrik Kiefer, Anton Klimek, Shane Carlson, Maximillian Weber, Otto Schullian und Amir Abbasi danke ich für die zahlreichen Gespräche und Anregungen.
Außerdem bedanke ich mich bei allen aktuellen und ehemaligen Mitgliedern der AG Netz und bei Annette Schumann-Welde für die angenehme Arbeitsatmosphäre.

Schließlich möchte ich der Liebe meines Lebens, meiner lieben Ehefrau Esra Önal-Ayaz, meinen tiefsten Dank dafür aussprechen, dass sie mich während der gesamten Promotion, selbst in den schwierigsten Phasen, unermüdlich unterstützt und ermutigt hat. Diese Leistung wäre ohne sie nicht möglich gewesen. Ich danke dir.

Zusammenfassung

Ausgehend von den Hamiltonschen Bewegungsgleichungen eines beliebigen, molekularen Mehrkörpersystems, leiten wir mithilfe von Projektionsoperatoren zunächst nicht-Markovsche Modelle in Form von verschiedenen generalisierten Langevin Gleichungen, im Englischen abgekürzt mit GLE, her. Die GLEs sind integro-Differentialgleichungen für Observablen, die beliebige Funktionen von atomistischen Positionen sind. Die Projektionsoperatoren sind so gewählt, dass sowohl nichtlineare Potenziale als auch nichtlineare Reibungseffekte in den GLEs berücksichtigt werden.

Im darauffolgenden Kapitel stellen wir numerische Methoden vor, um nichtlineare Reibungsfunktionen aus Zeitrehen zu bestimmen. Wir demonstrieren die numerische Extraktionsmethode anhand einer Zeitreihe für den Dihedralwinkel eines Butanmoleküls im Wasser, welche durch Molekulardynamik-Simulationen generiert wurde. Aus der Trajektorie berechnen wir mithilfe unserer Methode alle zuvor hergeleiteten GLEs und vergleichen diese. Wir finden für die untersuchte Dihedralwinkeldynamik des Butanmoleküls, dass eine positionsabhängige Masse zu nichtlinearen Memoryeffekten führen kann. Diese Effekte können beseitigt werden, indem man die Massenabhängigkeit des Potenzialterms berücksichtigt.

Im nächsten Teil konzentrieren wir uns auf die sogenannte approximative GLE, in der nichtlineare Reibungseffekte vernachlässigt werden. Wir diskutieren unter welchen Annahmen die approximative GLE aus einer nichtlinearen GLE hervorgeht. Durch die analytische Berechnung der Kramers-Moyal Koeffizienten der approximativen GLE zeigen wir, dass die Fokker-Planck Gleichung die Dynamik eines nicht-Markovschen Systems nicht vollständig beschreibt. Wir extrahieren den Reibungskernel des Polypeptids Alanine_9 aus Molekulardynamik-Simulationen, um die Wichtigkeit von Memoryeffekten bei Faltungsprozessen von Proteinen zu quantifizieren. Wir stellen fest, dass unser GLE-Modell die gemittelten Übergangszeiten solwohl der Faltungs-, als auch der Entfaltungsdynamik sehr gut reproduziert. Auch die Kramers-Moyal Koeffizienten und die mittlere quadratische Verschiebung, mit ausgeprägter anomaler Diffusion, werden durch die GLE sehr gut erfasst. Markovsche Modelle hingegen, die auf Langevin Gleichungen mit nichtlinearer Reibung basieren, sind nicht in der Lage, die Dynamik in beide Richtungen mit der gleichen Genauigkeit wiederzugeben. Hieraus schließen wir, dass eine konsistente Modellierung der Faltungsdynamik von Proteinen Memoryeffekte mitberücksichtigen muss.

Im letzten Teil der Arbeit geht es um die Markovsche Einbettung von GLEs mit nichtlinearer Reibung. Wir führen drei verschiedene Einbettungssysteme ein, die recheneffiziente Computersimulationen nichtlinearer GLEs ermöglichen. Das erste Einbettungssystem erlaubt die Simulation nichtlinearer Memoryeffekte bei konstanter effektiver Masse und, für den Fall, dass die Reibungsfunktion eine nicht verschwindende Komponente bestehend aus einer Deltafunktion der Zeit besitzt. Der Deltaanteil kann im zweiten Einbettungssystem vernachlässigt werden. Diese Einbettung leiten wir aus dem nichtlinearen Zwanzigmodell durch eine Störungsentwicklung her. Das dritte Einbettungssystem erlaubt GLE-Simulationen auch für den Fall, dass, zusätzlich zu einer nichtlinearen Reibungsfunktion, nun auch die effektive Masse von der Reaktionskoordinate abhängt. Diese Einbettung basiert nicht auf einer Näherung des Zwanzigmodells und setzt, ähnlich wie das erste System, einen Deltaanteil in der Reibungsfunktion voraus.

Abstract

Starting from the Hamiltonian equations of motion of an arbitrary molecular many-body system, we first derive non-Markovian models in the form of various generalized Langevin equations (GLEs) using projection operators. The derived GLEs are integro-differential equations for observables that are arbitrary functions of atomistic positions. We construct the projection operators to include nonlinear potentials and nonlinear memory functions in the GLEs. The primary motivation to introduce nonlinear GLEs is to move as much information as possible from the part of the GLE that ends up being modeled by a stochastic process to the deterministic part of the GLE. In this way, we ensure that one loses less information through the stochastic modeling of the exact GLE.

Following this chapter, we present numerical methods to determine nonlinear memory functions from time series data. We demonstrate the numerical extraction method using a trajectory for the dihedral angle of a butane molecule in water generated by molecular dynamics simulations. From the trajectory, we calculate all previously derived GLEs using our method and compare them. For the dihedral angle dynamics of the butane molecule, we find that a position-dependent mass can lead to nonlinear memory effects. This effect can be eliminated by adjusting the mass dependence of the potential term.

In the next part, we focus on the so-called approximate GLE, in which nonlinear memory effects are neglected. We discuss under which assumptions the approximate GLE emerges from a nonlinear GLE. By analytically computing the Kramers-Moyal coefficients of the approximate GLE, we show that the Fokker-Planck equation does not describe the dynamics of a non-Markovian system. We extract the friction kernel of the polypeptide Alanine_9 from molecular dynamics simulations to quantify the importance of memory effects in protein folding. After parameterizing our GLE, we use the Markovian embedding method to simulate the GLE. Our GLE model very well reproduces the mean first passage times of both the folding and unfolding dynamics. The Kramers-Moyal coefficients and the mean square displacement, with pronounced anomalous diffusion, are also very well captured by the GLE. On the other hand, Markovian models based on Langevin equations with nonlinear friction cannot reproduce the dynamics in both directions with the same accuracy. From this, we conclude that consistent modeling of protein folding dynamics must take into account memory effects.

The last part of the thesis is on the Markovian embedding of nonlinear GLEs. We introduce three different embedding systems that allow computationally efficient simulations of nonlinear GLEs. The first embedding system allows the simulation of nonlinear memory effects for a constant effective mass when the memory function has a nonvanishing component consisting of a delta function in time. The delta component can is not necessary for the second embedding system. We derive the second embedding from the nonlinear Zwanzig model by a perturbation expansion. The third embedding system also allows GLE simulations in the case that, in addition to a nonlinear memory function, the effective mass depends on the reaction coordinate. This embedding is not based on an approximation of the Zwanzig model and, like the first system, assumes a delta component in the memory function.

Table of Contents

List of Tables

List of Figures

1 Introduction

A central problem of modern-day physics is to predict the dynamics of many-body systems, whether numerical or analytical. Many-body problems refer to the motion of a large number of interacting particles. The interaction potentials of typical systems are of a form that precludes an analytical solution to the problem, the prime example being the three-body problem [6].

An important example incorporating many-body dynamics on the microscopic scale is the kinetic theory of gases. In 1738, David Bernoulli introduced the kinetic theory of gases to explain the pressure of a gas [84]. Here, the gas consists of many molecules with a size much smaller than the distance between molecules. The gas represents a many-body system in which the interaction is given by frequent, random collisions of the molecules with each other and with the walls of any container.
In 1857, Rudolf Clausius published an improved kinetic theory of gases [25] from which, two years later, James Clerk Maxwell formulated the first-ever statistical law in physics, the Maxwell distribution of molecular velocities [79, 78].
In 1868, Ludwig Boltzmann obtained the same velocity distribution from a theory that applies to a much broader class of systems than ideal gases, statistical mechanics [10]. Therefore, the molecular velocity distribution is referred to as the Maxwell-Boltzmann distribution. One of the essential statements of the Maxwell-Boltzmann distribution is given by

$$\langle v^2 \rangle = k_B T/m. \tag{1.1}$$

Equation (1.1) states that the second moment of the velocity v in a given direction is determined by the thermal energy $k_B T$ divided by the mass of a gas molecule, which is in line with experimental observations [85]. Although the kinetic theory of gases is able to predict the thermodynamic behavior of gases, Brownian motion is considered the first direct evidence of the microscopic nature of matter [95].

In 1827, Robert Brown discovered that the random motion of colloidal particles suspended in a liquid medium is caused by thermal molecular motion, referred to as Brownian motion [11]. Albert Einstein derived the diffusion coefficient of Brownian motion in 1905 using a probabilistic Ansatz within the kinetic theory of gases model, i.e., the motion arises from random collisions of particles [35]. The underlying assumption is that the mass of the colloidal particle (Brownian particle) is much larger than that of the liquid particles. As a consequence, the motion of the slow Brownian particle takes place on much larger time scales than the motion of the fast liquid particles, which is referred to as time scale separation.

Shortly after, in 1908, Paul Langevin developed what is known today as Langevin dynamics [68]. Langevin dynamics represents the advent of Markovian stochastic dynamics in the form of stochastic differential equations. Within Langevin dynamics, Brownian motion is explained in the following way:
The liquid environment influences the effective motion of the suspended Brownian particle

in two ways. First, the random collisions with the liquid particles lead to thermal molecular motion. Second, if one exerts a force on the Brownian particle, it experiences friction due to the same random collisions. The friction force that a spherical particle experiences during its motion through a viscous liquid was derived in 1851 by George Gabriel Stokes [105] referred to as Stoke's law which states that friction is proportional to the particle's velocity. The Langevin equation of Brownian motion can be written in the form

$$m\dot{\mathbf{v}}(t) = -\gamma \mathbf{v}(t) + \boldsymbol{\eta}(t), \tag{1.2}$$

$$\mathbb{E}[\eta_i(t)\eta_j(t')] = \delta_{ij}\, 2\gamma k_B T\, \delta(t-t'), \tag{1.3}$$

where m is the mass of the Brownian particle and γ is the friction that it experiences due to the liquid environment. The random force $\boldsymbol{\eta}(t)$ is the stochastic representation of the force due to the molecules of the liquid and $\mathbb{E}[\,]$ denotes an average over this noise. The amplitude $2\gamma k_B T$ establishes a connection between friction and driving forces since both are due to the same random collisions. Further, the factor $2\gamma k_B T$ ensures the validity of equation (1.1) within the Langevin model.
The delta correlation is a direct consequence of the time scale separation assumption and makes the random force $\boldsymbol{\eta}(t)$ a Markovian stochastic process [109]. The Markov hypothesis fails when time scale separation does not hold, which, in the case of a colloid particle suspended in a liquid medium, occurs when the mass of liquid particles is of the same order as the mass of the colloid particle [107].

In the 1950s, when digital machines using transistors became commercially available, numerical methods emerged as a way to solve many-body problems [37, 3]. In so-called classical molecular dynamics (MD) computer simulations, the classical equation of motion for each particle in the system is numerically integrated [41]. For this purpose, one models the interactions by force fields containing effective interaction potentials [33, 40]. These effective potentials are constructed to reproduce experimental measurements as accurately as possible. A many-body Hamiltonian for N interacting particles of a typical MD force field can be written as

$$H(\{\mathbf{r}_n, \mathbf{p}_n\}_{n=1}^N) = \sum_{n=1}^{N} \frac{\mathbf{p}_n^2}{2m_n} + V(\{\mathbf{r}_n\}_{n=1}^N), \tag{1.4}$$

where V denotes a general interaction potential that depends on particle positions $\mathbf{r}_n$. With the increase in computational power over the decades, even the folding dynamics of fast-folding proteins has become accessible in MD simulations [76]. Nevertheless, generating trajectories in MD simulations is computationally demanding and certain time scales are practically out of reach.

A remedy is provided by stochastic models, such as Langevin dynamics, and coarse-graining methods that can predict the dynamics beyond the time scale available in MD simulations [64]. The aim is to accurately reproduce the statistics of the MD system with only a fraction of the computational effort. In this context, the time scale separation assumption of Langevin dynamics becomes a critical approximation [60].

Non-Markovian stochastic systems in the form of generalized Langevin equations (GLEs) provide a solution to this problem. Different forms of GLEs were derived in the pioneering

works by Nakajima [89], Mori [88] and Zwanzig [113]. The Mori GLE for a colloidal particle suspended in a homogeneous liquid medium is given by

$$m\dot{\mathbf{v}}(t) = -\int_0^t \mathrm{d}s\, \Gamma(t-s)\mathbf{v}(s) + \mathbf{F}(t), \tag{1.5}$$

$$\langle F_i(t), F_j(s)\rangle = \delta_{ij}\, k_B T\, \Gamma(t-s). \tag{1.6}$$

In equation (1.5), the instateneous friction γ from the Langevin equation (1.2) is replaced by a memory kernel $\Gamma(t)$. As the name suggests, the memory kernel introduces a coupling to past values $v(s)$ at time $s \leq t$. The function $\Gamma(t-s)$ is usually maximal at $s = t$ and decreases as $t - s$ rises. The force $\mathbf{F}(t)$ depends on the initial state of the entire system, and $\langle *, * \rangle$ denotes an ensemble average over initial conditions. Thus, in equilibrium, the memory function only depends on the time difference. The GLE in equation (1.5) is an exact result derived from the Hamiltonian in equation (1.4) using projection operators. Solving all equations of motion in a many-body system reduces to solving the GLE for the coordinate of interest. However, $\mathbf{F}(t)$ preserves the complexity of the problem since it is uniquely determined by the initial conditions of the entire system. To circumvent the computation of $\mathbf{F}(t)$, one maps the GLE onto a stochastic system by replacing $\mathbf{F}(t)$ with a random, stochastic process.

To study barrier crossing kinetics in the presence of memory effects, Grote and Hynes introduced an approximate GLE which is essentially the stochastic Mori GLE supplemented by a general potential of mean force term [48, 47, 56] that ensures the correct stationary behavior [21]. Non-Markovian models based on the approximate GLE have proven to be a useful tool [26, 77, 82, 87, 74, 75, 32, 65, 12, 15]. Various data-based methods to compute the memory function in the approximate GLE from time series data have been introduced [51, 106, 16, 59, 26, 70, 5]. Once the memory function is known, Markovian embedding techniques allow for efficient computer simulations of GLE systems [103, 58]. With this, GLE predictions for the dynamics of a reaction coordinate can be obtained.

The validity of the approximate GLE has been the subject of recent discussions [67, 63, 43, 4, 110] leading to the advent of nonlinear friction GLEs. Nonlinear friction GLEs are the central subject of this thesis. A nonlinear friction GLE for an observable $A(t)$ can be written in the form

$$\ddot{A}(t) = F_{\text{eff}}\left[A(t)\right] - \int_0^t \mathrm{d}s\, \Gamma\left[A(s), \dot{A}(s), t-s\right] + F_{\text{R}}(t), \tag{1.7}$$

where F_{eff} is the effective force that typically contains the potential of mean force and F_{R} is the random force. The key quantity in a GLE is the memory kernel Γ. In the case of nonlinear friction GLEs, the memory kernel is a general function of the reaction coordinate A itself. Although the numerical estimation of nonlinear memory kernels is computationally more demanding, nonlinear friction GLEs are preferable over the approximate GLE since the deterministic part contains more information about the reference MD system. The deterministic part refers to everything but the random force $F_{\text{R}}(t)$ in equation (1.7). Therefore, one loses less information once the exact GLE is mapped onto a stochastic integro-differential equation.

One of the main goals of the research performed in this thesis is to shift as much information as possible from the random force into the deterministic part of the GLE. We accomplish this not on heuristic grounds but by exact derivations. Approximations only enter when simulating the GLE, not in the derivation and numerical extraction stage.
Furthermore, we want to study and quantify how accurately the approximate GLE is able to reproduce statistical quantities of an extremely relevant process, namely the folding dynamics of polypeptides.
Finally, we want to address the following problem: Once the parameters of a GLE with nonlinear friction kernels are numerically estimated, how does one use it to obtain a computationally efficient prediction of the reference dynamics, i.e., what kind of Markovian systems allow for a Markovian embedding of GLEs with nonlinear friction kernel?

We start with an introduction to projection operator methods in chapter 2. In addition, this chapter also introduces the notation used throughout the thesis, and important relations are derived. The introduction to projection operators is concluded with two examples, the Mori and Zwanzig projections.

In chapter 3, we apply projection operators to derive three different GLEs with nonlinear friction kernels for observables that are functions of atomistic positions. The first one is the hybrid GLE. Here, the memory function splits up into two terms. One memory term is proportional to the velocity, and the memory kernel depends on time only. This memory contribution resembles the form of the memory friction function in the approximate GLE. The second memory function corresponds to a correction in the presence of nonlinear memory effects. The mean force term in the first two GLEs is identical, but they differ in the structure of the memory function. In the second GLE, the entire memory function is proportional to the velocity and is nonlinear in the reaction coordinate itself. The memory function of the third GLE has the same structure, but the mean force term is adjusted to simulate a GLE with a position-dependent mass.

In chapter 4, we introduce a numerical extraction scheme based on *projected correlation functions* [16, 4] to compute nonlinear friction kernels from time series data. The extraction scheme is demonstrated by computing all memory functions for the three GLEs from chapter 3. The underlying observable is the dihedral angle of a butane molecule in water. The trajectory is generated by all-atom MD simulations.
The results for the hybrid GLE suggest that nonlinear memory effects are non-negligible. This is confirmed by the second GLE. The results from the third GLE, however, suggest that nonlinear memory effects become negligible, at least for the studied dihedral angle dynamics, if the dependency of the mean force on the position-dependent mass is adjusted. Thus, by comparing the results for different nonlinear GLEs, we conclude that position-dependent mass can introduce nonlinear memory effects in certain GLEs.

In chapter 5, we discuss under which conditions the approximate GLE arises from the hybrid GLE. Given the approximate GLE, we introduce a numerical extraction scheme based on Volterra equations to compute the running integral of the memory function from time series data. In order to perform GLE simulations, Markovian embedding of the approximate GLE is presented. Following this, we derive exact formulas for the Kramers-Moyal coefficients of the approximate GLE.

In the main part of chapter 5, we focus on the folding dynamics of the polypeptide Alanine_9. For this, we compute the potential of mean force and the time-dependent friction function, the two ingredients of the approximate GLE, from explicit-water MD simulations for a one-dimensional reaction coordinate based on the native H-bond distances.
Folding and unfolding times from numerical integration of the GLE for the reaction coordinate agree well with MD results, which demonstrates the robustness of our GLE-based non-Markovian model. Memory effects in the friction significantly speed up peptide kinetics and cause anomalous diffusion in configuration space.
In contrast, Markovian models do not accurately describe the peptide kinetics and, in particular, cannot reproduce the folding and unfolding kinetics simultaneously, even if a spatially dependent friction profile is fitted to the data. Our methods are applicable to any reaction coordinate and also to experimental trajectories. Our results demonstrate that a consistent description of protein folding dynamics must account for memory friction effects.

In the final chapter, we present three systems of coupled Markovian Langevin equations which allow us to perform efficient computer simulations of GLEs with nonlinear friction kernels. By making use of the projection operators with which the nonlinear friction GLEs are derived, we demonstrate that the mean dynamics of the Markovian embedding system is equivalent to the dynamics of the GLE, generated by the projection operator. In this way, we ensure that the numerical estimation of the GLE parameters and the Markovian embedding of the GLE are self-consistent.
The first two embedding methods work for GLEs with a constant effective mass. One of the two allows for the embedding of a memory function in which the time component contains a delta contribution and is otherwise a sum of exponentials. In the second embedding system, no delta contribution is needed in the time component which is now a sum of exponentially decaying oscillations.
The third embedding system works for a GLE in which the mass and the friction kernel are configuration-dependent, i.e., general nonlinear functions of the reaction coordinate.
The projection operators used to generate the nonlinear GLEs are incorporated into previously discussed numerical extraction schemes for the data-based reconstruction of GLE parameteters to validate the Markovian embedding schemes.

2 Projection Operator Method

The primary method in this thesis is the projection operator method (POM) [45]. We will use projection operators to derive nonlinear friction generalized Langevin equations (GLEs) and to numerically estimate GLE parameters. Projection operators will also be crucial when introducing Markovian embedding techniques for nonlinear friction GLEs.
In this chapter, we give a concise introduction to projection operators. First, we lay the foundation by introducing the type of many-body system we are considering, and we present the notation used throughout this work with some useful properties.
We proceed with the introduction of projection operators and end the chapter with two important examples, namely the Mori projection and the Zwanzig projection.

2.1 Hamiltonian and Liouville dynamics

We denote the phase space of a system of N interacting particles by Ω. One specific microstate, i.e., a point in Ω, is denoted by ω which is a $6N$ vector given by $\omega = (\mathbf{R}, \mathbf{P}) = (\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N, \mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_N)$ with cartesian positions $\mathbf{r}_n = (r_n^x \; r_n^y \; r_n^z)$ and the corresponding conjugate momenta $\mathbf{p}_n = (p_n^x \; p_n^y \; p_n^z)$ of all $n = 1, 2, \ldots, N$ particles in the system. The Hamiltonian of the system is an invariant of motion and of the form

$$H(\omega) = \sum_{n=1}^{N} \frac{\mathbf{p}_n^2}{2m_n} + V(\mathbf{R}). \tag{2.1}$$

The potential $V(\mathbf{R})$ contains all interactions between the particles and possible external potentials. The only assumption on V is that it is a function of the positions $\mathbf{R}$ only. The time evolution of a point ω in phase space is determined by Hamilton's equation of motion which can be written in the form

$$\dot{\omega}_t = \mathcal{L}\omega_t, \tag{2.2}$$

where ω_t is the location of the system in phase space at time t, given the system was initially at ω_0. For the sake of clarity, we denote time dependencies of phase space coordinates by a subscript. In eq. (2.2), $\mathcal{L}$ is the Liouville operator given by

$$\mathcal{L} = \sum_{n=1}^{N} \left(\frac{\mathbf{p}_n}{m_n} \cdot \nabla_{\mathbf{r}_n} - (\nabla_{\mathbf{r}_n} V) \cdot \nabla_{\mathbf{p}_n} \right). \tag{2.3}$$

All operators we consider in this work, including the Liouville operator $\mathcal{L}$, act on the initial phase space point ω_0. From eq. (2.2), it follows that the system is propagated in time by the operator $e^{t\mathcal{L}}$, i.e., $e^{t\mathcal{L}}\omega_0 = \omega_t$. We consider observables that are real-valued functions of phase space coordinates only and depend on time implicitly via the phase space coordinates. In order to keep the equations as short as possible, we denote the time dependency of such an observable by a subscript too, i.e., $A(\omega_t) = A(\omega_0, t) \equiv A_t$. Using the chain rule for

differentiation, it follows that the time evolution of observable A_t is also governed by the Liouville equation

$$\dot{A}_t = \mathcal{L} A_t. \tag{2.4}$$

Thus, the time propagation operator of an observable $A(\omega_0) \equiv A_0$ is also given by $e^{t\mathcal{L}}$. From this it follows that

$$A(\omega_{t+t'}) = e^{(t+t')\mathcal{L}} A(\omega_0) = e^{t\mathcal{L}} A(\omega_{t'}) = A(\omega_{t'}, t). \tag{2.5}$$

Eq. (2.5) is a property of observables that are propagated in time by $e^{t\mathcal{L}}$ and will be used when deriving the numerical scheme later. All observables of the system are elements of a Hilbert space, i.e., a vector space equipped with an inner product. Let A and B denote two observables of the system. For the inner product, we choose

$$\langle A_t, B_{t'} \rangle = \int_\Omega \mathrm{d}\omega_0 \, \rho_{\mathrm{eq}}(\omega_0) A(\omega_0, t) B(\omega_0, t'), \tag{2.6}$$

where $\rho_{\mathrm{eq}}(\omega_0) = \frac{1}{Z} e^{-\beta H(\omega_0)}$ is the canonical Boltzmann distribution with the inverse thermal energy $\beta = 1/k_B T$ and the partition function $Z = \int_\Omega \mathrm{d}\omega_0 \, e^{-\beta H(\omega_0)}$. The inner product in eq. (2.6) thus corresponds to an equilibrium time correlation function which establishes the link to statistical mechanics. Because of the form of the Hamiltonian in eq. (2.1), the Boltzmann distribution factorizes into a position and a momentum-dependent part

$$\rho_{\mathrm{eq}}(\omega_0) = \frac{1}{Z} e^{-\beta H(\omega_0)} = \rho_{\mathrm{kin}}(\mathbf{P}_0) \, \rho_{\mathrm{pot}}(\mathbf{R}_0), \tag{2.7}$$

where $\rho_{\mathrm{kin}}(\mathbf{P}_0) \propto \prod_{i=1}^{N} \exp\left(-\beta \, \mathbf{p}_{i,0}^2 / 2m_i\right)$ is a Gaussian with zero mean. In addition to averages over the entire phase space Ω, eq. (2.6), we will also use conditional averages, i.e., averages over a hyper surface in phase space on which an observable of choice, A_0, takes a constant value a. A conditional average of an observable B_t is defined by [46, 21]

$$\langle B_t \rangle_{A_0} = \langle B_t \rangle_{A(\omega_0)} = \frac{\langle \delta(A(\widehat{\omega}_0) - A(\omega_0)), B(\widehat{\omega}_0, t) \rangle}{\langle \delta(A(\widehat{\omega}_0) - A(\omega_0)) \rangle}. \tag{2.8}$$

In eq. (2.8), phase space variables with a hat inside inner products, i.e. $\widehat{\omega}_0$, are integrated out and the result depends on the initial condition ω_0.

2.2 Useful properties

In the following, we give some relations which will be used later on. First, the conditional average is a function of the initial value A_0, therefore its time propagation is given by

$$e^{t\mathcal{L}} \langle B_{t'} \rangle_{A_0} = \langle B_{t'} \rangle_{A_t}. \tag{2.9}$$

Further, with respect to the inner product in eq. (2.6), the Liouville operator as defined in eq. (2.3) is anti self-adjoint [115]

$$\langle \mathcal{L} A_t, B_{t'} \rangle = -\langle A_t, \mathcal{L} B_{t'} \rangle. \tag{2.10}$$

Acting with the Liouville operator on the delta function gives [53]

$$\mathcal{L}\delta(A_0 - a) = -\dot{A}_0 \frac{\mathrm{d}}{\mathrm{d}a}\delta(A_0 - a). \tag{2.11}$$

The quantity $\mathbb{P}(a) = \langle \delta(A_0 - a)\rangle$ gives the probability that the observable A_0 has the value a. It can be used to compute the potential of mean force (PMF) [29]

$$U_{\mathrm{PMF}}(a) = -k_B T \ln \mathbb{P}(a). \tag{2.12}$$

Using the definition in eq. (2.8) together with the relations in eq. (2.10), eq. (2.11) and the PMF in eq. (2.12), one finds the important relation

$$\langle \mathcal{L}B_t\rangle_{A_0} = \frac{\mathrm{d}}{\mathrm{d}A_0}\langle \dot{A}_0, B_t\rangle_{A_0} - \beta\langle \dot{A}_0, B_t\rangle_{A_0}\frac{\mathrm{d}}{\mathrm{d}A_0}U_{\mathrm{PMF}}(A_0). \tag{2.13}$$

To obtain eq. (2.13), we start with

$$\langle \mathcal{L}B_t\rangle_{A_0} = \frac{\langle \delta(A(\widehat{\omega}_0) - A(\omega_0)), \mathcal{L}B(\widehat{\omega}_0, t)\rangle}{\langle \delta(A(\widehat{\omega}_0) - A(\omega_0))\rangle}, \tag{2.14}$$

where the average is over variables with a hat and consequently, the Liouville operator $\mathcal{L}$ only acts on variables with a hat. To keep the equations as short as possible, we use $A(\widehat{\omega}_0) = A_0$ and $A(\omega_0) = a$ and consider the r.h.s. in eq. (2.14)

$$\frac{\langle \delta(A_0 - a), \mathcal{L}B_t\rangle}{\langle \delta(A_0 - a)\rangle} = -\frac{\langle \mathcal{L}\delta(A_0 - a), B_t\rangle}{\langle \delta(A_0 - a)\rangle} \tag{2.15a}$$

$$= \frac{\langle \dot{A}_0 \frac{\mathrm{d}}{\mathrm{d}a}\left[\delta(A_0 - a)\right], B_t\rangle}{\langle \delta(A_0 - a)\rangle}, \tag{2.15b}$$

where we used eq. (2.10) and eq. (2.11). We next pull out the derivative with respect to a in eq. (2.15b) from the inner product and use the product rule of differentiation, which gives

$$\frac{\langle \delta(A_0 - a), \mathcal{L}B_t\rangle}{\langle \delta(A_0 - a)\rangle} = \frac{\mathrm{d}}{\mathrm{d}a}\frac{\langle \delta(A_0 - a)\dot{A}_0, B_t\rangle}{\langle \delta(A_0 - a)\rangle} + \frac{\langle \delta(A_0 - a)\dot{A}_0, B_t\rangle}{\langle \delta(A_0 - a)\rangle}\frac{\mathrm{d}}{\mathrm{d}a}\ln\langle \delta(A_0 - a)\rangle. \tag{2.16}$$

Finally, we use the definition of the PMF in eq. (2.12) and the definition of conditional averages in eq. (2.8) to obtain eq. (2.13) from eq. (2.16).

2.3 The projection operator method

We now demonstrate how the equation of motion for an observable A_t can be reformulated as a GLE. A projection $\mathcal{P}$ is a linear, idempotent operator, i.e., for arbitrary scalars $c_1, c_2 \in \mathbb{R}$ and for arbitrary observables A_t, B_t, it fulfills the properties

$$\mathcal{P}(c_1 A_t + c_2 B_t) = c_1 \mathcal{P}A_t + c_2 \mathcal{P}B_t, \tag{2.17a}$$

$$\mathcal{P}^2 A_t = \mathcal{P}A_t. \tag{2.17b}$$

The projection

$$\mathcal{Q} = 1 - \mathcal{P} \tag{2.18}$$

is the projection onto the *complementary* subspace with 1 being the identity operator. From the idempotency, it follows that

$$\mathcal{P}\mathcal{Q} = \mathcal{Q}\mathcal{P} = 0. \tag{2.19}$$

We will refer to the subspace onto which $\mathcal{P}$ projects as the *relevant* subspace. The operators $\mathcal{P}$ and $\mathcal{Q}$ can be used to decompose the Liouville equation $\ddot{A}_t = \mathcal{L}\dot{A}_t$ for the observable $\dot{A}_t$ in the following way

$$\ddot{A}_t = e^{t\mathcal{L}}(\mathcal{P} + \mathcal{Q})\mathcal{L}\dot{A}_0 = e^{t\mathcal{L}}\mathcal{P}\mathcal{L}\dot{A}_0 + e^{t\mathcal{L}}\mathcal{Q}\mathcal{L}\dot{A}_0. \tag{2.20}$$

To obtain an equation of motion for A_t from eq. (2.20), we introduce the operator

$$\Phi(t) = e^{t\mathcal{L}}\mathcal{Q}. \tag{2.21}$$

$\Phi(t)$ propagates in time the part of an observable that lies in the complementary subspace. For $\Phi(t)$ we find

$$\frac{\mathrm{d}}{\mathrm{d}t}e^{t\mathcal{L}}\mathcal{Q} = e^{t\mathcal{L}}\mathcal{L}\mathcal{Q} = e^{t\mathcal{L}}\mathcal{Q}\mathcal{L}\mathcal{Q} + e^{t\mathcal{L}}\mathcal{P}\mathcal{L}\mathcal{Q}, \tag{2.22a}$$

$$\dot{\Phi}(t) = \Phi(t)\mathcal{L}\mathcal{Q} + e^{t\mathcal{L}}\mathcal{P}\mathcal{L}\mathcal{Q}. \tag{2.22b}$$

Eq. (2.22b) is an inhomogenous differential equation of first order. Using $\Phi(0) = \mathcal{Q}$, as follows from eq. (2.21), the solution reads

$$\Phi(t) = \mathcal{Q}e^{t\mathcal{L}\mathcal{Q}} + \int_0^t \mathrm{d}u\, e^{uL}\mathcal{P}\mathcal{L}\mathcal{Q}e^{(t-u)\mathcal{L}\mathcal{Q}}. \tag{2.23}$$

By using $\mathcal{Q}e^{t\mathcal{L}\mathcal{Q}} = e^{t\mathcal{Q}\mathcal{L}}\mathcal{Q}$ and the substitution $s = t - u$ in eq. (2.23), one finds

$$\Phi(t) = e^{t\mathcal{L}}\mathcal{Q} = e^{t\mathcal{Q}\mathcal{L}}\mathcal{Q} + \int_0^t \mathrm{d}s\, e^{(t-s)\mathcal{L}}\mathcal{P}\mathcal{L}e^{s\mathcal{Q}\mathcal{L}}\mathcal{Q}. \tag{2.24}$$

The first term on the right-hand side of eq. (2.24) stays in the complementary subspace for all times. The second term describes the coupling between the complementary subspace and the relevant subspace. By factoring out the operator $\mathcal{Q}$ on the r.h.s. of eq. (2.24), one obtains the *Dyson decomposition* [34, 38, 36] of the propagator $e^{t\mathcal{L}}$

$$e^{t\mathcal{L}} = e^{t\mathcal{Q}\mathcal{L}} + \int_0^t \mathrm{d}s\, e^{(t-s)\mathcal{L}}\mathcal{P}\mathcal{L}e^{s\mathcal{Q}\mathcal{L}}. \tag{2.25}$$

Note that, to obtain eq. (2.25), we assumed that $\mathcal{Q}^{-1}$ exists. However, by taking the time derivative of eq. (2.25), one finds that the left-hand side and the right-hand side both fulfill the differential equation $\dot{\mathcal{O}}(t) = \mathcal{L}\mathcal{O}(t)$ with $\mathcal{O}(0) = 1$ which establishes the equality in eq. (2.25) also for non-invertible $\mathcal{Q}$.

Replacing $e^{t\mathcal{L}}\mathcal{Q}$ in eq. (2.20) by eq. (2.24) leads to a GLE-like equation for A_t in terms of a general projection [113, 88, 115]

$$\ddot{A}_t = e^{t\mathcal{L}}\mathcal{P}\mathcal{L}\dot{A}_0 + \int_0^t \mathrm{d}s\, e^{(t-s)\mathcal{L}}\mathcal{P}\mathcal{L}\, F_{\mathrm{R}}(s) + F_{\mathrm{R}}(t), \tag{2.26a}$$

$$F_{\mathrm{R}}(t) = e^{t\mathcal{Q}\mathcal{L}}\mathcal{Q}\mathcal{L}\dot{A}_0 = \mathcal{Q}e^{t\mathcal{L}\mathcal{Q}}L\dot{A}_0. \tag{2.26b}$$

The function $F_{\mathrm{R}}(t)$ stays in the complementary subspace for all times. It is a function of the initial state of the entire system, i.e., $F_{\mathrm{R}}(t) = F_{\mathrm{R}}(\omega_0, t)$. Hence, for large systems, it is a fluctuating function of time. For the sake of brevity, we will write the ω_0 dependency of $F_{\mathrm{R}}(t)$ only when it improves clarity.

The second term in eq. (2.26a) is due to the coupling of the relevant part, i.e., the subspace onto which $\mathcal{P}$ projects, with the function $\mathcal{L}F_{\mathrm{R}}(\omega_0, t)$. Clearly, the explicit form of eq. (2.26a) depends on the specific form of the projection. Before we introduce projection operators that lead to GLEs with a nonlinear friction kernel, we will shortly revisit the results of Mori and Zwanzig. Hence, we will present two GLEs. One being generated by the Mori projection $\mathcal{P}_M$ and the other by the Zwanzig projection $\mathcal{P}_Z$.

2.3.1 The Mori projection

The Mori projection of an observable A_t is given by [88]

$$\mathcal{P}_M A_t = \langle A_t, \mathbf{B}_0\rangle \cdot \langle \mathbf{B}_0, \mathbf{B}_0\rangle^{-1} \cdot \mathbf{B}_0, \tag{2.27}$$

where $\langle *, *\rangle$ denotes the inner product given in eq. (2.6). The observable one projects onto, i.e., the vector $\mathbf{B}_0 \in \mathbb{R}^n$ in eq. (2.27), is referred to as the *projection function*, or the set of projection functions. The inner product $\langle \mathbf{B}_0, \mathbf{B}_0\rangle$ denotes an $n \times n$ matrix [115]. If one chooses the set of projection functions to be $B_0 \in \mathbb{R}$ and its time derivative $\dot{B}_0$, the Mori projection reads

$$\mathcal{P}_M A_t = \frac{\langle A_t, B_0\rangle}{\langle B_0^2\rangle} B_0 + \frac{\langle A_t, \dot{B}_0\rangle}{\langle \dot{B}_0^2\rangle}\dot{B}_0. \tag{2.28}$$

The projection in eq. (2.28) maps any observable A_t onto the subspace of all functions linear in the observables $B_0, \dot{B}_0$. $\mathcal{P}_M$ is self-adjoint with respect to the inner product in eq. (2.6), i.e., for two arbitrary observables A_t, C_t, the relation

$$\langle \mathcal{P}_M A_t, C_{t'}\rangle = \langle A_t, \mathcal{P}_M C_{t'}\rangle \tag{2.29}$$

holds [115]. Thus, it is an orthogonal projection since all functions $\mathcal{P}_M A$ and $\mathcal{Q}_M C$ are orthogonal, i.e.,

$$\langle \mathcal{P}_M B_t, \mathcal{Q}_M C_{t'}\rangle = 0, \tag{2.30}$$

as follows from eq. (2.29). With $\mathcal{P} \equiv \mathcal{P}_M$ being the Mori projection and choosing the projection function to be $B_t \equiv A_t$, i.e., projecting onto the observable of interest itself, eq. (2.26a) takes the form

$$\ddot{A}_t = -\mathrm{K}\, A_t - \int_0^t \mathrm{d}s\, \Gamma^M(s)\dot{A}_{t-s} + F_\mathrm{R}(t), \tag{2.31a}$$

$$\mathrm{K} = \frac{\langle \dot{A}_0^2 \rangle}{\langle A_0^2 \rangle}, \quad \Gamma^M(t) = \frac{\langle F_\mathrm{R}(t), F_\mathrm{R}(0) \rangle}{\langle \dot{A}_0^2 \rangle}, \tag{2.31b}$$

where $\Gamma^M(s)$ is the memory kernel obtained from the Mori projection. Eq. (2.31) is an exact decomposition of the Liouville equation into three terms:
The first term on the r.h.s. of eq. (2.31a) is a force due to a potential of quadratic form, the second term includes the memory kernel which is related to the fluctuating function $F_\mathrm{R}(t)$ via eq. (2.31b). The exact form of the memory kernel can only be computed for some simple models. In general, it is infeasible to compute for practical applications since the fluctuating term $F_\mathrm{R}(t)$, given in eq. (2.26b), is a function of the entire system.

Instead, one models the function $F_\mathrm{R}(t)$ as a stochastic process with zero mean and the second moment given in eq. (2.31b). There is no further information on higher-order correlations of $F_\mathrm{R}(t)$ obtained from the formalism. Therefore, it is mostly assumed to be a Gaussian process. In general, this assumption will not hold since $F_\mathrm{R}(t)$ contains all nonlinearities in A_t. Thus, imposing $F_\mathrm{R}(t)$ to be a Gaussian process becomes a critical approximation for nonlinear systems. For this reason, $F_\mathrm{R}(t)$ should contain as little information as possible so that its impact on the statistics is as little as possible.

This is the primary motivation to introduce a nonlinear potential term into the GLE that accounts for nonlinearities in A_t and, further, ensures that the mean behavior of the reduced dynamics, as described by the GLE, follows the mean dynamics of the complete system [21].

Before we continue with the Zwanzig projection, it is worth mentioning that the Mori formalism can be extended to systems far from equilibrium. By using a time-dependent distribution instead of the Boltzmann distribution in eq. (2.6), one can obtain a non-equilibrium GLE with a structure similar to equation (2.31) [83, 81].

2.3.2 The Zwanzig projection

Contrary to the Mori projection, the Zwanzig projection $\mathcal{P}_Z$ of an observable A_t is nonlinear in the projection function [113]. This is best explained by choosing the projection function to be the position and linear momentum of a tagged particle, i.e., $B_t \to \mathbf{r}_t$, $\dot{B}_t \to \mathbf{p}_t$. Then, the Zwanzig projection reads [21]

$$\mathcal{P}_Z A_t = \frac{\langle \delta(\widehat{\mathbf{r}}_0 - \mathbf{r}_0)\delta(\widehat{\mathbf{p}}_0 - \mathbf{p}_0), A_t \rangle}{\langle \delta(\widehat{\mathbf{r}}_0 - \mathbf{r}_0)\delta(\widehat{\mathbf{p}}_0 - \mathbf{p}_0) \rangle} = \langle A_t \rangle_{\mathbf{r}_0, \mathbf{p}_0}, \tag{2.32}$$

where variables with a hat inside inner products, e.g., $\widehat{\mathbf{r}}_0$, are integrated over. The Zwanzig projection is a conditional average similar to eq. (2.8). The condition is now the position $\mathbf{r}_0$

and momentum $\mathbf{p}_0$ of a tagged particle. With this choice of projection $P \equiv P_Z$ and setting $\dot{A}_0 = \mathbf{p}_0$, eq. (2.26a) becomes [30]

$$\dot{\mathbf{p}}_t = -\nabla_r U_{\mathrm{PMF}}(\mathbf{r}_t) + \int_0^t \mathrm{d}s \left[\left(\frac{\nabla_p}{\beta} - \frac{\mathbf{p}_s}{m} \right)^{\mathrm{T}} \cdot \Gamma^Z(t-s, \mathbf{r}_s, \mathbf{p}_s) \right]^T + \mathbf{F}^R(\omega_0, t), \quad (2.33a)$$

$$\Gamma^Z_{ij}(t, \mathbf{r}, \mathbf{p}) = \beta \langle F^R_i(0), F^R_j(t) \rangle_{\mathbf{r},\mathbf{p}}. \quad (2.33b)$$

Here, $U_{\mathrm{PMF}}(\mathbf{r}) = -k_B T \ln\langle \delta(\hat{\mathbf{r}}_0 - \mathbf{r}) \rangle$ denotes the potential of mean force (PMF) also given in eq. (2.12). The PMF determines the mean force $-\nabla U_{\mathrm{PMF}}$ that acts on the tagged particle when it has the position $\mathbf{r}$. $\Gamma^Z \in \mathbb{R}^{3\times3}$ denotes the memory matrix, it is the result of a conditional correlation function for given values $\mathbf{r}, \mathbf{p}$. Thus, Γ^Z has a position and momentum dependency by construction.

Via the Zwanzig projection, the PMF is introduced into the equation of motion. This ensures the correct mean behavior once we switch to a stochastic description by replacing the fluctuating force $F_{\mathrm{R}}(t)$ by a stochastic process with zero mean [21].

The drawback in eq. (2.33a) is that the memory kernel in general has a position and momentum dependency which is difficult to deal with in applications. Therefore, it is usually neglected by an ad-hoc assumption and the memory kernel is assumed to be independent of the particle position and momentum. This assumption leads to a GLE to which we refer to as the approximate GLE, it reads

$$\dot{\mathbf{p}}_t = -\nabla_r U_{\mathrm{PMF}}(\mathbf{r}_t) - \int_0^t \mathrm{d}s\, \Gamma^{\mathrm{app}}(s) \cdot \frac{\mathbf{p}_{t-s}}{m} + \mathbf{F}^R(t), \quad (2.34a)$$

$$\langle F^R_i(t) \rangle = 0, \quad (2.34b)$$

$$\Gamma^{\mathrm{app}}_{ij}(t) = \beta \langle F^R_i(0), F^R_j(t) \rangle. \quad (2.34c)$$

3 GLEs with Nonlinear Friction Kernel

In this chapter, assuming a Hamiltonian of the form given in eq. (2.1) and making use of the properties discussed in chapter 2.2, we derive three different GLEs containing nonlinear memory functions, i.e., the memory functions not only depend on time but also on the value of the observable A_t itself. For this, we introduce three projection operators that in general lead to different nonlinear memory functions.

The first projection operator is a combination of the Mori and Zwanzig projection operators presented in chapter 2.3.1 and chapter 2.3.2, respectively. We refer to the GLE generated by the first projection operator as the *hybrid GLE* [4] which we derive in section 3.1. The hybrid GLE is suited as an analytical tool to quantify how important nonlinear memory effects are.

The second projection operator is introduced in section 3.2. It leads to a GLE in which the memory term is linear in the velocity [110]. Therefore, we refer to this GLE as *linear velocity GLE*. We will utilize the linear velocity GLE in the final chapter to perform GLE simulations via Markovian embedding.

The third projection operator incorporates a term quadratic in the velocity into the Markovian part of the GLE. The quadratic velocity term turns out to be important when one performs GLE simulations with position dependent mass, as we will show in the final chapter. The memory term of the third GLE is also linear in the velocity, only the mass dependence of the Markovian part changes. Thus, we refer to the third GLE, derived in section 3.3, as *modified linear velocity GLE*.

3.1 Hybrid GLE

The hybrid projection $\mathcal{P}^{\mathrm{H}}$ is a combination of the Mori and Zwanzig projection. $\mathcal{P}^{\mathrm{H}}$ is of the form $\mathcal{P}^{\mathrm{H}} = \mathcal{P}^{\mathrm{H}}_{\mathrm{M}} + \mathcal{P}^{\mathrm{H}}_{\mathrm{Z}}$. For an observable $A_t = A(\omega_t)$ and using a projection function $B_t = B(\mathbf{R}_t)$, i.e., a function of atomistic positions only, it is defined as follows

$$\mathcal{P}^{\mathrm{H}} A_t = \mathcal{P}^{\mathrm{H}}_{\mathrm{Z}} A_t + \mathcal{P}^{\mathrm{H}}_{\mathrm{M}} A_t, \tag{3.1a}$$

$$\mathcal{P}^{\mathrm{H}}_{\mathrm{Z}} A_t = \langle A_t \rangle_{B_0}, \tag{3.1b}$$

$$\mathcal{P}^{\mathrm{H}}_{\mathrm{M}} A_t = \frac{\langle \dot{B}_0, A_t \rangle}{\langle \dot{B}_0^2 \rangle} \dot{B}_0. \tag{3.1c}$$

The projection $\mathcal{P}^{\mathrm{H}}_{\mathrm{Z}}$ is the conditional average given in eq. (2.8) with the difference that the observable we project onto is a function of positions $\mathbf{R}$ only. This restriction is necessary in order for $\mathcal{P}^{\mathrm{H}}$ to be idempotent.

3.1.1 Properties of the hybrid projection

Clearly, we have $(\mathcal{P}_{\mathrm{M}}^{\mathrm{H}})^2 = \mathcal{P}_{\mathrm{M}}^{\mathrm{H}}$ and $(\mathcal{P}_{\mathrm{Z}}^{\mathrm{H}})^2 = \mathcal{P}_{\mathrm{Z}}^{\mathrm{H}}$. Therefore, in order to confirm the idempotency of the hybrid projection, one has to check that $\mathcal{P}_{\mathrm{M}}^{\mathrm{H}}\mathcal{P}_{\mathrm{Z}}^{\mathrm{H}}\, A_t = \mathcal{P}_{\mathrm{Z}}^{\mathrm{H}}\mathcal{P}_{\mathrm{M}}^{\mathrm{H}}\, A_t = 0$ for an arbitrary observable $A_t = A(\omega_t)$. This is true because we project onto observables of positions only, i.e., onto $B_0 = B(\mathbf{R}_0)$. Thus, the velocity $\dot{B}_0$ is linearly proportional to the particle momenta

$$\dot{B}_0 = \mathcal{L}B_0 = \sum_{n=1}^{N} \frac{\mathbf{p}_n}{m_n} \cdot \nabla_{r_n} B_0. \tag{3.2}$$

The operator $\mathcal{P}_{\mathrm{M}}^{\mathrm{H}}$ maps any function onto the subspace of functions that are linear in the observable velocity $\dot{B}_0$, which is linear in the particle momenta $\mathbf{p}_n$. From this we see that

$$\mathcal{P}_{\mathrm{Z}}^{\mathrm{H}}\mathcal{P}_{\mathrm{M}}^{\mathrm{H}}\, A_t \propto \mathcal{P}_{\mathrm{Z}}^{\mathrm{H}}\dot{B}_0 = 0, \tag{3.3}$$

since the operator $\mathcal{P}_{\mathrm{Z}}^{\mathrm{H}}$ involves an integral over the particle momenta but adds no momentum dependence.

$\mathcal{P}_{\mathrm{Z}}^{\mathrm{H}}$ maps any observable onto a function which depends on particle positions only. Since $\dot{B}_0$ is linearly proportional to the particle momenta, $\mathcal{P}_{\mathrm{M}}^{\mathrm{H}}$ applied on a function that depends on particle positions only gives zero. Therefore, it follows that

$$\mathcal{P}_{\mathrm{M}}^{\mathrm{H}}\mathcal{P}_{\mathrm{Z}}^{\mathrm{H}} A_t = 0. \tag{3.4}$$

Thus, we can use $\mathcal{P}^{\mathrm{H}}$ in eq. (2.26a) to derive a GLE for observables that are functions of positions, such as, center of masses, distances and angles. In the remainder, we set the projection function $B(\mathbf{R}_t) \equiv A(\mathbf{R}_t)$, i.e., we choose the observable A in eq. (2.26a) to be a function of positions only and we project onto A_0 itself. From this, it follows that $\mathcal{P}^{\mathrm{H}}$ projects out the observable A_0 and its velocity $\dot{A}_0$

$$\mathcal{P}^{\mathrm{H}} A_0 = A_0, \qquad\qquad \mathcal{P}^{\mathrm{H}} \dot{A}_0 = \dot{A}_0. \tag{3.5}$$

Furthermore, the projection $\mathcal{P}^{\mathrm{H}}$ in eq. (3.1) is self-adjoint w.r.t. the inner product in eq. (2.6), i.e., for any observables $A_t = A(\omega_t)$ and $C_{t'} = C(\omega_{t'})$, we have $\langle A_t, \mathcal{P}^{\mathrm{H}} C_{t'}\rangle = \langle \mathcal{P}^{\mathrm{H}} A_t, C_{t'}\rangle$. To show this, we consider the projection operators $\mathcal{P}_{\mathrm{M}}^{\mathrm{H}}$ and $\mathcal{P}_{\mathrm{Z}}^{\mathrm{H}}$ separately. Using the definition in eq. (3.1c), we find

$$\left\langle A_t, \mathcal{P}_{\mathrm{M}}^{\mathrm{H}} C_{t'} \right\rangle = \left\langle A_t, \frac{\langle \dot{B}_0, C_{t'}\rangle}{\langle \dot{B}_0^2\rangle} \dot{B}_0 \right\rangle \tag{3.6a}$$

$$= \langle A_t, \dot{B}_0\rangle \frac{\langle \dot{B}_0, C_{t'}\rangle}{\langle \dot{B}_0^2\rangle} \tag{3.6b}$$

$$= \left\langle \frac{\langle A_t, \dot{B}_0\rangle}{\langle \dot{B}_0^2\rangle} \dot{B}_0, C_{t'} \right\rangle \tag{3.6c}$$

$$= \langle \mathcal{P}_{\mathrm{M}}^{\mathrm{H}} A_t, C_{t'}\rangle. \tag{3.6d}$$

Now, consider the projection $\mathcal{P}_{\mathrm{Z}}^{\mathrm{H}}$ in eq. (3.1b) together with the conditional average defined in eq. (2.8)

$$\langle A_t, \mathcal{P}_{\mathrm{Z}}^{\mathrm{H}} C_{t'} \rangle = \left\langle A(\omega_0', t), \int \mathrm{d}a\, \delta(B(\mathbf{R}_0') - a) \frac{\langle \delta(B(\widehat{\mathbf{R}}_0) - a), C(\widehat{\omega}_0, t') \rangle}{\mathbb{P}(a)} \right\rangle, \tag{3.7}$$

where variables with a prime, i.e., ω_0', are integrated over in the outer inner product and variables with a hat, i.e., $\widehat{\omega}_0$, are integrated over in the projection $\mathcal{P}_{\mathrm{Z}}^{\mathrm{H}}$. From eq. (3.7), we obtain

$$\langle A_t, \mathcal{P}_{\mathrm{Z}}^{\mathrm{H}} C_{t'} \rangle = \int \mathrm{d}a\, \langle A(\omega_0', t), \delta(B(\mathbf{R}_0') - a) \rangle \frac{\langle \delta(B(\widehat{\mathbf{R}}_0) - a), C(\widehat{\omega}_0, t') \rangle}{\mathbb{P}(a)} \tag{3.8a}$$

$$= \left\langle \int \mathrm{d}a\, \frac{\langle A(\omega_0', t), \delta(B(\mathbf{R}_0') - a) \rangle}{\mathbb{P}(a)} \delta(B(\widehat{\mathbf{R}}_0) - a), C(\widehat{\omega}_0, t') \right\rangle \tag{3.8b}$$

$$= \langle \mathcal{P}_{\mathrm{Z}}^{\mathrm{H}} A_t, C_{t'} \rangle. \tag{3.8c}$$

This means that the hybrid projection $\mathcal{P}^{\mathrm{H}}$ in eq. (3.1) is self-adjoint and thus is an orthogonal projection, i.e.,

$$\langle \mathcal{P}^{\mathrm{H}} A_t, \mathcal{Q}^{\mathrm{H}} C_{t'} \rangle = 0, \tag{3.9}$$

for arbitrary observables A_t and $C_{t'}$.

Another very important property of the hybrid projection is the following: The equilibrium average of any observable vanishes if it lies in the complementary subspace at all times. To prove this, we must show that for an arbitrary observable $A(\omega_t) = A_t$ that $\langle \mathcal{P}^{\mathrm{H}} A_t \rangle = \langle A_t \rangle$ holds. First, from the definition of $\mathcal{P}_{\mathrm{M}}^{\mathrm{H}}$ in eq. (3.1), it follows that

$$\langle \mathcal{P}_{\mathrm{M}}^{\mathrm{H}} A_t \rangle \propto \langle \dot{B}_0 \rangle = 0, \tag{3.10}$$

since our projection function $B_0 = B(\mathbf{R}_0)$ is a function of positions only and therefore, its velocity $\dot{B}_0 = \mathcal{L} B_0$ is linear in the momenta (see eq. (3.2)). For the $\mathcal{P}_{\mathrm{Z}}^{\mathrm{H}}$ projection operator we find

$$\langle \mathcal{P}_{\mathrm{Z}}^{\mathrm{H}} A_t \rangle = \langle \langle A_t \rangle_{B_0} \rangle = \langle A_t \rangle. \tag{3.11a}$$

From this, it immediately follows that $\langle \mathcal{P}^{\mathrm{H}} A_t \rangle = \langle A_t \rangle$ and thus all equilibrium averages in the complementary subspace vanish, i.e.,

$$\langle \mathcal{Q}^{\mathrm{H}} A_t \rangle = \langle (1 - \mathcal{P}^{\mathrm{H}}) A_t \rangle = \langle A_t \rangle - \langle \mathcal{P}^{\mathrm{H}} A_t \rangle = 0. \tag{3.12}$$

In particular, the equilibrium average of the random force vanishes at all times, i.e.,

$$\langle F_{\mathrm{R}}^{\mathrm{H}}(t) \rangle = \langle \mathcal{Q}^{\mathrm{H}} F_{\mathrm{R}}^{\mathrm{H}}(t) \rangle = 0. \tag{3.13}$$

$F_{\mathrm{R}}^{\mathrm{H}}(t)$ is the random force that follows from the hybrid projection, i.e, in eq. (2.26b) $\mathcal{Q}$ is replaced by $\mathcal{Q}^{\mathrm{H}}$. It completely lies in the complementary subspace for all times and, therefore, has by construction a vanishing equilibrium average.

3.1.2 Derivation of the hybrid GLE

Using the specific form of the projection $\mathcal{P}^{\mathrm{H}}$ in eq. (3.1), we find for the first term on the right-hand side of eq. (2.26a)

$$e^{t\mathcal{L}}\mathcal{P}^{\mathrm{H}}\mathcal{L}\dot{A}_0 = e^{t\mathcal{L}}\left(\mathcal{P}_{\mathrm{Z}}^{\mathrm{H}}\mathcal{L}\dot{A}_0 + \mathcal{P}_{\mathrm{M}}^{\mathrm{H}}\mathcal{L}\dot{A}_0\right), \tag{3.14a}$$

$$\mathcal{P}_{\mathrm{M}}^{\mathrm{H}}\mathcal{L}\dot{A}_0 \propto \langle \dot{A}_0, \mathcal{L}\dot{A}_0\rangle = -\langle \mathcal{L}\dot{A}_0, \dot{A}_0\rangle = 0, \tag{3.14b}$$

$$\begin{aligned}\mathcal{P}_{\mathrm{Z}}^{\mathrm{H}}\mathcal{L}\dot{A}_0 &= \langle \mathcal{L}\dot{A}_0\rangle_{A_0} \\ &= \frac{\mathrm{d}}{\mathrm{d}A_0}\langle \dot{A}_0^2\rangle_{A_0} - \beta\langle \dot{A}_0^2\rangle_{A_0}\frac{\mathrm{d}}{\mathrm{d}A_0}U_{\mathrm{PMF}}(A_0),\end{aligned} \tag{3.14c}$$

where we used the relation in eq. (2.13) to obtain eq. (3.14c). The function $\langle \dot{A}_0^2\rangle_{A_0}$ is strictly positive. We can use this to further simplify eq. (3.14c). We define

$$\langle \dot{A}_0^2\rangle_{A_0} \equiv k_B T/M(A_0), \tag{3.15}$$

in which the function $M(A_0)$ has units of a generalized mass. Using $M(A_0)$, eq. (3.14c) simplifies to

$$\mathcal{P}_{\mathrm{Z}}^{\mathrm{H}}\mathcal{L}\dot{A}_0 = -\frac{1}{M(A_0)}\frac{\mathrm{d}U_{\mathrm{eff}}}{\mathrm{d}A_0}, \tag{3.16a}$$

$$U_{\mathrm{eff}}(A_0) = U_{\mathrm{PMF}}(A_0) + k_B T \ln M(A_0). \tag{3.16b}$$

In eq. (3.16), U_{PMF} is the potential of mean force introduced in eq. (2.12). Applying the operator $e^{t\mathcal{L}}$ onto eq. (3.16) propagates all A_0 to A_t, which represents the final form for the first term on the right-hand side of eq. (2.26a).

The second term on the right-hand side of eq. (2.26a) couples the state at time t to all past states via the integrand $e^{(t-s)\mathcal{L}}(\mathcal{P}_{\mathrm{Z}}^{\mathrm{H}} + \mathcal{P}_{\mathrm{M}}^{\mathrm{H}})\mathcal{L}F_{\mathrm{R}}^{\mathrm{H}}(s)$. The $\mathcal{P}_{\mathrm{M}}^{\mathrm{H}}$ projection leads to a memory function of the same form as in the Mori projection

$$\mathcal{P}_{\mathrm{M}}^{\mathrm{H}}\mathcal{L}F_{\mathrm{R}}^{\mathrm{H}}(s) = \frac{\langle \dot{A}_0, \mathcal{L}F_{\mathrm{R}}^{\mathrm{H}}(s)\rangle}{\langle \dot{A}_0^2\rangle}\dot{A}_0 \tag{3.17a}$$

$$= -\frac{\langle F_{\mathrm{R}}^{\mathrm{H}}(0), F_{\mathrm{R}}^{\mathrm{H}}(s)\rangle}{\langle \dot{A}_0^2\rangle}\dot{A}_0 \tag{3.17b}$$

$$= -\Gamma_{\mathrm{M}}^{\mathrm{H}}(s)\dot{A}_0, \tag{3.17c}$$

where we defined the memory function due to the $\mathcal{P}_{\mathrm{M}}^{\mathrm{H}}$ projection by

$$\Gamma_{\mathrm{M}}^{\mathrm{H}}(s) = \frac{\langle F_{\mathrm{R}}^{\mathrm{H}}(0), F_{\mathrm{R}}^{\mathrm{H}}(s)\rangle}{\langle \dot{A}_0^2\rangle}. \tag{3.18}$$

The memory function due to the $\mathcal{P}_{\mathrm{Z}}^{\mathrm{H}}$ projection is directly obtained using the relation in eq. (2.13)

$$\mathcal{P}_{\mathrm{Z}}^{\mathrm{H}}\mathcal{L}F_{\mathrm{R}}^{\mathrm{H}}(s) = \langle \mathcal{L}F_{\mathrm{R}}^{\mathrm{H}}(s)\rangle_{A_0} \equiv \Gamma_{\mathrm{Z}}^{\mathrm{H}}(A_0, s), \tag{3.19a}$$

$$\Gamma_{\mathrm{Z}}^{\mathrm{H}}(A, s) = \frac{\mathrm{d}}{\mathrm{d}A}D^{\mathrm{H}}(A, s) - \beta D^{\mathrm{H}}(A, s)\frac{\mathrm{d}}{\mathrm{d}A}U_{\mathrm{PMF}}(A), \tag{3.19b}$$

$$D^{\mathrm{H}}(A, s) = \langle \dot{A}_0, F_{\mathrm{R}}^{\mathrm{H}}(s)\rangle_A. \tag{3.19c}$$

With the results in eq. (3.16), eq. (3.17) and eq. (3.19), eq. (2.26a) becomes

$$\ddot{A}_t = -\frac{\mathrm{d}U_{\text{eff}}(A_t)/\mathrm{d}A_t}{M(A_t)} - \int_0^t \mathrm{d}s\, \Gamma_{\mathrm{M}}^{\mathrm{H}}(s)\dot{A}_{t-s} + F_{\mathrm{R}}^{\mathrm{H}}(t) + \int_0^t \mathrm{d}s\, \Gamma_{\mathrm{Z}}^{\mathrm{H}}(A_{t-s}, s), \tag{3.20}$$

where the following exact relations hold

$$\Gamma_{\mathrm{M}}^{\mathrm{H}}(t) = \frac{\langle F_{\mathrm{R}}^{\mathrm{H}}(t), F_{\mathrm{R}}^{\mathrm{H}}(0)\rangle}{\langle \dot{A}_0^2\rangle}, \qquad \langle F_{\mathrm{R}}^{\mathrm{H}}(t)\rangle = 0, \qquad \langle F_{\mathrm{R}}^{\mathrm{H}}(t)\rangle_A = 0, \tag{3.21a}$$

$$\langle \dot{A}_0, F_{\mathrm{R}}^{\mathrm{H}}(t)\rangle = 0, \qquad \langle f(A_0), F_{\mathrm{R}}^{\mathrm{H}}(t)\rangle = 0, \qquad \langle f(A_0), F_{\mathrm{R}}^{\mathrm{H}}(t)\rangle_A = 0, \tag{3.21b}$$

Here, $f(A_0)$ is an arbitrary function of A_0.

In the rightmost term of eq. (3.20), we introduce an additional term due to the memory function $\Gamma_{\mathrm{Z}}^{\mathrm{H}}(A,t)$ which has an A dependency. This additional term therefore represents a correction to the approximate GLE (eq. (2.34)) in the presence of nonlinear memory effects. According to eq. (3.19), $\Gamma_{\mathrm{Z}}^{\mathrm{H}}(A,t)$ vanishes if the conditional correlation $\langle \dot{A}_0, F_{\mathrm{R}}^{\mathrm{H}}(t)\rangle_{A_0}$ vanishes. Eq. (3.20) is our first result.

3.1.3 The position-dependent mass

If the generalized mass $M(A)$ in eq. (3.15) is constant, i.e., if the variance of the velocity $\dot{A}$ is independent of the position A, one obtains $\mathrm{d}U_{\text{eff}}/\mathrm{d}A = \mathrm{d}U_{\mathrm{PMF}}/\mathrm{d}A$ and the first three terms on the right-hand side of eq. (3.20) reduce to the approximate GLE in eq. (2.34). For some observables, it can be shown analytically that the corresponding $M(A)$ is constant. These are, for example, the center of mass [43], which is a linear combination of positions, but also for certain nonlinear observables, such as distances, it can be shown that the generalized mass is constant, i.e., $M(A) = M$. In this case, the force term $\mathrm{d}U_{\text{eff}}/\mathrm{d}A$ in eq. (3.20) reduces to $\mathrm{d}U_{\mathrm{PMF}}/\mathrm{d}A$.

Assuming a Hamiltonian of the form as given in eq. (2.1), we can write down a formula for the position dependent mass in terms of the observable A. Noting that the velocity

$$\dot{A}_0 = \sum_{n=1}^{N} \frac{\mathbf{p}_n}{m_n} \cdot \nabla_{r_n} A_0 \tag{3.22}$$

is linear in the momenta $\mathbf{p}_n$, it follows that, in equilibrium, $\dot{A}_0$ is Gaussian distributed for given positions $\mathbf{R}_0$, or fixed A_0. According to eq. (3.15), the position dependent mass is given by the variance of the velocity $\dot{A}_0$ for given A_0. We can write down the variance $\langle \dot{A}_0^2\rangle_{A_0}$ by using relations for the sum of independent, identically distributed Gaussian random variables.

Each p_n^i for $i = x, y, z$ is rescaled with the factor $(1/m_n)(\partial A_0/\partial r_n^i)$. From this, it follows that the rescaled p_n^i has zero mean and variance $(k_B T/m_n)(\partial A_0/\partial r_n^i)^2$ [109]. $\dot{A}_0$ is the sum of these rescaled p_n^i. Thus, its variance is given by

$$\langle \dot{A}_0^2\rangle_{A_0} = \sum_{n=1}^{N} \frac{k_B T}{m_n} \left(\nabla_{r_n} A_0\right)^2 . \tag{3.23}$$

And therefore, the mass is given by

$$M(A) = \left(\sum_{n=1}^{N} \frac{1}{m_n} \left(\nabla_{r_n} A \right)^2 \right)^{-1}. \tag{3.24}$$

Note that, the same formula for the mass $M(A)$ was obtained in ref. [29] by performing a coordinate transformation in the Lagrangian and computing the generalized mass from the conjugate momentum.

As an example, we consider the hydrogen-bond distance between a nitrogen atom (donor) with initial position $\mathbf{r}_0^N$ and an oxygen atom (acceptor) with initial position $\mathbf{r}_0^O$ that are located four residues apart on the backbone of a polypeptide. The observable is thus given by

$$A_0 = A(\mathbf{R}_0) = \sqrt{(\mathbf{r}_0^N - \mathbf{r}_0^O)^2}. \tag{3.25}$$

Using eq. (3.24), we find

$$M(A_0) = M = \frac{m_N \, m_O}{m_N + m_O}, \tag{3.26}$$

which is the reduced mass of the nitrogen-oxygen distance coordinate. A similar derivation can also be done for a linear combination of distances. For example, consider the mean hydrogen-bond distance between N_R donor nitrogen atoms and N_R acceptor oxygen atoms that are located four residues apart along the backbone of a polypeptide. We define the observable as

$$A_0 = \frac{1}{N_R} \sum_{n=1}^{N_R} A_{n,0}, \tag{3.27}$$

with $A_{n,0}$ being the initial value of the n-th distance. Eq. (3.24) becomes

$$\langle \dot{A}_0^2 \rangle_{A_0} = \left\langle \left(\frac{1}{N_R} \sum_{n=1}^{N_R} \dot{A}_{n,0} \right)^2 \right\rangle_{A_0} = \frac{1}{N_R^2} \sum_{n=1}^{N_R} \langle \dot{A}_{n,0}^2 \rangle_{A_0}. \tag{3.28}$$

In analogy to eq. (3.26), the effective mass is constant also for a linear combination of spatial distances.

3.1.4 The multi-dimensional hybrid GLE

We generalize eq. (3.20) for a multidimensional observable that is a function of particle positions only, i.e., the case in which the reaction coordinate is a vector

$$\mathbf{A}(\mathbf{R}_t) = (A_1(\mathbf{R}_t), A_2(\mathbf{R}_t), \ldots, A_n(\mathbf{R}_t)), \tag{3.29}$$

containing a set of observables. We assume that the observables $A_i(\mathbf{R}_t)$ are functions of a disjoint set of positions. As before, all observables implicitly depend on time only

via the positions $\mathbf{R}_t$. We denote components as $A_k(\mathbf{R}_t) \equiv A_{k,t}$ and $A_k(\mathbf{R}_0) \equiv A_{k,0}$. In the multi-dimensional case, the hybrid projection operator reads for a vectorial projection function $\mathbf{A}_0$

$$\mathcal{P}^{\mathrm{H}} B_{m,t} = (\mathcal{P}^{\mathrm{H}}_{\mathrm{M}} + \mathcal{P}^{\mathrm{H}}_{\mathrm{Z}}) B_{m,t} = \sum_{k=1}^{n} \frac{\langle B_{m,t},\, \dot{A}_{k,0} \rangle}{\langle \dot{A}^2_{k,0} \rangle} \dot{A}_{k,0} + \langle B_{m,t} \rangle_{\mathbf{A}_0}. \tag{3.30}$$

Choosing $\mathbf{B}_0 = \mathbf{A}_0$, as before, the projection in eq. (3.30) leads to the following potential term

$$e^{t\mathcal{L}} \mathcal{P}^{\mathrm{H}}_{\mathrm{Z}} \mathcal{L} \dot{\mathbf{A}}_0 = k_B T \left(\nabla^T_A \cdot M^{-T}(\mathbf{A}_t) \right)^T - M^{-1}(\mathbf{A}_t) \cdot \nabla_A U_{\mathrm{PMF}}(\mathbf{A}_t). \tag{3.31}$$

where we introduced the inverse generalized mass matrix

$$M^{-1}_{kl}(\mathbf{A}) = \beta \langle \dot{A}_{k,0}, \dot{A}_{l,0} \rangle_{\mathbf{A}}. \tag{3.32}$$

The computation of the memory function proceeds similarly as in the main text and the multi-dimensional hybrid GLE reads

$$\begin{aligned} \ddot{\mathbf{A}}_t = {} & k_B T \left(\nabla^T_A \cdot M^{-T}(\mathbf{A}_t) \right)^T - M^{-1}(\mathbf{A}_t) \cdot \nabla_A U_{\mathrm{PMF}}(\mathbf{A}_t) \\ & - \int_0^t \mathrm{d}s\, \Gamma^{\mathrm{H}}_{\mathrm{M}}(t-s) \cdot \dot{\mathbf{A}}_s + \int_0^t \mathrm{d}s\, \mathbf{\Gamma}^{\mathrm{H}}_{\mathrm{Z}}(\mathbf{A}_{t-s}, s) + \mathbf{F}^{\mathrm{H}}_{\mathrm{R}}(t), \end{aligned} \tag{3.33}$$

where the following relations hold

$$\begin{aligned} &\langle (F^{\mathrm{H}}_{\mathrm{R}})_k(t) \rangle = 0, \qquad \langle (F^{\mathrm{H}}_{\mathrm{R}})_k(t), (F^{\mathrm{H}}_{\mathrm{R}})_l(0) \rangle = \langle \dot{A}^2_0 \rangle \, (\Gamma^{\mathrm{H}}_{\mathrm{M}})_{kl}(t), \\ &\langle (F^{\mathrm{H}}_{\mathrm{R}})_k(t), A_{l,0} \rangle = 0, \qquad \langle (F^{\mathrm{H}}_{\mathrm{R}})_k(t), \dot{A}_{l,0} \rangle = 0, \end{aligned} \tag{3.34}$$

for all $k, l = 1, 2, \ldots, n$. The k-th component of the vectorial nonlinear memory friction function $\mathbf{\Gamma}^{\mathrm{H}}_{\mathrm{Z}}(\mathbf{A}, s)$ is given by

$$(\Gamma^{\mathrm{H}}_{\mathrm{Z}})_k(\mathbf{A}, s) = \mathcal{P}^{\mathrm{H}}_{\mathrm{Z}} \mathcal{L} (F^{\mathrm{H}}_{\mathrm{R}})_k(s) = \left[\nabla_A \cdot \mathbf{D}_k(\mathbf{A}, s) - \beta\, \mathbf{D}_k(\mathbf{A}, s) \cdot \nabla_A U_{\mathrm{PMF}}(\mathbf{A}) \right], \tag{3.35a}$$

$$\mathbf{D}_k(\mathbf{A}, s) = \langle \dot{\mathbf{A}}_0,\, (F^{\mathrm{H}}_{\mathrm{R}})_k(s) \rangle_{\mathbf{A}}. \tag{3.35b}$$

3.2 Linear velocity GLE

The hybrid GLE from section 3.1 enables one to probe for nonlinear memory effects directly by checking if the function $D^{\mathrm{H}}(A, t)$ from eq. (3.19c) vanishes. At the time this thesis was written, a computationally efficient method to perform simulations of the hybrid GLE was still missing. Therefore, we derive in the following a second nonlinear GLE that can be simulated using Markovian embedding techniques (cf. chapter 6).

The projection operator $\mathcal{P}^{\mathrm{L}}$ is a modification of the hybrid projection and is given by

$$\mathcal{P}^{\mathrm{L}} B_t = \mathcal{P}^{\mathrm{H}}_{\mathrm{Z}} B_t + \mathcal{P}^{\mathrm{L}}_{\mathrm{M}} B_t, \tag{3.36a}$$

$$\mathcal{P}^{\mathrm{L}}_{\mathrm{M}} B_t = \frac{\langle \dot{A}_0, B_t \rangle_{A_0}}{\langle \dot{A}^2_0 \rangle_{A_0}} \dot{A}_0. \tag{3.36b}$$

$\mathcal{P}^{\mathrm{L}}$ is also an orthogonal projection w.r.t. the inner product in eq. (2.6) and leads to a GLE with a nonlinear friction kernel of the same form as discussed in ref. [110].

The first projection operator $\mathcal{P}_{\mathrm{Z}}^{\mathrm{H}}$ on the r.h.s. of eq. (3.36a) was introduced in the hybrid projection in eq. (3.1b). As demonstrated in the derivation of the hybrid GLE, $\mathcal{P}_{\mathrm{Z}}^{\mathrm{H}}$ is responsible for the potential of mean force term in the projected dynamics. However, contrary to the hybrid GLE where $\mathcal{P}_{\mathrm{Z}}^{\mathrm{H}}$, when combined with $\mathcal{P}_{\mathrm{M}}^{\mathrm{H}}$, generates the nonlinear kernel $\Gamma_{\mathrm{Z}}^{\mathrm{H}}$, it has no contribution to the memory term when combined with $\mathcal{P}_{\mathrm{M}}^{\mathrm{L}}$ in eq. (3.36b). This is due to the fact that the random force $F_{\mathrm{R}}^{\mathrm{L}}(t) = e^{t\mathcal{Q}^{\mathrm{L}}\mathcal{L}}\mathcal{Q}^{\mathrm{L}}\mathcal{L}\dot{A}_0$ is now orthogonal to all functions of the form $cf(A_0) + g(A_0)\dot{A}_0$, where f and g are arbitrary functions of A_0, and therefore

$$\mathcal{P}_{\mathrm{Z}}^{\mathrm{H}}\mathcal{L}F_{\mathrm{R}}^{\mathrm{L}}(t) = \langle \mathcal{L}F_{\mathrm{R}}^{\mathrm{L}}\rangle_A = -\frac{\langle \dot{A}_0 \frac{\partial}{\partial A_0}\left[\delta(A_0 - A)\right],\, F_{\mathrm{R}}^{\mathrm{L}}(t)\rangle}{\langle \delta(A_0 - A)\rangle} = 0. \tag{3.37}$$

This means that the memory term is only due to the second projection $\mathcal{P}_{\mathrm{M}}^{\mathrm{L}}$ which generates functions linear in the velocity $\dot{A}_0$ and nonlinear in A_0. The nonlinearity in A_0 follows from the fact that equilibrium averages are replaced by conditional averages, i.e.,

$$\mathcal{P}_{\mathrm{M}}^{\mathrm{L}}\mathcal{L}F_{\mathrm{R}}^{\mathrm{L}}(t) = \frac{\langle \dot{A}_0, \mathcal{L}F_{\mathrm{R}}^{\mathrm{L}}(t)\rangle_{A_0}}{\langle \dot{A}_0^2\rangle_{A_0}}\dot{A}_0. \tag{3.38}$$

Using the relation in eq. (2.16) in combination with the product rule of differentiation, we obtain the memory kernel

$$\frac{\langle \dot{A}_0, \mathcal{L}F_{\mathrm{R}}^{\mathrm{L}}(t)\rangle_{A_0}}{\langle \dot{A}_0^2\rangle_{A_0}}\dot{A}_0 = -\Gamma^{\mathrm{L}}(A_0, t)\dot{A}_0 \tag{3.39a}$$

$$\Gamma^{\mathrm{L}}(A,t) = \beta D^{\mathrm{L}}(A,t)\frac{\mathrm{d}U_{\mathrm{eff}}(A)}{\mathrm{d}A} - \frac{\mathrm{d}D^{\mathrm{L}}(A,t)}{\mathrm{d}A} + \frac{\langle \ddot{A}_0, F_{\mathrm{R}}^{\mathrm{L}}(t)\rangle_A}{\langle \dot{A}_0^2\rangle_A}, \tag{3.39b}$$

$$D^{\mathrm{L}}(A,t) = \frac{\langle \dot{A}_0^2, F_{\mathrm{R}}^{\mathrm{L}}(t)\rangle_A}{\langle \dot{A}_0^2\rangle_A}. \tag{3.39c}$$

Finally, we note that $\mathcal{P}_{\mathrm{M}}^{\mathrm{L}}$ does not contribute to the potential term, i.e.,

$$\mathcal{P}_{\mathrm{M}}^{\mathrm{L}}\mathcal{L}\dot{A}_0 = \frac{\langle \dot{A}_0, \mathcal{L}\dot{A}_0\rangle_{A_0}}{\langle \dot{A}_0^2\rangle_{A_0}}\dot{A}_0 = 0. \tag{3.40}$$

To show this, we consider the numerator in eq. (3.40)

$$\langle \dot{A}_0, \mathcal{L}\dot{A}_0\rangle_A \propto \langle \delta(A_0 - A)\dot{A}_0, \mathcal{L}\dot{A}_0\rangle = -\langle \mathcal{L}\delta(A_0 - A)\dot{A}_0, \dot{A}_0\rangle \tag{3.41a}$$

$$= -\langle \mathcal{L}[\delta(A_0 - A)]\dot{A}_0, \dot{A}_0\rangle - \langle \delta(A_0 - A)\mathcal{L}\dot{A}_0, \dot{A}_0\rangle \tag{3.41b}$$

$$= \left\langle \frac{\partial}{\partial A}[\delta(A_0 - A)]\dot{A}_0^2, \dot{A}_0 \right\rangle - \langle \delta(A_0 - A)\mathcal{L}\dot{A}_0, \dot{A}_0\rangle \tag{3.41c}$$

$$= \left\langle \frac{\partial}{\partial A}[\delta(A_0 - A)]\dot{A}_0^3 \right\rangle - \langle \delta(A_0 - A)\mathcal{L}\dot{A}_0, \dot{A}_0.\rangle \tag{3.41d}$$

The first term on the r.h.s. of eq. (3.41d) contains a sum of odd powers of momenta $\mathbf{p}_n$ and thus, vanishes. From this, it follows

$$\langle \delta(A_0 - A)\dot{A}_0, \mathcal{L}\dot{A}_0 \rangle = -\langle \delta(A_0 - A)\mathcal{L}\dot{A}_0, \dot{A}_0 . \rangle, \tag{3.42}$$

$$\Rightarrow \langle \delta(A_0 - A)\dot{A}_0, \mathcal{L}\dot{A}_0 \rangle = 0. \tag{3.43}$$

And the GLE reads

$$\ddot{A}_t = -\frac{1}{M(A_t)} \frac{\mathrm{d}U_{\text{eff}}(A_t)}{\mathrm{d}A_t} - \int_0^t \mathrm{d}s\, \Gamma^{\mathrm{L}}(A_s, t-s)\dot{A}_s + F_{\mathrm{R}}^{\mathrm{L}}(t), \tag{3.44}$$

where $U_{\text{eff}}(A)$ and the position dependent mass $M(A)$ were introduced in eq. (3.16) and eq. (3.15), respectively.
Finally, since $\mathcal{P}^{\mathrm{L}}$ is an orthogonal projection, the following exact relations hold

$$\langle F_{\mathrm{R}}^{\mathrm{L}}(t) \rangle_A = 0, \qquad \langle \dot{A}_0, F_{\mathrm{R}}^{\mathrm{L}}(t) \rangle_A = 0. \tag{3.45a}$$

3.3 Modified linear velocity GLE

Our motivation to introduce the linear velocity GLE in section 3.2 was that we will present a Markovian embedding for this GLE in chapter 6. We demonstrate now why the linear velocity GLE in in eq. (3.44) must be adjusted when the position-dependency of the effective mass cannot be neglected.
We consider a model Hamiltonian in which a generalized coordinate q has a configuration-dependent mass $M(q)$. We denote the conjugate momentum of q by $p = M(q)v$ with $v = \dot{q}$. In the Hamiltonian, we also include a confining potential $U(q)$, i.e,

$$H(q,p) = \frac{p^2}{2M(q)} + U(q). \tag{3.46}$$

Hamilton's equation of motion for p read

$$\dot{p} = \frac{M'(q)}{2M^2(q)} p^2 - U'(q), \tag{3.47a}$$

$$M'(q)v^2 + M(q)\dot{v} = \frac{M'(q)}{2} v^2 - U'(q). \tag{3.47b}$$

Solving eq. (3.47b) for $\dot{v}$, we find the equation of motion

$$\dot{v} = -\frac{1}{M(q)} \left(U'(q) + \frac{M'(q)}{2} v^2 \right). \tag{3.48}$$

From eq. (3.48), it follows that any Langevin equation with a position-dependent mass and with a stationary Boltzmann distribution requires the presence of a quadratic velocity term. In the following, we incorporate a term quadratic in the velocity into the linear velocity GLE by introducing a new projection.

To derive the hybrid GLE in eq. (3.20) and the linear velocity GLE in eq. (3.44), we started with eq. (2.26a). To obtain eq. (2.26a), we first decomposed the Liouville equation and, afterwards, the propagator $e^{t\mathcal{L}}$ using the Dyson operator identity in eq. (2.25). We used the same projection operator for both decompisitions, i.e., in case of the hybrid GLE $\mathcal{P}^{\mathrm{H}}$ and in case of the linear velocity GLE $\mathcal{P}^{\mathrm{L}}$.

Here, we choose different projection operators to decompose the Liouville equation $\ddot{A}_0 = \mathcal{L}\dot{A}_0$ and the propagator $e^{t\mathcal{L}}$. To decompose the Liouville equation, we use the following projection operator

$$\mathcal{P}^{\mathrm{K}} = \mathcal{P}^{\mathrm{L}} + \mathcal{P}^{\mathrm{K}}_{\mathrm{M}}, \tag{3.49a}$$

$$\mathcal{P}^{\mathrm{K}}_{\mathrm{M}} B_t = \frac{\left\langle \left(\dot{A}_0^2 - \langle \dot{A}_0^2 \rangle_{A_0}\right), B_t \right\rangle_{A_0}}{\left\langle \left(\dot{A}_0^2 - \langle \dot{A}_0^2 \rangle_{A_0}\right)^2 \right\rangle_{A_0}} \left(\dot{A}_0^2 - \langle \dot{A}_0^2 \rangle_{A_0}\right). \tag{3.49b}$$

The projection operator $\mathcal{P}^{\mathrm{K}}$ is an extension of the projection $\mathcal{P}^{\mathrm{L}}$ given in eq. (3.36). $\mathcal{P}^{\mathrm{K}}$ is an orthogonal projection w.r.t. the inner product in eq. (2.6). It includes the projection $\mathcal{P}^{\mathrm{K}}_{\mathrm{M}}$ onto functions quadratic in the velocity which will introduce a term quadratic in the velocity into the GLE (similar to eq. (3.48)). We use $\mathcal{P}^{\mathrm{K}}$ to decompose the Liouville equation for the velocity $\dot{A}_t$, i.e.,

$$\ddot{A}_t = e^{t\mathcal{L}}(\mathcal{P}^{\mathrm{K}} + \mathcal{Q}^{\mathrm{K}})\mathcal{L}\dot{A}_0. \tag{3.50}$$

In the Dyson decomposition, we use the projection operator $\mathcal{P}^{\mathrm{L}}$ in eq. (3.36) which we used to derive the linear velocity GLE, i.e.,

$$e^{t\mathcal{L}} = e^{t\mathcal{Q}^{\mathrm{L}}\mathcal{L}} + \int_0^t \mathrm{d}s\, e^{(t-s)\mathcal{L}}\mathcal{P}^{\mathrm{L}}\mathcal{L}e^{s\mathcal{Q}^{\mathrm{L}}\mathcal{L}}. \tag{3.51}$$

Inserting eq. (3.51) into eq. (3.50), we obtain

$$\ddot{A}_t = e^{t\mathcal{L}}\mathcal{P}^{\mathrm{K}}\mathcal{L}\dot{A}_0 + \int_0^t \mathrm{d}s\, e^{(t-s)\mathcal{L}}\mathcal{P}^{\mathrm{L}}\mathcal{L}F^{\mathrm{K}}_{\mathrm{R}}(s) + F^{\mathrm{K}}_{\mathrm{R}}(t), \tag{3.52a}$$

$$F^{\mathrm{K}}_{\mathrm{R}}(t) = e^{t\mathcal{Q}^{\mathrm{L}}\mathcal{L}}\mathcal{Q}^{\mathrm{K}}\mathcal{L}\dot{A}_0. \tag{3.52b}$$

Note that the projection $\mathcal{P}^{\mathrm{K}}$ determines the functional form of the first term $e^{t\mathcal{L}}\mathcal{P}^{\mathrm{K}}\mathcal{L}\dot{A}_0$ and the initial value $F^{\mathrm{K}}_{\mathrm{R}}(0) = \mathcal{Q}^{\mathrm{K}}\mathcal{L}B_0$, while the projection $\mathcal{P}^{\mathrm{L}}$ determines the functional form of the integrand $\mathcal{P}^{\mathrm{L}}\mathcal{L}F^{\mathrm{K}}_{\mathrm{R}}(s)$ and the propagation of $F^{\mathrm{K}}_{\mathrm{R}}(0)$ in time via $e^{t\mathcal{Q}^{\mathrm{L}}\mathcal{L}}$. Therefore, to compute the memory term, we can make use of the derivation of the linear velocity GLE. By noting that we have $\mathcal{Q}^{\mathrm{K}}B_t = \mathcal{Q}^{\mathrm{L}}\mathcal{Q}^{\mathrm{K}}B_t$, it follows that the random force $F^{\mathrm{K}}_{\mathrm{R}}(t) = \mathcal{Q}^{\mathrm{L}}F^{\mathrm{K}}_{\mathrm{R}}(t)$ is orthogonal to the subspace $\mathcal{P}^{\mathrm{L}}$ projects onto. Meaning, the derivation of the memory term is completely analogous to the previous section on the linear velocity GLE.

3.3.1 Derivation of the potential term

The derivation of the potential term is more elaborate and requires the computation of

$$e^{t\mathcal{L}}\mathcal{P}^{\mathrm{K}}\mathcal{L}\dot{A}_0 = e^{t\mathcal{L}}\left(\mathcal{P}^{\mathrm{H}}_{\mathrm{Z}} + \mathcal{P}^{\mathrm{L}}_{\mathrm{M}} + \mathcal{P}^{\mathrm{K}}_{\mathrm{M}}\right)\mathcal{L}\dot{A}_0. \tag{3.53}$$

The contribution due to $\mathcal{P}_{\mathrm{Z}}^{\mathrm{H}}\mathcal{L}\dot{A}_0$ was derived in eq. (3.16). Further, in eq. (3.40), eq. (3.41) and eq. (3.42), we showed that $\mathcal{P}_{\mathrm{M}}^{\mathrm{L}}\mathcal{L}\dot{A}_0 = 0$. Here, we derive the contribution due to

$$\mathcal{P}_{\mathrm{M}}^{\mathrm{K}}\mathcal{L}\dot{A}_0 = \frac{\left\langle \left(\dot{A}_0^2 - \langle \dot{A}_0^2 \rangle_{A_0}\right), \mathcal{L}\dot{A}_0 \right\rangle_A}{\left\langle \left(\dot{A}_0^2 - \langle \dot{A}_0^2 \rangle_{A_0}\right)^2 \right\rangle_A} \left(\dot{A}_0^2 - \langle \dot{A}_0^2 \rangle_{A_0}\right) \tag{3.54a}$$

$$= \frac{\left\langle \left(\dot{A}_0^2 - \frac{k_BT}{M(A_0)}\right), \mathcal{L}\dot{A}_0 \right\rangle_A}{\left\langle \left(\dot{A}_0^2 - \frac{k_BT}{M(A_0)}\right)^2 \right\rangle_A} \left(\dot{A}_0^2 - \frac{k_BT}{M(A_0)}\right) \tag{3.54b}$$

$$= \frac{\left\langle \delta(A_0 - A)\left(\dot{A}_0^2 - \frac{k_BT}{M(A_0)}\right), \mathcal{L}\dot{A}_0 \right\rangle}{\left\langle \delta(A_0 - A)\left(\dot{A}_0^2 - \frac{k_BT}{M(A_0)}\right)^2 \right\rangle} \left(\dot{A}_0^2 - \frac{k_BT}{M(A_0)}\right). \tag{3.54c}$$

$M(A) = k_BT/\langle \dot{A}_0^2 \rangle_A$ denotes the position dependent effective mass of the reaction coordinate discussed in section 3.1.3. In section 3.1.3, we used the fact that the velocity $\dot{A}_0$ is Gaussian distributed for given A_0 to derive a formula for the mass $M(A)$. We now further exploit the Gaussian nature of the velocity to simplify the denominator in eq. (3.54c).

$$\mathcal{N}(A) \equiv \left\langle \delta(A_0 - A)\left(\dot{A}_0^2 - \frac{k_BT}{M(A_0)}\right)^2 \right\rangle \tag{3.55a}$$

$$= \langle \delta(A_0 - A)\dot{A}_0^4 \rangle - 2\frac{k_BT}{M(A)}\langle \delta(A_0 - A)\dot{A}_0^2 \rangle + \left(\frac{k_BT}{M(A)}\right)^2 \langle \delta(A_0 - A) \rangle \tag{3.55b}$$

$$= \langle \delta(A_0 - A) \rangle \left(\frac{\langle \delta(A_0 - A)\dot{A}_0^4 \rangle}{\langle \delta(A_0 - A) \rangle} - 2\frac{k_BT}{M(A)} \frac{\langle \delta(A_0 - A)\dot{A}_0^2 \rangle}{\langle \delta(A_0 - A) \rangle} + \left(\frac{k_BT}{M(A)}\right)^2 \right) \tag{3.55c}$$

$$= \mathbb{P}(A) \left(\langle \dot{A}_0^4 \rangle_A - \left(\frac{k_BT}{M(A)}\right)^2 \right), \tag{3.55d}$$

where we used $\mathbb{P}(A) \equiv \langle \delta(A_0 - A) \rangle$. The fourth moment of the velocity is given by $\langle \dot{A}_0^4 \rangle_A = 3(k_BT/M(A))^2$. Thus, we find

$$\mathcal{N}(A) = 2\left(\frac{k_BT}{M(A)}\right)^2 \mathbb{P}(A). \tag{3.56}$$

With the result of eq. (3.56), we go back to eq. (3.54c)

$$\mathcal{P}_{\mathrm{M}}^{\mathrm{K}}\mathcal{L}\dot{A}_0 = \frac{\left\langle \delta(A_0 - A)\left(\dot{A}_0^2 - \frac{k_BT}{M(A_0)}\right), \mathcal{L}\dot{A}_0 \right\rangle}{2\left(\frac{k_BT}{M(A)}\right)^2 \mathbb{P}(A)} \left(\dot{A}_0^2 - \frac{k_BT}{M(A_0)}\right), \tag{3.57}$$

and we focus on the fraction on the r.h.s. of eq. (3.57)

$$\frac{\left\langle \delta(A_0 - A)\left(\dot{A}_0^2 - \frac{k_BT}{M(A_0)}\right), \mathcal{L}\dot{A}_0 \right\rangle}{2\left(\frac{k_BT}{M(A)}\right)^2 \mathbb{P}(A)} = \frac{\left\langle \dot{A}_0^2, \mathcal{L}\dot{A}_0 \right\rangle_A}{2\left(\frac{k_BT}{M(A)}\right)^2} - \frac{\left\langle \mathcal{L}\dot{A}_0 \right\rangle_A}{2\frac{k_BT}{M(A)}}. \tag{3.58}$$

The term $\langle \mathcal{L}\dot{A}_0 \rangle_A$ was computed in eq. (3.14c) and defined as $-(1/M(A))(\mathrm{d}U_{\text{eff}}/\mathrm{d}A)$ in eq. (3.16), i.e.,

$$-\frac{\left\langle \mathcal{L}\dot{A}_0 \right\rangle_A}{2\frac{k_BT}{M(A)}} = \frac{1}{2k_BT}\frac{\mathrm{d}U_{\text{eff}}(A)}{\mathrm{d}A}. \tag{3.59}$$

The last term remaining is the first term on the r.h.s. of eq. (3.58). For this term, we find

$$\frac{\left\langle \dot{A}_0^2, \mathcal{L}\dot{A}_0 \right\rangle_A}{2\left(\frac{k_BT}{M(A)}\right)^2} = \frac{\left\langle \delta(A_0 - A)\dot{A}_0^2, \mathcal{L}\dot{A}_0 \right\rangle}{2\left(\frac{k_BT}{M(A)}\right)^2 \mathbb{P}(A)} \tag{3.60a}$$

$$= \frac{\frac{\mathrm{d}}{\mathrm{d}A}\left\langle \delta(A_0 - A)\dot{A}_0^4 \right\rangle}{2\left(\frac{k_BT}{M(A)}\right)^2 \mathbb{P}(A)} - 2\frac{\left\langle \delta(A_0 - A)\dot{A}_0^2, \mathcal{L}\dot{A}_0 \right\rangle}{2\left(\frac{k_BT}{M(A)}\right)^2 \mathbb{P}(A)}, \tag{3.60b}$$

from which follows that

$$\frac{\left\langle \dot{A}_0^2, \mathcal{L}\dot{A}_0 \right\rangle_A}{2\left(\frac{k_BT}{M(A)}\right)^2} = \frac{1}{3}\frac{\frac{\mathrm{d}}{\mathrm{d}A}\left\langle \delta(A_0 - A)\dot{A}_0^4 \right\rangle}{2\left(\frac{k_BT}{M(A)}\right)^2 \mathbb{P}(A)} \tag{3.61a}$$

$$= \frac{1}{3}\frac{\frac{\mathrm{d}}{\mathrm{d}A}\left(3\frac{k_BT}{M(A)}\right)^2 \mathbb{P}(A)}{2\left(\frac{k_BT}{M(A)}\right)^2 \mathbb{P}(A)} \tag{3.61b}$$

$$= -\frac{\mathrm{d}}{\mathrm{d}A}\left(\ln M(A) - \frac{1}{2}\ln \mathbb{P}(A)\right) \tag{3.61c}$$

$$= -\frac{\mathrm{d}}{\mathrm{d}A}\left(\ln M(A) + \frac{1}{2k_BT}U_{\text{PMF}}(A)\right) \tag{3.61d}$$

$$= -\frac{1}{2k_BT}\frac{\mathrm{d}U_{\text{eff}}(A)}{\mathrm{d}A} - \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}A}\ln M(A). \tag{3.61e}$$

Inserting the result in eq. (3.59) and eq. (3.61e) into eq. (3.58), we get

$$\frac{\left\langle \delta(A_0 - A)\left(\dot{A}_0^2 - \frac{k_BT}{M(A_0)}\right), \mathcal{L}\dot{A}_0 \right\rangle}{2\left(\frac{k_BT}{M(A)}\right)^2 \mathbb{P}(A)} = -\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}A}\ln M(A) \tag{3.62}$$

This leads us to the following result for the potential term of the GLE

$$\mathcal{P}^{\mathrm{K}}\mathcal{L}\dot{A}_0 = \mathcal{P}_{\mathrm{Z}}^{\mathrm{H}}\mathcal{L}\dot{A}_0 + \mathcal{P}_{\mathrm{M}}^{\mathrm{K}}\mathcal{L}\dot{A}_0 \tag{3.63a}$$

$$= -\frac{1}{M(A)}\frac{\mathrm{d}}{\mathrm{d}A}\left(U_{\text{PMF}}(A) + \frac{k_BT}{2}\ln M(A) + \frac{M(A)}{2}\dot{A}^2\right). \tag{3.63b}$$

3.3.2 Final form of the GLE

With the result of eq. (3.63), the GLE reads

$$\ddot{A}_t = -\frac{1}{M(A_t)}\frac{\mathrm{d}U(A_t, \dot{A}_t)}{\mathrm{d}A_t} - \int_0^t \mathrm{d}s\, \Gamma^{\mathrm{K}}(A_s, t-s)\dot{A}_s + F_{\mathrm{R}}^{\mathrm{K}}(t), \tag{3.64a}$$

with

$$U(A,\dot{A}) = U_{\mathrm{PMF}}(A) + \frac{k_B T}{2}\ln M(A) + \frac{M(A)}{2}\dot{A}^2, \tag{3.64b}$$

$$\Gamma^{\mathrm{K}}(A,t) = \beta D^{\mathrm{K}}(A,t)\frac{\mathrm{d}U_{\mathrm{eff}}(A)}{\mathrm{d}A} - \frac{\mathrm{d}D^{\mathrm{K}}(A,t)}{\mathrm{d}A} + \frac{\langle \ddot{A}_0, F^{\mathrm{K}}_{\mathrm{R}}(t)\rangle_A}{\langle \dot{A}_0^2\rangle_A}, \tag{3.64c}$$

$$D^{\mathrm{K}}(A,t) = \frac{\langle \dot{A}_0^2, F^{\mathrm{K}}_{\mathrm{R}}(t)\rangle_A}{\langle \dot{A}_0^2\rangle_A}. \tag{3.64d}$$

Again, $M(A) = k_B T/\langle \dot{A}_0^2\rangle_A$ denotes the position dependent effective mass and $U_{\mathrm{eff}}(A)$ is given in eq. (3.16). Since the projection $\mathcal{P}^{\mathrm{K}}$ is an orthogonal projection, the following relations hold

$$\langle F^{\mathrm{K}}_{\mathrm{R}}(t)\rangle_A = 0, \qquad \langle \dot{A}_0, F^{\mathrm{K}}_{\mathrm{R}}(t)\rangle_A = 0. \tag{3.65}$$

An algorithm to extract GLE parameters from time series data by explicitly computing the random force $F_R(t)$ was introduced in refs. [16, 72] and extended to GLEs in the presence of finite, in general nonlinear, potentials and nonlinear friction in ref. [4]. Numerical extraction schemes will be the topic of the next chapter.

3.4 Conclusion

We introduce three projection operators and use them to incorporate nonlinear memory effects into the generalized Langevin equation.

Using a hybrid-projection scheme that combines linear Mori projection on the reaction coordinate velocities and nonlinear conditional Zwanzig projection on the reaction coordinates themselves, we derive a GLE that contains the nonlinear potential of mean force and a nonlinear memory friction contribution that is a function of the reaction coordinate A_t but not of its velocity $\dot{A}_t$. The complete memory friction splits into two parts. One part is linear in the reaction coordinate velocity and reflects linear friction proportional to a memory kernel $\Gamma^{\mathrm{H}}_{\mathrm{M}}(t)$. The memory kernel $\Gamma^{\mathrm{H}}_{\mathrm{M}}(t)$ is related to the fluctuating force $F^{\mathrm{H}}_{\mathrm{R}}(t)$, defined in eq. (2.26b), by a relation that resembles a fluctuation-dissipation theorem, eq. (3.18).

The nonlinear memory friction function $\Gamma^{\mathrm{H}}_{\mathrm{Z}}(A,t)$ accounts for nonlinear dependencies of friction on A_t and is connected to the fluctuating force $F^{\mathrm{H}}_{\mathrm{R}}(t)$ by a conditional correlation function, given in eq. (3.19c). Thus, when modeling $F^{\mathrm{H}}_{\mathrm{R}}(t)$ as a stochastic variable, it simultaneously has to fulfill both relations, eq. (3.18) and eq. (3.19c).

The approximate GLE in eq. (2.34) is obtained from our GLE in eq. (3.20) for the case when the memory function $\Gamma^{\mathrm{H}}_{\mathrm{Z}}(A,t)$ vanishes. From eq. (3.19c), it follows that this happens whenever the conditional correlation between the velocity and the random force in eq. (3.19c) vanishes. Thus, eq. (3.19c) establishes a firm criterion for the validity of the approximate GLE.

The second projection operator generates a nonlinear friction kernel of the form

$$\Gamma^{\mathrm{L}}(A,t) = \beta D^{\mathrm{L}}(A,t)\frac{\mathrm{d}U_{\mathrm{eff}}(A)}{\mathrm{d}A} - \frac{\mathrm{d}D^{\mathrm{L}}(A,t)}{\mathrm{d}A} + \frac{\langle \ddot{A}_0, F_{\mathrm{R}}^{\mathrm{L}}(t)\rangle_A}{\langle \dot{A}_0^2\rangle_A}. \tag{3.66}$$

The first two terms on the r.h.s. of eq. (3.66) resemble the structure of $\Gamma_{\mathrm{Z}}^{\mathrm{H}}(A,t)$ in eq. (3.19b). The rightmost term in eq. (3.66) can be written as

$$\frac{\langle \ddot{A}_0, F_{\mathrm{R}}^{\mathrm{L}}(s)\rangle_A}{\langle \dot{A}_0^2\rangle_A} = \frac{\langle \mathcal{Q}^{\mathrm{L}}\mathcal{L}\dot{A}_0, F_{\mathrm{R}}^{\mathrm{L}}(s)\rangle_A}{\langle \dot{A}_0^2\rangle_A} = \frac{\langle F_{\mathrm{R}}^{\mathrm{L}}(0), F_{\mathrm{R}}^{\mathrm{L}}(s)\rangle_A}{\langle \dot{A}_0^2\rangle_A}, \tag{3.67}$$

i.e., it has the structure of a fluctuation-dissipation theorem and resembles the form of $\Gamma_{\mathrm{M}}^{\mathrm{H}}$ in the hybrid GLE, but with conditional correlations.

In order to obtain the approximate GLE from the linear velocity GLE in eq. (3.44), not only the function D^{L} has to vanish, but one would also need to be able to replace the conditional correlations by non-conditional correlations, i.e.,

$$\frac{\langle F_{\mathrm{R}}^{\mathrm{L}}(0), F_{\mathrm{R}}^{\mathrm{L}}(s)\rangle_A}{\langle \dot{A}_0^2\rangle_A} = \frac{\langle F_{\mathrm{R}}^{\mathrm{L}}(0), F_{\mathrm{R}}^{\mathrm{L}}(s)\rangle}{\langle \dot{A}_0^2\rangle}. \tag{3.68}$$

The memory function of the third nonlinear friction GLE has the same form since it is also generated by $\mathcal{P}^{\mathrm{L}}$. But the potential term incorporates a kinetic term proportional to the velocity squared. As pointed out in the discussion leading to eq. (3.48), the velocity squared term is important when mapping the exact GLE onto a stochastic model (cf. chapter 6).

4 Numerical Estimation of GLE Parameters via Projected Correlation Functions

In this chapter, we discuss a numerical method to extract GLE parameters, e.g., nonlinear memory functions, from time series data. The method is based on the idea of *forward orthogonal propagation* to compute *projected correlation functions* [16, 72, 4].

The memory term in all GLEs presented in chapter 3 included a correlation function of the form

$$\langle * \, , \mathcal{L} F_{\mathrm{R}}(t) \rangle = -\langle \mathcal{L} *, F_{\mathrm{R}}(t) \rangle, \tag{4.1}$$

where $*$ represents a placeholder for the projection function. Since the random force is propagated by the projected propagator $e^{t\mathcal{Q}\mathcal{L}}$, i.e., $F_{\mathrm{R}}(t) = e^{t\mathcal{Q}\mathcal{L}} F_{\mathrm{R}}(0)$, the correlation function in eq. (4.1) is referred to as a projected correlation function [16].
In the first section, we discuss how to utilize projected correlation functions to compute nonlinear friction kernels in GLEs.

In the second section, we validate our numerical extraction scheme by computing the memory functions of all GLEs introduced in chapter 3. The underlying time series data is a trajectory for the dihedral angle of a butane molecule in water, obtained from all-atom MD simulations.
Interestingly, we find that the significance of nonlinear memory effects depends on how one accounts for the position-dependent mass.

4.1 GLE parameters from projected correlation functions

In the absence of a potential and in the absence of nonlinear friction, Carof et al. presented iterative algorithms to compute the random force trajectory and the linear friction kernel from a trajectory of the reaction coordinate [16, 72]. Their derivations explicitly use the Mori projection, so, strictly speaking, the results are only valid for the Mori GLE in eq. (2.31).

Given a trajectory of an arbitrary observable, we now introduce a method to compute the random force trajectory $F_{\mathrm{R}}(\omega_0, t)$ and from that a memory function that may depend on the value of the reaction coordinate itself. For this, let us consider the projected propagator $e^{t\mathcal{Q}\mathcal{L}}$ introduced in eq. (2.25). From the Dyson decomposition in eq. (2.25), we obtain by rearranging

$$e^{t\mathcal{Q}\mathcal{L}} = e^{t\mathcal{L}} - \int_0^t \mathrm{d}s \, e^{(t-s)\mathcal{L}} \mathcal{P}\mathcal{L} e^{s\mathcal{Q}L}. \tag{4.2}$$

Applying eq. (4.2) on the initial random force $F_{\mathrm{R}}(\omega_0, 0)$ and using eq. (2.26b), we find

$$F_{\mathrm{R}}(\omega_0, t) = e^{t\mathcal{L}} F_{\mathrm{R}}(\omega_0, 0) - \int_0^t \mathrm{d}s \, e^{(t-s)\mathcal{L}} \mathcal{P}\mathcal{L} F_{\mathrm{R}}(\omega_0, s). \tag{4.3}$$

Now, we consider eq. (4.2) at time $t + \Delta t$

$$e^{(t+\Delta t)\mathcal{Q}\mathcal{L}} = e^{t\mathcal{L}} e^{\Delta t \mathcal{L}} - \int_0^{t+\Delta t} \mathrm{d}s\, e^{(t-s)\mathcal{L}} e^{\Delta t \mathcal{L}} \mathcal{P}\mathcal{L} e^{s\mathcal{Q}\mathcal{L}}. \tag{4.4}$$

Splitting up the integral on the r.h.s. into two parts gives

$$e^{(t+\Delta t)\mathcal{Q}\mathcal{L}} = e^{t\mathcal{L}} e^{\Delta t \mathcal{L}} - \int_0^{t} \mathrm{d}s\, e^{(t-s)\mathcal{L}} e^{\Delta t \mathcal{L}} \mathcal{P}\mathcal{L} e^{s\mathcal{Q}\mathcal{L}} - \int_0^{\Delta t} \mathrm{d}s\, e^{(\Delta t-s)\mathcal{L}} \mathcal{P}\mathcal{L} e^{(t+s)\mathcal{Q}\mathcal{L}}, \tag{4.5}$$

where we used the substitution $s \to s - t$ in the second integral. Acting with the operator in eq. (4.5) on the initial random force $F_{\mathrm{R}}(\omega_0, 0)$ leads to

$$\begin{aligned} F_{\mathrm{R}}(\omega_0, t + \Delta t) = e^{\Delta t \mathcal{L}} \left(e^{t\mathcal{L}} F_{\mathrm{R}}(\omega_0, 0) - \int_0^t \mathrm{d}s\, e^{(t-s)\mathcal{L}} \mathcal{P}\mathcal{L} F_{\mathrm{R}}(\omega_0, s) \right) \\ - \int_0^{\Delta t} \mathrm{d}s\, e^{(\Delta t - s)\mathcal{L}} \mathcal{P}\mathcal{L} F_{\mathrm{R}}(\omega_0, t + s). \end{aligned} \tag{4.6}$$

Comparing eq. (4.6) with eq. (4.3), we see that the first two terms on the r.h.s. of eq. (4.6) are equal to $e^{\Delta t \mathcal{L}} F_{\mathrm{R}}(\omega_0, t) = F_{\mathrm{R}}(\omega_{\Delta t}, t)$, where we used eq. (2.5). Hence, we find

$$F_{\mathrm{R}}(\omega_0, t + \Delta t) = F_{\mathrm{R}}(\omega_{\Delta t}, t) - \int_0^{\Delta t} \mathrm{d}s\, \Gamma(A_{\Delta t - s}, \dot{A}_{\Delta t - s}, t + s). \tag{4.7}$$

Here, $\Gamma(A_{\Delta t-s}, \dot{A}_{\Delta t-s}, t + s) = e^{(\Delta t-s)\mathcal{L}} \mathcal{P}\mathcal{L} F_{\mathrm{R}}(\omega_0, t + s)$ denotes a general memory function.
For given trajectories $A_t, \dot{A}_t$ and given random force $F_{\mathrm{R}}(\omega_{\Delta t}, t)$ as a function of the phase space configuration $\omega_{\Delta t}$, eq. (4.7) gives the random force $F_{\mathrm{R}}(\omega_0, t + \Delta t)$ one time step Δt later as a function of the phase space configuration ω_0 one time step Δt before. To obtain an iterative scheme for the random force, eq. (4.7) is discretized in time and A-space.

First, we consider the easiest discretization, namely, the left rectangular rule to discretize the time integrals. The random force is discretized as $F_{\mathrm{R}}(\omega_{t'}, t) = F_{\mathrm{R}}(\omega_{i\Delta t}, j\Delta t) \equiv F_{\mathrm{R}}(i, j)$. The bounded A-space, $A \in [a, a + N_A \Delta A]$, is divided into N_A bins with width ΔA, the bin intervals are labeled by $I_\alpha = [a + \alpha \Delta A, a + (\alpha + 1)\Delta A]$ with $\alpha = 0, 1, 2, \ldots, N_A - 1$. For the sake of clarity, we give the discretization explicitly for the hybrid GLE presented in chapter 3.1, i.e., for the case

$$\Gamma(A_{\Delta t-s}, \dot{A}_{\Delta t-s}, t + s) = -\Gamma^{\mathrm{H}}_{\mathrm{M}}(t + s) \dot{A}_{\Delta t-s} + \Gamma^{\mathrm{H}}_{\mathrm{Z}}(A_{\Delta t-s}, t + s), \tag{4.8}$$

The memory kernel $\Gamma^{\mathrm{H}}_{\mathrm{M}}, \Gamma^{\mathrm{H}}_{\mathrm{Z}}$ are given in eq. (3.18) and eq. (3.19), respectively.

$$F^{\mathrm{H}}_{\mathrm{R}}(i, j + 1) = F^{\mathrm{H}}_{\mathrm{R}}(i + 1, j) + \Delta t\, \Gamma^{\mathrm{H}}_{\mathrm{M}}(j) \dot{A}_{i+1} - \Delta t\, \Gamma^{\mathrm{H}}_{\mathrm{Z}}(A_{i+1}, j), \tag{4.9a}$$

$$\Gamma^{\mathrm{H}}_{\mathrm{M}}(j) = \frac{\sum_{i=0}^{N_{\mathrm{traj}}-j-1} F^{\mathrm{H}}_{\mathrm{R}}(i, 0) F^{\mathrm{H}}_{\mathrm{R}}(i, j)}{\sum_{i=0}^{N_{\mathrm{traj}}-j-1} \dot{A}_i^2}, \tag{4.9b}$$

$$\Gamma^{\mathrm{H}}_{\mathrm{Z}}(A_{i+1}, j) = \left[\hat{\Delta}_\alpha D^{\mathrm{H}}(\alpha, j) - \beta\, D^{\mathrm{H}}(\alpha, j)\, \hat{\Delta}_\alpha U_{\mathrm{PMF}}(\alpha) \right]_{A_{i+1} \in I_\alpha}, \tag{4.9c}$$

$$D^{\mathrm{H}}(\alpha, j) = \frac{\sum_{\substack{i \le N_{\mathrm{traj}}-j-1 \\ A_i \in I_\alpha}} \dot{A}_i F^{\mathrm{H}}_{\mathrm{R}}(i, j)}{\sum_{\substack{i \le N_{\mathrm{traj}}-j-1 \\ A_i \in I_\alpha}} 1}. \tag{4.9d}$$

In eq. (4.9c), $\hat{\Delta}_\alpha$ denotes a numerical first derivative w.r.t. α. We use the central difference derivative, i.e.,

$$\hat{\Delta}_\alpha D^{\mathrm{H}}(\alpha, j) = \frac{D^{\mathrm{H}}(\alpha+1, j) - D^{\mathrm{H}}(\alpha-1, j)}{2\Delta A}. \tag{4.10}$$

If the observable A_t has at time $t = i\Delta t$ a value in the interval I_α, we write $A_i \in I_\alpha$; $\sum_{A_i \in I_\alpha}$ denotes the sum over all times i for which A_i is in the interval I_α, which is used to compute conditional averages in eq. (4.9). N_{traj} denotes the total length of the A_t trajectory. The sums run from $i = 0$ to $N_{\mathrm{traj}} - j - 1$, because for given j, the iterative scheme has only determined the random force at times up to $N_{\mathrm{traj}} - j - 1$, as follows from eq. (4.9a). The sums in the denominator extend over the same interval as in the numerator in order to increase the numerical stability [16, 72].

The iterative scheme in eq. (4.9) works as follows: First, note from eq. (2.26a) that the initial value of the random force is given by

$$F_{\mathrm{R}}(\omega_0, 0) = \ddot{A}(\omega_0) - \mathcal{P}\mathcal{L}\dot{A}(\omega_0), \tag{4.11}$$

where the rightmost term denotes the potential term, i.e., for the hybrid GLE the initial random force is

$$F_{\mathrm{R}}^{\mathrm{H}}(i, 0) = \ddot{A}_i + \frac{1}{M(A_i)} \frac{\mathrm{d}}{\mathrm{d}A_i} \left[U_{\mathrm{PMF}}(A_i) + k_B T \ln M(A_i) \right]. \tag{4.12}$$

The random force at time $t = 0$ equals the acceleration plus the force from the effective potential for all possible initial times $i\Delta t$ for $i = 0, 1, 2, \ldots, N_{\mathrm{traj}} - 1$. This, together with $\dot{A}_i$, can be obtained directly from a given trajectory of the observable A. Then, $F_{\mathrm{R}}^{\mathrm{H}}(i, 0)$, A_i, and $\dot{A}_i$ are inserted into eq. (4.9) to compute $F_{\mathrm{R}}^{\mathrm{H}}(i, 1)$ for $i = 0, 1, 2, \ldots, N_{\mathrm{traj}} - 2$. $F_{\mathrm{R}}^{\mathrm{H}}(i, 1)$ is then used to compute $F_{\mathrm{R}}^{\mathrm{H}}(i, 2)$ for $i = 0, 1, 2, \ldots, N_{\mathrm{traj}} - 3$ and so forth. While computing $F_{\mathrm{R}}^{\mathrm{H}}(i, j)$, the memory friction functions $\Gamma_{\mathrm{M}}^{\mathrm{H}}(j)$ and $\Gamma_{\mathrm{Z}}^{\mathrm{H}}(A, j)$ are computed simultaneously.

If our only goal is to compute the memory friction functions, we can stop the computation of $F_{\mathrm{R}}^{\mathrm{H}}(i, j)$ as soon as the memory functions have dropped to zero. As an example, if the memory functions decay to zero after N_{mem} time steps, we can abort the computation of the random force at $F_{\mathrm{R}}^{\mathrm{H}}(i, N_{\mathrm{mem}})$. At that point, we generated $N_{\mathrm{traj}} - N_{\mathrm{mem}} - 1$ distinct random-force trajectories of length N_{mem} each. Since the memory functions are computed simultaneously, the generated random-force trajectories only need to be stored if one is interested in the random-force statistics, in which case one could extend the length of the random-force trajectories.

Stabilizing the numerical extraction scheme

The discretization in eq. (4.9a) is based on the left rectangular rule. The numerical accuracy can be improved by using the trapezoidal rule to discretize the time integral in eq. (4.7). This leads to

$$F_{\mathrm{R}}(\omega_i, j+1) = F_{\mathrm{R}}(\omega_{i+1}, j) - \frac{\Delta t}{2} \Gamma(A_{i+1}, \dot{A}_{i+1}, j) - \frac{\Delta t}{2} \Gamma(A_i, \dot{A}_i, j+1). \tag{4.13}$$

The drawback of the trapezoidal rule is that the term $\Gamma(A_i, \dot{A}_i, j+1)$ is unknown at iteration step j. To estimate $\Gamma(A_i, \dot{A}_i, j+1)$, we use the left-point derivative, i.e.,

$$\begin{aligned}\Gamma(A_i, \dot{A}_i, j+1) &\approx \Gamma(A_i, \dot{A}_i, j) + \Delta t \frac{\Gamma(A_i, \dot{A}_i, j) - \Gamma(A_i, \dot{A}_i, j-1)}{\Delta t} \\ &= 2\,\Gamma(A_i, \dot{A}_i, j) - \Gamma(A_i, \dot{A}_i, j-1)\,. \end{aligned} \tag{4.14}$$

With this, the iterative scheme for the random force becomes

$$\begin{aligned}F_{\mathrm{R}}(\omega_i, j+1) = F_{\mathrm{R}}(\omega_{i+1}, j) - \frac{\Delta t}{2}\,\Gamma(A_{i+1}, \dot{A}_{i+1}, j) - \Delta t \Gamma(A_i, \dot{A}_i, j) \\ + \frac{\Delta t}{2}\Gamma(A_i, \dot{A}_i, j-1)\,. \end{aligned} \tag{4.15}$$

We set the derivative at $j = 0$ to be zero.

To stabilize the numerical estimation even further, we make use of the following relation which holds for all discussed GLEs in chapter 3

$$\langle F_{\mathrm{R}}(t) \rangle_A = 0. \tag{4.16}$$

For values A with a large number of occurences in the trajectory A_t, i.e., in bins with a high frequency of occurence, the relation in eq. (4.16) holds sufficiently well. It gets problematic at boundary regions that are rarely visited along the trajectory. To counteract instabilities and in order to estimate averages better in those critical regions, we compute the conditional mean of the random force, i.e., the mean in each bin I_α, at each step j. If the mean in a bin is larger than a chosen threshold value ε, we subtract the mean from all values that lie in this bin, i.e.,

$$\langle F_{\mathrm{R}}(t) \rangle_A > \varepsilon \Rightarrow F_{\mathrm{R}}(\omega, t) = F_{\mathrm{R}}(\omega, t) - \langle F_{\mathrm{R}}(t) \rangle_A,\ \forall \omega \text{ with } A(\omega) = A. \tag{4.17}$$

4.2 Numerical results for the dihedral angle in Butane

In this section, we analyze dihedral angle trajectories of a single butane molecule in water generated from all-atom force field molecular dynamics (MD) simulations. We apply the numerical extraction scheme based on projected correlation functions from section 4.1 on these trajectories to compute the memory functions discussed in chapter 3.

MD simulations are performed using the Gromacs MD package (version 2020-Modified) [97]. For the MD simulation of the butane molecule, we use the GROMOS53A6 force field [94] with the TIP4P/2005 rigid water model [1]. The simulation box has side lengths of 3.35 nm and contains 1250 water molecules and one butane molecule. We constrain the butane bond lengths and angles using the SHAKE algorithm [99]. For long-range electrostatic interactions, we use the particle-mesh Ewald [28], with a cut-off of 1 nm. The simulation time step is 2 fs. We generate 50 trajectories, each of length 20 ns, i.e., the total simulation time is 1 μs. All simulations are performed in the NVT ensemble with a temperature of 300 K, controlled with a velocity rescaling thermostat [14]. Input files of

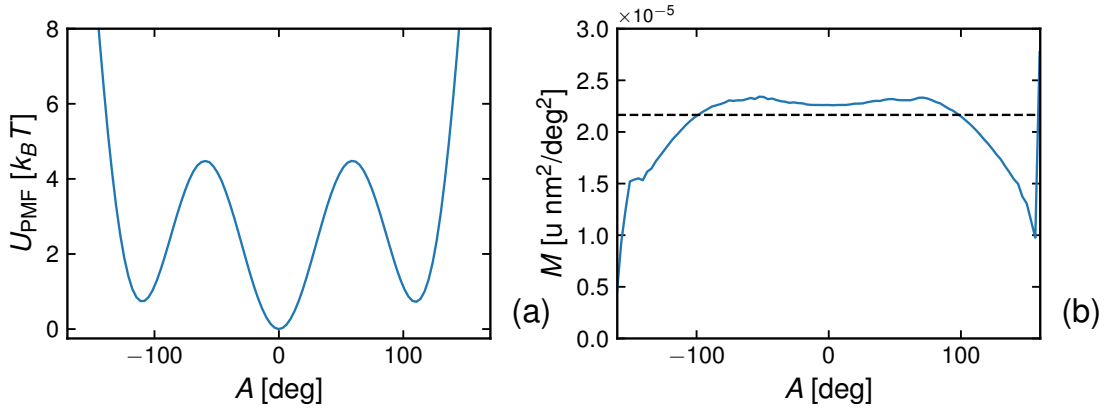

Figure 4.1: **(a)**: The potential of mean force computed from 50 dihedral angle trajectories of a single butane molecule in water. The trajectories are generated from all-atom force field MD simulations. **(b)**: The position-dependent mass according to eq. (3.15) for the dihedral angle reaction coordinate (blue solid line), computed using the same data as in (a). The broken black line shows the constant mass that follows from the equipartition theorem, i.e., $M = k_B T / \langle \dot{A}_0^2 \rangle$.

the MD simulations are available upon request. Our Python codes for extracting the GLE parameters and running Langevin simulations are also available upon request.

The dihedral angle of butane is a conceptually simple yet relevant observable and provides a simple scenario to study conformational transitions in polymers and proteins that is both theoretically [19] and experimentally [112] accessible.

In fig. 4.1(a), we show the potential of mean force $U_{\mathrm{PMF}}(A) = -k_B T \ln \mathbb{P}(A)$ for the dihedral angle. The PMF does not equal the effective potential U_{eff} in eq. (3.16) since the effective mass $M(A)$ of the dihedral angle, as shown in fig. 4.1(b), depends on the value of the dihedral angle. Between the values $A \in [-140, 140]$ deg, the mass varies between $M(A) \in [1.61, 2.34] \cdot 10^{-5}$ u nm^2/deg^2.

In fig. 4.2(a), we compare the memory kernel $\Gamma_{\mathrm{M}}^{\mathrm{H}}(t)$ from eq. (3.18) with the memory kernel $\Gamma^{\mathrm{A}}(t)$ from the approximate GLE in the presence of a position-dependent mass in eq. (5.5). The approximate GLE with a position-dependent mass is introduced in chapter 5 and contains no nonlinear memory contributions. The memory function $\Gamma^{\mathrm{A}}(t)$ is computed using the extraction scheme in chapter 5.2 [26, 66, 5] by making use of the relation in eq. (5.3). The deviations between the two memory functions are particularly well visible in the running integrals shown in the inset of fig. 4.2(a). $\Gamma_{\mathrm{M}}^{\mathrm{H}}(t)$ is defined within the exact GLE in eq. (3.20) and determined numerically from the MD trajectory via eq. (4.15). The pronounced deviations already suggest that nonlinear friction effects, not captured by $\Gamma^{\mathrm{A}}(t)$, are present. This is confirmed in fig. 4.2(b) where we show that the correlation function $D^{\mathrm{H}}(A, t)$, defined in eq. (3.19c), does not vanish for all times t. In fig. 4.2(b), $D^{\mathrm{H}}(A, t)$ is shown for a few different fixed times. Thus, a nonlinear memory contribution due to $\Gamma_{\mathrm{Z}}^{\mathrm{H}}(A, t)$, defined in eq. (3.19), is present in the hybrid GLE. As mentioned before, from the definition of $D^{\mathrm{H}}(A, t)$ in eq. (3.19c), it follows that $D^{\mathrm{H}}(A, t)$ vanishes at time $t = 0$, i.e., $D^{\mathrm{H}}(A, 0) = 0$, as confirmed by fig. 4.2(b). For finite time, $D^{\mathrm{H}}(A, t)$ rises in amplitude

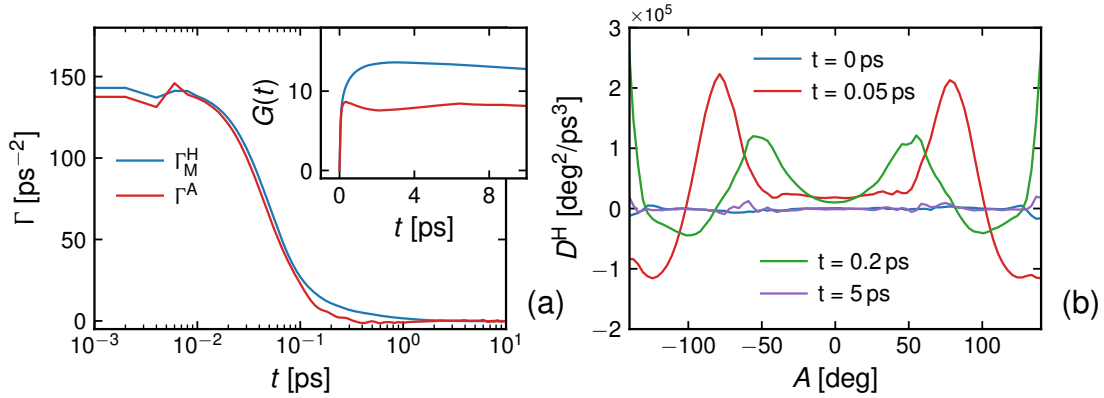

Figure 4.2: **(a)**: Comparison of the memory kernel $\Gamma^{\mathrm{H}}_{\mathrm{M}}(t)$ from the hybrid GLE in eq. (3.20) with the kernel $\Gamma^{\mathrm{A}}(t)$ from the approximate GLE given in eq. (5.5). $\Gamma^{\mathrm{H}}_{\mathrm{M}}(t)$ is computed using the extraction scheme in eq. (4.15), $\Gamma^{\mathrm{A}}(t)$ is computed using the Volterra equation presented in chapter 5.2. The inset shows the respective running integrals $G(t) = \int_0^t \mathrm{d}s\, \Gamma(s)$. In the running integrals, the deviations are particularly visible. **(b)**: The correlation function $D^{\mathrm{H}}(A,t)$ from eq. (3.19c) as a function of the dihedral angle A at fixed times t.

before decaying to zero after a time corresponding to the memory time of $\Gamma^{\mathrm{H}}_{\mathrm{M}}(t)$ in fig. 4.2(a), which is about 1 ps [4].

In fig. 4.3, we compare the memory kernel $\Gamma^{\mathrm{L}}(A,t)$, defined within the GLE in eq. (3.44), with the memory kernel $\Gamma^{\mathrm{K}}(A,t)$, defined within the GLE in eq. (3.64). As discussed in chapter 3.3, the position dependent mass generates an additional term in eq. (3.64) that is quadratic in the velocity, which is missing in eq. (3.44). Since both GLEs are exact descriptions of the same dynamics, the quadratic velocity term must be accounted for by the random force in eq. (3.44). Thus, the memory kernel $\Gamma^{\mathrm{L}}(A,t)$ must depend on the mass in order to account for its position-dependency. On the other hand, we expect that the position dependency of the mass has a smaller impact on the random force in eq. (3.64), and with that on the memory kernel $\Gamma^{\mathrm{K}}(A,t)$, due to the presence of the quadratic velocity term in eq. (3.64).

With that in mind, we consider the memory kernel $\Gamma^{\mathrm{L}}(A,t)$ and $\Gamma^{\mathrm{K}}(A,t)$ shown in fig. 4.3 for the dihedral angle in butane. The underlying trajectories are the same as in fig. 4.2 and the numerical extraction is again performed using eq. (4.15). The memory kernel $\Gamma^{\mathrm{L}}(A,t)$ from eq. (3.44) is shown in fig. 4.3(a) as a function of time at fixed angles A and in fig. 4.3(c) as a function of the dihedral angle at different times t, in both figures, $\Gamma^{\mathrm{L}}(A,t)$ is depicted by solid lines. In fig. 4.3(a) and (c), the position dependency of $\Gamma^{\mathrm{L}}(A,t)$ is clearly visible. In contrast, the position dependency of the memory kernel $\Gamma^{\mathrm{K}}(A,t)$, shown in fig. 4.3(b) as solid lines and in fig. 4.3(c) by connected markers, is less significant compared to $\Gamma^{\mathrm{L}}(A,t)$. This is rationalized by our discussion in the previous chapter leading to eq. (3.48). Since the Markovian part in eq. (3.64) contains more information on the position dependent mass, its impact on the memory term is less significant.

Nevertheless, the pronounced difference between the two memory kernels in fig. 4.3(c), in

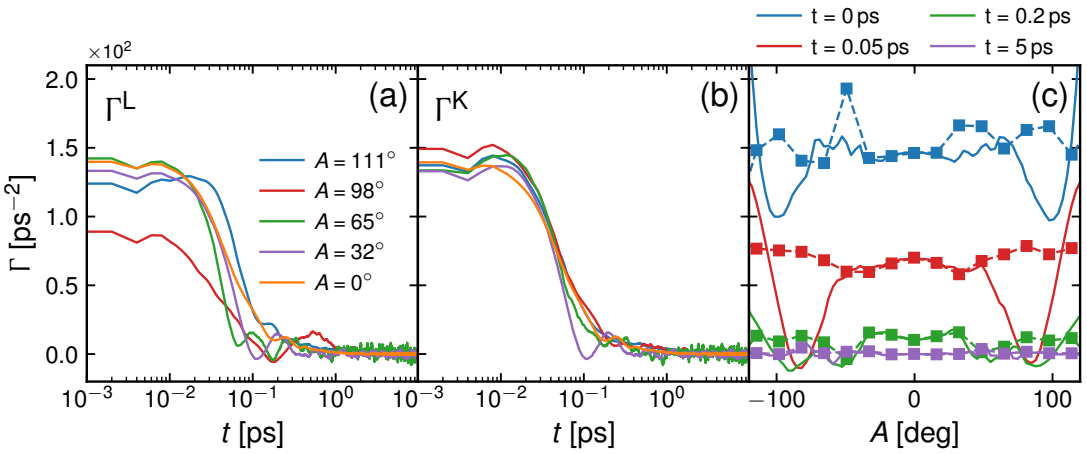

Figure 4.3: **(a)**: The memory kernel $\Gamma^{\mathrm{L}}(A,t)$ from the GLE in eq. (3.44) as a function of time for different values of the dihedral angle A. **(b)**: The same is shown for the memory kernel $\Gamma^{\mathrm{K}}(A,t)$ from the GLE in eq. (3.64). The colors correspond to the same dihedral angle values as in (a). **(c)**: Comparison of the two memory functions shown in (a) and (b), both are plotted against the dihedral angle A at fixed times t. Solid lines correspond to $\Gamma^{\mathrm{L}}(A,t)$ and the connected markers to $\Gamma^{\mathrm{K}}(A,t)$.

particular for the data at $t = 0.05\,\mathrm{ps}$ (red line and markers), is suprising since it can only be due to the position dependent mass. This is a consequence of the fact that the random forces in eq. (3.44) and eq. (3.64) are given by

$$F_{\mathrm{R}}^{\mathrm{L}}(t) = e^{t\mathcal{Q}^{\mathrm{L}}\mathcal{L}}\mathcal{Q}^{\mathrm{L}}\mathcal{L}\dot{A}_0, \qquad F_{\mathrm{R}}^{\mathrm{K}}(t) = e^{t\mathcal{Q}^{\mathrm{L}}\mathcal{L}}\mathcal{Q}^{\mathrm{K}}\mathcal{L}\dot{A}_0, \tag{4.18}$$

i.e., both random forces are propagated in time by $e^{t\mathcal{Q}^{\mathrm{L}}\mathcal{L}}$ and they are in general different functions because they differ at the initial time, $F_{\mathrm{R}}^{\mathrm{L}}(0) \neq F_{\mathrm{R}}^{\mathrm{K}}(0)$. However, when the mass is constant, it immediately follows from the GLEs in eq. (3.44) and eq. (3.64) that $F_{\mathrm{R}}^{\mathrm{L}}(0) = F_{\mathrm{R}}^{\mathrm{K}}(0)$ and, therefore, $F_{\mathrm{R}}^{\mathrm{L}}(t) = F_{\mathrm{R}}^{\mathrm{K}}(t)$ for all times t. This suggests that the deviations between $\Gamma^{\mathrm{L}}(A,t)$ and $\Gamma^{\mathrm{K}}(A,t)$ we observe in fig. 4.3 is mainly due to the position dependent mass shown in fig. 4.1(b).

The numerical extraction scheme in section 4.1 allows us to compute random force trajectories $F_R(\omega_0,t)$ in ω_0. Using these trajectories, we can estimate the distribution of the random force in each time step. This is shown in fig 4.4(a) for the random force from the hybrid GLE in eq. (3.20). The distribution of the random force is stationary. In fig. 4.4(b), we compare the random force distributions at $t = 1\,\mathrm{ps}$ for the three GLEs in eq. (3.20), eq. (3.44) and eq. (3.64). No significant deviations between the distributions can be discerned.

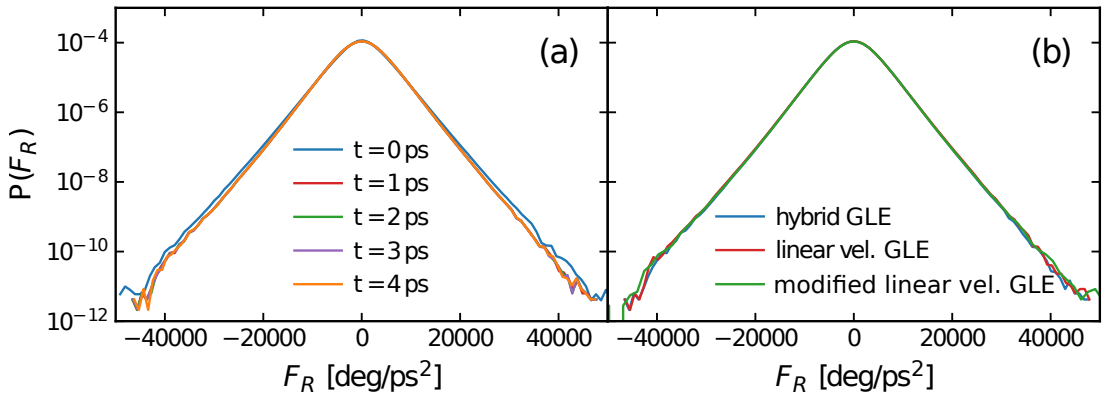

Figure 4.4: **(a)**: It is shown the random force distribution $\mathrm{P}(F_{\mathrm{R}}^{\mathrm{H}}(t))$ at different times t for the hybrid GLE in eq. (3.20), computed from the dihedral angle MD trajectories of a single butane molecule in water. The distribution is stationary. **(b)**: The random force distribution at $t = 1\,\mathrm{ps}$ for the three GLEs in eq. (3.20) (blue solid line), eq. (3.44) (red solid line) and eq. (3.64) (green solid line). The distributions are identical.

4.3 Conclusion

We introduce a numerical scheme to compute all parameters of nonlinear friction GLEs from a given trajectory A_t and apply the numerical scheme on a dihedral angle trajectory of butane in water, obtained from atomistic MD simulations.

We find that the effective mass of the dihedral angle depends on the value of the dihedral angle and that the nonlinear memory friction contribution in the hybrid GLE is finite and non-negligible. In order to estimate the importance of the nonlinear memory friction from the hybrid GLE, we have to compare the linear memory kernel $\Gamma_{\mathrm{M}}^{\mathrm{H}}(t)$ and the nonlinear memory function $\Gamma_{\mathrm{Z}}^{\mathrm{H}}(A,t)$. For this we multiply the linear-friction memory kernel at time zero, $\Gamma_{\mathrm{M}}^{\mathrm{H}}(0)$, by the root mean square velocity and obtain $\Gamma_{\mathrm{M}}^{\mathrm{H}}(0)\sqrt{\langle \dot{A}^2 \rangle} = 160\,\mathrm{ps}^{-3}$, which can be directly compared with the maximal value of the nonlinear memory friction function $\Gamma_{\mathrm{Z}}^{\mathrm{H}}(\tilde{A},\tilde{t}) = 104\,\mathrm{ps}^{-3}$, which is obtained for $\tilde{A} = 81^\circ$ and $\tilde{t} = 0.043$ ps. The value of $\Gamma_{\mathrm{Z}}^{\mathrm{H}}(\tilde{A},\tilde{t})$ thus turns out to be roughly half the value of $\Gamma_{\mathrm{M}}^{\mathrm{H}}(0)\sqrt{\langle \dot{A}^2 \rangle}$, which means that nonlinear memory friction effects are not negligible in the hybrid GLE in eq. (3.20).

We obtain a similar result from the linear velocity GLE in eq. (3.44). The memory function Γ^{L} changes significantly with A, which also suggestes that nonlinear memory effects cannot be neglected.

However, when we modify the mean force term by incorporating a term quadratic in the velocity, which accounts for a position-dependent mass, the dependency of the memory function Γ^{K} within the GLE in eq. (3.64) on A is less significant. We show that the difference between Γ^{L} and Γ^{K} can only be due to the presence of a position dependent mass and that Γ^{L} and Γ^{K} must be equal if the mass is constant.

Finally, we show that the equal-time random force distributions in all three GLEs exhibit small but detectable deviations from a Gaussian distribution, but the different memory and mean force terms seem to have no significant effect on this distribution.

All these results lead us to conclude that the nonlinear GLEs derived from projection schemes are practically useful and allow to detect and model nonlinear friction effects that have been neglected in previous applications of the approximate GLE with linear memory friction.

5 The Approximate GLE

In this chapter, we focus on the approximate GLE given in eq. (2.34) for an observable consisting of atomistic distances. First, we briefly dicuss under which conditions it arises from the hybrid GLE.

Following this, we present methods to numerically extract its parameters from time series data, and using the extracted parameters, perform computer simulations of the approximate GLE to generate trajectories from the GLE itself. The numerical extraction scheme presented in this chapter is different from the one discussed in chapter 4. We will use an extraction scheme based on inverting Volterra equations [51, 106, 26] and derive from this approach an iterative scheme to compute the running integral of the memory kernel [66, 5]. The GLE simulations are performed using Markovian embedding techniques [103, 5].

In the subsequent section, we study the Kramers Moyal expansion of the approximate GLE and show that all higher orders than first order Kramers Moyal coefficients of the approximate GLE vanish [5].

Finally, we study how accurately the approximate GLE can reproduce the folding dynamics of the polypeptide 9-alanine [5].

5.1 Approximate GLE as an approximation of the hybrid GLE

We consider the hybrid GLE in eq. (3.20)

$$\ddot{A}_t = -\frac{1}{M(A)}\frac{\mathrm{d}U_{\text{eff}}}{\mathrm{d}A_t} - \int_0^t \mathrm{d}s\, \Gamma_{\mathrm{M}}^{\mathrm{H}}(t-s)\dot{A}_s + F_{\mathrm{R}}^{\mathrm{H}}(t) + \int_0^t \mathrm{d}s\, \Gamma_{\mathrm{Z}}^{\mathrm{H}}(A_s, t-s), \quad (5.1a)$$

$$U_{\text{eff}}(A) = U_{\mathrm{PMF}}(A) + k_B T \ln M(A). \quad (5.1b)$$

The first three terms on the r.h.s. of eq. (5.1) correspond to the approximate GLE in eq. (2.34) for a reaction coordinate A that has a position-dependent mass $M(A)$.
The nonlinear memory term involving the kernel $\Gamma_{\mathrm{Z}}^{\mathrm{H}}$ represents a correction term in the presence of nonlinear memory effects. The memory function $\Gamma_{\mathrm{Z}}^{\mathrm{H}}$ is given in eq. (3.19), it vanishes if the following conditional correlation function vanishes

$$D^{\mathrm{H}}(A,t) = \langle \dot{A}_0, F_{\mathrm{R}}^{\mathrm{H}}(t)\rangle_A. \quad (5.2)$$

Since the hybrid projection is an orthogonal projection (chapter 3.1.1), we know that

$$\langle \dot{A}_0, F_{\mathrm{R}}^{\mathrm{H}}(t)\rangle = 0, \quad (5.3)$$

and also that $D^{\mathrm{H}}(A,0) = 0$. Further, for large many particle systems, we expect the random force and the velocity to become uncorrelated at any given value A when $t \to \infty$, i.e.,

$$\lim_{t\to\infty} D^{\mathrm{H}}(A,t) = \langle \dot{A}_0\rangle_A \langle F_{\mathrm{R}}^{\mathrm{H}}(t)\rangle_A = 0. \quad (5.4)$$

In summary, D^{H} is zero at time $t = 0$, then it increases in amplitude and decays to zero when t becomes larger than the memory time of the dynamics (see chapter 4.2).

According to the hybrid GLE, the approximate GLE is obtained by neglecting conditional correlations between the random force $F_{\mathrm{R}}^{\mathrm{H}}(t)$ and the velocity $\dot{A}_0$, i.e, by assuming that eq. (5.3) holds at any given value A and $D^{\mathrm{H}}(A, t) = 0$ at all times t. Then the hybrid GLE reads

$$\ddot{A}_t = -\frac{1}{M(A)} \frac{\mathrm{d}U_{\mathrm{eff}}}{\mathrm{d}A_t} - \int_0^t \mathrm{d}s\, \Gamma^{\mathrm{A}}(t-s)\dot{A}_s + F_{\mathrm{R}}^{\mathrm{A}}(t), \tag{5.5}$$

with $U_{\mathrm{eff}}(A)$ given in eq. (5.1). Eq. (5.5) together with the relation in eq. (5.3) is the approximate GLE in the presence of a position-dependent mass. It is the approximate GLE we used in chapter 4.2 when studying the dihedral dynamics in a butane molecule.

In the remainder of this chapter, we consider a reaction coordinate whose effective mass is constant in the reaction coordinate, for example, the mean native distance for a polypeptide (see chapter 3.1.3).
Given that $D^{\mathrm{H}}(A, t) = 0$, multiplying the hybrid GLE by the constant mass M, we recover the approximate GLE in eq. (2.34) for the observable A

$$M\ddot{A}_t = -\frac{\mathrm{d}U_{\mathrm{PMF}}(A_t)}{\mathrm{d}A_t} - \int_0^t \mathrm{d}s\, \Gamma^{\mathrm{A}}(t-s)\dot{A}_s + F_{\mathrm{R}}^{\mathrm{A}}(t), \tag{5.6a}$$

$$\langle F_{\mathrm{R}}^{\mathrm{A}}(0), F_{\mathrm{R}}^{\mathrm{A}}(t)\rangle = k_B T\, \Gamma^{\mathrm{A}}(t), \tag{5.6b}$$

$$\langle \dot{A}_0, F_{\mathrm{R}}^{\mathrm{A}}(t)\rangle = 0, \tag{5.6c}$$

$$\langle A_0, F_{\mathrm{R}}^{\mathrm{A}}(t)\rangle = 0, \tag{5.6d}$$

where we identified $M\Gamma_{\mathrm{M}}^{\mathrm{H}}(t) = \Gamma^{\mathrm{A}}(t)$ and $MF_{\mathrm{R}}^{\mathrm{H}}(t) = F_{\mathrm{R}}^{\mathrm{A}}(t)$.

5.2 Numerical estimation of GLE parameters by inverting a Volterra equation

Given the structure of the approximate GLE together with the relations in eq. (5.6), we derive an equation that allows us to compute the running integral of the memory kernel, $G(t) = \int_0^t \mathrm{d}s\, \Gamma^{\mathrm{A}}(s)$, directly from time-series data. The final equation generalizes the method to extract $G(t)$ in a harmonic potential derived in [66] to an arbitrary potential U_{PMF}.
We multiply the GLE in eq. (5.6) by the initial velocity $\dot{A}_0$ and ensemble average the result. Using the orthogonality in eq. (5.6b), this leads to

$$M\,\frac{\mathrm{d}}{\mathrm{d}t}C^{vv}(t) = C^{Fv}(t) - \int_0^t \mathrm{d}s\, \Gamma^{\mathrm{A}}(s)C^{vv}(t-s), \tag{5.7}$$

where $C^{vv}(t) = \langle \dot{A}_t \dot{A}_0\rangle$ denotes the velocity auto-correlation function and $C^{Fv}(t)$ is the correlation function of the mean force $F(t) = -U'_{\mathrm{PMF}}(A_t)$ with the velocity $\dot{A}_0$. Time

integration of (5.7) gives a continuous equation for the running integral $G(t)$

$$MC^{vv}(t) = -C^{FA}(t) + C^{FA}(0) + MC^{vv}(0) - \int_0^t \mathrm{d}s\, C^{vv}(s)G(t-s). \tag{5.8}$$

Multiplying the GLE in eq. (5.6a) by the initial position A_0, ensemble averaging and using the orthogonality relations in eq. (5.6d), we find

$$M\frac{\mathrm{d}}{\mathrm{d}t}C^{AA}(t) = C^{FA}(t) - \int_0^t \mathrm{d}s\Gamma^{\mathrm{A}}(s)\, C^{vA}(t-s). \tag{5.9}$$

Evaluating (5.9) at $t = 0$ and inserting $\frac{\mathrm{d}}{\mathrm{d}t}C^{AA}(t) = -C^{vv}(t)$ gives

$$MC^{vv}(0) = -C^{FA}(0). \tag{5.10}$$

Hence, (5.8) becomes

$$-\frac{C^{vv}(t)}{C^{vv}(0)}C^{FA}(0) = -C^{FA}(t) - \int_0^t \mathrm{d}s'\, C^{vv}(s')G(t-s). \tag{5.11}$$

Discretizing (5.11) by the use of the trapezoidal rule gives

$$\frac{C_n^{vv}}{C_0^{vv}}C_0^{FA} = C_n^{FA} + \frac{\Delta t}{2}G_nC_0^{vv} + \Delta t\sum_{i=1}^{n-1} G_{n-i}C_i^{vv}, \tag{5.12}$$

where we used $G(0) = 0$. Solving (5.12) for $G_n = G(t)$ leads to an iterative equation for G_n

$$G_n = \frac{2}{\Delta t\, C_0^{vv}}\left(-C_n^{FA} + \frac{C_0^{FA}}{C_0^{vv}}C_n^{vv} - \Delta t\sum_{i=1}^{n-1} G_{n-i}C_i^{vv}\right). \tag{5.13}$$

We apply the Volterra scheme only on the approximate GLE in eq. (5.6). An application of the Volterra scheme on GLEs with nonlinear friction kernel was discussed in ref. [110].

5.3 Simulating the approximate GLE

Since the approximate GLE does not contain memory functions nonlinear in A and $\dot{A}$, it is, from a numerical point of view, preferable over the nonlinear GLEs discussed in chapter 3. In the following, we present a Markovian embedding that can be used to generate trajectories from the approximate GLE in eq. (5.6).

For this, we first demonstrate how an Hamiltonian system can be mapped onto a stochastic Langevin system. Then we show that this Langevin system corresponds to a GLE that has the same structure as the approximate GLE in eq. (5.6), thus, we can use it to perform GLE simulations.

5.3.1 Mapping the nonlinear Zwanzig Hamiltonian system onto a stochastic Langevin system

The Hamiltonian of the nonlinear Zwanzig model is given by [114, 5]

$$H(x,p,\{q_n,v_n\}_{n=1}^N) = \frac{p^2}{2M} + U(x) + \sum_{n=1}^{N}\left(\frac{v_n^2}{2m_n} + \frac{k_n}{2}(g(x)-q_n)^2\right), \tag{5.14}$$

where we are only interested in the motion of the particle with position x and momentum p. A Nonlinear function $g(x)$ enters the coupling to other particles with positions q_n and momenta p_n. This represents a symmetry breaking in the interactions since the interactions do not depend on the distance $|q_n - x|$, but rather on the explicit value of x. However, the symmetrical case where the interactions only depend on the distance is easily recovered by setting $g(x) = x$. As we will show here, a nonlinear $g(x)$ induces nonlinear memory friction in the corresponding GLE. The equations of motion for the Hamiltonian in eq. (5.14) read

$$M\ddot{x}_t = -U'(x_t) - \sum_n k_n g'(x_t)\left(g(x_t) - q_n(t)\right), \tag{5.15a}$$

$$m_n\ddot{q}_{n,t} = -k_n\left(q_n(t) - g(x_t)\right), \tag{5.15b}$$

where the prime superscript denotes a derivative w.r.t. the argument, i.e., $U'(x) = \mathrm{d}U/\mathrm{d}x$. Eq. (5.15b) can be solved to give

$$q_n(t) = q_n(0)\cos(\mu_n t) + \frac{v_n(0)}{m_n\mu_n}\sin(\mu_n t) + \mu_n\int_0^t \mathrm{d}s\, \sin(\mu_n(t-s))\, g(x_s), \tag{5.16}$$

where $\mu_n^2 = k_n/m_n$. By partial integration, the solution in eq. (5.16) can be written in the form

$$\begin{aligned} q_n(t) = (q_n(0) - g(x_0))\cos(\mu_n t) + \frac{v_n(0)}{m_n\mu_n}\sin(\mu_n t) \\ - \int_0^t \mathrm{d}s\, \cos(\mu_n(t-s))\, g'(x_s)\dot{x}_s + g(x_t). \end{aligned} \tag{5.17}$$

Inserting eq. (5.17) into eq. (5.15a) leads to a GLE for x, i.e.,

$$M\ddot{x}_t = -U'(x_t) - \int_0^t \mathrm{d}s\, \Gamma\left[t-s, x_t, x_s\right]\dot{x}_s + F_R(t,x_t) \tag{5.18a}$$

$$F_R(t,x_t) = \sum_n g'(x_t)\Big(\mu_n v_n(0)\sin(\mu_n t) + k_n\left(q_n(0) - g(x_0)\right)\cos(\mu_n t)\Big), \tag{5.18b}$$

with the memory function

$$\Gamma(t-s,x_t,x_s) = \sum_{n=1}^{N} k_n g'(x_t)g'(x_s)\cos(\mu_n(t-s)). \tag{5.18c}$$

We now show how to map the nonlinear Zwanzig Hamiltonian system defined in eq. (5.14) onto a Markovian stochastic system of equations for which numerical simulations can be

efficiently performed. The results obtained here include the harmonic model by setting $g(x) = x$.

Consider the memory function in eq. (5.18c). It contains a sum over cosines. For $N \to \infty$, this represents a Fourier series of an even, periodic function in time with Fourier coefficients k_n. In the continuous limit, i.e., $k_n \to k(\mu)\mathrm{d}\mu/2\pi$, this defines an arbitrary even function $f(t)$ which we can choose to be an oscillating-exponential

$$\sum_{n=1}^{\infty} k_n \cos(\mu_n t) \to \int_{-\infty}^{\infty} \frac{\mathrm{d}\mu}{2\pi}\, k(\mu)\cos(\mu t) = f(t) \tag{5.19a}$$

$$= K\, e^{-|t|/\tau}\left(\cos\left(\frac{2\pi}{T}t\right) + c\sin\left(\frac{2\pi}{T}|t|\right)\right), \tag{5.19b}$$

with an exponential decay time τ and parameters T, K, c to be determined below. The function $k(\mu)$ follows from the Fourier transform as

$$k(\mu) = \int_{-\infty}^{\infty} \mathrm{d}t\, \cos(\mu t)\, f(t). \tag{5.20}$$

The memory function in eq. (5.18c) becomes

$$\Gamma(t-s, x_t, x_s) = g'(x_t)g'(x_s)\, K\, e^{-|t-s|/\tau}\left(\cos\left(\frac{2\pi}{T}(t-s)\right) + c\sin\left(\frac{2\pi}{T}|t-s|\right)\right). \tag{5.21}$$

Now, consider the random force in eq. (5.18b). For $N \to \infty$, it can be rewritten as

$$F_R(t, x_t) = g'(x_t)\tilde{F}_R(\{q_n(0), v_n(0)\}, t), \tag{5.22}$$

in terms of the random function $\tilde{F}_R$ which depends on the initial conditions of all auxiliary variables. In the stochastic interpretation of the GLE, it is sufficient to know the distribution of the initial conditions of the complementary variables. For the Hamiltonian in eq. (5.14), the distribution is given by the Boltzmann distribution. Thus, the initial values $q_n(0), v_n(0)$ are Gaussian distributed random variables with

$$\langle (q_n(0) - g(x_0))\rangle = 0, \qquad \langle v_n(0)\rangle = 0, \tag{5.23a}$$

$$\langle g(x_0), v_n(0)\rangle = 0, \qquad \langle q_n(0), v_n(0)\rangle = 0, \tag{5.23b}$$

$$\langle v_n(0), v_{m,0}\rangle = \delta_{n,m} m_n k_B T, \tag{5.23c}$$

$$\langle (q_n(0) - g(x_0)), (q_{m,0} - g(x_0))\rangle = \delta_{n,m}\frac{k_B T}{k_n}. \tag{5.23d}$$

From this, it follows that $\tilde{F}_R$ is a stationary Gaussian process fulfilling

$$\langle \tilde{F}_R(t)\rangle = 0, \tag{5.24a}$$

$$\langle \tilde{F}_R(t), \tilde{F}_R(0)\rangle = k_B T \sum_n k_n \cos(\mu_n t) \to k_B T f(t). \tag{5.24b}$$

The equal sign in eq. (5.24b) follows from the explicit form given in eq. (5.18b) and from the relation in eq. (5.23d) where the average is a Boltzmann average over the initial conditions

$\{q_n(0), v_n(0)\}$. A Markovian stochastic system which leads to a memory function of the form given in eq. (5.21) reads

$$M\ddot{x}_t = -U'(x_t) - kg'(x_t)\left(g(x_t) - y(t)\right), \tag{5.25a}$$

$$m_y\ddot{y}(t) = -k\left(y(t) - g(x_t)\right) - \gamma\dot{y}(t) + \sqrt{2k_BT\gamma}\eta(t), \tag{5.25b}$$

$$\mathbb{E}[\eta(t)] = 0, \qquad \mathbb{E}[\eta(t)\eta(s)] = \delta(t-s), \tag{5.25c}$$

with η being white noise and $\mathbb{E}[\cdot]$ denotes an average over the noise. The relations between the parameters in eq. (5.25) and the parameters in eq. (5.19) are given by

$$\nu^2 = 4\tau_\Gamma^2\mu^2 - 1, \tag{5.26a}$$

$$T = \frac{4\pi}{\nu}\tau_\Gamma, \qquad \tau = 2\tau_\Gamma, \qquad K = k, \tag{5.26b}$$

$$\tau_\Gamma = \frac{m_y}{\gamma}, \qquad \mu^2 = \frac{k}{m_y} \qquad c = \frac{1}{\nu}. \tag{5.26c}$$

By solving eq. (5.25b) and inserting the result into eq. (5.25a), we find the random force

$$\begin{aligned}\tilde{F}^R(t) = {} & ke^{-|t|/2\tau}\left(\cos\left(\frac{2\pi}{T}t\right) + \frac{1}{\nu}\sin\left(\frac{2\pi}{T}t\right)\right)\left(y(0) - g(x_0)\right) \\ & + \frac{2}{\gamma\nu}e^{-|t|/2\tau}\sin\left(\frac{2\pi}{T}t\right)p_y(0) \\ & + \sqrt{2k_BT\gamma}\int_0^t \mathrm{d}s\, 2e^{-(t-s)/2\tau}\frac{\tau}{\nu}\sin\left(\frac{2\pi}{T}(t-s)\right)\eta(s),\end{aligned} \tag{5.27}$$

where the variables $y(0)$ and $q_n(0)$ are identically distributed, the same is true for $p_y(0) = m_y\dot{y}(0)$ and $v_n(0)$. The stochastic equivalence of $\tilde{F}^R$ in eq. (5.22) and $\tilde{F}_R$ in eq. (5.27) follows from the fact that their first and second moments are the same when $t \gg \tau$, i,e., for sufficiently long trajectories.

5.3.2 Markovian embedding of the approximate GLE

We have mapped the nonlinear Hamiltonian Zwanzig model defined by eq. (5.14) onto the set of coupled Markovian stochastic equations in eq. (5.25), which can be used to perform numerical simulations.
Eq. (5.25) with N auxiliary variables $\{y_n(t)\}_{n=1}^N$ reads [103]

$$M\ddot{x}_t = -U'(x_t) - \sum_{n=1}^{N} k_n g_n'(x_t)\left(g_n(x_t) - y_n(t)\right), \tag{5.28a}$$

$$m_n\ddot{y}_n(t) = -k_n\left(y_n(t) - g_n(x_t)\right) - \gamma_n\dot{y}_n(t) + \sqrt{2k_BT\gamma_n}\eta_n(t), \tag{5.28b}$$

$$\mathbb{E}[\eta_n(t)] = 0, \qquad \mathbb{E}[\eta_n(t)\eta_l(s)] = \delta_{nl}\delta(t-s). \tag{5.28c}$$

The overdamped case of eq. (5.28) is obtained by considering the limit $\frac{m_n}{\gamma_n\tau_D} \ll 1$, where τ_D denotes the diffusive time scale. With the elimination of fast varibles method [108], one

finds that in the overdamped limit the inertial term in eq. (5.28b) can be neglected, i.e.,

$$M\ddot{x}_t = -U'(x_t) - \sum_{n=1}^{N} k_n g_n'(x_t)\left(g_n(x_t) - y_n(t)\right), \tag{5.29a}$$

$$\dot{y}_n(t) = -\frac{k_n}{\gamma_n}\left(y_n(t) - g_n(x_t)\right) + \sqrt{\frac{2k_BT}{\gamma_n}}\eta_n(t), \tag{5.29b}$$

$$\mathbb{E}[\eta_n(t)] = 0, \qquad \mathbb{E}[\eta_n(t)\eta_l(s)] = \delta_{nl}\delta(t-s). \tag{5.29c}$$

In the same way we obtained the GLE in eq. (5.18a), a GLE is found that describes eq. (5.29) by solving eq. (5.29b) for $y_n(t)$ and inserting the result into eq. (5.29a). The corresponding memory kernel for eq. (5.29) is given by

$$\Gamma(t-s, x_t, x_s) = \sum_{n=1}^{N} g_n'(x_t)g_n'(x_s)\frac{\gamma_n}{\tau_n}e^{-\frac{(t-s)}{\tau_n}}, \tag{5.30}$$

with the memory time given by $\tau_n = \gamma_n/k_n$. Note that, when $g_n(x) = x$, the memory kernel in eq. (5.30) only depends on $t-s$. Thus, the Markovian system in eq. (5.29) can be used to perform efficient simulations of the approximate GLE in eq. (5.6) with exponentially decaying memory functions.

5.4 Kramers-Moyal coeffiecients for the approximate GLE

We consider a general multidimensional stochastic process described by the trajectory $\mathbf{q}(t)$. For example, $\mathbf{q}(t)$ could denote the two-dimensional process given by $\mathbf{q}(t) = (A_t, \dot{A}_t) \equiv (q, v)$. The Kramers-Moyal coefficients are derived from the transition probability $W(\mathbf{q}, t|\mathbf{q}', t')$, which is the probability that $\mathbf{q}(t) = \mathbf{q}$ given that $\mathbf{q}(t') = \mathbf{q}'$, for $t > t'$. The transition probability is a well-defined quantity for Markov and non-Markov processes. However, in contrast to Markov processes, for non-Markov processes $W(\mathbf{q}, t|\mathbf{q}', t')$ does not contain the full information of the dynamics, i.e., we can not describe the stochastic process completely based on the transition probabilities alone [98]. For non-Markov processes, further information is required [60].

5.4.1 The Kramers-Moyal Expansion

To consider the multi-dimensional Kramers-Moyal expansion, it is useful to introduce the multi-index notation [100] which employs the multi-index $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_N) \in \mathbb{N}^N$. The convention is as follows: For some vector $\mathbf{r} \in \mathbb{R}^N$, the product of coefficients is defined as $\mathbf{r}^\alpha \equiv r_1^{\alpha_1} r_2^{\alpha_2} \cdots r_N^{\alpha_N}$ and we define $\alpha! \equiv \alpha_1!\alpha_2! \ldots \alpha_N!$. The absolute value of the multi-index is defined as $|\alpha| = \sum_{i=1}^{N} \alpha_i \equiv n$. Using the above described multi-index notation, the n-th multi-dimensional Kramers-Moyal coefficient with a finite lag time Δt can be defined as

$$D_\alpha^{(n)}(\mathbf{q}, t, \Delta t) = \frac{1}{n!\Delta t}\left\langle \left(\mathbf{q}(t+\Delta t) - \mathbf{q}(t)\right)^\alpha \right\rangle_{\mathbf{q}(t)=\mathbf{q}}. \tag{5.31}$$

In eq. (5.31), $\langle \quad \rangle_{\mathbf{q}(t)=\mathbf{q}}$ denotes the conditional expectation value that at time t, $\mathbf{q}(t) = \mathbf{q}$. We start with the joint probability $P_2(\mathbf{q}, t+\Delta t; \mathbf{q}', t)$ of observing $\mathbf{q}$ at time $t+\Delta t$ and $\mathbf{q}'$ at time t with $\Delta t > 0$. For Markov and non-Markov processes, P_2 can be written as [98]

$$P_2(\mathbf{q}, t+\Delta t; \mathbf{q}', t) = W(\mathbf{q}, t+\Delta t|\mathbf{q}', t)P_1(\mathbf{q}', t). \tag{5.32}$$

By integrating over $\mathbf{q}'$ we obtain

$$P_1(\mathbf{q}, t+\Delta t) = \int \mathrm{d}\mathbf{q}'\, W(\mathbf{q}, t+\Delta t|\mathbf{q}', t)P_1(\mathbf{q}', t), \tag{5.33}$$

where $\mathrm{d}\mathbf{q}' = \mathrm{d}q_1'\mathrm{d}q_2' \ldots \mathrm{d}q_N'$. Next we substitute $\Delta = \mathbf{q} - \mathbf{q}' \Rightarrow \mathrm{d}\Delta = -\mathrm{d}\mathbf{q}'$ and obtain

$$P_1(\mathbf{q}, t+\Delta t) = \int \mathrm{d}\Delta\, W(\mathbf{q}, t+\Delta t|\mathbf{q}-\Delta, t)P_1(\mathbf{q}-\Delta, t). \tag{5.34}$$

Now, we expand the integrand in eq. (5.34) in $\mathbf{q}$ around $\mathbf{q}+\Delta$

$$P_1(\mathbf{q}, t+\Delta t) = \int \mathrm{d}\Delta \sum_{|\alpha|\geq 0} \frac{\Delta^\alpha}{\alpha!}\left(-\frac{\partial}{\partial \mathbf{q}}\right)^\alpha W(\mathbf{q}+\Delta, t+\Delta t|\mathbf{q}, t)P_1(\mathbf{q}, t), \tag{5.35}$$

where the sum $\sum_{|\alpha|\geq 0}$ goes over all possible permutations and $|\alpha| = 0, 1, 2, \ldots, \infty$. Rearranging eq. (5.35) gives

$$P_1(\mathbf{q}, t+\Delta t) = \sum_{|\alpha|\geq 0} \frac{n!}{\alpha!}\left(-\frac{\partial}{\partial \mathbf{q}}\right)^\alpha P_1(\mathbf{q}, t)\Delta t D_\alpha^{(n)}(\mathbf{q}, t, \Delta t), \tag{5.36}$$

where the equivalent definition of the finite-time Kramers-Moyal coefficients has been used

$$D_\alpha^{(n)}(\mathbf{q}, t, \Delta t) = \int \mathrm{d}\Delta\, \frac{1}{n!\Delta t}\Delta^\alpha W(\mathbf{q}+\Delta, t+\Delta t|\mathbf{q}, t). \tag{5.37}$$

From eq. (5.37), it follows that $\Delta t D_\alpha^{(0)}(\mathbf{q}, t, \Delta t) = 1$. Therefore, we can write

$$P_1(\mathbf{q}, t+\Delta t) = P_1(\mathbf{q}, t) + \sum_{|\alpha|\geq 1} \frac{n!}{\alpha!}\left(-\frac{\partial}{\partial \mathbf{q}}\right)^\alpha P_1(\mathbf{q}, t)\Delta t D_\alpha^{(n)}(\mathbf{q}, t, \Delta t). \tag{5.38}$$

After rearranging we find

$$\frac{P_1(\mathbf{q}, t+\Delta t) - P_1(\mathbf{q}, t)}{\Delta t} = \sum_{|\alpha|\geq 1} \frac{n!}{\alpha!}\left(-\frac{\partial}{\partial \mathbf{q}}\right)^\alpha D_\alpha^{(n)}(\mathbf{q}, t, \Delta t)P_1(\mathbf{q}, t), \tag{5.39a}$$

$$= \sum_{n=1}^{\infty} \sum_{|\alpha|=n} \frac{n!}{\alpha!}\left(-\frac{\partial}{\partial \mathbf{q}}\right)^\alpha D_\alpha^{(n)}(\mathbf{q}, t, \Delta t)P_1(\mathbf{q}, t), \tag{5.39b}$$

where the sum $\sum_{|\alpha|=n}$ goes over all α with $|\alpha| = n$. From eq. (5.37), it follows that $D_\alpha^{(n)}(\mathbf{q}, t, 0) = 0$ for $n > 0$. Therefore, taking the limit $\Delta t \to 0$ on both sides gives the standard Kramers-Moyal expansion of the Fokker-Planck equation

$$\frac{\partial P_1(\mathbf{q}, t)}{\partial t} = \sum_{n=1}^{\infty} \sum_{|\alpha|=n} \frac{n!}{\alpha!}\left(-\frac{\partial}{\partial \mathbf{q}}\right)^\alpha \lim_{\Delta t\to 0} D_\alpha^{(n)}(\mathbf{q}, t, \Delta t)P_1(\mathbf{q}, t) = L_{KM}P_1(\mathbf{q}, t), \tag{5.40}$$

where L_{KM} denotes the Kramers-Moyal operator.

5.4.2 Pawula theorem

According to the Pawula Theorem, the expansion in eq. (5.40) either stops after the second term or it contains an infinite number of terms. This can be derived by using the generalized Schwartz inequality [98]. For a non-negative function $P(\mathbf{q}) \geq 0$ and arbitrary functions $f(\mathbf{q})$ and $g(\mathbf{q})$, one finds

$$\left(\int \mathrm{d}\mathbf{q}\, P(\mathbf{q}) f(\mathbf{q}) g(\mathbf{q})\right)^2 \leq \int \mathrm{d}\mathbf{q}\, P(\mathbf{q})\, f(\mathbf{q})^2 \int \mathrm{d}\mathbf{q}\, P(\mathbf{q})\, g(\mathbf{q})^2. \tag{5.41}$$

When $P(\mathbf{q})$ is the transition probability $W(\mathbf{q}+\Delta, t+\Delta t|\mathbf{q}, t)$, and $f(\mathbf{q}) \to \Delta^\beta$, $g(\mathbf{q}) \to \Delta^\gamma$, where the multi-indices β and γ are a partition of α, i.e., $|\alpha| = |\beta| + |\gamma| = n + (n+m)$ and $\alpha! = \beta!\gamma!$, the Schwartz inequality implies for the Kramers Moyal coefficients [98]

$$\left(D_\alpha^{(2n+m)}\right)^2 \leq \frac{(2n)!\,(2(n+m))!}{[(2n+m)!]^2}\, D_\beta^{(2n)} D_\gamma^{(2(n+m))}, \qquad \text{for } n, m \geq 1. \tag{5.42}$$

The factorial prefactors follow from the definition eq. (5.31). If $D_\beta^{(2n)} = 0$, we find from eq. (5.42) that $D_\alpha^{(2n+1)} = D_\alpha^{(2n+2)} = \cdots = 0$, i.e.,

$$D_\beta^{(2n)} = 0 \;\Rightarrow\; D_\alpha^{(2n+m)} = 0, \qquad \text{for } m \geq 1. \tag{5.43}$$

From eq. (5.42), it also follows that if $D_\gamma^{(2(n+m))} = D_\gamma^{(2r)} = 0$, the coefficients $D_\alpha^{(2n+m)} = D_\alpha^{(2(n+m)-m)} = D_\alpha^{(2r-m)}$ vanish for all combinations of n, m with $n + m = r$. Meaning

$$D_\gamma^{(2r)} = 0 \;\Rightarrow\; D_\alpha^{(2r-m)} = 0 \qquad \text{for } m = 1, 2, \ldots, r-1. \tag{5.44}$$

Combining eq. (5.43) with a repeated use of eq. (5.44) gives in summary the following non-trivial properties:

1. The expansion stops after the first order, which means that there is no stochastic part in the corresponding Langevin equation. This becomes clear by noting that, if all second order KMCs vanish, the amplitude of the random forces in the Langevin equation also vanish. The fact that the expansion may stop after the first order follows from eq. (5.43) for $n = 1$. If the second order KMCs are zero, so are all higher orders.

2. The expansion stops after the second order, i.e, the Kramers Moyal expansion becomes the ordinary Fokker Planck equation, which describes a diffusion process. This follows from eq. (5.43) and eq. (5.44). If we consider the case $n = 2$ in eq. (5.43), it follows that, if the fourth order KMCs are zero, then all orders higher than four are also zero. On the other hand, eq. (5.44) says that, if the fourth order KMCs are zero, the third order KMCs must also be zero. Hence, all orders larger than two must vanish if the fourth order KMCs turn out to be zero.

3. The expansion contains an infinite number of terms. This can be seen by setting the third order KMCs to be non-zero. Then, according to eq. (5.44), the fourth order KMCs can not be zero. Furthermore, if the fourth order KMCs are non-zero, eq. (5.44) again implies that the sixth order KMCs must also be non-zero. In this way, the repeated use of eq. (5.44) implies that the Kramers Moyal expansion contains an infinite number of terms.

5.4.3 Kramers-Moyal coefficients of a stationary Gaussian process

Here, we derive the KMCs for a stationary Gaussian process, i.e. for the GLE in eq. (5.6) in the absence of a potential U. The results of this subsection are used below to approximately compute the KMCs of the approximate GLE. To proceed, we consider the velocity $v(\cdot)$ as an N dimensional stationary Gaussian process,

$$v(\cdot) \;\to\; \mathbf{v} = (v_1, v_2, \ldots, v_N)^T, \tag{5.45}$$

$$P(\mathbf{v}) = \frac{1}{\sqrt{(2\pi)^N \,\det \sigma}}\, e^{-\frac{1}{2}\mathbf{v}^T\cdot\sigma^{-1}\cdot\mathbf{v}}, \tag{5.46}$$

where $\sigma_{ij} = \langle v_i, v_j\rangle = \langle v_{|i-j|}, v_0\rangle$ denotes the covariance matrix of the N dimensional stationary Gaussian distribution $P(v)$. We can use eq. (5.46) to compute the conditional probability distribution. For this, we partition $P(\mathbf{v})$ using

$$\mathbf{v}^T\cdot\sigma^{-1}\cdot\mathbf{v} = (\mathbf{v}_{N-1}\, v_N)\cdot\begin{pmatrix} M & O \\ O^T & N\end{pmatrix}\cdot\begin{pmatrix}\mathbf{v}_{N-1}\\ v_N\end{pmatrix}, \tag{5.47}$$

where we introduce the vectors $O \in \mathbb{R}^{N-1}$, $\mathbf{v}_{N-1} \in \mathbb{R}^{N-1}$ and the matrix $M \in \mathbb{R}^{(N-1)\times(N-1)}$. $v_N = v$ denotes the value of the velocity at time t. The partitioning in eq. (5.47) leads to

$$P(\mathbf{v}) = P_1(v_N)P_{N-1}(\mathbf{v}_{N-1}|v_N). \tag{5.48}$$

$P_{N-1}(\mathbf{v}_{N-1}|v_N)$ is the conditional probability distribution we must use when computing conditional averages. It has the form

$$P_{N-1}(\mathbf{v}_{N-1}|v_N) = \sqrt{\frac{\det M}{(2\pi)^N}}\, e^{-\frac{1}{2}\left(\mathbf{v}_{N-1}+M^{-1}Ov_N\right)^T\cdot M\cdot\left(\mathbf{v}_{N-1}+M^{-1}Ov_N\right)}. \tag{5.49}$$

As can be seen in eq. (5.49), the mean of the distribution is shifted. We find

$$\langle v_i\rangle_{v_N=v} = -(M^{-1}Ov_N)_i = \frac{\langle v_{|N-i|}, v_0\rangle}{\langle v_0^2\rangle}\, v. \tag{5.50}$$

In the continuous case, eq. (5.50) becomes

$$\langle v(s)\rangle_{v(t)=v} = \frac{C_{vv}(|t-s|)}{C_{vv}(0)}\, v, \tag{5.51}$$

where C_{vv} denotes the velocity autocorrelation function. We use eq. (5.51) in the conditional expectation $\langle v(t+\Delta t) - v(t)\rangle_v = \langle v(t+\Delta t)\rangle_v - \langle v(t)\rangle_v$, and find

$$\langle v(t+\Delta t) - v(t)\rangle_v = \frac{C_{vv}(\Delta t) - C_{vv}(0)}{C_{vv}(0)}\, v. \tag{5.52}$$

Inserting the last result into eq. (5.31), we obtain the first order KMC for a stationary Gaussian process

$$\widetilde{D}_v(v,\Delta t) = \frac{C_{vv}(\Delta t) - C_{vv}(0)}{\Delta t}\,\frac{v}{C_{vv}(0)} = \Delta C(\Delta t)\, v, \tag{5.53a}$$

where $\widetilde{D}_v(v, \Delta t)$ denotes the KMC in the absence of a potential U and

$$\Delta C(\Delta t) \equiv (C_{vv}(\Delta t) - C_{vv}(0))/(\Delta t\, C_{vv}(0)). \tag{5.53b}$$

To compute D_{vv} we start from

$$\left\langle (v(t+\Delta t) - v(t))^2 \right\rangle_v = \left\langle v(t+\Delta t)^2 \right\rangle_v + \left\langle v(t)^2 \right\rangle_v - 2\left\langle v(t+\Delta t) v(t) \right\rangle_v . \tag{5.54}$$

For the conditional velocity autocorrelation function, we obtain similar to the conditional average velocity in eq. (5.51)

$$\langle v(s), v(s') \rangle_{v(t)=v} = C_{vv}(s-s') - \frac{C_{vv}(t-s)C_{vv}(t-s')}{C_{vv}(0)} + \frac{C_{vv}(t-s)C_{vv}(t-s')}{C_{vv}(0)^2} v^2. \tag{5.55}$$

Inserting this into eq. (5.54) gives

$$\begin{aligned} \left\langle (v(t+\Delta t) - v(t))^2 \right\rangle_v = &- (C_{vv}(\Delta t) + C_{vv}(0)) \frac{C_{vv}(\Delta t) - C_{vv}(0)}{C_{vv}(0)} \\ &+ \left(\frac{C_{vv}(\Delta t) - C_{vv}(0)}{C_{vv}(0)}\, v \right)^2 . \end{aligned} \tag{5.56}$$

From this, the second order KMC immediately follows as

$$\begin{aligned} \widetilde{D}_{vv}(v, \Delta t) &= -\frac{(C_{vv}(\Delta t) + C_{vv}(0))}{2} \frac{(C_{vv}(\Delta t) - C_{vv}(0))}{\Delta t} \frac{1}{C_{vv}(0)} \\ &\quad + \frac{\Delta t}{2} \left(\frac{C_{vv}(\Delta t) - C_{vv}(0)}{\Delta t} \frac{v}{C_{vv}(0)} \right)^2 && \text{(5.57a)} \\ &= -\bar{C}(\Delta t)\, \Delta C(\Delta t) + \Delta t\, (\Delta C(\Delta t)\, v)^2 /2, && \text{(5.57b)} \end{aligned}$$

where we defined

$$\bar{C}(\Delta t) \equiv (C_{vv}(\Delta t) + C_{vv}(0))/2, \tag{5.57c}$$

and $\Delta C(\Delta t)$ is given in eq. (5.53b).

5.4.4 KMCs in the presence of a potential

To obtain analytical approximations for the KMCs from the GLE in eq. (5.6), we use the general results for stationary Gaussian processes derived above in subsection 5.4.3. To approximate $D_q(q, v)$, we integrate the relation $\dot{q}(t) = v(t)$ from t to $t + \Delta t$, which gives

$$q(t+\Delta t) - q(t) = \int_t^{t+\Delta t} \mathrm{d}s\, v(s). \tag{5.58}$$

For small Δt, the integral on the r.h.s. can be approximated by the trapezoidal rule as

$$\begin{aligned} q(t+\Delta t) - q(t) &= \frac{v(t+\Delta t)+v(t)}{2}\,\Delta t + \mathcal{O}(\Delta t^3) \\ &= v(t)\Delta t + \frac{v(t+\Delta t)-v(t)}{2}\Delta t + \mathcal{O}(\Delta t^3). \end{aligned} \tag{5.59}$$

By taking the conditional average on both sides and dividing by Δt, we can make use of the definition of D_q given by

$$D_q(q,v,\Delta t) = \frac{1}{\Delta t}\left\langle q(t+\Delta t) - q(t)\right\rangle_{\substack{q(t)=q \\ v(t)=v}}, \tag{5.60}$$

which is a special case of eq. (5.31). In this way, we find D_q in terms of D_v for finite Δt

$$D_q(q,v,\Delta t) = v + \frac{\Delta t}{2} D_v(q,v,\Delta t) + \mathcal{O}(\Delta t^2). \tag{5.61}$$

To compute D_v in the presence of a potential, we integrate the GLE in eq. (5.6) from t to $t+\Delta t$, using $v(t) = \dot{A}_t$, $q(t) = A_t$ and $U_{\mathrm{PMF}}(A_t) = U[q(t)]$, we obtain

$$v(t+\Delta t) - v(t) = \int_t^{t+\Delta t} \mathrm{d}s\, \left(-\frac{U'[q(s)]}{M} - \int_0^s \mathrm{d}s'\, \Gamma(s-s')v(s') + F_R(s) \right). \tag{5.62}$$

The function $U'(q(s))/M$ is expanded around $s = t$, to leading order in Δt, the expansion yields

$$\begin{aligned} v(t+\Delta t) - v(t) = &-\Delta t \frac{U'[q(t)]}{M} + \mathcal{O}(\Delta t^2) \\ &+ \int_t^{t+\Delta t} \mathrm{d}s\, \left(-\int_0^s \mathrm{d}s'\, \Gamma(s-s')v(s') + F_R(s) \right). \end{aligned} \tag{5.63}$$

Again, we take the conditional average on both sides and divide by Δt. The conditional average of the integral part on the r.h.s. is approximated by eq. (5.53a), i.e., by the result for D_v in the absence of a potential U. For this approximation to work, the velocity distribution should be independent of the position q so that we can neglect the condition on q in the conditional average of the velocity in eq. (5.63). From chapter 3.1.3, we know that for an observable which is a linear combination of distances, the mass is constant and $P(v|q) \approx P(v)$ is a Gaussian. Therefore, approximating conditional averages over q and v by conditional averages over v alone is a valid approximation for such observables. By further neglecting the impact of the potential U on the memory kernel, we find

$$D_v(q,v,\Delta t) \approx -\frac{U'(q)}{M} + \widetilde{D}_v(v,\Delta t) + \mathcal{O}(\Delta t), \tag{5.64}$$

where $\widetilde{D}_v(v,\Delta t)$ denotes the drift coefficient in the absence of a potential U and is given in eq. (5.53a). To compute D_{qq}, we square eq. (5.59) and find

$$\left(q(t+\Delta t) - q(t)\right)^2 = \frac{\left(v(t+\Delta t) - v(t)\right)^2}{4}\Delta t^2 + v(t)v(t+\Delta t)\Delta t^2 + \mathcal{O}(\Delta t^6). \tag{5.65}$$

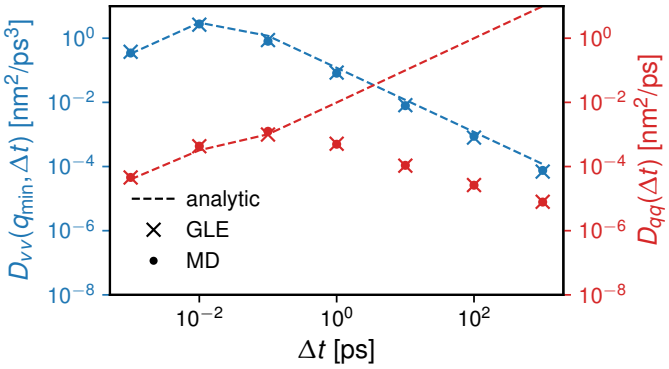

Figure 5.1: The second-order KMCs $D_{qq}(\Delta t)$ and $D_{vv}(q_{\min}, \Delta t)$ as a function of the lag time Δt. D_{vv} is averaged over v and is evaluated at the PMF minimum at $q_{\min} = 0.32\,\text{nm}$ (cf. figure 5.2). D_{qq} is averaged over q and v. The data points show the KMCs computed from the MD (circles) and GLE (crosses) trajectories of Alanine$_9$ (cf. section 5.5). The dashed lines show the analytic expressions in eq. (5.70g) (blue) and eq. (5.70c) (red).

Taking the conditional average, dividing by $2\Delta t$ and using the definitions of D_{qq} and D_{vv} gives

$$D_{qq}(q, v, \Delta t) = \frac{\Delta t^2}{4} D_{vv}(q, v, \Delta t) + \frac{\Delta t}{2} \langle v(t + \Delta t) v(t) \rangle_{q,v} + \mathcal{O}(\Delta t^5). \tag{5.66}$$

The conditional autocorrelation function on the r.h.s. is approximated by eq. (5.55). This is again motivated by assuming $P(v)$ to be a Gaussian with vanishing mean and constant variance, i.e., by the fact that $P(v|q) \approx P(v)$, which yields

$$D_{qq}(q, v, \Delta t) = \frac{\Delta t^2}{4} D_{vv}(q, v, \Delta t) + \frac{\Delta t}{2} \frac{C_{vv}(\Delta t)}{C_{vv}(0)} v^2 + \mathcal{O}(\Delta t^5). \tag{5.67}$$

In an analogous manner, $D_{vv}(q, v)$ is approximated by squaring eq. (5.63) and using eq. (5.57b), which gives

$$D_{vv}(q, v, \Delta t) = \frac{\Delta t}{2} \left(\frac{U'(q)}{M} \right)^2 + \widetilde{D}_{vv}(v, \Delta t) + \mathcal{O}(\Delta t^2). \tag{5.68}$$

If we take the limit $\Delta t \to 0$, only the first order KMCs do not vanish, as follows from eq. (5.61), eq. (5.64), eq. (5.67) and eq. (5.68), so that we obtain

$$\lim_{\Delta t \to 0} D_q(q, v, \Delta t) = v, \qquad \lim_{\Delta t \to 0} D_v(q, v, \Delta t) = -\frac{U'(q)}{M}, \tag{5.69a}$$

$$\lim_{\Delta t \to 0} D_{qq}(q, v, \Delta t) = 0, \qquad \lim_{\Delta t \to 0} D_{vv}(q, v, \Delta t) = 0. \tag{5.69b}$$

Hence, in the limit of Δt going to zero, the KMCs of a GLE only reproduce the deterministic part of the dynamics and contain no information on the stochastic part. By averaging over q

with a weight function $\propto e^{-\beta U(q)}$ or over v with a weight function $\propto e^{-\beta M v^2/2}$, the results for finite Δt in eq. (5.61), eq. (5.64), eq. (5.67) and eq. (5.68) become to leading order

$$D_q(q,\Delta t) = -\frac{\Delta t}{2}\frac{U'(q)}{m} + \mathcal{O}(\Delta t^2) \tag{5.70a}$$

$$D_q(v,\Delta t) = v + \mathcal{O}(\Delta t), \tag{5.70b}$$

$$D_{qq}(q,\Delta t) = \frac{\Delta t}{2}C_{vv}(\Delta t) - \frac{\Delta t^2}{4}\bar{C}(\Delta t)\Delta C(\Delta t) + \mathcal{O}(\Delta t^3), \tag{5.70c}$$

$$D_{qq}(v,\Delta t) = \frac{\Delta t}{2}\frac{C_{vv}(\Delta t)}{C_{vv}(0)}v^2 - \frac{\Delta t^2}{4}\bar{C}(\Delta t)\Delta C(\Delta t) + \mathcal{O}(\Delta t^3), \tag{5.70d}$$

$$D_v(q,\Delta t) = -\frac{U'(q)}{m} + \mathcal{O}(\Delta t), \tag{5.70e}$$

$$D_v(v,\Delta t) = v\Delta C(\Delta t) + \mathcal{O}(\Delta t), \tag{5.70f}$$

$$D_{vv}(q,\Delta t) = -\bar{C}(\Delta t)\Delta C(\Delta t) + \frac{\Delta t}{2}\left(\frac{U'(q)}{m}\right)^2 + \mathcal{O}(\Delta t^2), \tag{5.70g}$$

$$D_{vv}(v,\Delta t) = -\bar{C}(\Delta t)\Delta C(\Delta t) + \frac{\Delta t}{2}\left\langle\left(\frac{U'(q)}{m}\right)^2\right\rangle + \mathcal{O}(\Delta t^2). \tag{5.70h}$$

The quantities $\bar{C}(\Delta t)$ and $\Delta C(\Delta t)$ are defined in eq. (5.57c) and in eq. (5.53b), respectively.

In fig. 5.1, we compare eq. (5.70c) and eq. (5.70g) with numerically computed KMCs as a function of the lag time Δt, obtained from MD and GLE simulations of Alanine_9 (see section 5.5). In order to compare D_{qq} and D_{vv} as a function of Δt, we average the numerical D_{vv} over v and evaluate it in the vicinity of the minimum at $q = 0.32\,\text{nm}$ of $U_{\text{PMF}}(q)$, i.e., where the mean force vanishes (see figure 5.2). The numerical D_{qq} is averaged over q and v. As can be seen in fig. 5.1, the analytical approximations describe the KMCs very well for small Δt, as expected. This also confirms that the numerical computation of the KMCs is accurate.

5.4.5 Numerical computation of KMCs: Kernel density estimators

To compute the KMCs numerically, we divide our parameter space $(A, \dot{A})$ into a grid. In each cell of the grid, we must find an estimate for the conditional probability distribution $P(A, \dot{A}, t+\Delta t|A', \dot{A}', t)$, which is usually accomplished by histograms. When the statistics in a cell are insufficient, i.e., when the number of data points in a cell is too small, the estimate of the probability distribution via a histogram depends significantly on the choice of the boundaries of the cell. To avoid such ambiguities, we use kernel density estimators [44]. Kernel density estimators have the advantage that each point of occurrence within a cell contributes to the estimate of the density at its point of occurrence and not in a range of a bin. This means that, for a given data set $\mathbf{q} \in \mathbb{R}^N$, a p.d.f. $P(q)$ is estimated by $\widehat{P}(q)$ via

$$\widehat{P}(q) = \frac{1}{Nh}\sum_{i=1}^{N} K\left(\frac{q-q_i}{h}\right). \tag{5.71}$$

Here K denotes the kernel with bandwidth h. For example, one could use a Gaussian kernel

$$K\left(\frac{q-q_i}{h}\right) \propto e^{-\frac{1}{2}\left(\frac{q-q_i}{h}\right)^2}. \tag{5.72}$$

In the results shown in this PhD-thesis, we used an Epanechnikov kernel

$$K\left(\frac{q-q_i}{h}\right) \propto \max\left(1-\left(\frac{q-q_i}{h}\right)^2, 0\right), \tag{5.73}$$

where the optimal bandwidth h was estimated by the inverse of the grid size. For small time steps $\Delta t = 0.001$ ps, the grid size in q direction is 5000 and in v direction it is 300. Hence, the respective bandwidths are 0.0002 and 0.003. For larger time steps $\Delta t \geq 0.001$ ps, the grid size in q direction is decreased to 500 and therefore the bandwidth is 0.002.

5.5 Application to protein folding dynamics

Biological macromolecular function relies on coupled processes that take place on widely different time scales, this fact makes the theoretical description of such systems challenging. For proteins, the topic of this chapter, folding occurs in the range of microseconds to many minutes or even hours and involves bond vibrations and hydration water motion on sub-picosecond times [13, 73, 31]. In order to enable large-scale simulations as well as meaningful theories, which should concentrate on the essential features of such *slow* processes, several methods for the elimination of degrees of freedom have been introduced.

For the classical dynamics of an interacting many-body system, the approximate GLE in eq. (5.6) has proven to be useful tool [26, 102, 70, 66]. Instead of $6N$ equations of motion for all positions and momenta of an N-particle system, the dynamics is described by few equations for the observables of interest. The potential of mean force $U_{\mathrm{PMF}}(A)$ in eq. (5.6) corresponds for proteins to the folding free-energy landscape. The elimination of degrees of freedom introduces non-Markovian effects in terms of the memory function $\Gamma^{\mathrm{A}}(t)$, which describes time-dependent friction and thereby couples the present dynamics to the past states, and stochastic effects in terms of the random force $F_{\mathrm{R}}^{\mathrm{A}}(t)$.
From the hybrid GLE, we obtained the approximate GLE together with the fluctuation dissipation theorem which connects $F_{\mathrm{R}}^{\mathrm{A}}(t)$ and $\Gamma^{\mathrm{A}}(t)$ by neglecting $D^{\mathrm{H}}(A,t)$. Thus, for a given reaction coordinate that is a nonlinear function in the microscopic coordinates, the validity of eq. (5.6) rather has to be tested which is among the goals of this chapter.

Clearly, there is no guarantee that a given reaction coordinate, which could be an experimental observable such as the distance between two attached fluorophores, is a good reaction coordinate, meaning that it leads to a Markovian description of the folding process. Different reaction coordinates have been proposed for the efficient description of protein-folding simulations [52]; schemes to construct reaction coordinates that optimally yield the transition state, which separates unfolded and folded basins of attraction from each other, have been developed [9].
As an alternative to continuous reaction coordinates, Markov models describe protein dynamics in terms of a set of metastable states [93, 20], for which full access to the underlying

microscopic coordinates is typically needed.
These works have in common that descriptions are sought which minimize memory effects, so that stochastic Markovian theory applies. In the opposite direction, various methods were developed to extract the memory function $\Gamma^{\mathrm{A}}(t)$ from time-series data for a given reaction coordinate [51, 106, 55, 30, 59, 26], but the complexities of the GLE, in particular for a nonlinear protein-folding free energy in combination with a numerically determined memory function, prevented predictions of protein folding times from the GLE, with the notable exception of di-alanine [70].

This is why, in protein folding theory, the Markovian Langevin equation, where the memory integral is replaced by an instantaneous friction term, is predominantly used. Such a Markovian theory yields many useful insights into protein folding dynamics and culminated in the comparison of transition-path times and mean-folding times [23, 24]. However, the success of free-energy folding theory on the Markovian level relies partly on the fact that the friction, which determines the prefactor of the Kramers folding time, is normally used as a fitting parameter. Even when the friction is allowed to vary with the reaction coordinate and is extracted from simulations, it is typically computed from folding or reconfiguration times, which by construction leads to self-consistent predictions of the kinetics [8, 54].
In fact, recent experiments revealed significant inconsistencies when comparing directly measured free-energy barrier heights with those inferred from transition-path and folding times [91] which were suggested to be due to memory effects [101, 102]. The same inconsistencies are obtained when the friction of a reaction coordinate is not fitted to folding times but rather extracted directly from simulation trajectories and used in the framework of Markovian theory, as we demonstrate here.

In our approach, instead of searching for a good reaction coordinate, we employ a standard one-dimensional coordinate in terms of the sum of the separations between native contacts. We use the numerical method presented in chapter 5.2 for extracting all parameters of the GLE from molecular dynamics (MD) simulations for the helix-forming polypeptide Ala_9 in water. The free energy $U_{\mathrm{PMF}}(A)$ shows multiple minima separated by low barriers, indicative of the sequential formation of the helix, while the decay time of the multi-exponential memory function $\Gamma^{\mathrm{A}}(t)$ is of the order of the unfolding time: These properties render Ala_9 as a very sensitive test of kinetic theory.

We simulate the resulting GLE by Markovian embedding technique discussed in chapter 5.3. By comparison of the MD and GLE results for the mean folding and unfolding times, we demonstrate that the one-dimensional GLE is an accurate and practical tool for the description of protein folding dynamics described by distances between particles.
On the other hand, the Markovian version of the overdamped GLE cannot describe the folding and unfolding kinetics of the peptide as long as the friction is not a fitting parameter but rather taken as extracted from the MD simulations. This stays true even when the friction coefficient is allowed to depend on the reaction coordinate.
In fact, memory typically accelerates barrier crossing, where the acceleration magnitude depends primarily on the ratio of the memory time and the distance between the minimum and the barrier in reaction coordinate space [61, 62, 69]. Since for Ala_9 the barrier state is closer to the folded state than to the unfolded state, a memory-induced asymmetry between folding and unfolding kinetics results, which cannot be captured by a Markovian model and

which is testable experimentally. Memory acceleration is confirmed by theoretical models such as Grote-Hynes and Pollack-Grabert-Hänggi [48, 49, 96].

Our results are corroborated by a systematic Kramers-Moyal coefficient (KMC) analysis, which shows that higher order quartic KMCs are non-negligible and that the linear and quadratic KMCs vanish in the short-time limit, as expected in the presence of non-Markovian effects.

5.5.1 MD simulations and GLE parameter extraction

The effective GLE is constructed from a 10 μs long MD trajectory for Ala_9 in water, which is the simplest polypeptide that forms an α-helix [57]. We use the all-atom Amber03 force field field [33] with SPC/E water [7]. The cubic simulation box has side lengths of 4.95 nm and contains 4023 water molecules. The LJ interactions are cut-off after 1.0 nm. For long-range electrostatic interactions, we use the Particle Mesh Ewald Method (PME) [28]. The simulation time step is 1 fs and the total simulation time is $10\,\mu$s. All simulations are performed in the NVT ensemble using the Gromacs 2019 MD package [2]. In the GLE simulations, we use the same time step and simulation time as in the MD simulations.

The procedure in the MD simulations is as follows: We first minimize the energy of the system using the steepest descent method. Then, we let the system equilibrate at 300 K in the NPT ensemble. During the equilibration, the polypeptide is positionally restrained to ensure that no conformational transitions occur at this stage. We do this by applying harmonic potentials with a force constant of $k = 1000$ kJ/mol/nm to all Ala_9 atoms. The equilibration time is 1 ns, i.e., 10^6 time steps. To speed up the equilibration process, we assign to each particle an initial velocity drawn from a Maxwell-Boltzmann distribution at the desired temperature. We perform production runs in the NVT ensemble using a modified Berendsen thermostat with a time constant of 0.1 ps [14].
We constrain all bond lengths using the LINCS algorithm. The bond angles are unconstrained.

The Ala_9 polypeptide was build using the open-source molecular builder software Avogadro 1.2.0. [50]. The N terminus is NH_3^+ and the C terminus is CO_2^-.

As a reaction coordinate, we use the summed separations between the H-bond donor nitrogen of residue n and the acceptor oxygen of residue n+4,

$$A_t = \frac{1}{3}\sum_{i=2}^{4} \|\mathbf{r}_i^N(t) - \mathbf{r}_{i+4}^O(t)\|, \tag{5.74}$$

which characterizes the left-handed α-helical conformation. The index $i = 1$ is the C-terminus of the ploypeptide. In the α-helical state, A has a value around 0.3 nm, the mean H-bond length between nitrogen and oxygen.
The free energies $U_{\mathrm{PMF}}(q)$ in fig. 5.2(b) for different simulation lengths demonstrate that the simulation is fully converged after about $6\,\mu$s. The free energy displays several metastable states, which are also discernable in the trajectory in fig. 5.2(a) and make this simple polypeptide challenging for theoretical description.

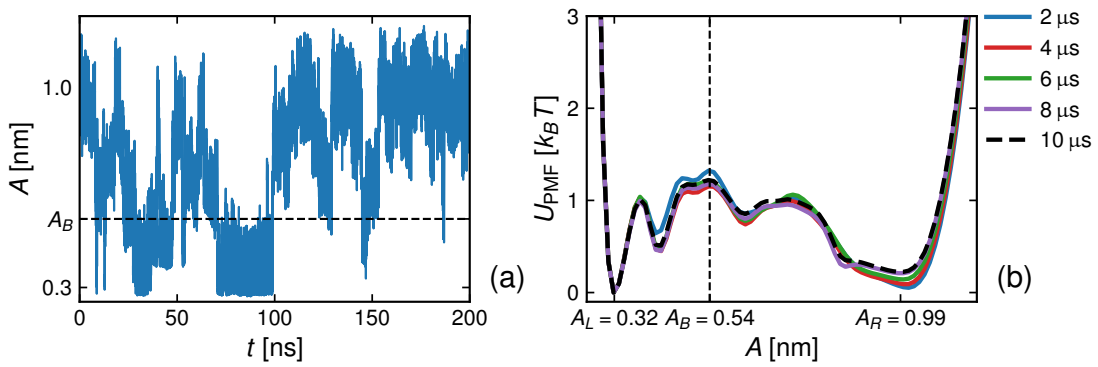

Figure 5.2: **(a)**: A 200 ns long segment of the trajectory is shown. **(b)**: The potential of mean force $U_{\mathrm{PMF}}(A)$ for the hydrogen-bond-distance reaction coordinate of Ala9 for different simulation lengths. The barrier used for the calculation of unfolding and folding times is positioned at $A_B = 0.54\,\mathrm{nm}$.

Using a generalization of earlier methods [66], we extract the running integral $G(t) = \int_0^t ds\Gamma^{\mathrm{A}}(s)$ (see section 5.2 for details), from which the memory function $\Gamma^{\mathrm{A}}(t)$ is obtained via a numerical derivative and fitted using least-square methods to a multi-exponential of the form

$$\Gamma^{\mathrm{A}}(t) = \sum_{n=1}^{5} \frac{\gamma_n}{\tau_n} e^{-t/\tau_n}. \tag{5.75}$$

The extracted $G(t)$ (gray line) is compared with the corresponding fit (red line) in fig. 5.3(a), no significant deviations can be discerned. The comparison of the extracted and fitted memory function $\Gamma^{\mathrm{A}}(t)$ in Fig. 5.3(b) reveals oscillations below a picosecond, which are not reproduced by the exponential fit function but also do not play a role for the kinetics, as will be shown below.
The fitted memory times τ_n and friction coefficients γ_n are presented in table 5.1, the typical reconfiguration time, which can be qualitatively inferred from the trajectory in fig. 5.2(a), is of the order of the longest decay time $\tau_5 \approx 5\,\mathrm{ns}$. This means that the reaction coordinate is not particularly good, since it exhibits pronounced non-Markovian effects, and thus constitutes a suitable test of our methods.

In chapter 3.1.3, the effective mass of the reaction coordinate in eq. (5.74) was found to be configuration independent. Thus, we can compute the effective mass using the equipartition theorem, i.e, $M = k_B T/\langle \dot{A}_0^2 \rangle$ which turns out to be $M = 31.3\,\mathrm{u}$. This value deviates from the expected value of $M \approx 22\,\mathrm{u}$ from eq. (3.24). This is due to the applied bond constraints in the MD simulation. Bond constraints have the effect that the backbone of the polypeptide becomes a rigid rotor.

The motion described by the GLE is expected to become diffusive after the inertial time $\tau_m = M/\bar{\gamma}$, where the total friction coefficient is given by $\bar{\gamma} = \sum_n \gamma_n = 3.5 \cdot 10^5\,\mathrm{u/ps}$, see table 5.1. It follows that $\tau_m = 0.1\,\mathrm{fs}$, even shorter than the MD integration time step, thus inertial effects are completely negligible. Nevertheless, the acceleration term in eq. (5.6) is

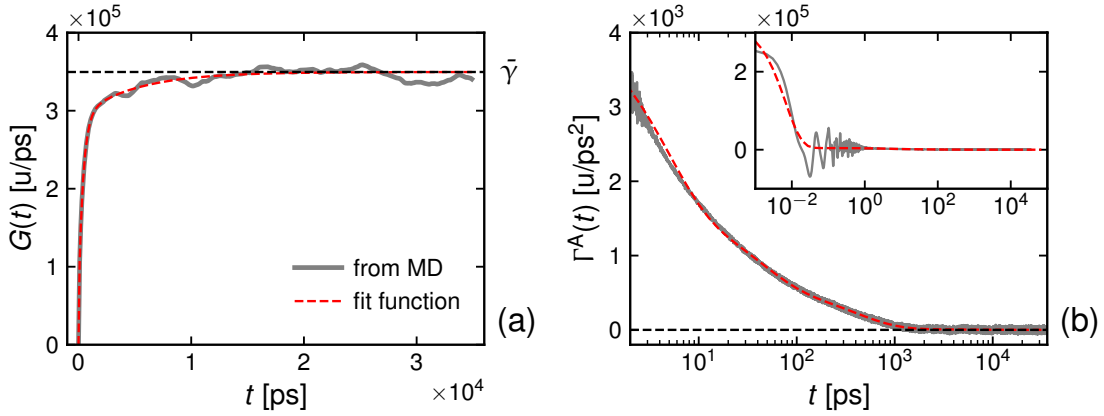

Figure 5.3: **(a)**: Running integral $G(t)$ over the memory function. The horizontal dashed line denotes the total friction coefficient $\bar{\gamma}$. **(b)**: Memory function $\Gamma^{\mathrm{A}}(t)$, the inset includes short times. Gray lines correspond to the numerical data, red lines correspond to the multi-exponential fit according to eq. (5.75) with the fit parameters given in tab. 5.1.

kept in the GLE simulations, as it stabilizes the numerical integration.
In order to estimate the importance of memory effects, one compares the memory times τ_n with the diffusion time scale $\tau_D = \beta\bar{\gamma}L^2/2$ [61], which is the time it takes a free Brownian particle to diffuse over a length L in reaction coordinate space, where $\beta = 1/k_BT$ is the inverse thermal energy. For $L = 0.22\,\mathrm{nm}$, the distance between the folded minimum at $A = 0.32\,\mathrm{nm}$ and the barrier at $A = 0.54\,\mathrm{nm}$ in one obtains $\tau_D = 6.8\,\mathrm{ns}$, which is of the order of the longest memory time τ_5. This places the system in the so-called memory-acceleration regime, where memory effects are relevant and significantly accelerate barrier crossing [61, 62, 69].

5.5.2 Comparison of MD and GLE simulations

Numerical integration of the GLE is straightforwardly achieved by Markovian embedding, i.e., by transforming the GLE into a system of linearly coupled Langevin equations (see

Table 5.1: Fitted memory function parameters from eq. (5.75).

n	γ_n [u/ps]	τ_n [ps]
1	$2.2 \cdot 10^3$	0.007
2	$1.2 \cdot 10^4$	4.6
3	$4.2 \cdot 10^4$	40.3
4	$2.4 \cdot 10^5$	399
5	$5.7 \cdot 10^4$	4970
$\bar{\gamma} = \sum_n \gamma_n$	$3.5 \cdot 10^5$	

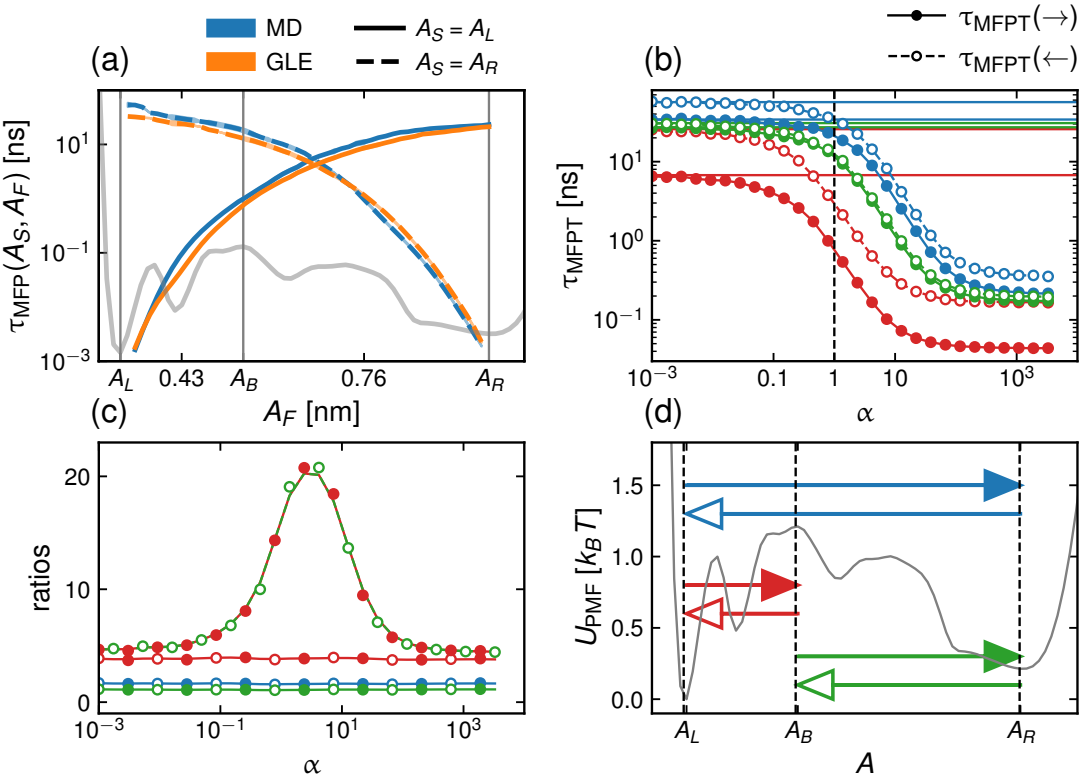

Figure 5.4: **(a)**: Comparison of unfolding and folding MFPTs from MD (blue) and GLE (orange) simulations as a function of the final position A_F for start positions $A_S = A_L = 0.32\,\text{nm}$ (solid lines) and $A_S = A_R = 0.99\,\text{nm}$ (broken lines). The gray curve shows the folding free energy $U_{\text{PMF}}(A)$. **(b)**: Dependence of various MFPTs from GLE simulations on the memory time rescaling factor α. The corresponding transitions are illustrated in (d). Open/filled circles correspond to open/filled arrows of the same color. The horizontal dotted lines denote the overdamped Markov limit from eq. (5.82). **(c)**: The ratios of the MFPTs shown in (b) are computed. As expected from detailed balance, ratios of MFPTs of the same color do not depend on the memory time. The ratio of the unfolding and folding times, i.e., $\tau_{\text{MFPT}}(A_R, A_B)/\tau_{\text{MFPT}}(A_L, A_B)$ (open green arrow and filled red arrow), shows a clear dependency on the memory time (red green). **(d)**: Illustration of the transitions considered in (b).

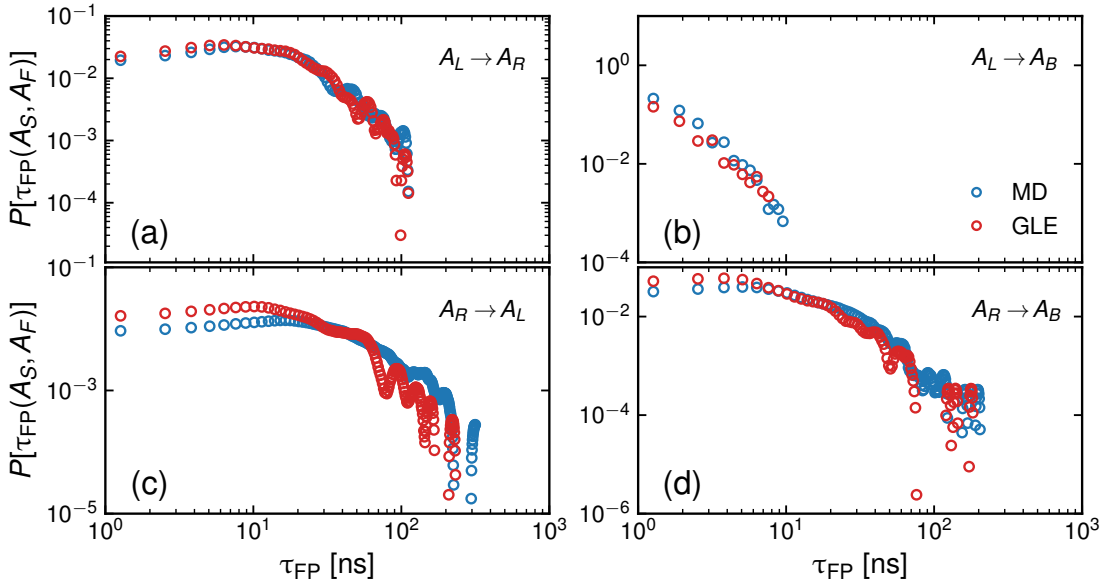

Figure 5.5: First-passage time distributions between the positions $A_L = 0.32\,\text{nm}$, $A_B = 0.54\,\text{nm}$ and $A_R = 0.99\,\text{nm}$. We compare distributions computed from the MD system (blue) and the corresponding GLE system (red). **(a)**: For the transition $A_L \to A_R$. **(b)**: For the transition $A_L \to A_B$. **(c)**: For the transition $A_R \to A_L$. **(d)**: For the transition $A_R \to A_B$.

chapter 5.3).

In fig. 5.4, we show profiles of the mean first-passage time (MFPT) $\tau_{\text{MFPT}}(A_S, A_F)$ for unfolding (start position $A_S = A_L = 0.32\,\text{nm}$, represented by solid lines) and folding kinetics (start position $A_S = A_R = 0.99\,\text{nm}$, broken lines) as a function of the final position A_F.
Statistical errors are determined accounting for data correlations [39] and are shown as shaded areas in the corresponding color. The errors computed are so small that they are barely visible in fig. 5.4.
MD and GLE simulation results (blue and orange lines) agree nicely, this demonstrates that GLE-based non-Markovian modeling of protein folding is feasible and accurate.

First passage time distributions for some selected start and end positions are shown in fig. 5.5. No significant deviation can be discerned between the MD (blue circles) and the GLE system (red circles) in the distributions shown in fig. 5.5(a), (b) and (d). In Fig. fig. 5.5(c), it can be seen that the MD system has higher probabilities towards larger first-passage times compared to the GLE system. The impact of this on the MFPT can be seen in fig. 5.4(a).

Beyond reproducing MD results, the GLE is a diagnostic tool that allows to quantify the importance of memory effects. In order to modulate memory effects in the GLE, we rescale the memory times according to $\tau_n \to \alpha\tau_n$ for $n = 2, 3, 4, 5$ while keeping the memory time τ_1 of the fastest exponential contribution fixed. Since $\tau_1 = 7\,\text{fs}$ is above the simulation time step of 1 fs, this ensures that in the limit $\alpha \to 0$ we obtain a regularized model that, as we will show below, corresponds to the Markovian limit.

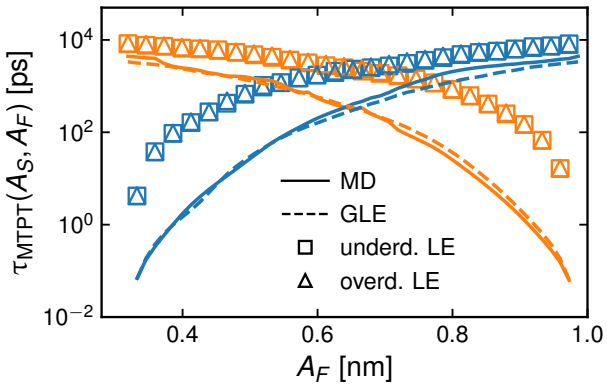

Figure 5.6: We show the mean transition path time (MTPT) between $A_S = 0.32, 0.99\,\mathrm{nm}$ (blue, orange) and A_F as a function of A_F. We compute the MTPT from the MD and GLE trajectories as well as trajectories generated using an underdamped (ULE) and overdamped Langevin equation (OLE) with the friction $\bar{\gamma} = \int_0^\infty \Gamma^{\mathrm{A}}(s)$ and the PMF $U_{\mathrm{PMF}}(A)$.

In fig. 5.4(b), we consider MFPTs between the three positions $A_L = 0.32\,\mathrm{nm}$, $A_B = 0.54\,\mathrm{nm}$ and $A_R = 0.98\,\mathrm{nm}$ as a function of the rescaling factor α from GLE simulations of hb4. The starting and end points of the MFPTs are illustrated in fig. 5.4(d). Filled/open circles in (b) correspond to filled/open arrows of the same color in (d). We see that reducing the memory time increases all MFPTs, in other words, memory accelerates barrier crossing [61]. As expected, the GLE results approach the overdamped Markov limit, denoted by the horizontal dotted lines in the corresponding color and calculated from the exact expression in eq. (5.82) without adjustable parameters, as α tends to zero.
Interestingly, for folding (open green circles) the MFPT for $\alpha = 1$ and the Markovian limit for $\alpha \to 0$ differ only by a factor of around 2.5. On the other hand, for unfolding (filled red circles), the $\alpha \to 0$ and $\alpha = 1$ MFPTs differ by a factor of around 9. This means that even when treating the total friction coefficient $\bar{\gamma}$ as a free parameter, the Markovian overdamped theory eq. (5.82), because it is linear in the friction, can reproduce either the folding or unfolding MD times, but not both simultaneously. This is not due to inertial effects, since the overdamped Markovian theory works perfectly for $\alpha \to 0$, as seen in fig. 5.4(b). Rather, memory effects influence folding and unfolding times differently; this is demonstrated by the plot of the MFPT ratios over α in fig. 5.4(c) (which furthermore implies that the law-of-mass action has an apparent memory-dependence that depends subtly on the definition of states and transition times). The reason is that the memory effect on barrier crossing dynamics depends primarily on the ratio of the memory and diffusion times, τ/τ_D, where $\tau_D \sim L^2$ and L is the distance between free energy minimum and barrier [61]. Since the folded state is closer to the barrier, memory effects are therefore expected to be more pronounced for unfolding than for folding.

In fig. 5.6, we demonstrate that the memory-induced speedup is even more pronounced for transition path times compared to folding and unfolding times, in agreement with previous findings [91, 101, 102].

The high accuracy of GLE simulations is reflected by the good agreement of the mean-

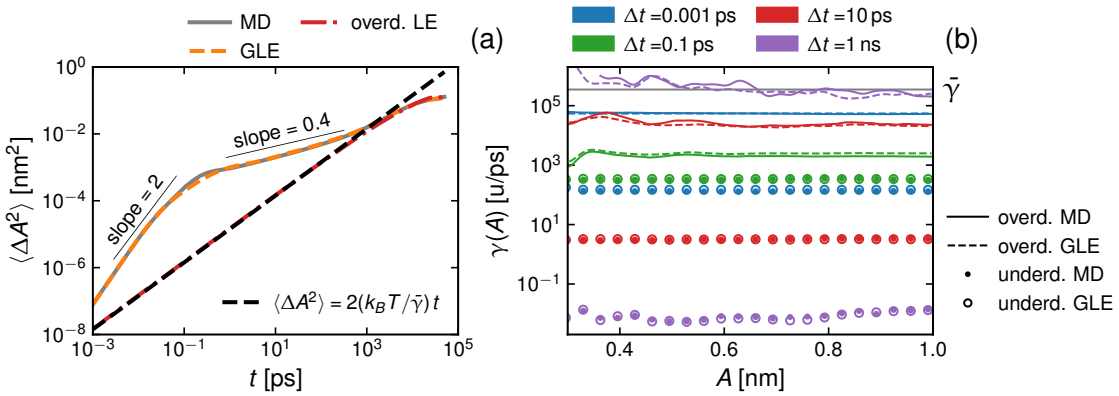

Figure 5.7: **(a)**: Mean-square displacement of the reaction coordinate, MD (gray line) and GLE (orange broken line) simulation results agree perfectly and exhibit superdiffusion for times up to 0.1 ps and subdiffusion up to 1 ns. Overdamped Markovian Langevin (red dashdotted line) simulations agree perfectly with theoretical prediction (broken black line) but miss the anomalous diffusion. **(b)**: Friction coefficient profiles $\gamma(A)$ from Kramers-Moyal coefficient analysis for different lag times Δt (different colors) for the underdamped Langevin model, eq. (5.80), from MD (filled circles) and GLE simulations (open circles), and for the overdamped Langevin model, eq. (5.81), from MD (solid lines) and GLE simulations (broken lines). The gray horizontal line shows the total friction coefficient $\bar{\gamma}$ extracted from MD simulations given in tab. 5.1.

square displacement $\langle \Delta A(t)^2 \rangle = \langle (A(t'+t) - A(t'))^2 \rangle$ from MD and GLE simulations in fig. 5.7(a), which exhibits pronounced sub-diffusive behavior with an exponent 0.4 for times between 1 ps and 1 ns. Anomalous diffusion is often modeled by fractional theories [80, 101], fig. 5.7(a) shows that it is accurately reproduced by multi-exponential memory and that it disappears when memory effects are eliminated, in line with recent theoretical analysis [86].

Furthermore, the good agreements between the MD system and the GLE model in terms of mean first-passage times, transition path times and the MSD suggest that the assumptions for the GLE in eq. (5.6) are sufficiently fulfilled; at least for a reaction coordinate that is a linear combination of distances.

5.5.3 Reaction-coordinate dependent friction

So far, we demonstrated that the GLE in eq. (5.6) reproduces the MD simulation kinetics and that memory effects are significant. We now investigate whether reaction-coordinate dependent friction effects, which are not included in the approximate GLE, are relevant.

The Markovian Langevin equation (LE) that incorporates a friction function $\gamma(A)$ has been amply used to describe protein folding dynamics [18, 8, 54]. In the underdamped version it

reads

$$M\ddot{A}(t) = -U'(A) - \gamma(A)\dot{A}(t) + \sqrt{k_B T \gamma(A)}\,\eta(t), \tag{5.76}$$

which for general $U(A)$ unfortunately is analytically intractable. The overdamped version

$$0 = -U'(A) - \gamma(A)\dot{A}(t) - \frac{k_B T}{2}\frac{\gamma'(A)}{\gamma(A)} + \sqrt{k_B T \gamma(A)}\,\eta(t) \tag{5.77}$$

is much more useful, since the MFPTs can be calculated analytically. In these expressions, the random force $\eta(t)$ has vanishing mean and its correlator is given by $\mathbb{E}[\eta(t)\eta(t')] = 2\delta(t-t')$. For constant friction, the underdamped LE ((5.76)) results from eq. (5.6) when the memory kernel $\Gamma^{\mathrm{A}}(t-s)$ is sharply peaked at $s=t$, i.e., only contributes to the integral around $s=t$.
The overdamped LE ((5.77)) follows from eq. (5.76) by neglecting the inertia term [108], the term proportional to the gradient $\gamma'(A)$ cancels a spurious drift term and follows by mapping on the Fokker-Planck equation [98].

Various methods to extract $\gamma(A)$ from experimental or simulated trajectories have been proposed, a systematic approach involves the Kramers Moyal Coefficients (KMCs) discussed in section 5.4, which for the overdamped case and for finite lag time Δt read

$$D_k(A) = \frac{1}{k!}\frac{1}{\Delta t}\left\langle (A(t+\Delta t) - A(t))^k \right\rangle_{A(t)=A}. \tag{5.78}$$

The Fokker-Planck equation for the time-dependent probability distribution $P(A,t)$ in terms of the KMCs follows in the limit $\Delta t \to 0$ as [98]

$$\frac{\partial P(A,t)}{\partial t} = \sum_{k=1}^{\infty} \frac{\partial^k}{\partial A^k}\left[D_k(A)P(A,t)\right]. \tag{5.79}$$

According to the Pawula theorem (section 5.4.2), for a Markovian process, all KMCs with $k > 2$ vanish for $\Delta t \to 0$ and eq. (5.79) takes the standard form of a second-order partial differential equation [98]. For a non-Markovian process, i.e. if the memory function $\Gamma^{\mathrm{A}}(t)$ in eq. (5.6) has a finite range, all KMCs with $k > 1$ vanish for $\Delta t \to 0$ and thus the stochastic properties of the process cannot be described by a partial differential equation for $P(A,t)$ at all. We demonstrate this in chapter 5.4. For the underdamped LE, the relation between the second-order velocity KMC D_{vv} and the friction profile $\gamma_{\mathrm{UD}}(A)$ reads [98]

$$D_{vv}(A) = \frac{1}{2\Delta t}\langle(\dot{A}(t+\Delta t) - \dot{A}(t))^2\rangle_{A(t)=A} = k_B T\frac{\gamma_{\mathrm{UD}}(A)}{m^2}. \tag{5.80}$$

For the overdamped LE, $\gamma_{\mathrm{OD}}(A)$ follows from the second-order position KMC D_{qq} as

$$D_{qq}(A) = \frac{1}{2\Delta t}\langle(A(t+\Delta t) - A(t))^2\rangle_{A(t)=A} = \frac{k_B T}{\gamma_{\mathrm{OD}}(A)} \tag{5.81}$$

For the numerical computation of the KMCs, we use kernel density estimators [44]. In fig. 5.7(b), we show the friction profiles $\gamma_{\mathrm{UD}}(A)$ (circles) and $\gamma_{\mathrm{OD}}(A)$ (lines) computed from the KMCs for different lag times Δt, a number of points are noteworthy:

i) We find no significant deviations between the friction profiles extracted from MD (solid lines and filled circles) and GLE (broken lines and open circles) trajectories, this reverberates that the GLE describes the protein dynamics very faithfully.

ii) The underdamped and overdamped friction profiles $\gamma_{\mathrm{UD}}(A)$ and $\gamma_{\mathrm{OD}}(A)$ disagree for all lag times Δt, which very clearly demonstrates an inconsistency in the Markovian description of protein folding. In fact, in the limit $\Delta t \to 0$, both D_{qq} and D_{vv} vanish and thus $\gamma_{\mathrm{OD}}(A)$ diverges while $\gamma_{\mathrm{UD}}(A)$ goes to zero (see section 5.4.)

iii) While the underdamped friction $\gamma_{\mathrm{UD}}(q)$ never reaches a realistic value close to $\bar{\gamma}$, the overdamped friction $\gamma_{\mathrm{OD}}(q)$ approaches $\bar{\gamma}$ for $\Delta t \approx 1\,\mathrm{ns}$. This shows that lag times of the order of the longest memory time have to be used in order to generate realistic friction values.

iv) The friction profiles extracted from the GLE simulations are position dependent, seen most clearly in $\gamma_{\mathrm{OD}}(q)$ for $\Delta t = 1\,\mathrm{ns}$ (purple broken line); this is clearly a spurious effect since the GLE has no position-dependent friction. We conclude that the mapping of a non-Markovian process onto a Markovian Langevin equation produces spurious position-dependent friction effects. Presumably, the effective friction of proteins will in general exhibit a dependence on the reaction coordinate, but the extraction of friction profiles would have to account for memory effects in order to avoid spurious effects. The capability of the GLE eq. (5.6) to very accurately reproduce the MD simulation kinetics suggests that for the present case of Ala_9, the spatial dependence of friction is negligible.

An alternative way to determine a friction profile $\gamma(A)$ in the overdamped limit uses the one-to-one relation between the MFPT profiles in fig. 5.4(a) and $\gamma(A)$. For the overdamped LE given in eq. (5.77), the MFPT to reach from a start position A_S the final position A_F for the first time reads for $A_S < A_F$ [111]

$$\tau_{\mathrm{MFPT}}(A_S, A_F) = \beta \int_{A_S}^{A_F} \mathrm{d}A\, e^{\beta U(A)} \gamma(A) \int_{A_{min}}^{A} \mathrm{d}A'\, e^{-\beta U(A')} \tag{5.82a}$$

and for $A_S > A_F$

$$\tau_{\mathrm{MFPT}}(A_S, A_F) = \beta \int_{A_F}^{A_S} \mathrm{d}A\, e^{\beta U(A)} \gamma(A) \int_{A}^{A_{max}} \mathrm{d}A'\, e^{-\beta U(A')}. \tag{5.82b}$$

Taking the derivative of eq. (5.82) w.r.t. A_F gives the friction profile $\gamma(A_F)$ as [54]

$$\gamma_{\mathrm{unf}}(A_F) = k_B T \frac{e^{-\beta U(A_F)}}{Z_1} \frac{\partial \tau_{\mathrm{MFPT}}}{\partial A_F} \qquad \text{for } A_S < A_F, \tag{5.83a}$$

$$\gamma_{\mathrm{fol}}(A_F) = -k_B T \frac{e^{-\beta U(A_f)}}{Z_2} \frac{\partial \tau_{\mathrm{MFPT}}}{\partial A_F} \qquad \text{for } A_S > A_F, \tag{5.83b}$$

where $Z_1 = \int_{A_{min}}^{A_F} \mathrm{d}A\, e^{-\beta U(A)}$ and $Z_2 = \int_{A_F}^{A_{max}} \mathrm{d}A\, e^{-\beta U(A)}$.

In fig. 5.8(a), we show $\gamma_{\mathrm{unf}}(A_F)$ and $\gamma_{\mathrm{fol}}(A_F)$ computed from unfolding and folding MFPTs from MD simulations for start positions $A_S = A_L$ and $A_S = A_R$, respectively. Not suprisingly, the profiles $\gamma_{\mathrm{unf}}(A_F)$ and $\gamma_{\mathrm{fol}}(A_F)$ are rather close to $\bar{\gamma}$ extracted from the MD

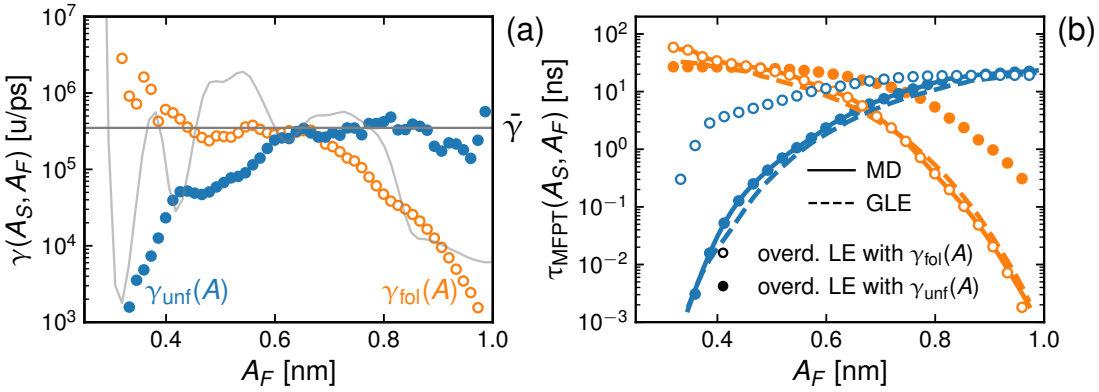

Figure 5.8: **(a)**: Friction profiles computed from the MD MFPT profiles in fig. 5.4(a) using eq. (5.83). $\gamma_{\mathrm{unf}}(A_F)$ follows from the unfolding MFPTs for start position $A_S = 0.32\,\mathrm{nm}$, $\gamma_{\mathrm{fol}}(A_F)$ follows from folding MFPTs for $A_S = 0.99\,\mathrm{nm}$. The gray horizontal line denotes the friction coefficient extracted from MD simulations. The gray curve in the background shows the folding free energy $U_{\mathrm{PMF}}(A)$. **(b)**: MFPTs from MD and GLE simulations are compared with overdamped Markovian predictions according to eq. (5.82) using $\gamma_{\mathrm{unf}}(A_F)$ and $\gamma_{\mathrm{fol}}(A_f)$ from (a).

simulations, which is shown as a gray horizontal line in fig. 5.8(a), but differ significantly from each other. This suggests that a single friction profile cannot describe folding and unfolding of Ala_9 simultaneously. In fact, the values of $\gamma_{\mathrm{unf}}(A_F)$ and $\gamma_{\mathrm{fol}}(A_F)$ go down as A_F moves to the respective start positions, i.e., as the folding and unfolding times become shorter. This reflects that memory effects particularly accelerate fast transitions, as discussed above.

To demonstrate the limitations of the friction profiles in fig. 5.8(a), we show in fig. 5.8(b) folding and unfolding MFPT profiles that are calculated according to eq. (5.82) from $\gamma_{\mathrm{unf}}(A)$ (filled circles) and $\gamma_{\mathrm{fol}}(A)$ (open circles). By construction, the MFPTs using $\gamma_{\mathrm{unf}}(A)$ reproduce the unfolding simulation data while the MFPTs using $\gamma_{\mathrm{fol}}(A)$ reproduce the folding simulation data.
In contrast, the MFPTs using $\gamma_{\mathrm{unf}}(A)$ fail to reproduce the simulated folding times and the MFPTs using $\gamma_{\mathrm{fol}}(A)$ fail to reproduce the simulated unfolding times, in particular when the folding/unfolding times become smaller than about 10 ns.

The GLE model (broken lines) reproduces both folding and unfolding MD dynamics (solid lines). This underlines that there is no consistent way of describing the complete folding/unfolding dynamics with a Markovian model.

5.6 Summary and conclusions

Using the hybrid GLE derived in chapter 3.1, we discuss under which conditions the approximate GLE in eq. (5.6) arises from the hybrid GLE in eq. (3.20). Given this structure of the approximate GLE, we introduce a numerical extraction scheme to compute the running integral of the memory kernel from time series data and, once the memory kernel is computed, we present how to perform efficient computer simulations of the GLE via Markovian embedding.

We derive expressions for the Kramers Moyal coefficients of a non-Markovian process described by the approximate GLE. In this way, we are able to show that all Kramers Moyal coefficients of higher order than one vanish in the limit of Δt going to zero.

By extracting the time-dependent friction for the polypeptide Ala_9 from explicit-water MD simulations, we demonstrate that the approximate GLE model reproduces the folding and unfolding kinetics captured by the reaction coordinate hb4 very accurately.

Decreasing the memory time in the GLE while keeping the friction coefficient, i.e., the integral over the memory function, constant, the folding kinetics changes significantly and differently for folding and unfolding events. This shows that memory effects are important even for the formation kinetics of a single α-helix and depend on the distance between the folded and unfolded state to the barrier, as predicted by non-Markovian reaction rate theory [61].

In contrast, the Markovian Langevin equation cannot reproduce the full Ala_9 reconfiguration dynamics, even with a fitted friction profile; this follows from the comparison of the folding and unfolding kinetics, which would need to be modeled with different friction profiles in order to reproduce the MD simulation kinetics.

In this chapter, we have mostly used the approximate GLE model as a diagnostic tool to understand and quantify non-Markovian effects; since non-Markovian simulations are rather inexpensive, they can also be used as an efficient tool to simulate the response of proteins to environmental changes, e.g. externally applied forces. In fact, our extraction technique for the memory function can also be applied to trajectories from single-molecule experiments [91], which would enable to perform non-Markovian GLE simulations on experimental systems directly, without the need of atomistic MD simulations. Because of the limited time resolution of typical experimental data, suitable extraction techniques would have to be used [59, 87].

6 Self-Consistent Markovian Embedding of Nonlinear Friction GLEs

In chapter 3 and chapter 4, we discussed the derivation of GLEs with nonlinear friction kernels and the data-based numerical estimation of these kernels. For configuration-independent memory functions, we demonstrated in chapter 5 how to extract the memory kernel from time series data and perform GLE simulations via Markovian embedding.

In this chapter, we focus on performing efficient computer simulations of GLEs with given configuration-dependent memory functions, i.e., generating data from the nonlinear GLE itself. For this, we discuss three systems of coupled Markovian Langevin equations whose mean dynamics are shown to be equivalent to the dynamics of corresponding GLEs with nonlinear friction kernels generated by specified projection operators. Self-consistency is accomplished by applying the projection operator formalism to the system of Markovian stochastic equations and, by doing so, obtaining closed relations between the parameters in the embedding system of Markovian Langevin equations and the parameters appearing in the GLE. Self-consistency is further explained below.
The first Markovian embedding system allows for the simulation of a memory function in which the time component contains a delta contribution and, else, is a sum of exponentials. In the second system, no delta contribution is needed in the time component, which is now a sum of exponentially decaying oscillations. In the first two embedding methods, the effective mass of the reaction coordinate is assumed to be constant.
In the third system, we consider a nonlinear GLE where the mass and the friction kernel are configuration-dependent, i.e., general nonlinear functions of the reaction coordinate.

To understand our approach, one needs to know what self-consistency means. Thus, before continuing with Markovian embedding, we motivate our approach by discussing self-consistency so that, in the following sections, it becomes clear how we ensure it.

6.1 Self-consistency

In a nutshell, self-consistency means that one gets *out* what one puts *in*. Here, we talk about *input* parameters and *output* parameters.
Input parameters refer to the functions in a GLE, such as the position-dependent mass, the potential of mean force (PMF), and the memory functions. These are computed from a *reference trajectory* using a numerical extraction scheme like the one presented in chapter 4 and then used as input in the Markovian embedding system to generate a *predicted trajectory*. To obtain the *output* parameters, one now computes the mass, the PMF, and the memory functions from the *predicted trajectory* using the same extraction scheme.
The embedding scheme is self-consistent if the output parameters, extracted from the predicted trajectory, equal the input parameters, extracted from the reference trajectory.

We exemplify self-consistency using a model given by a GLE with constant mass M and single-exponential memory,

$$\ddot{x}_t = -\frac{1}{M}U'(x_t) - \int_0^t \mathrm{d}s\, \Gamma_0 e^{-(t-s)/\tau}\dot{x}_s + F_\mathrm{R}(t), \tag{6.1a}$$

$$\langle F_\mathrm{R}(t)\rangle = 0, \qquad \langle F_\mathrm{R}(t), F_\mathrm{R}(0)\rangle = \frac{k_B T}{M}\,\Gamma_0 e^{-t/\tau}, \tag{6.1b}$$

where the brackets $\langle\,\rangle$ in eq. (6.1b) are averages over initial conditions. Eq. (6.1) can be generated from the following coupled system of Markovian Langevin equations

$$\dot{x}_t = v_t, \tag{6.2a}$$

$$\dot{v}_t = -\frac{1}{M}U'(x_t) + \gamma_1\, u_t, \tag{6.2b}$$

$$\dot{u}_t = -\gamma_2\, u_t - \gamma_1\, v_t + \sqrt{2k_B T\gamma_2}\,\eta(t), \tag{6.2c}$$

$$\mathbb{E}[\eta(t)] = 0, \qquad \mathbb{E}[\eta(t)\eta(t')] = \delta(t-t'), \tag{6.2d}$$

if one chooses the input parameters γ_1, γ_2 to be

$$\gamma_1 = \sqrt{\Gamma_0}, \qquad\qquad \gamma_2 = \frac{1}{\tau}. \tag{6.3}$$

In eq. (6.2), $\eta(t)$ is white noise and $\mathbb{E}[\,]$ denotes an average over the noise. When it comes to numerical simulations, eq. (6.2) is preferable over eq. (6.1) because eq. (6.1) is an integro-differential equation and thus numerically much more demanding. Suppose the memory kernel $\Gamma(t) = \Gamma_0 e^{-t/\tau}$ is known, e.g., the parameters Γ_0 and τ have been estimated from data. Then, eq. (6.3) determines how to choose the parameters of eq. (6.2) in order to simulate the GLE in eq. (6.1). Self-consistency then means that, after performing computer simulations of the system in eq. (6.2) and numerically extracting GLE parameters from the simulated trajectory using the same method employed to estimate Γ_0 and τ from data, one obtains values of Γ_0 and τ that are identical to the values used in the simulations. This means that the numerical extraction of GLE parameters and the simulation procedure are consistent.

6.2 Embedding for constant mass

For the Markovian embedding of nonlinear friction GLEs (nfGLEs) with constant mass, we consider the nfGLE in chapter 3.2. In the following, we recapitulate the derivation of the nfGLE from a general classical Hamiltonian system. For further details, we refer to chapter 3. We consider the N particle Hamiltonian

$$H(\omega) = \frac{1}{2}\mathbf{P}^T\hat{\mathbb{M}}^{-1}\mathbf{P} + V(\mathbf{R}), \tag{6.4}$$

where $\omega = (\mathbf{R}, \mathbf{P})$ is a $6N$ state vector in phase space $\Omega = \mathbb{R}^{6N}$ with positions $\mathbf{R} \in \mathbb{R}^{3N}$, $\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)$, and momenta $\mathbf{P} \in \mathbb{R}^{3N}$, $\mathbf{P} = (\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_N)$. The diagonal mass

matrix $\hat{\mathrm{M}}$ contains the masses of all particles. The interaction potential V is a function of positions only, meaning velocities and positions are decoupled. A point in phase space at time $t \geq 0$ is denoted by ω_t, given the initial state ω_0 at time $t = 0$. Observables of phase space are denoted by $B_t = B(\omega_t) = B(\omega_0, t)$. The inner product of two observables B and C is the equilibrium average over the Boltzmann distribution $\rho_{\mathrm{eq}}(\omega_0) = \exp(-\beta H(\omega_0))/Z$, i.e.,

$$\langle B_t, C_{t'} \rangle = \int_\Omega \mathrm{d}\omega_0 \, \rho_{\mathrm{eq}}(\omega_0) B(\omega_0, t) C(\omega_0, t'), \tag{6.5}$$

with $Z = \int_\Omega \mathrm{d}\omega_0 \, e^{-\beta H(\omega_0)}$ being the partition function. The conditional correlation between two observables B and C is defined by

$$\langle B_t, C_{t'} \rangle_{A_0} = \frac{\langle \delta(A(\hat{\omega}_0) - A(\omega_0)), B(\hat{\omega}_0, t) C(\hat{\omega}_0, t') \rangle}{\langle \delta(A(\hat{\omega}_0) - A(\omega_0)) \rangle}, \tag{6.6}$$

In eq. (6.6), variables with a hat, i.e., $\hat{\omega}_0$, are integrated over. The condition in eq. (6.6) is that the observable A has initially the value $A_0 = A(\omega_0)$.
In the remainder, we choose the observable of interest, or reaction coordinate, to be A, and we assume it to be a function of particle positions only, i.e., $A(\omega_t) = A(\mathbf{R}_t) = A_t$. Remind you that, with this choice, the Liouville equation for the velocity reads

$$\dot{A}_t = (\hat{\mathrm{M}}^{-1} \mathbf{P}_t) \cdot \nabla_R A_t. \tag{6.7}$$

Hence, the velocity $\dot{A}_t$ is linear in particle momenta. The projection operator is given by [110, 4]

$$\mathcal{P} B_t = \langle B_t \rangle_{A_0} + \frac{\langle \dot{A}_0, B_t \rangle_{A_0}}{\langle \dot{A}_0^2 \rangle_{A_0}} \dot{A}_0 \, . \tag{6.8}$$

The projection $\mathcal{P}$ in eq. (6.8) has two terms. The first term is independent of the velocity $\dot{A}_0$ and is responsible for the potential of mean force term in the projected dynamics. The second projection is linear in the velocity and generates the nonlinear friction kernel. Using the Dyson operator identity [36], the Liouville equation

$$\ddot{A}_t = \mathcal{L} \dot{A}_t, \tag{6.9a}$$

$$\mathcal{L} = \sum_{n=1}^{N} \left(\frac{\partial H}{\partial \mathbf{p}_n} \cdot \frac{\partial}{\partial \mathbf{r}_n} - \frac{\partial H}{\partial \mathbf{r}_n} \cdot \frac{\partial}{\partial \mathbf{p}_n} \right), \tag{6.9b}$$

where $\mathcal{L}$ denotes the Liouville operator, is decomposed into [88, 113]

$$\ddot{A}_t = e^{t\mathcal{L}} \mathcal{P} \mathcal{L} \dot{A}_0 + \int_0^t \mathrm{d}s \, e^{(t-s)\mathcal{L}} \mathcal{P} \mathcal{L} F_{\mathrm{R}}(s) + F_{\mathrm{R}}(t), \tag{6.10a}$$

$$F_{\mathrm{R}}(t) = e^{t\mathcal{Q}\mathcal{L}} \mathcal{Q} \mathcal{L} \dot{A}_0. \tag{6.10b}$$

Here, we introduced the projection $\mathcal{Q} = 1 - \mathcal{P}$ onto the complementary subspace [115]. Inserting the projection in eq. (6.8) into eq. (6.10), one obtains the nfGLE [110, 4] (see

chapter 3)

$$\ddot{A}_t = -\frac{1}{M(A_t)}\frac{\mathrm{d}U_{\text{eff}}(A_t)}{\mathrm{d}A_t} - \int_0^t \mathrm{d}s\, \Gamma(A_s, t-s)\dot{A}_s + F_{\mathrm{R}}(t), \tag{6.11a}$$

$$U_{\text{eff}}(A) = U_{\mathrm{PMF}}(A) + k_B T \ln M(A), \tag{6.11b}$$

$$\Gamma(A,t) = \beta D(A,t)\frac{\mathrm{d}U_{\text{eff}}(A)}{\mathrm{d}A} - \frac{\mathrm{d}D(A,t)}{\mathrm{d}A} + \frac{\langle \ddot{A}_0, F_{\mathrm{R}}(t)\rangle_A}{\langle \dot{A}_0^2\rangle_A}, \tag{6.11c}$$

$$D(A,t) = \frac{\langle \dot{A}_0^2, F_{\mathrm{R}}(t)\rangle_A}{\langle \dot{A}_0^2\rangle_A}, \tag{6.11d}$$

with the position dependent mass

$$M(A) = k_B T / \langle \dot{A}_0^2\rangle_A, \tag{6.11e}$$

and the potential of mean force (PMF)

$$U_{\mathrm{PMF}}(A) = -k_B T \ln\langle \delta(A_0 - A)\rangle. \tag{6.11f}$$

An algorithm to extract GLE parameters from time series data by explicitly computing the random force $F_{\mathrm{R}}(t)$ was introduced in refs. [16, 72] for the Mori GLE and extended to nfGLEs in the presence of a potential of mean force in chapter 4 (cf. ref. [4]). The random force (and simultaneously the memory kernel) is computed by discretizing the iterative equation

$$F_{\mathrm{R}}(\omega_0, t+\Delta t) = F_{\mathrm{R}}(\omega_{\Delta t}, t) + \int_0^{\Delta t} \mathrm{d}s\, \Gamma(A_{\Delta t - s}, t+s)\dot{A}_{\Delta t - s}, \tag{6.12}$$

where $\omega_{\Delta t}$ corresponds to the configuration obtained by propagating ω_0 with the Liouville propagator $e^{\Delta t L}$, while $F_{\mathrm{R}}(\omega_0, t+\Delta t)$ follows from an orthogonal propagation with $e^{\Delta t Q L}$. One can also use the Volterra scheme discussed in ref. [110].

6.2.1 Coupling via velocity

We now return to the main topic of this chapter. We present the first Markovian embedding scheme for the nfGLE in eq. (6.11) where the effective mass $M(A) \equiv M$ is constant. As discussed in chapter 3.1.3, for example, the effective mass is constant for linear combinations of positions [43] and also for linear combinations of distances [4]. The case in which the mass is also a function of the reaction coordinate is discussed in section 6.3.

The first embedding consists of a system of underdamped Langevin equations that are coupled via their velocities [71]:

$$\dot{x}_t = v_t, \tag{6.13a}$$

$$\hat{\mathrm{M}} \cdot \dot{\boldsymbol{\nu}}_t = \mathbf{F}(x_t) - \hat{\boldsymbol{\gamma}}(x_t)\cdot\boldsymbol{\nu}_t + \hat{\boldsymbol{\sigma}}(x_t)\cdot\boldsymbol{\eta}(t). \tag{6.13b}$$

In eq. (6.13), the velocity vector is given by $\boldsymbol{\nu}_t = (v_t, u_1(t), u_2(t), \ldots, u_N(t))$, meaning, there are N auxiliary velocities $\{u_n\}$. The mass matrix $\hat{\mathrm{M}} \in \mathbb{R}^{(N+1)\times(N+1)}$ is a diagonal

matrix with entries $M, m_1, m_2, \ldots, m_N$. The friction matrix $\hat{\boldsymbol{\gamma}} \in \mathbb{R}^{(N+1)\times(N+1)}$ is only non-zero in the first row, the first column and in the diagonal entries, i.e., it is given by

$$\hat{\boldsymbol{\gamma}}(x_t) = \begin{pmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} & \gamma_{14} & \cdots \\ \gamma_{21}(x_t) & \gamma_{22} & 0 & 0 & \\ \gamma_{31}(x_t) & 0 & \gamma_{33} & 0 & \\ \gamma_{41}(x_t) & 0 & 0 & \gamma_{44} & \\ \vdots & & & & \ddots \end{pmatrix}. \tag{6.14}$$

The multiplicative noise matrix $\boldsymbol{\sigma}(x_t) \in \mathbb{R}^{(N+1)\times 2N}$ is given by

$$\boldsymbol{\sigma}(x_t) = \begin{pmatrix} s_{11} & 0 & s_{21} & 0 & s_{31} & 0 & \cdots \\ s_{12}(x_t) & s_{13}(x_t) & 0 & 0 & 0 & 0 & \\ 0 & 0 & s_{22}(x_t) & s_{23}(x_t) & 0 & 0 & \\ 0 & 0 & 0 & 0 & s_{32}(x_t) & s_{33}(x_t) & \\ \vdots & & & & & & \ddots \end{pmatrix}. \tag{6.15}$$

The noise vector is $\boldsymbol{\eta}(t) \in \mathbb{R}^{2N}$ and consists of independent white noises, i.e.,

$$\mathbb{E}[\eta_i(t)] = 0, \qquad \mathbb{E}[\eta_i(t), \eta_j(t')] = \delta_{ij}\delta(t-t'), \tag{6.16}$$

for all $i, j = 1, 2, \ldots, 2N$ where $\mathbb{E}[\cdot]$ denotes an average over the noise. The force vector $\mathbf{F}(x_t) \in \mathbb{R}^{N+1}$ in eq. (6.13) is given by $\mathbf{F}(x_t) = (-U'(x_t), 0, 0, \ldots, 0)^T$ containing an arbitrary potential U.

In order for the system in eq. (6.13) to be in equilibrium, the fluctuation-dissipation theorem (FDT) must hold [71]

$$\left(\hat{\boldsymbol{\gamma}}(x_t) + \hat{\boldsymbol{\gamma}}^T(x_t)\right)/\beta = \hat{\boldsymbol{\sigma}}(x_t) \cdot \hat{\boldsymbol{\sigma}}^T(x_t), \tag{6.17}$$

where $\beta = 1/k_B T$ is the inverse thermal energy. If the FDT in eq. (6.17) holds, it follows from the Fokker-Planck equation that the system in eq. (6.13) is Boltzmann distributed, i.e., it has the stationary distribution function

$$\rho_{st}(x, v, \mathbf{u}) = e^{-\beta U(x)} e^{-\beta M v^2/2} e^{-\beta \sum_n m_n u_n^2/2}/\mathcal{N}, \tag{6.18}$$

where $1/\mathcal{N}$ denotes the normalization factor.
To derive explicit expressions for the matrix components in eq. (6.15) using the FDT in eq. (6.17), we write $\boldsymbol{\sigma}(x_t)$ as a block matrix

$$\boldsymbol{\sigma}(x_t) = \begin{pmatrix} \hat{\mathrm{S}}_1(x_t) & \hat{\mathrm{S}}_2(x_t) & \hat{\mathrm{S}}_3(x_t) & \cdots & \hat{\mathrm{S}}_N(x_t) \end{pmatrix}, \tag{6.19}$$

where the matrices $\hat{\mathrm{S}}_n \in \mathbb{R}^{(N+1)\times 2}$ have non-zero entries only in the first row and in the

$(n+1)$-th row, i.e., they are given by

$$\hat{\mathrm{S}}_n(x_t) = \begin{pmatrix} s_{n1} & 0 \\ 0 & 0 \\ \vdots & \\ s_{n2}(x_t) & s_{n3}(x_t) \\ 0 & 0 \\ \vdots & \end{pmatrix} \leftarrow (n+1)\text{-th row}, \tag{6.20}$$

where the $s_{ij}(x_t)$ are the matrix entries in eq. (6.15). With the block matrix notation, the r.h.s. of the FDT in eq. (6.17) becomes

$$\hat{\boldsymbol{\sigma}}(x_t) \cdot \hat{\boldsymbol{\sigma}}^T(x_t) = \sum_{n=1}^{N} \hat{\mathrm{S}}_n(x_t)\hat{\mathrm{S}}_n^T(x_t). \tag{6.21}$$

We write the l.h.s. of eq. (6.17) also as a sum. For this, we set

$$\gamma_{11} = \sum_{n=1}^{N} g_n, \tag{6.22}$$

and write the friction matrix as a sum of matrices $\hat{\mathrm{G}}_n(x_t) \in \mathbb{R}^{(N+1)\times(N+1)}$

$$\hat{\boldsymbol{\gamma}}(x_t) = \sum_{n=1}^{N} \hat{\mathrm{G}}_n(x_t), \tag{6.23}$$

where the matrices $\hat{\mathrm{G}}_n(x_t)$ have non-zero entries only in the first and $(n+1)$-th row and in the $(n+1)$-th column, i.e.,

$$\hat{\mathrm{G}}_n(x_t) = \begin{pmatrix} g_n & 0 & \cdots & \gamma_{1,n+1} & 0 & \cdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots \\ \vdots & & & \vdots & & \\ \gamma_{n+1,1}(x_t) & 0 & \cdots & \gamma_{n+1,n+1} & 0 & \cdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots \\ \vdots & & & \vdots & & \end{pmatrix}. \tag{6.24}$$

With this, the FDT in eq. (6.17) becomes

$$\sum_{n=1}^{N} \left(\hat{\mathrm{G}}_n(x_t) + \hat{\mathrm{G}}_n^T(x_t)\right) = \sum_{n=1}^{N} \hat{\mathrm{S}}_n(x_t)\hat{\mathrm{S}}_n^T(x_t). \tag{6.25}$$

Eq. (6.25) is fulfilled when $(\hat{\mathrm{G}}_n(x_t) + \hat{\mathrm{G}}_n^T(x_t)) = S_n(x_t)\hat{\mathrm{S}}_n^T(x_t)$ holds, which leads to the relations

$$2\frac{g_n}{\beta} = s_{n1}^2, \tag{6.26a}$$

$$\frac{\gamma_{1,n+1} + \gamma_{n+1,1}(x_t)}{\beta} = s_{n1}(x_t)s_{n2}(x_t), \tag{6.26b}$$

$$2\frac{\gamma_{n+1,n+1}}{\beta} = s_{n2}^2(x_t) + s_{n3}^2(x_t). \tag{6.26c}$$

From the equations in (6.26), we find the entries of $\hat{\boldsymbol{\sigma}}(x_t)$ in terms of the friction coefficients

$$s_{n1} = \sqrt{2k_BTg_n}, \tag{6.27a}$$

$$s_{n2}(x_t) = k_BT\frac{\gamma_{1,n+1} + \gamma_{n+1,1}(x_t)}{s_{n1}}, \tag{6.27b}$$

$$s_{n3}(x_t) = \sqrt{2k_BT\gamma_{n+1,n+1} - s_{n2}^2(x_t)}. \tag{6.27c}$$

Inserting eq. (6.27b) into eq. (6.27c) leads to

$$\sqrt{2k_BT\gamma_{n+1,n+1} - s_{n2}^2(x_t)} = \sqrt{2k_BT\gamma_{n+1,n+1} - \left(k_BT\frac{\gamma_{1,n+1} + \gamma_{n+1,1}(x_t)}{s_{n1}}\right)^2} \tag{6.28a}$$

$$= \sqrt{\frac{k_BT}{2\,g_n}}\sqrt{4\gamma_{n+1,n+1}g_n - \left(\gamma_{1,n+1} + \gamma_{n+1,1}(x_t)\right)^2}. \tag{6.28b}$$

Meaning, in order to obtain finite, real valued multiplicative noise factors, the following inequalities

$$g_n > 0, \tag{6.29a}$$

$$4\gamma_{n+1,n+1}\,g_n \geq \left(\gamma_{1,n+1} + \gamma_{n+1,1}(x_t)\right)^2, \tag{6.29b}$$

must hold for all $n = 1, 2, \ldots, N$ and all values x_t. For each auxiliary velocity $u_n(t)$ and for given $\gamma_{1,n+1}, \gamma_{n+1,1}(x_t), \gamma_{n+1,n+1}$, eq. (6.29) puts a lower bound on the friction coefficient g_n.

Assuming the validity of eq. (6.29), by solving the equation for the velocity $u_n(t)$ in eq. (6.13) and inserting the result into the equation for v_t, one obtains the GLE

$$\dot{v}_t = -\frac{U'(x_t)}{M} - \int_0^t \mathrm{d}s\,\Gamma(x_s, t-s)v_s + \xi(t), \tag{6.30a}$$

$$\Gamma(x_s, t-s) = \frac{1}{M}\sum_{n=1}^{N}\left(g_n\delta(t-s) - \gamma_{1,n+1}e^{-\frac{\gamma_{n+1,n+1}}{m_n}(t-s)}\frac{\gamma_{n+1,1}(x_s)}{m_n}\right), \tag{6.30b}$$

$$\xi(t) = \frac{1}{M}\left(-\sum_{n=1}^{N}\frac{\gamma_{1,n+1}}{m_n}e^{-\frac{\gamma_{n+1,n+1}}{m_n}t}u_n(0) + \sum_{j=1}^{2N}\sigma_{1,j}(x_t)\eta_j(t)\right.$$
$$\left. - \sum_{n=1}^{N}\sum_{j=1}^{2N}\int_0^t \mathrm{d}s\,\frac{\gamma_{1,n+1}}{m_n}e^{-\frac{\gamma_{n+1,n+1}}{m_n}(t-s)}\sigma_{n+1,j}(x_s)\,\eta_j(s)\right). \tag{6.30c}$$

Eq. (6.30) has the structure of the GLE in eq. (6.11) for constant mass M. To show that eq. (6.13) can be used as a Markovian embedding of the GLE in eq. (6.11), we apply the projection in eq. (6.8) on the mean dynamics of eq. (6.13) and demonstrate that this produces the GLE in eq. (6.30), when averaged over the white noise. For this, we need the infinitesimal generator [71, 104] of the system in eq. (6.13).

To obtain the infinitesimal generator, we consider eq. (6.13) as an Ito diffussion process. An Ito diffusion process can be written in the general form [42]

$$\mathrm{d}\boldsymbol{\zeta}(t) = \mathbf{a}(\zeta(t))\,\mathrm{d}t + \hat{\mathrm{B}}(\boldsymbol{\zeta}(t)) \cdot \mathrm{d}\mathbf{W}(t). \tag{6.31}$$

In eq. (6.31), $\boldsymbol{\zeta}(t) \in \mathbb{R}^N$ is the vector of state variables at time t. The vector valued function $\mathbf{a} : \mathbb{R}^N \to \mathbb{R}^N$ is an arbitrary, analytic function of the state variables, and the multiplicative noise matrix $\hat{\mathrm{B}} : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^{N\times M}$ is an analytic, matrix valued function of the state variables. $\mathrm{d}\mathbf{W} \in \mathbb{R}^M$ denotes a vector of M increments of a Wiener process with properties

$$\mathbb{E}[\mathrm{d}W_i(t)] = 0, \qquad \mathbb{E}[\mathrm{d}W_i(t)\,\mathrm{d}W_j(s)] = \delta_{ij}\delta_{t,s}\,\mathrm{d}t. \tag{6.32}$$

Further, an Ito diffusion process is chractarized by the property [42]

$$\mathbb{E}[\mathrm{d}W_i(t)\zeta_j(t)] = \mathbb{E}[\mathrm{d}W_i(t)]\,\mathbb{E}[\zeta_j(t)] = 0. \tag{6.33}$$

We consider an observable $F(\boldsymbol{\zeta}(t))$ that is a function of the state variables $\boldsymbol{\zeta}$ in eq. (6.31) and write down its total differential up to second order in $\mathrm{d}\boldsymbol{\zeta}$

$$\mathrm{d}F(\boldsymbol{\zeta}) = \nabla_\zeta F \cdot \mathrm{d}\boldsymbol{\zeta} + \frac{1}{2}\,\mathrm{d}\boldsymbol{\zeta}^T \cdot \hat{\mathrm{H}}_F(\boldsymbol{\zeta}) \cdot \mathrm{d}\boldsymbol{\zeta} + \mathcal{O}(\mathrm{d}\boldsymbol{\zeta}^3), \tag{6.34}$$

where $\hat{\mathrm{H}}_F(\boldsymbol{\zeta})$ is the hessian of the function F. Inserting eq. (6.31) into eq. (6.34) and neglecting terms $\mathcal{O}(\mathrm{d}t^2)$, we find, using the Einstein summation convention,

$$\mathrm{d}F = \frac{\partial F}{\partial \zeta_i} a_i\,\mathrm{d}t + \frac{\partial F}{\partial \zeta_i} B_{ij}\,\mathrm{d}W_j(t) + \frac{1}{2}\,\mathrm{d}W_j(t) B_{ij} \frac{\partial^2 F}{\partial \zeta_i \partial \zeta_k} B_{kl}\,\mathrm{d}W_l(t). \tag{6.35}$$

Eq. (6.35) is known as *Ito's formula* [42], it is the chain rule of differentiation for Ito diffusion processes. We average eq. (6.35) over the noise and make use of the properties given in eq. (6.32) and eq. (6.33) to obtain

$$\mathrm{d}F = \left(\frac{\partial F}{\partial \zeta_i} a_i + \frac{1}{2}\, B_{ij} B_{kj} \frac{\partial^2 F}{\partial \zeta_i \partial \zeta_k} \right) \mathrm{d}t. \tag{6.36}$$

This can be written as

$$\frac{\mathrm{d}F}{\mathrm{d}t} = \left(a_i \frac{\partial}{\partial \zeta_i} + \frac{1}{2}\, B_{ij} B_{kj} \frac{\partial^2}{\partial \zeta_i \partial \zeta_k} \right) F. \tag{6.37}$$

Eq. (6.37) defines the infinitesimal generator of the mean dynamics since we performed an average to derive it where the average is taken with respect to the Wiener measure associated with the driving noise process [71], i.e., the white noise $\boldsymbol{\eta}(t)$ in the case of eq. (6.13). The infinitesimal generator reads

$$L^\dagger_{\mathrm{FP}} = a_i \frac{\partial}{\partial \zeta_i} + \frac{1}{2}\, B_{ij} B_{kj} \frac{\partial^2}{\partial \zeta_i \partial \zeta_k}. \tag{6.38}$$

We choose to denote the infinitesimal generator by $L^\dagger_{\mathrm{FP}}$ because its adjoint w.r.t. the L_2 inner product is the Fokker-Planck operator [71], i.e.,

$$\begin{aligned} \langle \rho, L^\dagger_{\mathrm{FP}} F \rangle_{L_2} &= \int \mathrm{d}^N\zeta\, \rho(\boldsymbol{\zeta}) L^\dagger_{\mathrm{FP}} F(\boldsymbol{\zeta}) = \int \mathrm{d}^N\zeta\, F(\boldsymbol{\zeta}) L_{\mathrm{FP}} \rho(\boldsymbol{\zeta}) \\ &= \langle L_{\mathrm{FP}} \rho, F \rangle_{L_2}, \end{aligned} \tag{6.39}$$

assuming that ρ vanishes at the integration boundaries and with

$$L_{\mathrm{FP}} = -\frac{\partial}{\partial \zeta_i} a_i + \frac{1}{2} \frac{\partial^2}{\partial \zeta_i \partial \zeta_k} B_{ij} B_{kj}. \tag{6.40}$$

For the Markovian embedding system in eq.(6.13), the infinitesimal generator reads

$$\begin{aligned} L^{\dagger}_{\mathrm{FP}} = v\frac{\partial}{\partial x} - \frac{1}{M}\left(U'(x) + \sum_{n=1}^{N}\left(g_n v + \gamma_{1,n+1} u_n\right)\right)\frac{\partial}{\partial v} \\ - \sum_{n=1}^{N} \frac{1}{m_n}\left(\gamma_{n+1,1}(x) v + \gamma_{n+1,n+1} u_n\right)\frac{\partial}{\partial u_n} \\ + \frac{1}{2}\sum_{i,j=1}^{N+1}\left(\hat{\boldsymbol{\sigma}}(x)\cdot\hat{\boldsymbol{\sigma}}^{T}(x)\right)_{ij}\frac{\partial^2}{\partial \nu_i \partial \nu_j}. \end{aligned} \tag{6.41}$$

Applying the corresponding Fokker-Planck operator L_{FP} on the distribution $\rho_{st}(x, v, \mathbf{u})$ in eq. (6.18), one finds

$$L_{\mathrm{FP}} \rho_{st}(x, v, \mathbf{u}) = 0, \tag{6.42}$$

which proves that ρ_{st} in eq. (6.18) is indeed the stationary distribution of the system in eq. (6.13).

On average, the time evolution of the velocity v in eq. (6.13) obeys the equation

$$\dot{v}_t = L^{\dagger}_{\mathrm{FP}} v_t. \tag{6.43}$$

This is the stochastic equivalent of the Liouville equation for deterministic systems. To find the GLE generated by the projection in eq. (6.8), we adjust the formalism to the system in eq. (6.13). The inner product is now an average over the stationary distribution in eq. (6.18) and the projection reads

$$\mathcal{P}\phi_t = \langle \phi_t \rangle_{x_0} + \frac{\langle v_0, \phi_t \rangle_{x_0}}{\langle v_0^2 \rangle_{x_0}} v_0. \tag{6.44}$$

By applying the decomposition given in eq. (6.10) in combination with the projection above, we can transform eq. (6.43) into the GLE in eq. (6.30), when the white noise terms in the random force $\xi(t)$ are averaged out.

According to eq. (6.10) and the adjoint Fokker-Planck operator $L^{\dagger}_{\mathrm{FP}}$ in eq. (6.41), the Markovian contribution is given by

$$e^{tL^{\dagger}_{\mathrm{FP}}} \mathcal{P} L^{\dagger}_{\mathrm{FP}} v = -\frac{1}{M}\left(\frac{\mathrm{d}U(x_t)}{\mathrm{d}x_t} + \sum_{n=1}^{N} g_n v_t\right). \tag{6.45}$$

Next, we compute the random force term using eq. (6.10)

$$\begin{aligned} F_{\mathrm{R}}(t) = e^{t\mathcal{Q}L^{\dagger}_{\mathrm{FP}}} \mathcal{Q} L^{\dagger}_{\mathrm{FP}} v = \sum_{n=0}^{\infty} \frac{t^n}{n!} (\mathcal{Q} L^{\dagger}_{\mathrm{FP}})^n \mathcal{Q} L^{\dagger}_{\mathrm{FP}} v \\ = -\frac{1}{M}\sum_{n=1}^{N} \gamma_{1,n+1} e^{-\frac{\gamma_{n+1,n+1}}{m_n} t} u_n, \end{aligned} \tag{6.46}$$

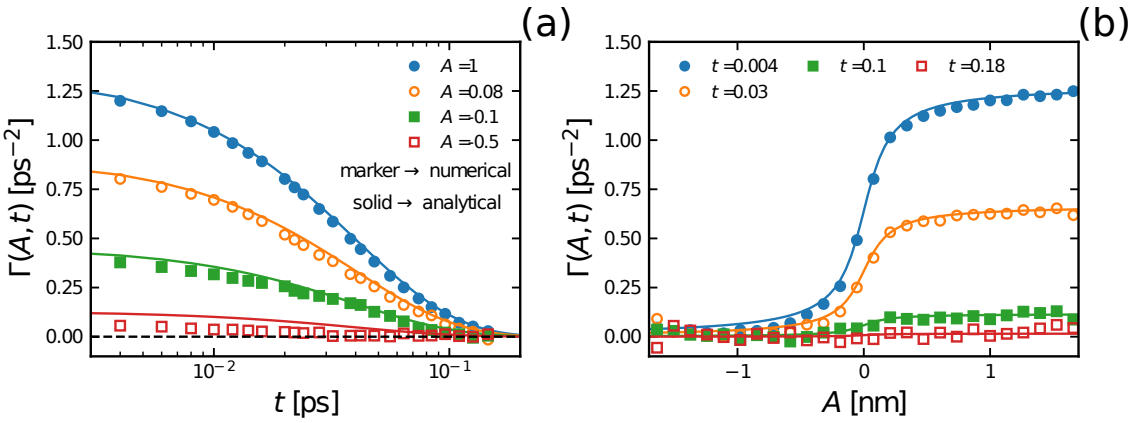

Figure 6.1: Comparison of the numerically extracted memory kernel in eq. (6.11c) (markers) with the theoretical memory kernel in eq. (6.30b) (solid lines). In **(a)**, as a function of time at different positions A, and in **(b)**, as a function of the position A at different times t. The underlying trajectory was generated via simulations of the system in eq. (6.13) with U being a double well potential given in eq. (6.48) and $\gamma_{21}(A)$ given in eq. (6.49). The numerical results are obtained using eq. (6.12) (cf. ref. [4]).

from which we obtain

$$\int_0^t \mathrm{d}s\, e^{(t-s)\mathcal{L}_{\mathrm{FP}}^\dagger} \mathcal{P} L_{\mathrm{FP}}^\dagger F_{\mathrm{R}}(s) = -\int_0^t \mathrm{d}s \sum_{n=1}^{N} \frac{\gamma_{1,n+1}}{M} e^{-\frac{\gamma_{n+1,n+1}}{m_n} t} \frac{\gamma_{n+1,1}(x_{t-s})}{m_n} v_{t-s}. \tag{6.47}$$

The results in eq. (6.45) and eq. (6.47) coincide with the deterministic part of the GLE in eq. (6.30). Note that eq. (6.46) reproduces the noise term in eq. (6.30) averaged over the noises η_j. This means, we recover the GLE in eq. (6.30), when averaged over the white noise $\boldsymbol{\eta}(t)$. Thus, the memory kernel $\Gamma(x,t)$, obtained from the Markovian embedding in eq. (6.13) by applying the projection operators in eq. (6.44), can be computed in a closed form and is equal to the one in eq. (6.30b), which was obtained by solving the Langevin equation in eq. (6.13).

The Markovian embedding is numerically demonstrated in fig. 6.1. Using eq. (6.13) and eq. (6.27), we generate a trajectory x_t in a double well potential

$$U(x_t) = U_0(x_t^2 - 1)^2, \tag{6.48}$$

with $U_0 = k_B T$ where the thermal energy has the value $k_B T = 2.494\,\mathrm{kJ/mol}$, which corresponds to the thermal energy at room temperature in MD units. The remaining parameter values are $M = 4\,\mathrm{u}$, $m = 10\,\mathrm{u}$, $\gamma_{11} = 0.9\,\mathrm{u/ps}$, $\gamma_{12} = -6\,\mathrm{u/ps}$, $\gamma_{22} = 100\,\mathrm{u/ps}$, and as the friction profile, we use

$$\gamma_{21}(x_t) = \pi + 9\arctan(2\pi x_t). \tag{6.49}$$

As solid lines in fig. 6.1, we show $\Gamma(x,t)$ given in eq. (6.30b). The markers in fig. 6.1 correspond to the numerically extracted $\Gamma(x,t)$ in eq. (6.11c), using eq. (6.12) [4].

The agreement between input functions and extracted output data is very good. Only for small values of the memory kernel, we observe slight deviations which are due to numerical errors. The presence of a delta peak in the memory kernel adds noise to the data which makes the numerical estimation more challenging. This will become more clear in the next section as we discuss the second embedding system where the delta peak in the memory kernel is absent.

6.2.2 Fast auxiliary variable limit

A drawback of the method discussed in the previous section is the presence of an instantaneous friction that leads to a delta contribution in the memory kernel, i.e., g_n in eq. (6.30b). It has been shown that the system in eq. 6.13 becomes equivalent to the nonlinear Zwanzig model (cf. chapter 5.3.1) with overdamped auxiliary variables [71, 114] when the instantaneous friction is set to zero, i.e., $g_n = 0$ and $\gamma_{1,n+1}(x) = -\gamma_{n+1,1}(x)$. However, when $\gamma_{1,n+1}(x) = -\gamma_{n+1,1}(x)$, one cannot derive a closed expression for the memory kernel of the nfGLE using the projection operator in eq. (6.44). Thus, it remains unclear how to choose the input functions in the embedding system in order to generate a GLE with the same output functions.
Again, we refer to input functions as the functions that are used in eq. (6.13) and output functions refer to the numerically extracted functions from simulations of the embedding system using the extraction scheme in eq. (6.12).

In this section, we will show that, in a certain limit, the nonlinear Zwanzig model from chapter 5.3.1 can be used as an alternative embedding for the nfGLE corresponding to the projection in eq. (6.44): Consider the N body Hamiltonian given by

$$H(x, p, \{x_i, p_i\}_{i=1}^N) = \frac{p^2}{2M} + \sum_{i=1}^{N} \frac{p_i^2}{m} + U(x) + \sum_{i=1}^{N} \frac{\kappa_i}{2}(\alpha_i(x) - x_i)^2, \tag{6.50}$$

where x_i, with $i = 1, 2, \ldots, N$, denote the 1D positions of auxiliary particles with momenta p_i and equal mass m. The pair (x, p) denotes the relevant coordinates of a tagged particle.

The problem we want to solve is the following: From chapter 5.3.1, we know that the nonlinear Zwanzig model generates a nfGLE with a memory kernel of the form $\Gamma(x_t, x_s, t - s)$. The memory kernel in eq. (6.11) is of a different structure than $\Gamma(x_t, x_s, t - s)$. Using the projection operator in eq. (6.44), we cannot obtain a closed relation between the memory kernel $\Gamma(x, t)$ in eq. (6.11) and the nonlinear functions $\alpha_i(x)$ in eq. (6.50). We will resolve this problem by considering the case

$$\epsilon = \sqrt{m/M} \ll 1. \tag{6.51}$$

Using the rescaled momentum $\tilde{p} = \epsilon\, p$, we can write the Liouville operator of the system in

eq. (6.50) as [107]

$$\mathcal{L} = \mathcal{L}_0 + \epsilon\mathcal{L}_1, \tag{6.52a}$$

$$\mathcal{L}_0 = \sum_{i=1}^{N}\left(\frac{p_i}{m}\frac{\partial}{\partial x_i} - \frac{\partial H}{\partial x_i}\frac{\partial}{\partial p_i}\right), \tag{6.52b}$$

$$\mathcal{L}_1 = \frac{\tilde{p}}{m}\frac{\partial}{\partial x} - \frac{\partial H}{\partial x}\frac{\partial}{\partial \tilde{p}}. \tag{6.52c}$$

Now, we consider the case in which there are N_α distinct nonlinear functions α_n, $n = 1, 2, \ldots, N_\alpha$ with $N_\alpha < N$ and $\sum_\alpha N_\alpha = N$. We divide the set of auxiliary variables into N_α subsets with the index sets I_n that have the function $\alpha_n(x)$ in common, i.e., if $n, k \in I_j$, then $\alpha_n(x) = \alpha_k(x) \equiv \alpha_j(x)$, and we can write

$$\sum_{i=1}^{N}\frac{\kappa_i}{2}(\alpha_i(x) - x_i)^2 = \sum_{n=1}^{N_\alpha}\sum_{i\in I_n}\frac{\kappa_i}{2}(\alpha_n(x) - x_i)^2. \tag{6.53}$$

To obtain a nfGLE for the nonlinear Zwanzig model using projection operators, we use the projection

$$\mathcal{P}B_t = \langle B_t\rangle_x + \frac{\langle\tilde{p}, B_t\rangle_x}{\langle\tilde{p}^2\rangle}\tilde{p}. \tag{6.54}$$

Note that the projection operator in eq. (6.54) is independent of the masses m and M. First, we compute the potential term in the GLE and find

$$e^{t\mathcal{L}}\mathcal{P}\mathcal{L}\tilde{p} = -\epsilon\frac{\mathrm{d}U}{\mathrm{d}x_t}. \tag{6.55}$$

Next, we compute the random force to leading order in ϵ from

$$F_\mathrm{R}(t) = e^{t\mathcal{Q}\mathcal{L}}\mathcal{Q}\mathcal{L}\tilde{p} \tag{6.56a}$$

$$= e^{t\mathcal{Q}\mathcal{L}}\mathcal{Q}(\mathcal{L}_0 + \epsilon\mathcal{L}_1)\tilde{p} \tag{6.56b}$$

$$= \epsilon\, e^{t\mathcal{Q}\mathcal{L}}\mathcal{Q}\mathcal{L}_1\tilde{p}. \tag{6.56c}$$

For this, we use

$$e^{t\mathcal{Q}\mathcal{L}} = e^{\mathcal{Q}(\mathcal{L}_0+\epsilon\mathcal{L}_1)} \tag{6.57a}$$

$$= e^{t\mathcal{Q}\mathcal{L}_0} + \epsilon\int_0^t \mathrm{d}s\, e^{(t-s)\mathcal{Q}\mathcal{L}}\mathcal{Q}\mathcal{L}_1 e^{s\mathcal{Q}\mathcal{L}_0}. \tag{6.57b}$$

Inserting eq. (6.57b) into eq. (6.56c), we find to leading order in ϵ

$$F_\mathrm{R}(t) = \epsilon\, e^{\mathcal{Q}\mathcal{L}_0}\mathcal{Q}\mathcal{L}_1\tilde{p} + \mathcal{O}(\epsilon^2) \tag{6.58a}$$

$$= \epsilon\sum_{i=1}^{N}\Big(\kappa_i\alpha_i'(x)\,(x_i - \alpha(x))\cos(\omega_i t) + \kappa_i\omega_i p_i\alpha_i'(x)\sin(\omega_i t)\Big), \tag{6.58b}$$

where $\omega_i^2 = \kappa_i/m$. According to eq. (6.11), the memory kernel follows from eq. (6.58b) by computing

$$\mathcal{P}\mathcal{L}F_{\mathrm{R}}(t) = \epsilon\,\mathcal{P}\mathcal{L}_1 F_{\mathrm{R}}(t) \tag{6.59a}$$

$$= -\epsilon^2 \sum_{i=1}^{N} \left(\alpha_i'(x)\right)^2 \omega_i^2 \tilde{p}\cos(\omega_i t). \tag{6.59b}$$

The GLE then reads to leading order in ϵ

$$\dot{\tilde{p}} = -\epsilon\, U'(x_t) - \epsilon^2 \int_0^t \mathrm{d}s \sum_{i=1}^{N} \omega_i^2 \cos(\omega_i s) \left(\alpha_i'(x_{t-s})\right)^2 \tilde{p}_{t-s} + F_{\mathrm{R}}(t), \tag{6.60}$$

with $F_{\mathrm{R}}(t)$ given in eq. (6.58b). Rescaling back to $p = \tilde{p}/\epsilon$ and dividing by the mass M, we obtain

$$\dot{v}_t = -\frac{U'(x_t)}{M} - \int_0^t \mathrm{d}s \sum_{i=1}^{N} \kappa_i \cos(\omega_i s) \left(\alpha_i'(x_{t-s})\right)^2 v_{t-s} + \frac{1}{M}\tilde{F_{\mathrm{R}}}(t), \tag{6.61}$$

where $\tilde{F_{\mathrm{R}}}(t) = F_{\mathrm{R}}(t)/\epsilon$.

The memory kernel in eq. (6.60), which we obtained from $\dot{\tilde{p}} = \mathcal{L}\tilde{p}$ in combination with the projection (6.54), can be written as

$$\Gamma(x,t) = \epsilon^2 \sum_{n=1}^{N_\alpha} \frac{\alpha_n'(x)^2}{m} \sum_{i\in I_n} \kappa_i \cos(\omega_i t) + \mathcal{O}(\epsilon^4). \tag{6.62}$$

For small ϵ, eq. (6.62) gives a closed relation between the memory kernel $\Gamma(x,t)$ and the functions $\alpha_n(x)$. In order to perform efficient computer simulations of the Hamiltonian system in eq. (6.50) for $N \to \infty$, we proceed analogously to chapter 5.3 (cf. ref. [4]). We interpret the function of time $\sum_i \kappa_i \cos(\omega_i t)$ as the inverse Fourier transform of an exponentially decaying function, i.e.,

$$\sum_i \kappa_i \cos(\omega_i t) = k_n e^{-t/\tau_n} \left(\cos\left(\frac{\nu_n}{\tau_n}t\right) + \frac{1}{\nu_n}\sin\left(\frac{\nu_n}{\tau_n}t\right)\right). \tag{6.63}$$

This means, we map the Hamiltonian system in eq. (6.50) for $N \to \infty$ onto the stochastic nonlinear Zwanzig model given by [4]

$$\dot{x} = -\frac{p}{M}, \tag{6.64a}$$

$$\dot{p} = -U'(x) - \sum_n k_n\, \alpha_n'(x)\left(\alpha_n(x) - y_n\right), \tag{6.64b}$$

$$\dot{y}_n = -\frac{w_n}{m}, \tag{6.64c}$$

$$\dot{w}_n = -k_n\left(y_n - \alpha_n(x)\right) - \gamma_n w_n + \sqrt{2k_B T\gamma_n}\,\eta_n(t), \tag{6.64d}$$

with the pair (y_n, w_n) being the auxiliary variables and $\eta_n(t)$ being white noise

$$\mathbb{E}[\eta_n(t)] = 0, \qquad \mathbb{E}[\eta_i(t)\eta_j(t')] = \delta_{ij}\delta(t-t'), \tag{6.65}$$

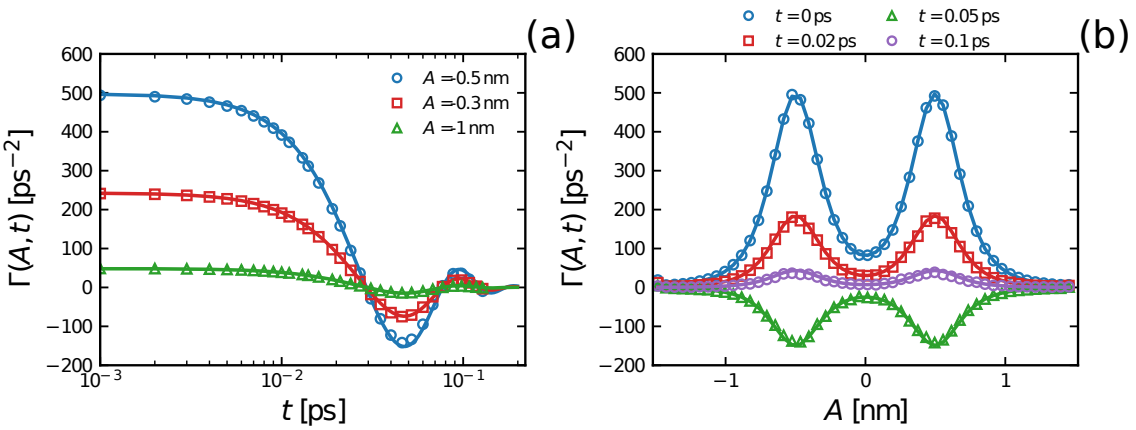

Figure 6.2: Comparison of the numerically extracted memory kernel in eq. (6.11c) (markers) with the theoretical memory kernel in eq. (6.67) and eq. (6.68) (solid lines). In **(a)**, as a function of time at different positions A, and in **(b)**, as a function of the position A at different times t. The underlying trajectory was generated via simulations of the system in eq. (6.64) in a double well potential (eq. (6.48)) and with two auxiliary variables where $\alpha_n(x)$ is given in eq. (6.69). The numerical extraction is accomplished using eq. (6.12) (cf. ref. [4]).

where $\mathbb{E}[\cdot]$ denotes an average taken over the noise. The relations between the parameters appearing in eq. (6.63) and eq. (6.64) are (see chapter 5.3.1)

$$\tau_n = 2\frac{m}{\gamma_n}, \tag{6.66a}$$

$$\nu_n^2 = \frac{2k_n\tau_n}{\gamma_n} - 1, \tag{6.66b}$$

and the position dependent friction kernel in the limit $\epsilon = \sqrt{m/M} \ll 1$ reads

$$\Gamma(x,t) = \sum_n \frac{k_n\,\alpha_n'(x)^2}{M} e^{-t/\tau_n}\left(\cos\left(\frac{\nu_n}{\tau_n}t\right) + \frac{1}{\nu_n}\sin\left(\frac{\nu_n}{\tau_n}t\right)\right). \tag{6.67}$$

In fig. 6.2, we show results for a simulation of the system in eq. (6.64) with the double well potential given in eq. (6.48) and a barrier height of $U_0 = 2k_BT$. We set $k_BT = 2.494$ kJ/mol, i.e. we use MD units (length in nm, time in ps and mass in u). The friction kernel in eq. (6.67) is shown in fig. 6.2 (a) as a function of time at different positions and, in fig. 6.2 (b), as a function of position at different times. The friction kernel is composed of two contributions, i.e., we use two auxiliary variables with

$$\alpha_n'(x)^2 = \left(\alpha_{n,0}\frac{l_n}{l_n^2 + (x-x_n)^2}\right)^2, \tag{6.68}$$

where $x_n = \pm 0.5$ nm, $\alpha_{n,0} = 1$ nm and $l_n = \pi$ nm for $n = 1, 2$.
In order to perform simulations of the friction kernel given in eq. (6.67) and eq. (6.68), one has to compute the functions $\alpha_n(x)$ from the extracted functions $\alpha_n'(x)^2$ and use them as

input in eq. (6.64). While it might not be possible to compute $\alpha_n(x)$ from $\alpha_n'(x)^2$ analytically, it is always possible numerically. For $\alpha_n'(x)^2$ in eq. (6.68), we know the analytical form of $\alpha_n(x)$. It reads

$$\alpha_n(x) = \alpha_{n,0} \arctan\left(\frac{x - x_n}{l_n}\right). \tag{6.69}$$

The shape of the time components of the memory kernel is determined by the parameters in eq. (6.66). In fig. 6.2, we use the same values for the two components, the values are $\tau_n = 0.04\,\text{ps}$ and $\nu_n^2 = 0.84$. The value of the effective mass is $M = 2\,\text{u}$, the value of the auxiliary variable masses is $m = 0.02\,\text{u}$. The mass m is chosen such that $\epsilon = \sqrt{m/M} \ll 1$, i.e., in fig. 6.2 $\epsilon = 0.1$. The value of the friction coefficient is $\gamma_n = 1\,\text{u/ps}$ and the value of the coupling parameter is $k_n = 100\,\text{u/ps}^2$. The amplitudes of the coupling functions $\alpha_n(x)$ are given by $\alpha_{0,n} = 1$.

We observe perfect agreement between the input functions (solid lines) and the output data, numerically extracted from simulated trajectories (markers), shown in fig. 6.2. Compared to the results in fig. 6.1, the perfect agreement in fig. 6.2 suggests that the numerical extraction is more stable in the absence of instanteneous friction.

Validity of eq. (6.67)

The memory function $\Gamma(A, t)$ in eq. (6.11) is invariant under a rescaling of the coordinate A. This follows directly from dimensional analysis of the nfGLE in eq. (6.11). Any rescaling in the reaction coordinate A leads to the same rescaling in the reaction coordinate velocity $\dot{A}$ and acceleration $\ddot{A}$, thus, the memory kernel in eq. (6.11) must be invariant under a rescaling of the reaction coordinate, to arbitrary order in ϵ. Meaning, the correction terms in the expansion in eq. (6.62) must be invariant under a rescaling in A. This is also reflected in the units of the memory kernel.
On the other hand, the effective mass $M = k_B T / \langle \dot{A}^2 \rangle$ scales inversely quadratic in the reaction coordinate velocity. Thus, without changing the memory function, we can rescale A such that $M \gg m$ is fulfilled.
However, the correction terms in the expansion in eq. (6.62) can become non-negligible for certain combinations of parameters, e.g., when $\alpha_{0,n}$ becomes too large. Therefore, one needs to check that the approximation in eq. (6.67) holds numerically. For this, we propose a simple test in the next section.

6.3 Accounting for position-dependent mass

In the previous section, we discussed GLE simulations via Markovian embedding for the nfGLE in eq. (6.11), i.e., when $\Gamma(A_s, \dot{A}_s, t - s) = \Gamma(A_s, t - s)\dot{A}_s$, but in which the mass M is assumed to be constant in A_t. In this section, we consider the more general case where the effective mass may also depend on A_t.

We proceed analogously to section 6.2.1, but now using the projection operator from chapter 3.3 that generates a GLE in which the mean force, containing the PMF, is adjusted for the purpose of simulating GLEs with a position dependent mass.
First, we briefly recapitulate the derivation of the nfGLE from chapter 3.3. Afterwards, we present a system of coupled Markovian Langevin equations and use the projection operator from chapter 3.3 to show that the presented Markovian Langevin system can be used to simulate the nfGLE.

Similar to the previous section, we consider the many-body Hamiltonian of the form in eq. (6.4). The interaction potential V is assumed to be a function of positions only such that there is no coupling between atomistic velocities and positions.
Observables of phase space are denoted by $B_t = B(\omega_t) = B(\omega_0, t)$. The inner product of two observables B and C is given by the equilibrium average over the stationary Boltzmann distribution $\rho_{\mathrm{eq}}(\omega_0) = \exp(-\beta H(\omega_0))/Z$ in eq. (6.5), with $Z = \int \mathrm{d}\omega_0\, e^{-\beta H(\omega_0)}$ being the partition function. The Liouville equation determines the time evolution of an observable B_t

$$\dot{B}_t = \mathcal{L} B_t, \tag{6.70a}$$

$$\mathcal{L} = \sum_{n=1}^{N} \left(\frac{\partial H}{\partial \mathbf{p}_n} \cdot \frac{\partial}{\partial \mathbf{r}_n} - \frac{\partial H}{\partial \mathbf{r}_n} \cdot \frac{\partial}{\partial \mathbf{p}_n} \right), \tag{6.70b}$$

where $\mathcal{L}$ is the Liouville operator. The general solution of the Liouville equation in eq. (6.70) is given by $B_t = e^{t\mathcal{L}} B_0$. Remind you that the Liouville operator $\mathcal{L}$ is anti self-adjoint with respect to the inner product in eq. (6.5), i.e.,

$$\langle \mathcal{L} B_t, C_{t'} \rangle = -\langle B_t, \mathcal{L} C_{t'} \rangle. \tag{6.71}$$

Using projection operators, the Liouville equation eq. (6.70a) can be decomposed into two terms. To derive the nfGLE in chapter 3.3, we used two projection operators. One of the two projection operators is used to decompose the Liouville equation in the following way

$$\dot{B}_t = \mathcal{L} B_t = e^{t\mathcal{L}} \mathcal{L} B_0 = e^{t\mathcal{L}} (\mathcal{P}_L + \mathcal{Q}_L) \mathcal{L} B_0. \tag{6.72}$$

We use the subscript $\mathcal{L}$ to highlight that the pair $\mathcal{P}_L$, $\mathcal{Q}_L$ is used to decompose the Liouville equation.
Using the Dyson decomposition [34, 38] of the propagation operator $e^{t\mathcal{L}}$, we can further decompose eq. (6.72) with a second projection operator. The Dyson decomposition is given by [36]

$$e^{t\mathcal{L}} = e^{t\mathcal{Q}_D \mathcal{L}} + \int_0^t \mathrm{d}s\, e^{(t-s)\mathcal{L}} \mathcal{P}_D \mathcal{L} e^{s\mathcal{Q}_D \mathcal{L}}. \tag{6.73}$$

The subscript D denotes that we use the pair $\mathcal{P}_D$, $\mathcal{Q}_D$, which in general can be different from $\mathcal{P}_L$, $\mathcal{Q}_L$, in the Dyson decomposition. Inserting eq. (6.73) into eq. (6.72), we obtain an equation with a similar structure to a GLE in terms of general projection operators $\mathcal{P}_L, \mathcal{Q}_L, \mathcal{P}_D, \mathcal{Q}_D$:

$$\dot{B}_t = e^{t\mathcal{L}} \mathcal{P}_L \mathcal{L} B_0 + \int_0^t \mathrm{d}s\, e^{(t-s)\mathcal{L}} \mathcal{P}_D \mathcal{L} F_{\mathrm{R}}(s) + F_{\mathrm{R}}(t), \tag{6.74a}$$

$$F_{\mathrm{R}}(t) = e^{t\mathcal{Q}_D \mathcal{L}} \mathcal{Q}_L \mathcal{L} B_0. \tag{6.74b}$$

Again, note that the projection $\mathcal{P}_L$ determines the functional form of the first term $e^{t\mathcal{L}}\mathcal{P}_L\mathcal{L}B_0$ and the initial value $F_{\mathrm{R}}(0) = \mathcal{Q}_L\mathcal{L}B_0$, while the projection $\mathcal{P}_D$ determines the functional form of the integrand $\mathcal{P}_D\mathcal{L}F_{\mathrm{R}}(s)$ and the propagation of $F_{\mathrm{R}}(0)$ in time via $e^{t\mathcal{Q}_D L}$.

Next, we specify the projections $\mathcal{P}_L$ and $\mathcal{P}_D$ in order to obtain the explicit form of the nfGLE in chapter 3.3. For this, we need the conditional correlation between two observables B and C, which is defined by

$$\langle B_t, C_{t'}\rangle_{A_0} = \frac{\langle \delta(A(\hat{\omega}_0) - A(\omega_0)), B(\hat{\omega}_0, t)C(\hat{\omega}_0, t')\rangle}{\langle \delta(A(\hat{\omega}_0) - A(\omega_0))\rangle}. \tag{6.75}$$

In eq. (6.75), the phase space position with a hat, i.e., $\hat{\omega}_0$, is integrated over. The condition in eq. (6.75) is that the observable A has initially the value $A_0 = A(\omega_0)$. We refer to the observable A, which we project onto, as the observable of interest, or reaction coordinate. As before, we assume that A is a function of particle positions only, i.e., $A(\omega_0) = A(\mathbf{R}_0) = A_0$. For example, A could be the center of mass of a cluster of particles, the mean native distance in a polymer or the dihedral angle. When A is a function of positions only, its velocity $\dot{A}_0 = (M^{-1}\mathbf{P}_0)\cdot\nabla_R A_0$ is linear in particle momenta, where $\hat{\mathrm{M}}$ is the diagonal mass matrix $M_{ij} = m_i\delta_{ij}$ in eq. (6.4). From this, and from eq. (6.4) and eq. (6.5), it follows that, for a given position $\mathbf{R}_0$, the velocity $\dot{A}_0$ is Gaussian distributed with zero mean.

Using conditional correlations, we define the first projection operator $\mathcal{P}_D$. It is a reformulation of the projection operators discussed in ref. [4, 110] and given by

$$\mathcal{P}_D B_t = \langle B_t\rangle_{A_0} + \frac{\langle \dot{A}_0, B_t\rangle_{A_0}}{\langle \dot{A}_0^2\rangle_{A_0}}\dot{A}_0\,. \tag{6.76}$$

The projection operator $\mathcal{P}_L$ is an extension of $\mathcal{P}_D$ and reads

$$\mathcal{P}_L = \mathcal{P}_D + \mathcal{P}_2, \tag{6.77a}$$

$$\mathcal{P}_2 B_t = \frac{\left\langle \left(\dot{A}_0^2 - \langle \dot{A}_0^2\rangle_{A_0}\right), B_t\right\rangle_{A_0}}{\left\langle \left(\dot{A}_0^2 - \langle \dot{A}_0^2\rangle_{A_0}\right)^2\right\rangle_{A_0}}\left(\dot{A}_0^2 - \langle \dot{A}_0^2\rangle_{A_0}\right). \tag{6.77b}$$

Both projections, $\mathcal{P}_L$ and $\mathcal{P}_D$, are orthogonal projections w.r.t. the inner product in eq. (6.5). Further, we have $\mathcal{Q}_L B_t = \mathcal{Q}_D\mathcal{Q}_L B_t$. To obtain the final form of the GLE, we set $B_t = \dot{A}_t$ in eq. (6.74) and use the projection operators defined in eq. (6.76) and eq. (6.77). This gives the GLE (see chapter 3.3)

$$\ddot{A}_t = -\frac{1}{M(A_t)}\frac{\mathrm{d}U(A_t, \dot{A}_t)}{\mathrm{d}A_t} - \int_0^t \mathrm{d}s\,\Gamma(A_s, t-s)\dot{A}_s + F_{\mathrm{R}}(t), \tag{6.78a}$$

with

$$U(A, \dot{A}) = U_{\mathrm{PMF}}(A) + \frac{M(A)}{2}\dot{A}^2 + k_B T \ln\sqrt{M(A)}, \tag{6.78b}$$

$$\Gamma(A, t) = \beta U'_{\mathrm{eff}}(A)D(A, t) - D'(A, t) + \frac{\langle \ddot{A}_0, F_{\mathrm{R}}(t)\rangle_A}{\langle \dot{A}_0^2\rangle_A}, \tag{6.78c}$$

$$U_{\mathrm{eff}}(A, t) = U_{\mathrm{PMF}}(A) + k_B T \ln M(A), \tag{6.78d}$$

$$D(A, t) = \frac{\langle \dot{A}_0^2, F_{\mathrm{R}}(t)\rangle_A}{\langle \dot{A}_0^2\rangle_A}. \tag{6.78e}$$

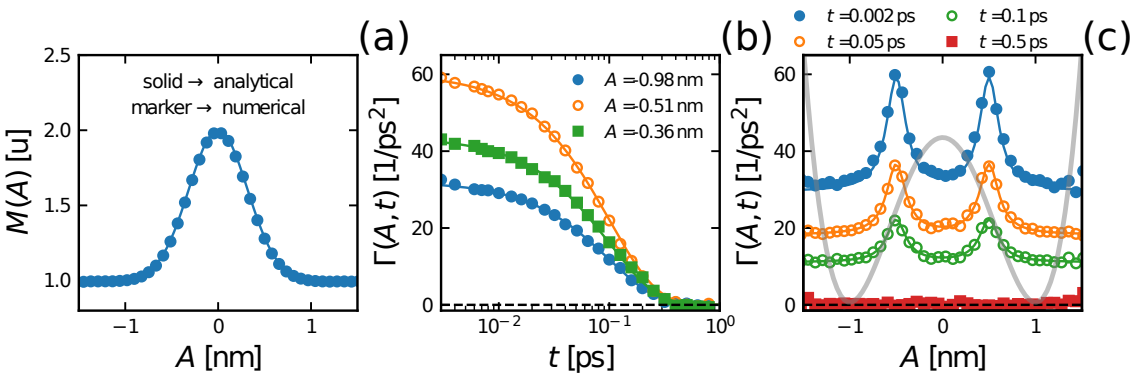

Figure 6.3: Comparison of the numerically extracted GLE parameters (markers) with the input functions (solid lines). In **(a)**, we show the position dependent mass $M(A)$, where we compare the numerically computed conditional average in eq. (6.84b) (markers) with the input function in eq. (6.95) (solid line). In **(b)** and **(c)**, we compare the numerically extracted memory kernel defined in eq. (6.78c) (markers) with the input memory kernel given in eq. (6.88c) and eq. (6.96) (solid lines). In **(b)**, $\Gamma(A,t)$ is shown as a function of time at different positions A, and in **(c)**, as a function of the position A at different times t. We show the shape of the PMF $U(A)$ in the background as a thick gray line. The underlying trajectory was generated via simulations of eq. (6.79) with U given in eq. (6.94) and $\gamma_{21}(A)$ given by eq. (6.96). The memory function $\Gamma(A,t)$ is extracted using eq. (6.12) (cf. ref. [4]).

Here, $M(A) = k_B T / \langle \dot{A}_0^2 \rangle_A$ denotes the position-dependent effective mass of the reaction coordinate and $U_{\mathrm{PMF}}(a) = -k_B T \ln \mathbb{P}(a)$ is the potential of mean force (PMF) with $\mathbb{P}(a) = \langle \delta(A(\hat{\omega}_0) - a) \rangle$ being the positional distribution.
The GLE parameters can be extracted from time series data by explicitly computing the random force $F_{\mathrm{R}}(t)$ using the numerical extraction discussed in chapter 4 and represented by eq. (6.12) [4, 16].

In the following, we present a system of Markovian Langevin equations that allows for a self-consistent Markovian embedding of a GLE with a configuration-dependent mass and nonlinear friction. Self-consistency is discussed in section 6.1.

The Markovian Langevin equations are given by

$$\dot{x}_t = v_t, \tag{6.79a}$$

$$\dot{v}_t = -\frac{1}{M(x_t)} \left(U'(x_t) + \frac{k_B T}{2} \frac{M'(x_t)}{M(x_t)} + \frac{M'(x_t)}{2} v_t^2 - \sum_{n=1}^{N+1} \gamma_{1n}\, \nu_n(t) + \sum_{n=1}^{2N} \sigma_{1n}(x_t)\, \eta_n(t) \right), \tag{6.79b}$$

$$\dot{u}_n(t) = \frac{1}{m_n} \left(- \sum_{j=1}^{N+1} \gamma_{n+1,j}(x_t)\, \nu_j(t) + \sum_{j=1}^{2N} \sigma_{n+1,j}(x_t)\, \eta_j(t) \right), \tag{6.79c}$$

$$\mathbb{E}[\eta_i(t)] = 0, \quad \mathbb{E}[\eta_i(t)\eta_j(t')] = \delta_{ij}\delta(t-t'), \tag{6.79d}$$

In eq. (6.79), we introduce a velocity vector $\boldsymbol{\nu}(t) = (v_t, u_1(t), u_2(t), \ldots, u_N(t))^T \in \mathbb{R}^{N+1}$ and the white noise vector $\boldsymbol{\eta}(t) \in \mathbb{R}^{2N}$. The $u_n(t)$ denote auxiliary velocity variables. The friction matrix $\hat{\boldsymbol{\gamma}} \in \mathbb{R}^{(N+1)\times(N+1)}$ is only non-zero in the first row, the first column and in the diagonal, i.e., it is given by

$$\begin{aligned}\hat{\boldsymbol{\gamma}}(x_t) &= \begin{pmatrix} \gamma_{11}(x_t) & \gamma_{12}(x_t) & \gamma_{13}(x_t) & \cdots \\ \gamma_{21}(x_t) & \gamma_{22} & 0 & \\ \gamma_{31}(x_t) & 0 & \gamma_{33} & \\ \vdots & & & \ddots \end{pmatrix} \\ &= \begin{pmatrix} M(x_t)\sum_{n=1}^{N} g_n & M(x_t)h_1 & M(x_t)h_2 & \cdots \\ \gamma_{21}(x_t) & \gamma_{22} & 0 & \\ \gamma_{31}(x_t) & 0 & \gamma_{33} & \\ \vdots & & & \ddots \end{pmatrix}.\end{aligned} \tag{6.80}$$

The multiplicative noise matrix $\hat{\boldsymbol{\sigma}}(x_t) \in \mathbb{R}^{(N+1)\times 2N}$ is given by

$$\hat{\boldsymbol{\sigma}}(x_t) = \begin{pmatrix} s_{11}(x_t) & 0 & s_{21}(x_t) & 0 & s_{31}(x_t) & 0 & \cdots \\ s_{12}(x_t) & s_{13}(x_t) & 0 & 0 & 0 & 0 & \\ 0 & 0 & s_{22}(x_t) & s_{23}(x_t) & 0 & 0 & \\ 0 & 0 & 0 & 0 & s_{32}(x_t) & s_{33}(x_t) & \cdots \\ \vdots & & & & & \vdots & \end{pmatrix}. \tag{6.81}$$

If the system in eq. (6.79) fulfills the following fluctuation-dissipation relation [71]

$$\left(\hat{\boldsymbol{\gamma}}(x_t) + \hat{\boldsymbol{\gamma}}^T(x_t)\right)/\beta = \hat{\boldsymbol{\sigma}}(x_t) \cdot \hat{\boldsymbol{\sigma}}^T(x_t), \tag{6.82}$$

it has the stationary probability distribution

$$\rho_{\mathrm{st}}(x, v, \mathbf{u}) \propto \sqrt{M(x)} e^{-\beta U(x)} e^{-\beta \frac{M(x)}{2} v^2} e^{-\beta \sum_n \frac{m_n}{2} (u_n)^2}, \tag{6.83}$$

with the desired properties

$$U(x) = -k_B T \ln \rho_{\mathrm{st}}(x), \tag{6.84a}$$

$$\langle v^2 \rangle_x = \frac{k_B T}{M(x)}, \tag{6.84b}$$

where $\rho_{\mathrm{st}}(x) = \int \mathrm{d}v \int \mathrm{d}^N u \, \rho_{\mathrm{st}}(x, v, \mathbf{u})$ is the marginal distribution in x. The fluctuation-dissipation theorem in eq. (6.82) holds when

$$2\frac{g_n M(x_t)}{\beta} = s_{n1}^2(x_t), \tag{6.85a}$$

$$\frac{h_n M(x_t) + \gamma_{n+1,1}}{\beta} = s_{n1}(x_t) s_{n2}(x_t), \tag{6.85b}$$

$$2\frac{\gamma_{n+1,n+1}}{\beta} = s_{n2}^2(x_t) + s_{n3}^2(x_t), \tag{6.85c}$$

and the entries s_{ij} of the multiplicative noise matrix $\hat{\boldsymbol{\sigma}}(x_t)$ are given by (see section 6.2.1)

$$s_{n1}(x_t) = \sqrt{2k_BTg_nM(x_t)}, \tag{6.86a}$$

$$s_{n2}(x_t) = k_BT\frac{h_nM(x_t) + \gamma_{n+1,1}}{s_{n1}(x_t)}, \tag{6.86b}$$

$$s_{n3}(x_t) = \sqrt{2k_BT\gamma_{n+1,n+1} - s_{n2}^2(x_t)}, \tag{6.86c}$$

where g_n and h_n are defined in eq. (6.80). In order to obtain finite, real valued multiplicative noise factors, eq. (6.86) requires the following inequalities to hold

$$g_n > 0, \tag{6.87a}$$

$$4\gamma_{n+1,n+1}\, g_nM(x_t) \geq \left(h_nM(x_t) + \gamma_{n+1,1}\right)^2, \tag{6.87b}$$

for all $n = 1, 2, \ldots, N$. Eq. (6.87) puts a lower bound on each g_n for given $M(x_t)$, h_n, $\gamma_{n+1,1}(x_t)$ and $\gamma_{n+1,n+1}$. Consequently, the presence of a delta contribution in the memory kernel is required. Solving eq. (6.79c) for $u_n(t)$ and inserting the result into eq. (6.79b) yields the GLE

$$\dot{v}_t = -\frac{1}{M(x_t)}\frac{\mathrm{d}\tilde{U}(x_t, v_t)}{\mathrm{d}x_t} - \int_0^t \mathrm{d}s\, \Gamma(x_s, t-s)v_s + \xi(t), \tag{6.88a}$$

with

$$\tilde{U}(x, v) = U(x) + \frac{M(x)}{2}v^2 + k_BT\ln\sqrt{M(x)}, \tag{6.88b}$$

$$\Gamma(x_s, t-s) = \sum_{n=1}^{N}\left(g_n\delta(t-s) - h_ne^{-\frac{\gamma_{n+1,n+1}}{m_n}(t-s)}\frac{\gamma_{n+1,1}(x_s)}{m_n}\right), \tag{6.88c}$$

$$\xi(t) = \sum_{n=1}^{N}\left(-\frac{h_n}{m_n}e^{-\frac{\gamma_{n+1,n+1}}{m_n}t}u_n(0) + \sum_{j=1}^{2N}\left[\sigma_{nj}(x_t)\eta_j(t) - \int_0^t \mathrm{d}s\,\frac{h_n}{m_n}e^{-\frac{\gamma_{n+1,n+1}}{m_n}(t-s)}\sigma_{n+1,j}(x_s)\eta_j(s)\right]\right). \tag{6.88d}$$

To show that the system of Markovian Langevin equations in eq. (6.79) can be used to perform a self-consistent Markovian embdedding of the GLE in eq. (6.78) in combination with the extraction scheme in eq. (6.12), we consider the mean dynamics of v_t in eq. (6.79).

The mean dynamics is obtained from Ito's lemma in eq. (6.35) which can be used to derive an equation for the differential $\mathrm{d}f(\boldsymbol{\zeta}_t)$ of a function $f(\boldsymbol{\zeta}_t)$ of state variables $\boldsymbol{\zeta}_t$, e.g., the chain rule for an Ito diffusion process [42]. Averaging the differential equation for $\mathrm{d}f(\boldsymbol{\zeta}_t)$ over the noise provides the mean dynamics. By setting $f(\boldsymbol{\zeta}_t) = v_t$, one obtains the mean dynamics of v_t in eq. (6.79)

$$\dot{v}_t = L_{\mathrm{FP}}^{\dagger}v_t, \tag{6.89}$$

where $L^\dagger_{\mathrm{FP}}$ is the infinitesimal generator of the system of stochastic differential equations in eq. (6.79) (cf. [71])

$$\begin{aligned} L^\dagger_{\mathrm{FP}} = v\frac{\partial}{\partial x} - \left(\frac{1}{M(x)}\frac{\mathrm{d}\tilde{U}(x,v)}{\mathrm{d}x} + \sum_{n=1}^{N}\left(g_n v + h_n u_n\right)\right)\frac{\partial}{\partial v} \\ - \sum_{n=1}^{N}\frac{1}{m_n}\left(\gamma_{n+1,1}(x)v + \gamma_{n+1,n+1}u_n\right)\frac{\partial}{\partial u_n} \\ + \frac{1}{2}\sum_{i,j=1}^{N+1}\left(\hat{\boldsymbol{\sigma}}(x)\cdot\hat{\boldsymbol{\sigma}}^T(x)\right)_{ij}\frac{\partial^2}{\partial\nu_i\partial\nu_j}. \end{aligned} \tag{6.90}$$

The L_2-adjoint of $L^\dagger_{\mathrm{FP}}$ is the Fokker-Planck operator, i.e., $L_{\mathrm{FP}}\rho_{st}(x,v,\mathbf{u}) = 0$, where the stationary distribution is given in eq. (6.83).

Starting from eq. (6.89), we can derive the GLE using the projection operators $\mathcal{P}_D$ and $\mathcal{P}_L$ in eq. (6.76) and eq. (6.77), respectively. The inner product in eq. (6.5) is obtained by replacing ρ_{eq} by the stationary distribution in eq. (6.83).

First, we calculate the Markovian contribution generated by the projection. According to eq. (6.74) and the adjoint Fokker-Planck operator $L^\dagger_{\mathrm{FP}}$ in eq. (6.90), the Markovian contribution is given by

$$e^{tL^\dagger_{\mathrm{FP}}}\mathcal{P}_L L^\dagger_{\mathrm{FP}} v = -\frac{1}{M(x)}\frac{\mathrm{d}\tilde{U}(x_t,v_t)}{\mathrm{d}x_t} - \sum_{n=1}^{N} g_n v_t. \tag{6.91}$$

Next, we compute the random force term using eq. (6.74b)

$$\begin{aligned} F_{\mathrm{R}}(t) &= e^{t\mathcal{Q}_D L^\dagger_{\mathrm{FP}}}\mathcal{Q}_L L^\dagger_{\mathrm{FP}} v = \sum_{n=0}^{\infty}\frac{t^n}{n!}(\mathcal{Q}_D L^\dagger_{\mathrm{FP}})^n \mathcal{Q}_L L^\dagger_{\mathrm{FP}} v \\ &= -\sum_{n=1}^{N} h_n e^{-\frac{\gamma_{n+1,n+1}}{m_n}t} u_n(0), \end{aligned} \tag{6.92}$$

from which we obtain

$$\int_0^t \mathrm{d}s\, e^{(t-s)\mathcal{L}^\dagger_{\mathrm{FP}}}\mathcal{P}_D L^\dagger_{\mathrm{FP}} F_{\mathrm{R}}(s) = -\int_0^t \mathrm{d}s \sum_{n=1}^{N} h_n e^{-\frac{\gamma_{n+1,n+1}}{m_n}t}\frac{\gamma_{n+1,1}(x_{t-s})}{m_n} v_{t-s}. \tag{6.93}$$

The results in eq. (6.91) and eq. (6.93) coincide with the deterministic part of the GLE in eq. (6.88). Note that eq. (6.92) reproduces the noise term in eq. (6.88d) averaged over the noises η_j. This means that we recover the GLE in eq. (6.88), when averaged over the white noise $\boldsymbol{\eta}(t)$. Thus, the memory kernel $\Gamma(A,t)$, obtained from the Markovian embedding in eq. (6.79) by applying the projection operators in eq. (6.8) and eq. (6.77), can be computed in closed-form and is equal to the one in eq. (6.88c), which was obtained by solving the Langevin equation in eq. (6.79c).

We numerically demonstrate the Markovian embedding of a nonlinear GLE with position dependent mass in fig. 6.3. Here, we take $k_B T = 2.494\,\mathrm{kJ/mol}$, i.e., we use molecular

dynamics (MD) units. Based on eq. (6.79) and the relations in eq. (6.86), we generate a trajectory x_t in a double well potential

$$U(x) = U_0(x^2 - 1)^2, \tag{6.94}$$

with $U_0 = 2\,k_BT$. The position dependent mass is taken to be

$$M(x) = M_0(1 + e^{-5x^2}), \tag{6.95}$$

with $M_0 = 1\,\text{u}$, which is shown as a solid line in fig. 6.3(a). We simulate $N = 2$ auxiliary variables with equal masses $m_n = 1\,\text{u}$. The friction constants are set to $g_n = 1\,\text{ps}^{-1}$, $h_n = -5\,\text{ps}^{-1}$ and $\gamma_{n+1,n+1} = 10\,\text{u/ps}$ for $n = 1, 2$ and for the friction profile, we take

$$\gamma_{n+1,1}(x) = \frac{\gamma_0}{1 + (\frac{x - x_n}{l_n})^2}, \tag{6.96}$$

with $\gamma_0 = 3\,\text{u/ps}$, $l_n = 0.125\,\text{nm}$ and the values $x_n = \pm 0.5\,\text{nm}$ for the two auxiliary variables. In other words, we simulate the case where the friction kernel increases at the inflection points of the potential. Such a friction behavior is motivated by our numerical results for the memory function of the dihedral angle in butane from fully atomistic MD simulations [4].
In fig. 6.3, we compare the input functions given in eq. (6.95) and eq. (6.88c) (solid lines) with the numerically extracted ones (markers). The position dependent mass is numerically computed from the simulated trajectories using eq. (6.84b), the friction kernel is numerically computed using the extraction scheme in eq. (6.12). In fig. 6.3, we observe perfect agreement between the input and the extracted functions. This constitutes a numerical validation of the embedding method and the extraction/simulation techniques used by us.

6.4 Modeling an arbitrary memory kernel

In applications, when a general $\Gamma(A, t)$ is given, there are several ways to obtain the parameter values for the embedding system, e.g., eq. (6.13). The most straightforward way is to expand $\Gamma(A, t)$ in a countable set of basis functions $\{f_n(A)\}$,.i.e.,

$$\Gamma(A, t) \approx \sum_{n=1}^{N} c_n(t) f_n(A). \tag{6.97}$$

The coefficients are computed from the inner product $c_n(t) = (\Gamma(t), f_n)$ which depends on the choice of basis functions $f_n(A)$. Then, the coefficients, which are functions of time only, are fitted by a sum of exponentials.
As an example, take the case in which it is possible to accurately capture $\Gamma(A, t)$ with 10 functions $f_n(A)$. To fit each coefficient $c_n(t)$, one uses, let us say, a sum of three exponentials. This means that one needs in total 30 auxiliary variables in the embedding system.
A way to reduce the number of functions f_n in the expansion is to look for dominating shapes in $\Gamma(A, t)$ when plotted as a function of A at different times. For example, if one

observes a nearly constant A-profile with peaks only at certain positions, it makes more sense to use fit functions $f_n(A)$ with this property, e.g., Gaussians. This way, the main features of $\Gamma(A,t)$ can be captured more efficiently, i.e., with a smaller number of auxiliary variables in the embedding. To obtain the coefficients $c_n(t)$ for this case, one could do a least-squares fit at each time step.

We turn to the perturbative embedding system in sec. 6.2.2. As can be seen from the discussion above, the fitting procedure for a memory kernel of two variables A and t can become time consuming. Since the embedding in section 6.2.2 is not exact and we do not know how the correction terms scale in the parameters κ_n, γ_n and $\alpha_n(x)$ in eq. (6.64), we propose a simple numerical test to check if eq. (6.62) holds:

Approximate the memory kernel $\Gamma(A,t)$ as a product of two functions, i.e.,

$$\Gamma(A,t) \approx c(t) f(A), \tag{6.98a}$$

$$f(A) = \frac{1}{\tau}\int_0^\tau \mathrm{d}s\, \Gamma(A,s), \tag{6.98b}$$

$$c(t) = \langle \Gamma(A,t) \rangle. \tag{6.98c}$$

Eq. (6.98b) is a time average of the memory kernel where τ denotes the largest memory time in Γ, i.e., the time after which Γ has decayed to less than 36% (exponential decay) of its initial amplitude at all positions A. Eq. (6.98c) is an ensemble average over all positions A. If the embedding in eq. (6.64) with eq. (6.67) works for the single product approximation $c(t)f(A)$, then we expect it to work for the real Γ too. Once this is ensured, one can start with a more involved fitting of Γ.

Another approach is to write

$$\Gamma(A,t) = \bar{\Gamma}(t) + \Big(\Gamma(A,t) - \bar{\Gamma}(t)\Big) \tag{6.99a}$$

$$= \bar{\Gamma}(t) + \Delta\Gamma(A,t), \tag{6.99b}$$

$$\max_t(\bar{\Gamma}(t)) = \bar{\Gamma}(0), \tag{6.99c}$$

$$\max_t(\Delta\Gamma(A,t)) = \Delta\Gamma(A,0), \tag{6.99d}$$

and determine $\bar{\Gamma}(t)$, if the data allows it, such that

$$\bar{\Gamma}(0) \gg \Delta\Gamma(A,0) \tag{6.100}$$

for all A. This is preferable since the memory kernel $\bar{\Gamma}(t)$ can be simulated with standard Markovian embedding methods [103] and the nonlinear friction kernel $\Delta\Gamma(A,t)$ is treated as a small correction term where the approximation in eq. (6.67) may hold.

6.5 Conclusion

In this chapter, we presented three Markovian Langevin systems whose mean dynamics are shown to be equivalent to nfGLEs generated by given projection operators.
The first and third systems are coupled via velocities, and the embedding works for any combination of parameters, though only in the presence of a delta contribution in the memory kernel. Here, the time component of the memory kernels consists of a sum of pure exponentials. The second system also works without a delta contribution, but it is applicable only for certain combinations of parameters since it is a perturbation expansion. The time component of this system is a sum of exponentially decaying oscillations. We propose a simple numerical test to check the validity of the approximations.
While the effective mass in the first two systems is assumed to be constant in the reaction coordinate, in the third system, we allow a position-dependent mass. We introduce a method to simulate a generalized Langevin equation with position-dependent mass and friction functions by Markovian embedding. Such GLEs were numerically extracted in chapter 4.2 (cf. [4]) from molecular dynamics simulation trajectories of the dihedral angle of butane in water. We introduce a projection operator for the Markovian embedding that allows for a self-consistent extraction/simulation procedure. The methods presented in this chapter will be useful for simulating general non-Markovian systems. In future work, it will be interesting to study the relative effects of position-dependent mass and memory on the kinetics of such systems.

7 References

[1] **Jose LF Abascal and Carlos Vega**. "A general purpose model for the condensed phases of water: TIP4P/2005". In: *The Journal of chemical physics* 123.23 (2005), p. 234505.

[2] **Mark James Abraham, Teemu Murtola, Roland Schulz, Szilárd Páll, Jeremy C Smith, Berk Hess, and Erik Lindahl**. "GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers". In: *SoftwareX* 1 (2015), pp. 19–25.

[3] **Berni J Alder and Thomas Everett Wainwright**. "Studies in molecular dynamics. I. General method". In: *The Journal of Chemical Physics* 31.2 (1959), pp. 459–466.

[4] **Cihan Ayaz, Laura Scalfi, Benjamin A. Dalton, and Roland R. Netz**. "Generalized Langevin equation with a nonlinear potential of mean force and nonlinear memory friction from a hybrid projection scheme". In: *Physical Review E* 105.5 (May 2022), p. 054138. DOI: 10.1103/PhysRevE.105.054138. URL: https://link.aps.org/doi/10.1103/PhysRevE.105.054138 (visited on 06/06/2022).

[5] **Cihan Ayaz, Lucas Tepper, Florian N Brünig, Julian Kappler, Jan O Daldrop, and Roland R Netz**. "Non-Markovian modeling of protein folding". In: *Proceedings of the National Academy of Sciences* 118.31 (2021), e2023856118.

[6] **June Barrow-Green**. *Poincaré and the three body problem*. 11. American Mathematical Soc., 1997.

[7] **HJC Berendsen, JR Grigera, and TP Straatsma**. "The missing term in effective pair potentials". In: *Journal of Physical Chemistry* 91.24 (1987), pp. 6269–6271.

[8] **Robert B Best and Gerhard Hummer**. "Coordinate-dependent diffusion in protein folding". In: *Proceedings of the National Academy of Sciences* 107.3 (2010), pp. 1088–1093.

[9] **Robert B Best and Gerhard Hummer**. "Reaction coordinates and rates from transition paths". In: *Proceedings of the National Academy of Sciences* 102.19 (2005), pp. 6732–6737.

[10] **Ludwig Boltzmann**. *Studien über das Gleichgewicht der lebendigen Kraft zwischen bewegten materiellen Punkten: vorgelegt in der Sitzung am 8. October 1868*. K. und k. Hof-und Staatsdr., 1868.

[11] **Robert Brown**. "XXVII. A brief account of microscopical observations made in the months of June, July and August 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies". In: *The philosophical magazine* 4.21 (1828), pp. 161–173.

[12] **Florian N Brünig, Otto Geburtig, Alexander von Canal, Julian Kappler, and Roland R Netz**. "Time-dependent friction effects on vibrational infrared frequencies and line shapes of liquid water". In: *The Journal of Physical Chemistry B* 126.7 (2022), pp. 1579–1589.

[13] **Joseph D Bryngelson, Jose Nelson Onuchic, Nicholas D Socci, and Peter G Wolynes**. "Funnels, pathways, and the energy landscape of protein folding: a synthesis". In: *Proteins: Structure, Function, and Bioinformatics* 21.3 (1995), pp. 167–195.

[14] **Giovanni Bussi, Davide Donadio, and Michele Parrinello**. "Canonical sampling through velocity rescaling". In: *The Journal of chemical physics* 126.1 (2007), p. 014101.

[15] **Shane Carlson, Florian N Brünig, Philip Loche, Douwe Jan Bonthuis, and Roland R Netz**. "Exploring the absorption spectrum of simulated water from MHz to infrared". In: *The Journal of Physical Chemistry A* 124.27 (2020), pp. 5599–5605.

[16] **Antoine Carof, Rodolphe Vuilleumier, and Benjamin Rotenberg**. "Two algorithms to compute projected correlation functions in molecular dynamics simulations". In: *The Journal of chemical physics* 140.12 (2014), p. 124103.

[17] **Michele Ceriotti, Giovanni Bussi, and Michele Parrinello**. "Langevin Equation with Colored Noise for Constant-Temperature Molecular Dynamics Simulations". In: *Physical Review Letters* 102.2 (Jan. 2009), p. 020601. DOI: 10.1103/PhysRevLett.102.020601. URL: https://link.aps.org/doi/10.1103/PhysRevLett.102.020601 (visited on 07/20/2022).

[18] **Jorge Chahine, Ronaldo J Oliveira, Vitor BP Leite, and Jin Wang**. "Configuration-dependent diffusion can shift the kinetic transition state and barrier height of protein folding". In: *Proceedings of the National Academy of Sciences* 104.37 (2007), pp. 14646–14651.

[19] **David Chandler**. "Statistical mechanics of isomerization dynamics in liquids and the transition state approximation". In: *The Journal of Chemical Physics* 68.6 (1978), pp. 2959–2970.

[20] **John D Chodera, Nina Singhal, Vijay S Pande, Ken A Dill, and William C Swope**. "Automatic discovery of metastable states for the construction of Markov models of macromolecular conformational dynamics". In: *The Journal of chemical physics* 126.15 (2007), 04B616.

[21] **Alexandre J Chorin, Ole H Hald, and Raz Kupferman**. "Optimal prediction and the Mori–Zwanzig representation of irreversible processes". In: *Proceedings of the National Academy of Sciences* 97.7 (2000), pp. 2968–2973.

[22] **Hoi Sung Chung and William A Eaton**. "Protein folding transition path times from single molecule FRET". In: *Current opinion in structural biology* 48 (2018), pp. 30–39.

[23] **Hoi Sung Chung and William A Eaton**. "Single-molecule fluorescence probes dynamics of barrier crossing". In: *Nature* 502.7473 (2013), pp. 685–688.

[24] **Hoi Sung Chung, Stefano Piana-Agostinetti, David E Shaw, and William A Eaton**. "Structural origin of slow diffusion in protein folding". In: *Science* 349.6255 (2015), pp. 1504–1510.

[25] **Rudolf Clausius**. "Über die Art der Bewegung, welche wir Wärme nennen". In: *Annalen der Physik* 176.3 (1857), pp. 353–380.

[26] **Jan O Daldrop, Julian Kappler, Florian N Brünig, and Roland R Netz**. "Butane dihedral angle dynamics in water is dominated by internal friction". In: *Proceedings of the National Academy of Sciences* 115.20 (2018), pp. 5169–5174.

[27] **Benjamin A Dalton, Cihan Ayaz, Lucas Tepper, and Roland R Netz**. "Protein folding is governed by memory-dependent friction". In: *arXiv preprint arXiv:2208.13842* (2022).

[28] **Tom Darden, Darrin York, and Lee Pedersen**. "Particle mesh Ewald: An N log (N) method for Ewald sums in large systems". In: *The Journal of chemical physics* 98.12 (1993), pp. 10089–10092.

[29] **Eric Darve and Andrew Pohorille**. "Calculating free energies using average force". In: *The Journal of chemical physics* 115.20 (2001), pp. 9169–9183.

[30] **Eric Darve, Jose Solomon, and Amirali Kia**. "Computing generalized Langevin equations and generalized Fokker–Planck equations". In: *Proceedings of the National Academy of Sciences* 106.27 (2009), pp. 10884–10889.

[31] **Ken A Dill and Justin L MacCallum**. "The protein-folding problem, 50 years on". In: *science* 338.6110 (2012), pp. 1042–1046.

[32] **Timo J Doerries, Sarah AM Loos, and Sabine HL Klapp**. "Correlation functions of non-Markovian systems out of equilibrium: Analytical expressions beyond single-exponential memory". In: *Journal of Statistical Mechanics: Theory and Experiment* 2021.3 (2021), p. 033202.

[33] **Yong Duan, Chun Wu, Shibasish Chowdhury, Mathew C Lee, Guoming Xiong, Wei Zhang, Rong Yang, Piotr Cieplak, Ray Luo, Taisung Lee, et al.** "A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations". In: *Journal of computational chemistry* 24.16 (2003), pp. 1999–2012.

[34] **Freeman J Dyson**. "The radiation theories of Tomonaga, Schwinger, and Feynman". In: *Physical Review* 75.3 (1949), p. 486.

[35] **Albert Einstein**. "Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen". In: *Annalen der physik* 4 (1905).

[36] **DJ Evans and GP Morriss**. *Statistical Mechanics of Nonequilibrium Liquids (Academic, London, 1990)*.

[37] **Enrico Fermi, P Pasta, Stanislaw Ulam, and Mary Tsingou**. *Studies of the nonlinear problems*. Tech. rep. Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 1955.

[38] **Richard P Feynman**. "An operator calculus having applications in quantum electrodynamics". In: *Physical Review* 84.1 (1951), p. 108.

[39] **Henrik Flyvbjerg**. "Error estimates on averages of correlated data". In: *Advances in computer simulation*. Springer, 1998, pp. 88–103.

[40] **Nicolas Foloppe and Alexander D MacKerell Jr**. "All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data". In: *Journal of computational chemistry* 21.2 (2000), pp. 86–104.

[41] **Daan Frenkel, Berend Smit, and Mark A Ratner**. *Understanding molecular simulation: from algorithms to applications*. Vol. 2. Academic press San Diego, 1996.

[42] **CW Gardiner**. "Handbook of stochastic methods for physics, chemistry and the natural sciences. Third. Vol. 13". In: *Springer Series in Synergetics. Berlin: Springer-Verlag* (2004), p. 13.

[43] **Fabian Glatzel and Tanja Schilling**. "The interplay between memory and potentials of mean force: A discussion on the structure of equations of motion for coarse-grained observables". In: *Europhysics Letters* 136.3 (2022), p. 36001.

[44] **Leonardo Rydin Gorjão and Francisco Meirinhos**. "kramersmoyal: Kramers–Moyal coefficients for stochastic processes". In: *arXiv preprint arXiv:1912.09737* (2019).

[45] **Hermann Grabert**. *Projection operator techniques in nonequilibrium statistical mechanics*. Vol. 95. Springer, 2006.

[46] **Hermann Grabert, Peter Hänggi, and Peter Talkner**. "Microdynamics and nonlinear stochastic processes of gross variables". In: *Journal of Statistical Physics* 22.5 (1980), pp. 537–552.

[47] **Richard F Grote and James T Hynes**. "Reactive modes in condensed phase reactions". In: *The Journal of Chemical Physics* 74.8 (1981), pp. 4465–4475.

[48] **Richard F Grote and James T Hynes**. "The stable states picture of chemical reactions. II. Rate constants for condensed and gas phase reaction models". In: *The Journal of Chemical Physics* 73.6 (1980), pp. 2715–2732.

[49] **Peter Hanggi and Fatemeh Mojtabai**. "Thermally activated escape rate in presence of long-time memory". In: *Physical Review A* 26.2 (1982), p. 1168.

[50] **Marcus Hanwell and Geoffrey Hutchison**. "AVOGADRO: free, open source, cross-platform computer program for building molecules and visualizing structure". In: *APS March Meeting Abstracts*. 2009, K1–006.

[51] **GD Harp and BJ Berne**. "Time-correlation functions, memory functions, and molecular dynamics". In: *Physical Review A* 2.3 (1970), p. 975.

[52] **Rainer Hegger and Gerhard Stock**. "Multidimensional Langevin modeling of biomolecular dynamics". In: *The Journal of chemical physics* 130.3 (2009), p. 034106.

[53] **Carmen Hijón, Pep Español, Eric Vanden-Eijnden, and Rafael Delgado-Buscalioni**. "Mori–Zwanzig formalism as a practical computational tool". In: *Faraday discussions* 144 (2010), pp. 301–322.

[54] **Michael Hinczewski, Yann von Hansen, Joachim Dzubiella, and Roland R Netz**. "How the diffusivity profile reduces the arbitrariness of protein folding free energies". In: *The Journal of chemical physics* 132.24 (2010), 06B615.

[55] **Illia Horenko, Carsten Hartmann, Christof Schütte, and Frank Noe**. "Data-based parameter estimation of generalized multidimensional Langevin processes". In: *Physical Review E* 76.1 (2007), p. 016706.

[56] **James T Hynes**. "Chemical reaction dynamics in solution". In: *Annual Review of Physical Chemistry* 36.1 (1985), pp. 573–597.

[57] **Gouri S Jas, William A Eaton, and James Hofrichter**. "Effect of viscosity on the kinetics of α-helix and β-hairpin formation". In: *The Journal of Physical Chemistry B* 105.1 (2001), pp. 261–272.

[58] **Gerhard Jung, Martin Hanke, and Friederike Schmid**. "Generalized Langevin dynamics: construction and numerical integration of non-Markovian particle-based models". In: *Soft matter* 14.46 (2018), pp. 9368–9382.

[59] **Gerhard Jung, Martin Hanke, and Friederike Schmid**. "Iterative Reconstruction of Memory Kernels". In: *Journal of Chemical Theory and Computation* 13.6 (June 2017), pp. 2481–2488. ISSN: 1549-9618. DOI: 10.1021/acs.jctc.7b00274. URL: https://doi.org/10.1021/acs.jctc.7b00274 (visited on 07/19/2022).

[60] **Nicolaas G van Kampen**. "Remarks on non-Markov processes". In: *Brazilian Journal of Physics* 28 (1998), pp. 90–96.

[61] **Julian Kappler, Jan O Daldrop, Florian N Brünig, Moritz D Boehle, and Roland R Netz**. "Memory-induced acceleration and slowdown of barrier crossing". In: *The Journal of Chemical Physics* 148.1 (2018), p. 014903.

[62] **Julian Kappler, Victor B Hinrichsen, and Roland R Netz**. "Non-Markovian barrier crossing with two-time-scale memory is dominated by the faster memory component". In: *The European Physical Journal E* 42.9 (2019), pp. 1–16.

[63] **Tomoyuki Kinjo and Shi-aki Hyodo**. "Equation of motion for coarse-grained simulation based on microscopic description". In: *Physical review E* 75.5 (2007), p. 051109.

[64] **Viktor Klippenstein, Madhusmita Tripathy, Gerhard Jung, Friederike Schmid, and Nico FA van der Vegt**. "Introducing memory in coarse-grained molecular simulations". In: *The Journal of Physical Chemistry B* 125.19 (2021), pp. 4931–4954.

[65] **Viktor Klippenstein and Nico FA van der Vegt**. "Cross-correlation corrected friction in (generalized) Langevin models". In: *The Journal of Chemical Physics* 154.19 (2021), p. 191102.

[66] **Bartosz Kowalik, Jan O Daldrop, Julian Kappler, Julius CF Schulz, Alexander Schlaich, and Roland R Netz**. "Memory-kernel extraction for different molecular solutes in solvents of varying viscosity in confinement". In: *Physical Review E* 100.1 (2019), p. 012126.

[67] **Oliver F Lange and Helmut Grubmüller**. "Collective Langevin dynamics of conformational motions in proteins". In: *The Journal of chemical physics* 124.21 (2006), p. 214903.

[68] **Paul Langevin**. "Sur la théorie du mouvement brownien". In: *Compt. Rendus* 146 (1908), pp. 530–533.

[69] **Laura Lavacchi, Julian Kappler, and Roland R Netz**. "Barrier crossing in the presence of multi-exponential memory functions with unequal friction amplitudes and memory times". In: *EPL (Europhysics Letters)* 131.4 (2020), p. 40004.

[70] **Hee Sun Lee, Surl-Hee Ahn, and Eric F Darve**. "The multi-dimensional generalized Langevin equation for conformational motion of proteins". In: *The Journal of chemical physics* 150.17 (2019), p. 174113.

[71] **Benedict Leimkuhler and Matthias Sachs**. "Ergodic properties of quasi-Markovian generalized Langevin equations with configuration dependent noise and non-conservative force". In: *International workshop on Stochastic Dynamics out of Equilibrium*. Springer. 2017, pp. 282–330.

[72] **Dominika Lesnicki, Rodolphe Vuilleumier, Antoine Carof, and Benjamin Rotenberg**. "Molecular hydrodynamics from memory kernels". In: *Physical review letters* 116.14 (2016), p. 147804.

[73] **Yaakov Levy and José N Onuchic**. "Water mediation in protein folding and molecular recognition". In: *Annual review of biophysics and biomolecular structure* 35.1 (2006), pp. 389–415.

[74] **Benjamin Lickert and Gerhard Stock**. "Modeling non-Markovian data using Markov state and Langevin models". In: *The Journal of Chemical Physics* 153.24 (2020), p. 244112.

[75] **Benjamin Lickert, Steffen Wolf, and Gerhard Stock**. "Data-driven Langevin modeling of nonequilibrium processes". In: *The Journal of Physical Chemistry B* 125.29 (2021), pp. 8125–8136.

[76] **Kresten Lindorff-Larsen, Stefano Piana, Ron O Dror, and David E Shaw**. "How fast-folding proteins fold". In: *Science* 334.6055 (2011), pp. 517–520.

[77] **Sarah AM Loos and Sabine HL Klapp**. "Fokker–Planck equations for time-delayed systems via Markovian embedding". In: *Journal of Statistical Physics* 177.1 (2019), pp. 95–118.

[78] **James Clerk Maxwell**. "II. Illustrations of the dynamical theory of gases". In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 20.130 (1860), pp. 21–37.

[79] **James Clerk Maxwell**. "V. Illustrations of the dynamical theory of gases.—Part I. On the motions and collisions of perfectly elastic spheres". In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 19.124 (1860), pp. 19–32.

[80] **Ralf Metzler, Jae-Hyung Jeon, Andrey G Cherstvy, and Eli Barkai**. "Anomalous diffusion models and their properties: non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking". In: *Physical Chemistry Chemical Physics* 16.44 (2014), pp. 24128–24164.

[81] **Hugues Meyer, Philipp Pelagejcev, and Tanja Schilling**. "Non-Markovian out-of-equilibrium dynamics: A general numerical procedure to construct time-dependent memory kernels for coarse-grained observables". In: *EPL (Europhysics Letters)* 128.4 (2020), p. 40001.

[82] **Hugues Meyer, Thomas Voigtmann, and Tanja Schilling**. "On the dynamics of reaction coordinates in classical, time-dependent, many-body processes". In: *The Journal of chemical physics* 150.17 (2019), p. 174118.

[83] **Hugues Meyer, Thomas Voigtmann, and Tanja Schilling**. "On the non-stationary generalized Langevin equation". In: *The Journal of chemical physics* 147.21 (2017), p. 214110.

[84] **GK Mikhailov**. "Daniel bernoulli, hydrodynamica (1738)". In: *Landmark Writings in Western Mathematics 1640-1940*. Elsevier, 2005, pp. 131–142.

[85] **RC Miller and P Kusch**. "Velocity distributions in Potassium and Thallium atomic beams". In: *Physical Review* 99.4 (1955), p. 1314.

[86] **Bernhard G Mitterwallner, Laura Lavacchi, and Roland R Netz**. "Negative friction memory induces persistent motion". In: *The European Physical Journal E* 43.10 (2020), pp. 1–11.

[87] **Bernhard G Mitterwallner, Christoph Schreiber, Jan O Daldrop, Joachim O Rädler, and Roland R Netz**. "Non-Markovian data-driven modeling of single-cell motility". In: *Physical Review E* 101.3 (2020), p. 032408.

[88] **Hazime Mori**. "Transport, collective motion, and Brownian motion". In: *Progress of theoretical physics* 33.3 (1965), pp. 423–455.

[89] **Sadao Nakajima**. "On quantum theory of transport phenomena: steady diffusion". In: *Progress of Theoretical Physics* 20.6 (1958), pp. 948–959.

[90] **Roland R Netz**. "Approach to equilibrium and nonequilibrium stationary distributions of interacting many-particle systems that are coupled to different heat baths". In: *Physical Review E* 101.2 (2020), p. 022120.

[91] **Krishna Neupane, Daniel AN Foster, Derek R Dee, Hao Yu, Feng Wang, and Michael T Woodside**. "Direct observation of transition paths during the folding of proteins and nucleic acids". In: *Science* 352.6282 (2016), pp. 239–242.

[92] **Krishna Neupane, Dustin B Ritchie, Hao Yu, Daniel AN Foster, Feng Wang, and Michael T Woodside**. "Transition path times for nucleic acid folding determined from energy-landscape analysis of single-molecule trajectories". In: *Physical review letters* 109.6 (2012), p. 068102.

[93] **Frank Noé, Illia Horenko, Christof Schütte, and Jeremy C Smith**. "Hierarchical analysis of conformational dynamics in biomolecules: Transition networks of metastable states". In: *The Journal of chemical physics* 126.15 (2007), 04B617.

[94] **Chris Oostenbrink, Alessandra Villa, Alan E Mark, and Wilfred F Van Gunsteren**. "A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6". In: *Journal of computational chemistry* 25.13 (2004), pp. 1656–1676.

[95] **J Perrin**. "Molecular agitation and the brownian movement". In: *CRAS* 146 (1908), pp. 967–970.

[96] **Eli Pollak, Hermann Grabert, and Peter Hänggi**. "Theory of activated rate processes for arbitrary frequency dependent friction: Solution of the turnover problem". In: *The Journal of chemical physics* 91.7 (1989), pp. 4073–4087.

[97] **Sander Pronk, Szilárd Páll, Roland Schulz, Per Larsson, Pär Bjelkmar, Rossen Apostolov, Michael R Shirts, Jeremy C Smith, Peter M Kasson, David Van Der Spoel, et al.** "GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit". In: *Bioinformatics* 29.7 (2013), pp. 845–854.

[98] **Hannes Risken**. "Fokker-planck equation". In: *The Fokker-Planck Equation*. Springer, 1996, pp. 63–95.

[99] **Jean-Paul Ryckaert, Giovanni Ciccotti, and Herman JC Berendsen**. "Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes". In: *Journal of computational physics* 23.3 (1977), pp. 327–341.

[100] **Xavier Saint Raymond**. *Elementary introduction to the theory of pseudodifferential operators*. Routledge, 2018.

[101] **Rohit Satija, Atanu Das, and Dmitrii E Makarov**. "Transition path times reveal memory effects and anomalous diffusion in the dynamics of protein folding". In: *The Journal of chemical physics* 147.15 (2017), p. 152707.

[102] **Rohit Satija and Dmitrii E Makarov**. "Generalized Langevin equation as a model for barrier crossing dynamics in biomolecular folding". In: *The Journal of Physical Chemistry B* 123.4 (2019), pp. 802–810.

[103] **Peter Siegle, Igor Goychuk, Peter Talkner, and Peter Hänggi**. "Markovian embedding of non-Markovian superdiffusion". In: *Physical Review E* 81.1 (2010), p. 011136.

[104] **Attal Stéphane, Alain Joye, and Claude-Alain Pillet**. *Open Quantum Systems: 2. The Markovian Approach*. 2006.

[105] **George Gabriel Stokes et al.** "On the effect of the internal friction of fluids on the motion of pendulums". In: (1851).

[106] **John E Straub, Michal Borkovec, and Bruce J Berne**. "Calculation of dynamic friction on intramolecular degrees of freedom". In: *Journal of Physical Chemistry* 91.19 (1987), pp. 4995–4998.

[107] **NG Van Kampen and I Oppenheim**. "Brownian motion as a problem of eliminating fast variables". In: *Physica A: Statistical Mechanics and its Applications* 138.1-2 (1986), pp. 231–248.

[108] **Nicolaas Godfried Van Kampen**. "Elimination of fast variables". In: *Physics Reports* 124.2 (1985), pp. 69–160.

[109] **Nicolaas Godfried Van Kampen**. *Stochastic processes in physics and chemistry*. Vol. 1. Elsevier, 1992.

[110] **Hadrien Vroylandt and Pierre Monmarché**. “Position-dependent memory kernel in generalized Langevin equations: Theory and numerical estimation”. In: *The Journal of Chemical Physics* 156.24 (June 2022), p. 244105. ISSN: 0021-9606. DOI: 10.1063/5.0094566. URL: https://aip.scitation.org/doi/full/10.1063/5.0094566 (visited on 07/20/2022).

[111] **George H Weiss**. “First passage time problems in chemical physics”. In: *Advances in Chemical Physics* 13 (1967), pp. 1–18.

[112] **Junrong Zheng, Kyungwon Kwak, Jia Xie, and Michael D Fayer**. “Ultrafast carbon-carbon single-bond rotational isomerization in room-temperature solution”. In: *Science* 313.5795 (2006), pp. 1951–1955.

[113] **Robert Zwanzig**. “Memory effects in irreversible thermodynamics”. In: *Physical Review* 124.4 (1961), p. 983.

[114] **Robert Zwanzig**. “Nonlinear generalized Langevin equations”. In: *Journal of Statistical Physics* 9.3 (1973), pp. 215–220.

[115] **Robert. Zwanzig**. *Nonequilibrium statistical mechanics [electronic resource] / Robert Zwanzig.* Oxford ; Oxford University Press, 2001. ISBN: 1-280-53110-X.