

Generalized Langevin Equation with a Non-Linear Potential of Mean Force and Non-Linear Memory Friction From a Hybrid Projection Scheme

Cihan Ayaz, Benjamin A Dalton, and Roland R. Netz
Fachbereich Physik, Freie Universität Berlin
(Dated: January 28, 2022)

We introduce a hybrid projection scheme that combines linear Mori projection and conditional Zwanzig projection techniques and use it to derive a Generalized Langevin Equation (GLE) for a general interacting many-body system. The resulting GLE includes i) explicitly the potential of mean force (PMF) that describes the equilibrium distribution of the system in the chosen space of reaction coordinates, ii) a random force term that explicitly depends on the initial state of the system, and iii) a memory friction contribution that splits into two parts: a part that is linear in the past reaction-coordinate velocity and a part that is in general non-linear in the past reaction coordinates but does not depend on velocities. Our hybrid scheme thus combines all desirable properties of the Zwanzig and Mori projection schemes. The non-linear memory friction contribution is shown to be related to correlations between the reaction-coordinate velocity and the random force. We present a numerical method to compute all parameters of our GLE, in particular the non-linear memory friction function and the random force distribution, from a trajectory in reaction coordinate space. We apply our method on the dihedral-angle dynamics of a butane molecule in water obtained from atomistic molecular dynamics simulations. For this example, we demonstrate that non-linear memory friction is present and that the random force exhibits significant non-Gaussian corrections. We also present the derivation of the GLE for multidimensional reaction coordinates that are general functions of all positions in the phase space of the underlying many-body system; this corresponds to a systematic coarse-graining procedure that preserves not only the correct equilibrium behavior but also the correct dynamics of the coarse-grained system.

I. INTRODUCTION

Most interesting physical systems are interacting many-body systems. When dealing with the kinetics of such systems, one is typically interested in the dynamics of a low-dimensional reaction coordinate, which is, however, generally influenced by the entire system [1]. Examples include the motion of a particle in a liquid [2–7], vibrational modes of a molecule in the gas phase or in a liquid [8–11], chemical or associative reactions between molecules [12–16] and protein folding [17–20]. To predict the dynamics of the reaction coordinate, one in principle has to solve the equation of motion of the underlying many-body system, which is typically analytically impossible and is only numerically possible for small systems and over short times. The very attractive idea of coarse-grained modeling is to replace the description of the full many-body system by a description in terms of the reaction coordinates only. The challenge is to derive the appropriate equation of motion that describes the dynamics of the reaction coordinates accurately while maintaining numerical efficiency. For some biologically relevant scenarios, such as the folding of a protein, sufficiently long simulations of the full system dynamics can be performed [21–23], but even for these cases, the interpretation of the results typically requires mapping onto a low-dimensional reaction coordinate.

Rigorous coarse-graining methods based on projection operator techniques were introduced by Zwanzig and Mori, which are directly applied to the Liouville equation that describes the dynamics of a classical many-body system governed by a time-independent Hamiltonian [24, 25]

(in fact, a similar approach applicable to quantum systems was developed by Nakajima even earlier [26]). The result of the projection is a coarse-grained equation of motion for the chosen set of reaction coordinates, the so-called generalized Langevin equation (GLE). It contains three distinct terms: a force term due to a potential that depends on the reaction coordinates, a memory friction contribution that involves the past time dynamics of the reaction coordinates, and a force that explicitly depends on the initial state of the entire many-body system and which is typically interpreted as a random or stochastic force. The GLE is therefore an integro-differential stochastic equation. It should be noted that the Zwanzig and Mori projection schemes give rise to fundamentally different GLEs for non-linear systems, which are both rigorous and reproduce the system dynamics described by the reaction coordinates exactly [24, 25]. However, except a few notable exceptions [27–32], the exact Zwanzig or Mori equation have rarely been used in practice for non-trivial, i.e. non-linear, systems, for different reasons: In the Mori framework, the force from the potential as well as the memory friction are linear in the reaction coordinate and their velocities, respectively, and therefore all non-linearities are accounted for by the random force, which thus becomes non-Gaussian and is difficult to parameterize; in the Zwanzig framework, the potential term in the GLE is in general non-linear and corresponds to the potential of mean force (PMF), which ensures the correct equilibrium distribution of the reaction coordinates [28], which is a desired property. On the other hand, the memory friction is a general function of both the reaction coordinates and their velocities, which poses

severe problems when estimating such a function from simulation or experimental data.

As a consequence, many previous works considered a simplified form of the GLE, which in this paper we refer to as the approximate GLE. It contains the non-linear PMF and a memory friction that is linear in the velocity of the reaction coordinate [6, 33–40]. In principle, this approximate GLE follows from the Zwanzig GLE, assuming that the friction memory depends only linearly on the past reaction coordinates and is independent of the reaction-coordinate velocities. The validity of this approximation can typically not be checked in a systematic manner. The applications of the approximate GLE range from non-Markovian rate theory [41–43], over protein folding dynamics [17–20] to molecular diffusion and conformational dynamics [6, 7, 37, 44]. Methods to derive memory functions from trajectory data for non-linear systems within the framework of the approximate GLE have been introduced and it was demonstrated that the resulting GLE correctly describes the multi-scale fractal dynamics of protein folding [20] and the vibrational spectra of molecules in non-linear bond-length and bond-angle potentials [45]. Although widely used, the validity of the approximate GLE in the presence of a non-linear potential is subject to ongoing discussions [46, 47].

In this paper, we introduce a projection method that is a hybrid of the Zwanzig and Mori projection schemes. As an advantage over the Mori projection scheme, the resulting GLE contains the force stemming from the generally non-linear PMF, which by itself guides the system into the correct equilibrium distribution in the long-time limit. As an advantage over the Zwanzig projection scheme, the generally non-linear memory friction does not depend on the velocity of the reaction coordinate but only on the reaction coordinate itself, which significantly simplifies the numerical estimation of the memory function from trajectory data. We develop the necessary framework to compute all parameters of the resulting GLE from trajectories of a reaction coordinate. Thus, we present data-based methods i) to derive the non-linear memory friction from simulation or experimental trajectories, ii) to thereby examine the validity of the approximate GLE, and iii) to study the distribution and correlation of the random force from trajectories. We also derive a multidimensional GLE in terms of a general set of reaction coordinates that are arbitrary functions of the positions of the underlying many-body system; this constitutes a rigorous derivation of the equations of motion that accurately describe the equilibrium and dynamic behavior of coarse-grained systems. For the explicit example of the dihedral-angle dynamics of a butane molecule in water, obtained from atomistic molecular dynamics simulations, we demonstrate that non-linear memory friction is present and that the random force exhibits significant non-Gaussian corrections. Therefore, we find that even for this simple molecular system, the approximate GLE, which neglects non-linear memory friction and assumes Gaussian random forces, does not correctly describe the

dynamics.

The paper is organized as follows: First, we introduce the Hamiltonian of the many-body system, as well as our notation, and we present important expressions for correlation functions and conditional averages. We then review the Mori and Zwanzig projection schemes and highlight practical problems of the resulting GLEs. After this, we introduce our hybrid projection scheme and derive the GLE that features a non-linear PMF and non-linear memory friction. In the subsequent section, we introduce an algorithm to extract all parameters of our GLE from trajectories. In the final section, we apply our formalism on two exactly solvable model systems and on MD trajectories for the dihedral angle dynamics of a butane molecule in water.

II. HAMILTONIAN MODEL, NOTATION AND USEFUL PROPERTIES

We denote the phase space of a system of N interacting particles in three-dimensional space by Ω . One specific microstate, i.e., a point in Ω , is denoted by $\omega = (\mathbf{R}, \mathbf{P}) = (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, \mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N)$ which is a $6N$ vector of the Cartesian positions $\mathbf{r}_i = (r_i^x, r_i^y, r_i^z)$, and the conjugate momenta $\mathbf{p}_i = (p_i^x, p_i^y, p_i^z)$ of all $i = 1, 2, \dots, N$ particles in the system. The Hamiltonian of the system is an invariant of motion and splits into a kinetic and a potential part

$$H(\omega) = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m_i} + V(\mathbf{R}). \quad (1)$$

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

$$\dot{\omega}_t = L\omega_t, \quad (2)$$

where ω_t is the location of the system in phase space at time t and $\dot{\omega}_t$ denotes the corresponding velocity, given the system was initially at ω_0 . For the sake of compact notation, we denote time dependencies of phase space coordinates by a subscript. In eq. (2), L is the Liouville operator given by

$$L = \sum_{n=1}^N \left(\frac{\mathbf{p}_n}{m_n} \cdot \nabla_{\mathbf{r}_n} - (\nabla_{\mathbf{r}_n} V(\mathbf{R})) \cdot \nabla_{\mathbf{p}_n} \right). \quad (3)$$

All of the operators that we consider in this work, including the Liouville operator L , act on the initial phase space position ω_0 . From eq. (2), it follows that the system is propagated in time by the operator e^{tL} , i.e., $e^{tL}\omega_0 = \omega_t$. We consider observables that are real-valued functions of phase-space coordinates only and that depend on time

implicitly via the time dependence of a trajectory moving in phase space. For the sake of notational brevity, we also denote the time dependency of observables by a subscript too, i.e., $A_t \equiv A(\omega_t) = A(\omega_0, t)$. Using the chain rule for differentiation, it follows that the time evolution of an observable A_t is also governed by the Liouville equation [48]

$$\dot{A}_t = LA_t, \quad (4)$$

where \dot{A}_t denotes the time derivative of A_t . Thus, the time propagation operator of an observable in the initial state $A(\omega_0) \equiv A_0$, is also given by e^{tL} . From this, it follows that

$$A(\omega_{t+t'}) = e^{(t+t')L}A(\omega_0) = e^{tL}A(\omega_{t'}) = A(\omega_{t'}, t). \quad (5)$$

Eq. (5) describes how observables are propagated in time by e^{tL} and will be used throughout our derivations. All observables are elements of a Hilbert space, i.e., a vector space equipped with an inner product. Let A and B denote two system observables. For the inner product, we choose

$$\langle A_t, B_{t'} \rangle \equiv \int_{\Omega} d\omega_0 \rho_{\text{eq}}(\omega_0) A(\omega_0, t) B(\omega_0, t'), \quad (6)$$

where $\rho_{\text{eq}}(\omega_0) = e^{-\beta H(\omega_0)}/Z$ is the canonical Boltzmann distribution with the inverse thermal energy $\beta = 1/k_B T$ and the partition function $Z = \int_{\Omega} d\omega_0 e^{-\beta H(\omega_0)}$. The inner product in eq. (6) thus corresponds to an equilibrium time correlation function which establishes the link to statistical mechanics. The average of a single observable B_t is given by $\langle B_t \rangle \equiv \langle B_t, 1 \rangle$ and does not depend on time. Because of the form of the Hamiltonian in eq. (1), the Boltzmann distribution factorizes into a position and a momentum-dependent part

$$\rho_{\text{eq}}(\omega_0) = \frac{1}{Z} e^{-\beta H(\omega_0)} = \rho_{\text{kin}}(\mathbf{P}_0) \rho_{\text{pot}}(\mathbf{R}_0), \quad (7)$$

where $\rho_{\text{kin}}(\mathbf{P}_0) \propto \prod_{i=1}^N \exp(-\beta \mathbf{p}_{i,0}^2/2m_i)$ is a Gaussian with zero mean. With respect to the inner product in eq. (6), the Liouville operator, as defined in eq. (3), is anti-self-adjoint [48]

$$\langle LA_t, B_{t'} \rangle = -\langle A_t, LB_{t'} \rangle. \quad (8)$$

A. Conditional Averages

In addition to time-correlation functions calculated over the entire phase space Ω , as in eq. (6), we will also use conditional time-correlation functions that result from averages over a hyper surface in phase space on which an observable of choice at the initial time $t = 0$, $A_0 = A(\hat{\omega}_0)$, takes a constant value $A(\omega_s)$. A conditional correlation of two observables $B_t = B(\omega_t) = B(\omega_0, t)$ and $C_{t'} = C(\omega_{t'}) = C(\omega_0, t')$ is defined by [27, 28]

$$\langle B_t, C_{t'} \rangle_{A_s} = \frac{\langle \delta[A(\hat{\omega}_0) - A(\omega_s)], B(\hat{\omega}_0, t) C(\hat{\omega}_0, t') \rangle}{\langle \delta[A(\hat{\omega}_0) - A(\omega_s)] \rangle}. \quad (9)$$

In eq. (9), the phase space variable with a hat, $\hat{\omega}_0$, is integrated over. The phase space variable ω_s is not. Therefore, $\langle B_t, C_{t'} \rangle_{A_s}$ is a function of ω_s , and the times t and t' . The conditional average of a single observable B_t is given by $\langle B_t \rangle_{A_s} \equiv \langle B_t, 1 \rangle_{A_s}$.

Finally, we give a few relations which will be frequently used later on. We repeat that a conditional average is a function of phase space via the conditional function $A_s = A(\omega_s)$ in eq. (9). The time propagation of a conditional average is thus given by

$$e^{tL} \langle B_{t'} \rangle_{A_0} = \langle B_{t'} \rangle_{A_t}. \quad (10)$$

The normalized probability that an observable A_t has the value a is given by $\mathbb{P}(a) \equiv \langle \delta(A_t - a) \rangle$, from which the potential of mean force (PMF) for an observable follows as [33]

$$U_{\text{PMF}}(a) \equiv -k_B T \ln \mathbb{P}(a). \quad (11)$$

Acting with the Liouville operator on a delta function gives [30]

$$L\delta(A_t - a) = -\dot{A}_t \frac{d}{da} \delta(A_t - a). \quad (12)$$

Using the definition in eq. (9) together with the relations in eq. (8), eq. (12) and the PMF defined in eq. (11), we derive in appendix A the important relation [30]

$$\langle LB_{t'} \rangle_{A_t} = \frac{d}{dA_t} \langle \dot{A}_0, B_{t'} \rangle_{A_t} - \beta \langle \dot{A}_0, B_{t'} \rangle_{A_t} \frac{d}{dA_t} U_{\text{PMF}}(A_t). \quad (13)$$

III. PROJECTION OPERATOR METHOD

We now derive the equation of motion for an arbitrary scalar observable A_t , which can of course also be the position of a single particle [48]. The derivation for a general multi-dimensional observable is given in appendix B. A projection P is a linear, idempotent operator, i.e., for arbitrary scalars c_1, c_2 , it fulfills the properties $P(c_1 A_t + c_2 B_t) = c_1 P A_t + c_2 P B_t$ and $P^2 = P$. The operator $Q = 1 - P$ projects onto the *complementary* subspace with 1 being the identity operator. For brevity, we will refer to the subspace onto which P projects as the *relevant* subspace. The operators P and Q can be used to decompose the Liouville equation $\dot{A}_t = L A_t$ for the observable \dot{A}_t as

$$\dot{A}_t = e^{tL} (P + Q) L \dot{A}_0 = e^{tL} P L \dot{A}_0 + e^{tL} Q L \dot{A}_0. \quad (14)$$

To obtain an equation of motion for A_t from eq. (14), we introduce the operator

$$\Phi(t) = e^{tL} Q. \quad (15)$$

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

$$\frac{d}{dt} e^{tL} Q = e^{tL} L Q = e^{tL} Q L Q + e^{tL} P L Q, \quad (16a)$$

$$\dot{\Phi}(t) = \Phi(t) L Q + e^{tL} P L Q. \quad (16b)$$

Eq. (16b) is an inhomogenous differential equation of first order. Using $\Phi(0) = Q$, as follows from eq. (15), the solution reads

$$\Phi(t) = Qe^{tLQ} + \int_0^t du e^{uL} PLQe^{(t-u)LQ}. \quad (17)$$

By using $Qe^{tLQ} = e^{tQL}Q$ and the substitution $s = t - u$ in eq. (17), we find

$$\Phi(t) = e^{tL}Q = e^{tQL}Q + \int_0^t ds e^{(t-s)L} PL e^{sQL}Q. \quad (18)$$

Since the operator $e^{tQL}Q$ exhibits a Q operator on the left side when the exponential is expanded, the first term on the r.h.s. of eq. (18) stays in the complementary subspace for all times. The second term describes the effect of the complementary subspace on the relevant subspace. By factoring out the operator Q on the r.h.s. of eq. (18), one obtains the *Dyson decomposition* [49–51] of the propagator e^{tL}

$$e^{tL} = e^{tQL} + \int_0^t ds e^{(t-s)L} PL e^{sQL}. \quad (19)$$

Replacing $e^{tL}Q$ in eq. (14) by eq. (18) leads to the GLE for A_t in terms of a general projection P [24, 25, 48]

$$\ddot{A}_t = e^{tL} PL \dot{A}_0 + \int_0^t ds e^{(t-s)L} PL F^R(s) + F^R(t), \quad (20a)$$

$$F^R(t) \equiv e^{tQL} QL \dot{A}_0 = Qe^{tLQ} L \dot{A}_0. \quad (20b)$$

The function $F^R(t)$ stays in the complementary subspace for all times and is an explicit function of the initial state of the entire system, i.e., $F^R(t) = F^R(\omega_0, t)$. Hence, for large systems, it can be interpreted as a random or stochastic function. For the sake of brevity, we will write out the ω_0 dependence of $F^R(t)$ only when it improves clarity. The first term on the r.h.s. of eq. (20a) represents the time evolution of the part of $\ddot{A}_0 = L \dot{A}_0$ which lies in the relevant subspace and reflects a deterministic force. The second term on the r.h.s. of eq. (20a) is due to the relevant part of $LF^R(\omega_0, t)$ and describes dissipative effects. Clearly, the explicit form of eq. (20a) depends on the specific form of the projection operator P . Before we introduce our hybrid projection scheme, we will present the GLE's generated by the Mori projection P_M and by the Zwanzig projection P_Z .

A. Mori Projection

The Mori projection applied on an observable A_t is given by [25]

$$P_M A_t = \frac{\langle A_t, B_0 \rangle}{\langle B_0^2 \rangle} B_0 + \frac{\langle A_t, \dot{B}_0 \rangle}{\langle \dot{B}_0^2 \rangle} \dot{B}_0, \quad (21)$$

and uses the inner product defined in eq. (6). The observables one projects onto, i.e., B_0 and \dot{B}_0 , are referred to as the *projection functions*. The projection in eq. (21) maps any observable A_t onto the subspace of all functions linear in the observables B_0 and \dot{B}_0 . In addition to being linear and idempotent, P_M is self-adjoint w.r.t. to the inner product in eq. (6), i.e., for two arbitrary observables $A_t, C_{t'}$, the relation

$$\langle P_M A_t, C_{t'} \rangle = \langle A_t, P_M C_{t'} \rangle \quad (22)$$

holds. Thus, it is an orthogonal projection, since all functions $P_M A_t$ and $P_M C_{t'}$ are orthogonal, i.e.,

$$\langle P_M A_t, P_M C_{t'} \rangle = 0, \quad (23)$$

as follows directly from eq. (22) and from the idempotence of P . For $P = P_M$ and choosing the projection functions to be $B_t = A_t$ and $\dot{B}_t = \dot{A}_t$, i.e., projecting onto the observable of interest itself, eq. (20a) takes the form [25, 48]

$$\ddot{A}_t = -K A_t - \int_0^t ds \Gamma^M(s) \dot{A}_{t-s} + F^R(\omega_0, t), \quad (24a)$$

$$K = \frac{\langle \dot{A}_0^2 \rangle}{\langle A_0^2 \rangle}, \quad \Gamma^M(t) = \frac{\langle F^R(\omega_0, t), F^R(\omega_0, 0) \rangle}{\langle \dot{A}_0^2 \rangle}, \quad (24b)$$

where $\Gamma^M(s)$ is the memory friction kernel obtained from the Mori projection. Eq. (24) is an exact decomposition of the Liouville equation into three terms: the first term is a generalized force due to a potential of quadratic form; the second term accounts for linear friction and includes the memory kernel $\Gamma^M(s)$, which is related via eq. (24b) to the second moment of the random force $F^R(\omega_0, t)$, defined in eq. (20b). The exact form of the memory function can only be computed for very simple models, for realistic systems and practical applications it is infeasible to compute since the fluctuating term $F^R(\omega_0, t)$ is an explicit function of the initial state of the entire system. Instead, one typically models the function $F^R(t)$ as a stochastic process with zero mean and a second moment given in eq. (24b). Although information on higher-order moments of $F^R(t)$ can be obtained from the Mori formalism, $F^R(t)$ is typically assumed to be Gaussian. In general, however, this assumption can not hold, since $F^R(t)$ contains all non-linearities that A_t may exhibit. Thus, imposing $F^R(t)$ to be a Gaussian variable becomes a bad approximation for non-linear systems, which reflects a fundamental short-coming of the Mori projection scheme for practical applications.

B. Zwanzig Projection

Contrary to the Mori projection, the Zwanzig projection P_Z of an observable A_t is non-linear in the projection

functions B_0 and \dot{B}_0 [24]

$$P_Z A_t = \frac{\langle \delta(B(\hat{\omega}_0) - B(\omega_0)) \delta(\dot{B}(\hat{\omega}_0) - \dot{B}(\omega_0)), A(\hat{\omega}_0, t) \rangle}{\langle \delta(B(\hat{\omega}_0) - B(\omega_0)) \rangle \langle \delta(\dot{B}(\hat{\omega}_0) - \dot{B}(\omega_0)) \rangle} \\ = \langle A_t \rangle_{B_0, \dot{B}_0}, \quad (25)$$

where we repeat that phase-space variables with a hat inside inner products, i.e., $\hat{\omega}_0$, are integrated over. The Zwanzig projection thus is a conditional average as defined in eq. (9) and is linear, idempotent and self-adjoint, similar to the Mori projection. The resulting GLE from the Zwanzig projection is best illustrated by choosing the observable of interest to be the momentum of a single particle, $\dot{A}_0 = \mathbf{p}_0$, and the projection functions as the position and the linear momentum of the same particle, i.e., $B_0 \rightarrow \mathbf{r}_0$, $\dot{B}_0 \rightarrow \mathbf{p}_0$. With this, eq. (20a) becomes [52]

$$\dot{\mathbf{p}}_t = -\nabla_{\mathbf{r}_t} U_{\text{PMF}}(\mathbf{r}_t) + \mathbf{F}^R(\omega_0, t) \\ + \int_0^t ds \left[\left(\frac{\nabla_{\mathbf{p}_s}}{\beta} - \frac{\mathbf{p}_s}{m} \right) \right]^T \cdot \Gamma^Z(t-s, \mathbf{r}_s, \mathbf{p}_s) \quad (26a)$$

with a memory friction kernel defined by

$$\beta \Gamma_{ij}^Z(t-s, \mathbf{r}_s, \mathbf{p}_s) = \langle F_i^R(0), F_j^R(t-s) \rangle_{\mathbf{r}_s, \mathbf{p}_s}. \quad (27)$$

Here, $U_{\text{PMF}}(\mathbf{r}) = -k_B T \ln \langle \delta(\mathbf{r}_0 - \mathbf{r}) \rangle$ denotes the potential of mean force (PMF) defined in eq. (11), which creates in the GLE a force on the particle that tends to establish the equilibrium positional distribution. This is the main advantage over the Mori projection, since this ensures the correct equilibrium behavior once we switch to a stochastic description and replace the fluctuating force $F^R(t)$ by a Gaussian stochastic variable with zero mean [28]. The memory friction kernel Γ^Z is a 3×3 matrix that, as a result of the conditional average, is a function of particle position \mathbf{r}_s and particle momentum \mathbf{p}_s . This is the main drawback of the GLE in eq. (26a), since the position and momentum dependence is difficult to deal with in applications. As a way out, one typically invokes the ad-hoc assumption that the memory function is independent of position and momentum, i.e., $\Gamma^Z(t-s, \mathbf{r}_s, \mathbf{p}_s) \approx \Gamma^{\text{app}}(t-s)$. This assumption leads to an approximate GLE that is amply used in literature [7, 17–20, 33, 34, 36–43, 45] and reads

$$\dot{\mathbf{p}}_t = -\nabla_{\mathbf{r}_t} U_{\text{PMF}}(\mathbf{r}_t) - \int_0^t ds \Gamma^{\text{app}}(t-s) \cdot \frac{\mathbf{p}_s}{m} + \mathbf{F}^R(t). \quad (28)$$

While for various applications the approximate GLE has been demonstrated to reproduce the full system dynamics very accurately [20, 45], it is difficult to check for realistic systems whether the ad-hoc assumption $\Gamma^Z(t-s, \mathbf{r}_s, \mathbf{p}_s) \approx \Gamma^{\text{app}}(t-s)$ is in fact valid. This is one motivation for our hybrid projection scheme, since it allows to derive all parameters of the exact GLE from trajectory data and thereby to access the validity of the approximate GLE explicitly.

IV. HYBRID GLE

Our projection operator P_H is a hybrid of the Mori and Zwanzig projection operators and is written in the form $P_H = P_x + P_p$. Here, we derive the GLE for a scalar observable $A_t = A(\omega_t)$, the derivation for a general multi-dimensional observable is given in appendix B. Using general projection functions $B_0 = B(\mathbf{R}_0)$, which is a function of positions only, and $\dot{B}_0 = \dot{B}(\mathbf{R}_0, \mathbf{P}_0)$, which in general is a function of positions and momenta, the hybrid projection operator is given by

$$P_H A_t = P_x A_t + P_p A_t, \quad (29a)$$

$$P_x A_t = \frac{\langle \delta(B(\hat{\mathbf{R}}_0) - B(\mathbf{R}_0)), A(\hat{\omega}_0, t) \rangle}{\langle \delta(B(\hat{\mathbf{R}}_0) - B(\mathbf{R}_0)) \rangle} = \langle A_t \rangle_{B_0}, \quad (29b)$$

$$P_p A_t = \frac{\langle \dot{B}_0, A_t \rangle}{\langle \dot{B}_0^2 \rangle} \dot{B}_0. \quad (29c)$$

The projection P_x is a conditional average, defined in eq. (9), onto the observable $B_0 = B(\mathbf{R}_0)$, which is a function of positions \mathbf{R} only. As a result, the conditional average is independent of momenta. In appendix C we show that $P_x P_p = P_p P_x = 0$, from which follows that $P_H^2 = P_H$, so that P_H is idempotent in addition to being linear and hence is a projection. In appendix D, we show that P_H is self-adjoint w.r.t the inner product defined in eq. (6), i.e., it fulfills the property in eq. (22). Therefore, P_H is an orthogonal projection. Again, we denote the projection onto the complementary subspace of P_H by $Q_H = 1 - P_H$, where 1 is the identity operator. In appendix E, we prove for the projections P_H, Q_H of an arbitrary observable A_t the important property

$$\langle P_H A_t \rangle = \langle A_t \rangle \Rightarrow \langle Q_H A_t \rangle = 0. \quad (30)$$

Hence, the equilibrium ensemble average of any observable that lies completely in the complementary subspace vanishes. As an important consequence, the random force $F^R(t)$ defined in eq. (20a) lies completely in the complementary subspace for all times and, therefore, has a vanishing equilibrium average. This property is also obtained for the Zwanzig projection, but not for the Mori projection.

In the remainder, we choose the observable of interest and the projection function to coincide, $B(\mathbf{R}_t) = A(\mathbf{R}_t)$. Therefore, the GLE we derive from our hybrid scheme describes observables that are functions of positions only, such as the center of mass position, distances and angles. As an important property, Our hybrid projection P_H projects the observable A_0 and its velocity \dot{A}_0 onto themselves, meaning that

$$P_H A_0 = A_0, \quad P_H \dot{A}_0 = \dot{A}_0. \quad (31)$$

With this choice for the projection function and the specific form of the projection P_H in eq. (29), we find for the

first term on the r.h.s. of eq. (20a),

$$e^{tL} P_H L \dot{A}_0 = e^{tL} (P_x L \dot{A}_0 + P_p L \dot{A}_0), \quad (32a)$$

$$P_p L \dot{A}_0 \propto \langle \dot{A}_0, L \dot{A}_0 \rangle = -\langle L \dot{A}_0, \dot{A}_0 \rangle = 0, \quad (32b)$$

$$P_x L \dot{A}_0 = \langle L \dot{A}_0 \rangle_{A_0} = \frac{d}{dA_0} \langle \dot{A}_0^2 \rangle_{A_0} - \beta \langle \dot{A}_0^2 \rangle_{A_0} \frac{d}{dA_0} U_{\text{PMF}}(A_0), \quad (32c)$$

where we used the relation in eq. (13) to obtain eq. (32c). Equation (32c) describes the force due to a potential. To show this, we make use of the fact that the expectation value $\langle \dot{A}_0^2 \rangle_{A_0}$ is strictly positive. Thus, we can use it via

$$\langle \dot{A}_0^2 \rangle_{A_0} \equiv k_B T / M(A_0), \quad (33)$$

to define the generalized mass $M(A_0)$, which in general is a function of A_0 . Using $M(A_0)$, eq. (32c) can be simplified to

$$P_x L \dot{A}_0 = -\frac{1}{M(A_0)} \frac{dU_{\text{eff}}(A_0)}{dA_0}, \quad (34)$$

where we defined the effective potential as

$$U_{\text{eff}}(A_0) = U_{\text{PMF}}(A_0) + k_B T \ln M(A_0). \quad (35)$$

The effective potential combines the effects of the PMF and the logarithmic effective mass.

The second term on the r.h.s. of eq. (20a) accounts for memory friction, the integrand for our hybrid projection reads $e^{(t-s)L} (P_x + P_p) L F^R(s)$. The P_p projection leads to a memory function of the same form as in the Mori projection

$$P_p L F^R(s) = \frac{\langle \dot{A}_0, L F^R(s) \rangle}{\langle \dot{A}_0^2 \rangle} \dot{A}_0 = -\frac{\langle F^R(0), F^R(s) \rangle}{\langle \dot{A}_0^2 \rangle} \dot{A}_0, \quad (36a)$$

$$e^{(t-s)L} P_p L F^R(s) = -\Gamma^p(s) \dot{A}_{t-s}, \quad (36b)$$

where we defined the memory kernel due to the P_p projection as

$$\Gamma^p(s) = \frac{\langle F^R(0), F^R(s) \rangle}{\langle \dot{A}_0^2 \rangle}. \quad (37)$$

The memory friction due to the P_x projection can, using eq. (10), be written as a conditional average

$$e^{(t-s)L} P_x L F^R(s) = \langle L F^R(s) \rangle_{A_{t-s}} \equiv \Gamma^x(A_{t-s}, s), \quad (38)$$

which, using the relation in eq. (13), can be rewritten as

$$\Gamma^x(A_t, s) = \frac{d}{dA_t} D(A_t, s) - \beta D(A_t, s) \frac{d}{dA_t} U_{\text{PMF}}(A_t). \quad (39)$$

Here, we introduced the conditional correlation function between the time derivative of the observable at the initial time, \dot{A}_0 , and the random force $F^R(s)$

$$D(A_t, s) = \langle \dot{A}_0, F^R(s) \rangle_{A_t}. \quad (40)$$

With the definition of the hybrid projection operator P_H in eq. (29) and the results in eq. (34), eq. (36b) and eq. (38), the general GLE in eq. (20a) takes the specific form

$$\ddot{A}_t = -\frac{1}{M(A_t)} \frac{d}{dA_t} (U_{\text{PMF}}(A_t) + k_B T \ln M(A_t)) - \int_0^t ds \Gamma^p(s) \dot{A}_{t-s} + \int_0^t ds \Gamma^x(A_{t-s}, s) + F^R(t), \quad (41)$$

which is the exact GLE that follows from our hybrid projection scheme and constitutes a main result of our paper. A few comments are in order: i) The PMF $U_{\text{PMF}}(A_t)$ appears explicitly in the equation of motion, similar to the Zwanzig projection scheme. ii) An inhomogeneous effective mass $M(A_t)$ gives rise to a drift term. If $M(A_t)$ is constant, i.e., if the variance of \dot{A}_t is independent of A_t , see eq. (33), this drift term vanishes. For an observable A_t that is a linear combination of positions, it follows directly that the effective mass is constant [47]. Even for certain non-linear observables, such as distances in position space, it can be shown that the generalized mass is constant, as demonstrated in appendix F. On the other hand, for angles, which are three-body terms, the effective mass will in general depend on A_t , as demonstrated for the dihedral angle of butane in section VIC. iii) The memory kernel $\Gamma^p(s)$ is determined via the unconditional average over the random-force correlations in eq. (37), similarly to the Mori projection, and therefore only depends on time. It thus describes the linear friction contribution. iv) The memory friction function $\Gamma^x(A_{t-s}, s)$ is a general function of the observable A_{t-s} , it therefore accounts for non-linear friction contributions. According to eq. (39), this contribution disappears if the conditional correlation function between the random force and the time derivative of the observable, $D(A_t, s)$, as defined in eq. (40), vanishes. This constitutes the exact condition for which the approximate GLE in eq. (28) is valid. v) The first moment of the random force vanishes, $\langle F^R(t) \rangle = 0$, as follows from the relation eq. (30). The second moment is determined by the memory kernel $\Gamma^p(s)$ via eq. (37). Higher cumulants do not necessarily vanish but are not expected to play a significant role since non-linear effects are already accounted for by the PMF $U_{\text{PMF}}(A_t)$. Indeed, in section VIC we demonstrate for the explicit example of the butane dihedral angle that the random-force distribution exhibits finite but moderate non-Gaussian contributions.

The multi-dimensional generalization of eq. (41), i.e., the case in which the observable is a vector $\mathbf{A}(\mathbf{R}_t) = (A_1(\mathbf{R}_t), A_2(\mathbf{R}_t), \dots, A_n(\mathbf{R}_t))$, is derived in appendix B.

V. NUMERICAL SCHEME FOR EXTRACTING RANDOM FORCES FROM TRAJECTORIES

In the absence of a potential and in the absence of non-linear 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 [6, 35]. Their derivations explicitly use the Mori projection, so the results are only valid for the Mori GLE in eq. (24).

We now introduce a method to compute the random force trajectory $F^R(\omega_0, t)$ and from that the memory kernel $\Gamma^p(t)$ and the non-linear memory function $\Gamma^x(A_s, t-s)$ as defined by our GLE, eq. (41), from a given trajectory of an arbitrary observable. For this, let us consider the projected propagator $e^{tQ_H L}$ based on our hybrid projection scheme eq. (29). From the Dyson decomposition in eq. (19), we obtain by rearranging

$$e^{tQ_H L} = e^{tL} - \int_0^t ds e^{(t-s)L} P_H L e^{sQ_H L}. \quad (42)$$

Applying eq. (42) on the initial random force $F^R(\omega_0, 0)$ and using eq. (20b) and the memory functions $\Gamma^p(t)$ and $\Gamma^x(A, t)$ defined in eq. (36b) and eq. (38), respectively, we find

$$\begin{aligned} F^R(\omega_0, t) &= e^{tL} F^R(\omega_0, 0) + \int_0^t ds \Gamma^p(s) e^{(t-s)L} \dot{A}_0 \\ &\quad - \int_0^t ds e^{(t-s)L} \Gamma^x(A_0, s). \end{aligned} \quad (43)$$

Now, we consider eq. (42) at time $t + \Delta t$

$$\begin{aligned} e^{(t+\Delta t)Q_H L} &= e^{tL} e^{\Delta t L} \\ &\quad - \int_0^{t+\Delta t} ds e^{(t-s)L} e^{\Delta t L} P_H L e^{sQ_H L}. \end{aligned} \quad (44)$$

By splitting up the integral on the r.h.s. into two parts,

we obtain

$$\begin{aligned} e^{(t+\Delta t)Q_H L} &= e^{tL} e^{\Delta t L} - \int_0^t ds e^{(t-s)L} e^{\Delta t L} P_H L e^{sQ_H L} \\ &\quad - \int_0^{\Delta t} ds e^{(\Delta t-s)L} P_H L e^{(t+s)Q_H L}, \end{aligned} \quad (45)$$

where we used the substitution $s \rightarrow s - t$ in the second integral. Acting with the operator in eq. (45) on the initial random force $F^R(\omega_0, 0)$ and using eq. (5) gives

$$\begin{aligned} F^R(\omega_0, t + \Delta t) &= e^{tL} F^R(\omega_{\Delta t}, 0) \\ &\quad + \int_0^t ds \Gamma^p(s) e^{(t-s)L} \dot{A}_{\Delta t} - \int_0^t ds e^{(t-s)L} \Gamma^x(A_{\Delta t}, s) \\ &\quad - \int_0^{\Delta t} ds e^{(\Delta t-s)L} P_H L F^R(\omega_0, t + s). \end{aligned} \quad (46)$$

Comparing eq. (46) with eq. (43), we see that the first three terms on the r.h.s. of eq. (46) are equal to $F^R(\omega_{\Delta t}, t)$. Hence, we find

$$\begin{aligned} F^R(\omega_0, t + \Delta t) &= F^R(\omega_{\Delta t}, t) + \int_0^{\Delta t} ds \Gamma^p(t + s) \dot{A}_{\Delta t-s} \\ &\quad - \int_0^{\Delta t} ds \Gamma^x(A_{\Delta t-s}, t + s). \end{aligned} \quad (47)$$

For given trajectories A_t, \dot{A}_t and given random force $F^R(\omega_{\Delta t}, t)$ as a function of the phase space configuration $\omega_{\Delta t}$, eq. (47) gives the random force $F^R(\omega_0, t + \Delta t)$ 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. (47) is discretized in time and A -space. For this, we use the left rectangular rule to discretize the time integrals. The random force is discretized as $F^R(\omega_{t'}, t) = F^R(\omega_{i\Delta t}, j\Delta t) \equiv F^R(i, j)$. The A -space is divided into N_A bins with width ΔA , the bin intervals are labeled by $I_\alpha = [\alpha\Delta A, (\alpha+1)\Delta A]$ with $\alpha = 1, 2, \dots, N_A$. The discretized versions of eqs. (37), (39), (40) and (47) read

$$F^R(i, j+1) = F^R(i+1, j) + \Delta t \Gamma^p(j) \dot{A}_{i+1} - \Delta t \Gamma^x(A_{i+1}, j) + \mathcal{O}(\Delta t^2), \quad (48a)$$

$$\Gamma^p(j) = \frac{\sum_{i=0}^{N_{\text{traj}}-j-1} F^R(i, 0) F^R(i, j)}{\sum_{i=0}^{N_{\text{traj}}-j-1} \dot{A}_i^2}, \quad (48b)$$

$$\Gamma^x(A_{i+1}, j) = \left[\frac{D(\alpha+1, j) - D(\alpha-1, j)}{2\Delta A} - \beta D(\alpha, j) \frac{U_{\text{PMF}}(\alpha+1) - U_{\text{PMF}}(\alpha-1)}{2\Delta A} \right]_{A_{i+1} \in I_\alpha}, \quad (48c)$$

$$D(\alpha, j) = \frac{\sum_{\substack{i \leq N_{\text{traj}}-j-1 \\ A_i \in I_\alpha}} \dot{A}_i F^R(i, j)}{\sum_{\substack{i \leq N_{\text{traj}}-j-1 \\ A_i \in I_\alpha}} 1}. \quad (48d)$$

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. (48). N_{traj} denotes the total length of the A_t trajectory used. The sums run from $i = 0$ to $N_{\text{traj}} - j - 1$, because for given j , the iterative scheme has only determined the random force at times up to $N_{\text{traj}} - j - 1$, as follows from eq. (48a). The sums in the denominator extend over the same interval as in the numerator in order to increase the numerical stability [6, 35]. The derivatives in A -space in eq. (48c) are computed using central differences. The iterative scheme in eq. (48) works as follows: First, note from eq. (41) that $F^R(i, 0) = \dot{A}_i + (1/M(A_i))d/dA_i [U_{\text{PMF}}(A_i) + k_B T \ln M(A_i)]$, i.e., 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 $i = 0, 1, 2, \dots, N_{\text{traj}} - 1$. This, together with \dot{A}_i , can be obtained directly from a given trajectory of the observable A . Then, $F^R(i, 0)$, A_i , and \dot{A}_i are inserted into eq. (48) to compute $F^R(i, 1)$ for $i = 0, 1, 2, \dots, N_{\text{traj}} - 2$. $F^R(i, 1)$ is then used to compute $F^R(i, 2)$ for $i = 0, 1, 2, \dots, N_{\text{traj}} - 3$ and so forth. While computing $F^R(i, j)$, the memory friction functions $\Gamma^p(i)$ and $\Gamma^x(A_i, i - j)$ are computed simultaneously. If our only goal is to compute the memory friction functions, we can stop the computation of $F^R(i, j)$ 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 $F^R(i, N_{\text{mem}})$. At that point, we generated $N_{\text{traj}} - N_{\text{mem}} - 1$ 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. In appendix G, we present an alternative discretization in time for eq. (47) using the trapezoidal rule.

VI. APPLICATIONS

We test our numerical algorithm in eq. (48) on three different systems: an exactly solvable harmonic Hamiltonian model which leads to a GLE without spatial dependencies in the memory friction term, the non-linear Hamiltonian version of the Zwanzig model [53], where spatial dependencies in the memory friction term are present, and finally, we discuss results obtained for the dihedral angle dynamics of a butane molecule in water from MD simulations.

A. Harmonic Hamiltonian Model

The exactly solvable harmonic model is defined by the Hamiltonian

$$H(x, p, \{q_n, v_n\}) = \frac{p^2}{2m} + \sum_{n=1}^N \frac{v_n^2}{2m_n} + U(x) + \sum_{n=1}^N \frac{k_n}{2} (x - q_n)^2. \quad (49)$$

The relevant coordinates are the one-dimensional position x and momentum p which are coupled to the auxiliary particle positions q_n and momenta v_n . If we choose the potential $U(x)$ to be a harmonic potential, i.e., $U(x) = kx^2/2$, we can use our hybrid projection formalism to exactly derive the GLE. For this we compute the random force $F^R(t)$ defined in eq. (20b) by an operator expansion to all orders

$$F^R(t) = e^{tQ_H L} Q_H L p_0 = \sum_n \frac{t^n}{n!} (Q_H L)^n Q_H L p_0, \quad (50)$$

as shown in appendix H. Once $F^R(t)$ is computed, the memory functions $\Gamma^p(t)$ and $\Gamma^x(x, t)$ are obtained from eq. (37) and eq. (39), respectively. In appendix I we show how to alternatively obtain a GLE without projection, namely by solving the equations of motion for the q_n variables and inserting the result back into the equation of motion for x , which works for general potential $U(x)$. The GLE's obtained from the projection formalism and the exact solution agree with one another and take the form of the approximate GLE in eq. (28),

$$\dot{p}_t = -U'_{\text{PMF}}(x_t) - \int_0^t ds \Gamma^p(t-s) p_s + F^R(t). \quad (51)$$

The memory friction kernel is given by

$$\Gamma^p(t) = \frac{1}{m} \sum_{n=1}^N k_n \cos(\mu_n t) \quad (52)$$

with $\mu_n = \sqrt{k_n/m_n}$. We note that the spatially dependent memory friction term $\Gamma^x(x, t)$ vanishes, as shown in appendix H and I.

In order to test our numerical scheme in eq. (48), we need to generate trajectories of x_t . To do this in a numerically efficient fashion, we identify eq. (52) as the Fourier series of an even function with Fourier coefficients k_n . In the limit of $N \rightarrow \infty$ and for a continuous frequency dependency, i.e., $k_n \rightarrow k(\mu)d\mu/2\pi$, we can choose the exponential-oscillating memory kernel

$$\Gamma^p(t) = \frac{K}{m} e^{-\frac{|t|}{2\tau_\Gamma}} \left(\cos\left(\frac{\nu}{2\tau_\Gamma} t\right) + \frac{1}{\nu} \sin\left(\frac{\nu}{2\tau_\Gamma} |t|\right) \right). \quad (53)$$

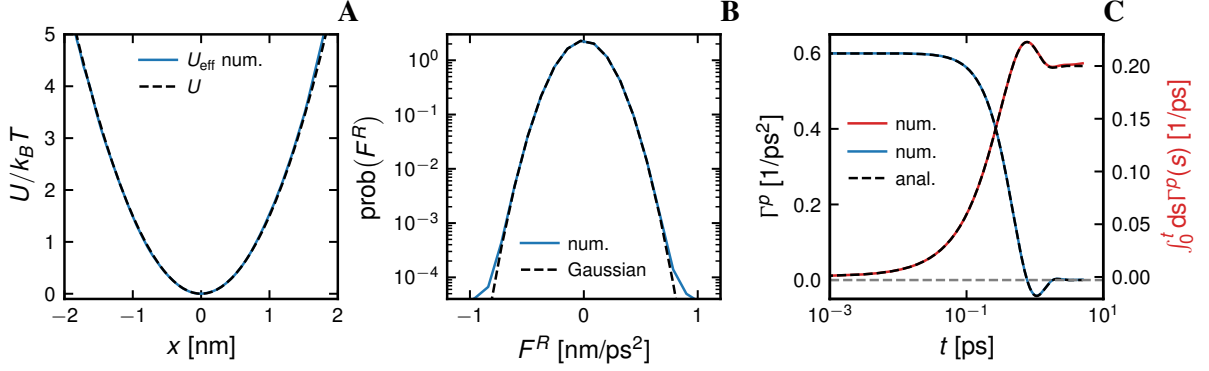


FIG. 1. Test of the numerical extraction scheme in eq. (48) using the harmonic Hamiltonian model defined in eq. (49). In **A**, we compare the input potential $U(x) = kx^2/2$ (broken line) with the numerically obtained effective potential $U_{\text{eff}}(x)$ defined in eq. (35) (solid blue line). In **B**, we confirm that the numerical extraction of the random force leads to the expected Gaussian distribution with zero mean and standard deviation $\sqrt{\langle p_0^2 \rangle \Gamma^p(0)}$. The shaded area in blue highlights the numerical error. In **C**, we check that the analytical memory function $\Gamma^p(t)$ given in eq. (53) and its running integral are accurately reproduced by the numerical results from the extraction scheme.

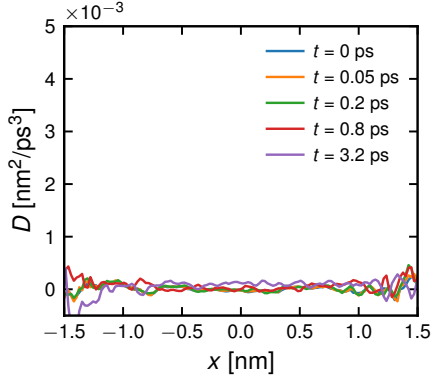


FIG. 2. Conditional velocity-random force correlation function $D(x, t)$ defined in eq. (40) for the harmonic model defined in eq. (49).

This maps the Hamiltonian system in eq. (49) onto the stochastic system of two linearly coupled Langevin equations

$$m\ddot{x}_t = -kx_t - K(x_t - y_t), \quad (54a)$$

$$m_y\ddot{y}_t = -K(y_t - x_t) - \gamma\dot{y}_t + \sqrt{2k_B T \gamma}\eta(t), \quad (54b)$$

$$\langle \eta(t) \rangle = 0, \quad \langle \eta(t), \eta(0) \rangle = \delta(t), \quad (54c)$$

where $\eta(t)$ in eq. (54) is a white noise variable, as derived in appendix J. The parameters in eq. (53) and eq. (54) are related by $\tau_\Gamma = m_y/\gamma$ and $\nu^2 = 4m_y K/\gamma^2 - 1$. The scalar y variable in eq. (54) is the stochastic representation of the Hamiltonian environment produced by the q_n variables in eq. (49). Using eq. (54), we numerically generate trajectories x_t for a system with thermal energy $k_B T = 2.5$ kJ/mol, which corresponds to $T = 300$ K [54]. The other parameters are chosen to be $m = 50$ u, $m_y = 2$ u, $K = 30$ kJ/mol/nm, $k = 7.5$ kJ/mol/nm,

$\gamma = 10$ u/ps and a time step of $dt = 0.001$ ps. The simulation time is 100 ns. The results shown in fig. 1B, fig. 1C and fig. 2 are obtained by averaging over the results of 100 independent trajectories.

In fig. 1 we compare analytical results with results derived from the numerically generated trajectories using the scheme in eq. (48), where the x -space is discretized using $N_A = 200$ bins of equal length. In fig. 1A we compare the input potential $U(x) = kx^2/2$ (broken line) with the numerically obtained effective potential $U_{\text{eff}}(x)$ defined in eq. (35) (solid blue line), both potentials are shifted so that they are zero at $x = 0$. The agreement is perfect, which in particular means that the effective mass $M(x)$ defined in eq. 33 is a constant, as expected. In fig. 1B we compare the analytical and the numerically determined random force distribution, which demonstrates that indeed non-Gaussian contributions are absent. In fig. 1C we compare the analytic memory kernel $\Gamma^p(t)$ in eq. (53) with the one extracted from the simulation trajectory and again obtain perfect agreement. This all shows that the numerical extraction scheme works perfectly on fluctuating trajectories. In fig. 2 we show the numerical result for the function $D(x, t)$ defined by eq. (38) for different times. As predicted in appendix H, $D(x, t)$ vanishes for all times.

B. The Non-Linear Zwanzig Model

As the second exactly solvable model we consider the Hamiltonian version of the non-linear Zwanzig model [53], for which non-linear friction effects are present and therefore the approximate GLE in eq. (28) is not valid

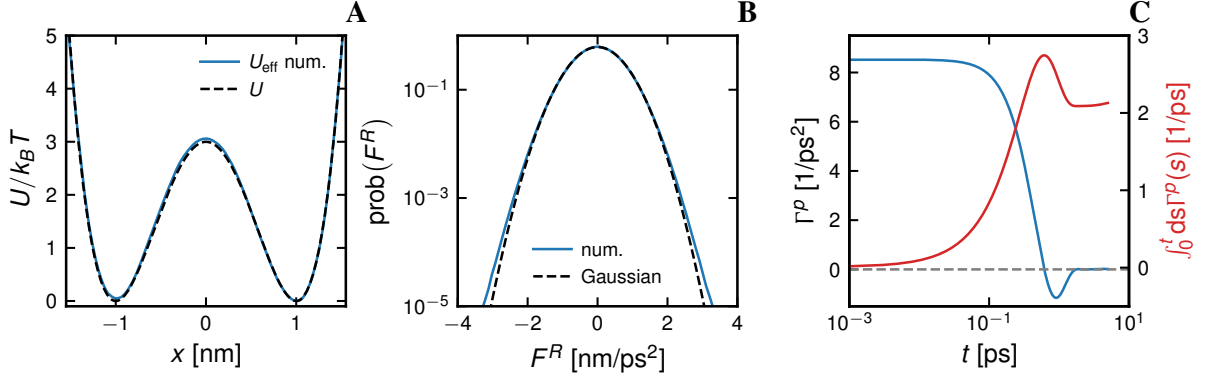


FIG. 3. Test of the numerical algorithm in eq. (48) for the non-linear Zwanzig model defined in eq. (55) using a double-well potential $U(x) = U_0(x^2 - 1)^2$. **A** Comparison of the input potential $U(x)$ (broken line) with the numerically obtained effective potential $U_{\text{eff}}(x)$ defined in eq. (35) (solid blue line). The two potentials coincide, which means that the effective mass is constant, as expected on analytic grounds. **B** The numerically extracted random force distribution (blue solid line) is well described by a Gaussian with vanishing mean and standard deviation $\sqrt{\langle \dot{x}_0^2 \rangle \Gamma^p(0)}$ (broken black line). The shaded area in blue highlights the numerical error. In **C**, we show the numerically extracted memory function $\Gamma^p(t)$ and its running integral.

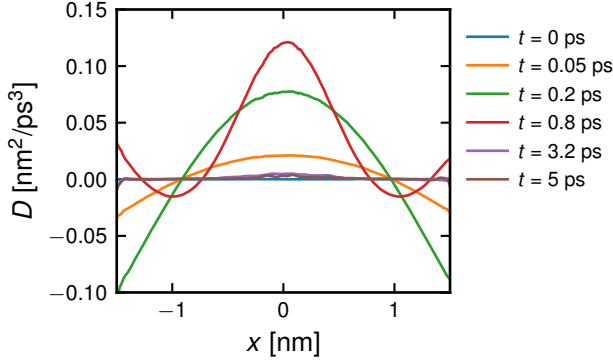


FIG. 4. Conditional velocity-random force correlation function $D(x, t)$ defined in eq. (40) for the non-linear Zwanzig model defined in eq. (55) as a function of the position x for different times.

anymore. This model is defined by the Hamiltonian

$$H(x, p, \{q_n, v_n\}) = \frac{p^2}{2m} + \sum_{n=1}^N \frac{v_n^2}{2m_n} + U(x) + \sum_{n=1}^N \frac{k_n}{2} (\alpha(x) - q_n)^2. \quad (55)$$

In eq. (55), a generally non-linear function $\alpha(x)$ determines the coupling between the relevant variable x and the auxiliary variables q_n . Note that for $\alpha(x) = x$, we obtain back the harmonic model defined in eq. (49). The GLE that follows from the Hamiltonian system in eq. (55) can not be calculated in closed form using our hybrid projection scheme for general $\alpha(x)$, we therefore cannot derive the exact form of $\Gamma^p(t)$. On the other hand, by solving the equations of motion for the q_n variables and

inserting the result into the equation for x , one finds a GLE of the form

$$m\ddot{x}_t = -U'(x_t) - \int_0^t ds \Gamma[t-s, x_t, x_s] \dot{x}_s + F_Z^R(t, x_t), \quad (56a)$$

with a (for $\alpha(x) \neq x$) non-linear memory friction function

$$\Gamma(t-s, x_t, x_s) = \alpha'(x_t)\alpha'(x_s) \sum_{n=1}^N k_n \cos(\mu_n(t-s)). \quad (56b)$$

Actually, the form of the memory kernel $\Gamma(t-s, x_t, x_s)$ in eq. (56b), and in particular its dependence on the trajectory x_t , is not compatible with the form of the memory function $\Gamma^x(x_{t-s}, s)$ or, equivalently, $\Gamma^x(x_s, t-s)$, in eq. (41). In fact, in appendix K we demonstrate that the GLEs given in eq. (56b) and in eq. (41) are equivalent in the sense that they produce, for identical initial conditions, identical trajectories x_t . This of course is expected, since they follow via exact derivations from the same Hamiltonian. This finding is similar to the fact that the Mori and Zwanzig GLEs are, in the absence of approximations, also equivalent and shows that even GLEs with identical PMFs and different friction memory and random force terms can be equivalent. For $\alpha(x) \neq x$, it is therefore interesting to extract the non-linear friction term $\Gamma^x(x_{t-s}, s)$, as defined by our GLE in eq. (56b), from simulation trajectories of x_t .

Similar to our approach to obtain eq. (54) for $N \rightarrow \infty$, we exploit the structure of eq. (56b), which is equivalent to a Fourier decomposition in the time domain, to map the Hamiltonian system in eq. (55) onto a system of non-linearly coupled Langevin equations given by (see

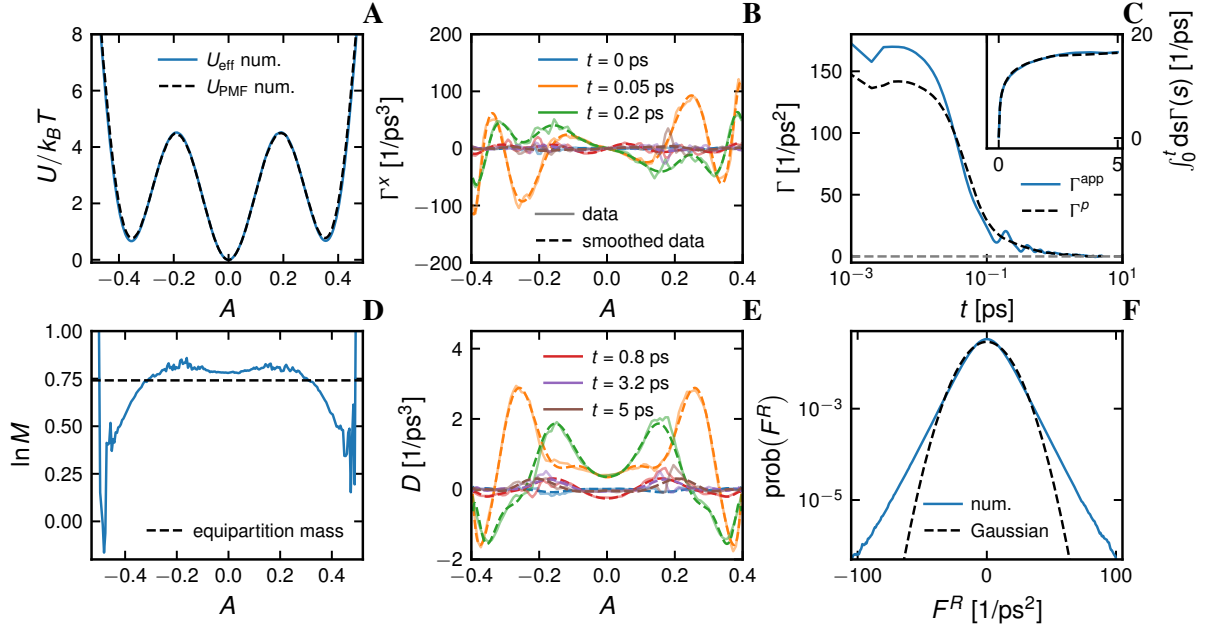


FIG. 5. Extraction of GLE parameters for MD data of the dihedral angle dynamics of butane in water. Here, $A = \phi/(\phi_{\max} - \phi_{\min}) \in (-0.5, 0.5)$ denotes the rescaled dihedral angle with ϕ being the dihedral angle and ϕ_{\max} , ϕ_{\min} being the largest and smallest value of ϕ along the trajectory, respectively. **A** A small deviation between the effective potential $U_{\text{eff}}(A)$ defined in eq. (35) (blue solid line) and the numerically computed potential of mean force $U_{\text{PMF}}(A)$ (broken line) is observed, which is explained by the dependence of the effective mass $M(A)$ on the dihedral angle, as shown in D. **B** Non-linear friction function $\Gamma^x(A, t)$ defined by eq. (39) for different times. Solid lines correspond to the original data, the dashed lines show smoothed data. **C** Comparison of the memory function $\Gamma^p(t)$ defined within the hybrid GLE eq. (41) and the memory function $\Gamma^{\text{app}}(t)$ defined within the approximate GLE in eq. (28). The inset shows the respective running integrals. **D** Logarithm of the effective mass $M(A)$ as defined by eq. (33). **E** Correlation function $D(A, t)$ defined in eq. (40) as a function of A for different times. Solid lines correspond to the original data, the dashed lines show smoothed data. **F** The extracted random force distribution (blue line) shows pronounced deviations from a Gaussian with vanishing mean and standard deviation of $\sqrt{\langle \dot{A}_0^2 \rangle \Gamma^p(0)}$ (broken line). The shaded area in blue highlights the numerical error and is of the order of the line width.

appendix J)

$$m\ddot{x}_t = -U'(x_t) - K\alpha'(x_t)(\alpha(x_t) - y_t), \quad (57a)$$

$$m_y\ddot{y}_t = -K(y_t - \alpha(x_t)) - \gamma\dot{y}_t + \sqrt{2k_B T \gamma} \eta(t), \quad (57b)$$

$$\langle \eta(t) \rangle = 0, \quad \langle \eta(t), \eta(0) \rangle = \delta(t). \quad (57c)$$

For $U(x) = kx^2/2$ and $\alpha(x) = x$ we recover eq. (54). Using eq. (57), we perform simulations for the parameter set $k_B T = 2.5 \text{ kJ/mol}$, $m = 50 \text{ u}$, $m_y = 2 \text{ u}$, $K = 30 \text{ kJ/mol/nm}$, $\gamma = 10 \text{ u/ps}$ to generate 100 trajectories x_t of 100 ns length each. For the potential we choose a non-linear double-well potential $U(x) = U_0(x^2 - 1)^2$ with $U_0 = 3k_B T$, as shown in fig. 3A, and for the non-linear coupling potential we choose a quadratic function $\alpha(x) = \alpha_0 x^2/2$ with $\alpha_0 = 4 \text{ nm}^{-1}$. The resulting trajectories are then used to compute via eq. (48) all parameters of the hybrid GLE in eq. (41), which are presented in fig. 3. In this calculation, the x -space is discretized using $N_A = 200$ bins of equal length.

The effective mass $M(A_t)$ for an observable A_t that is a linear function of atomic positions is constant [47], as follows directly from the fact that the velocity distribu-

tion function factorizes for Hamiltonians of the form in eq. (1). Indeed, in fig. 3A the numerically obtained effective potential $U_{\text{eff}}(x)$ defined in eq. (35) (solid blue line) is shown to agree perfectly with the input potential $U(x)$ (broken line) when both potentials are shifted so that they are zero at $x = 0$. In fig. 3B we compare the random force distribution obtained numerically via eq. (48) from the simulated trajectory (blue line) with a Gaussian with vanishing mean and a variance of $\sqrt{\langle \dot{x}_0^2 \rangle \Gamma^p(0)}$, as predicted by eq. (37), and obtain very good agreement; for the comparison, the value $\Gamma^p(0)$ is numerically extracted from the simulated trajectory. Note that eq. (37) does not imply that the distribution of the random force is a pure Gaussian, but the data in fig. 3B demonstrate that non-Gaussian contributions are either absent or very small.

In fig. 3C we show the memory kernel $\Gamma^p(t)$ extracted from the simulation trajectory, the result looks qualitatively similar to the result in fig. 1C for the harmonic model. In fig. 4, we show the correlation function $D(x, t)$ defined in eq. (40) for a few different fixed times. Note that $D(x, 0)$ vanishes at time 0, which is

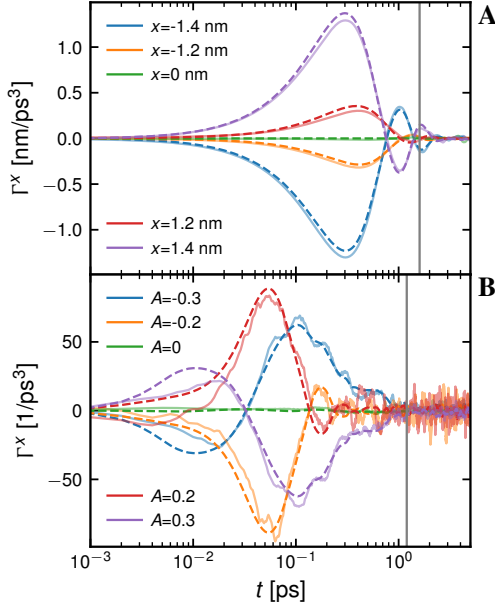


FIG. 6. Non-linear friction function $\Gamma^x(A, t)$ as a function of time for different values of the reaction coordinate. **A** Results for the non-linear Zwanzig model in eq. (55), here the reaction coordinate is given by the position of the relevant particle coordinate x . **B** Results for the rescaled dihedral angle of butane from MD simulations, here the reaction coordinate is given by the rescaled dihedral angle A . The dihedral angle data is more noisy compared to the non-linear Zwanzig model system because of the reduced simulation time. The vertical gray lines indicate the time after which the corresponding linear friction kernel Γ^p stays below 1% of its initial value. The dashed lines are obtained from a smoothing procedure, see appendix L.

true for general $A_0 = A(\mathbf{R}_0)$, since the product $\dot{A}_0 F^R(0)$ is odd in the momenta and thus the conditional average $\langle \dot{A}_0, F^R(0) \rangle_A = D(A, 0)$ vanishes. For finite time, $D(x, t)$ in fig. 4 rises before dropping back to zero in the long-time limit. The time after which $D(x, t)$ decays to zero is about 1 ps and thus comparable to the memory time of $\Gamma^p(t)$ in fig. 3C. Note that a finite correlation function $D(x, t)$ will via eq. (39) give rise to a non-linear friction memory function $\Gamma^x(x, t)$. The non-linear Zwanzig model defined by the Hamiltonian eq. (55) is thus represented by a constant effective mass term $M(A)$ but a non-vanishing non-linear friction memory.

C. Dihedral Angle Dynamics of Butane from MD Simulations

To test our algorithm for an observable that is a non-linear function of atomic positions, we consider the dihedral angle dynamics of a butane molecule in water as obtained from MD simulations. The dihedral angle ϕ of butane is a conceptually simple yet relevant observable and provides a simple scenario to study con-

formational transitions in polymers and proteins that is both theoretically [55] and experimentally [56] accessible. In fig. 5, we present results for the rescaled angle $A = \phi/(\phi_{\max} - \phi_{\min})$, where the maximal and minimal observed angles in the studied trajectory are $\phi_{\max} = 155^\circ$ and $\phi_{\min} = -157^\circ$. In fig. 5A, the effective potential $U_{\text{eff}}(A)$ defined in eq. (35) (blue solid line) shows small but significant deviations from the PMF $U_{\text{PMF}}(A)$ (broken line), which is explained by the dependence of the effective mass $M(A)$ on the dihedral angle, as shown in fig. 5D. In fig. 5C, the deviations between $\Gamma^p(t)$, as defined within the exact GLE eq. (41) and determined numerically from the MD trajectory via eq. (48), and $\Gamma^{\text{app}}(t)$, defined within the approximate GLE in eq. (28) and obtained via a Volterra scheme [20, 37], are pronounced and already suggest that non-linear friction effects, not captured by $\Gamma^{\text{app}}(t)$, are present. A closer look at the results in fig. 5C reveals that $\Gamma^p(t)$ and $\Gamma^{\text{app}}(t)$ have similar decay times, but $\Gamma^{\text{app}}(t)$ oscillates in time while $\Gamma^p(t)$ does not. These deviations between $\Gamma^p(t)$ and $\Gamma^{\text{app}}(t)$ must be due to non-linear memory effects, as confirmed in fig. 5E, where the correlation function $D(A, t)$ defined in eq. (40) is shown for a few different fixed times. Thus, a non-linear memory friction contribution $\Gamma^x(A, t)$, defined in eq. (39) and shown in fig. 5B, is present in the GLE. As mentioned before, from the definition of $D(A, t)$ in eq. (40) it follows that $D(A, t)$ vanishes at time $t = 0$, i.e., $D(A, 0) = 0$, from which it is easy to see via eq. (39) that $\Gamma^x(A, 0)$ vanishes, too, as indeed confirmed by the data in fig. 5B and E. For finite time, both $D(A, t)$ and $\Gamma^x(A, t)$ rise in amplitude before decaying to zero after a time corresponding to the memory time of $\Gamma^p(t)$ in fig. 5C, which is about 1 ps.

The rise and decay of non-linear friction effects is presented in fig. 6, where we show $\Gamma^x(A, t)$ as a function of time for different fixed values of A for the non-linear Zwanzig model in fig. 6A and for the butane dihedral angle dynamics in fig. 6B. The vertical gray lines indicate the time after which the linear friction kernel $\Gamma^p(t)$ for each system stays below 1% of its initial value $\Gamma^p(0)$. In addition to the raw numerical data (solid lines), we show smoothed curves which are obtained by fits to Legendre polynomials (broken lines), as described in appendix L.

VII. SUMMARY AND DISCUSSION

By using a hybrid projection scheme that combines linear Mori projection on the reaction coordinate velocities and non-linear conditional Zwanzig projection on the reaction coordinates themselves, we derive a GLE that contains the non-linear potential of mean force and a non-linear 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 then splits into two parts. One part is linear in the reaction coordinate velocity and reflects linear friction proportional to a memory kernel $\Gamma^p(t)$. The memory kernel $\Gamma^p(t)$ is related to the fluctuat-

ing force $F^R(t)$, defined in eq. (20b), by a relation that resembles a fluctuation-dissipation theorem, eq. (37). The non-linear memory friction function $\Gamma^x(A_{t-s}, s)$ accounts for non-linear dependencies of friction on A_{t-s} and is connected to the fluctuating force $F^R(t)$ by a conditional correlation function, given in eq. (40). Thus, when modeling $F^R(t)$ as a stochastic variable, it has to fulfill both relations, eq. (37) and eq. (40). The approximate GLE in eq. (28) is obtained from our GLE in eq. (41) only when the memory friction function $\Gamma^x(A_{t-s}, s)$ vanishes, which thus establishes a firm criterion for the validity of the approximate GLE.

We also introduce a numerical scheme to compute all parameters of our GLE from a given trajectory A_t and apply it on numerically determined trajectories for a harmonic and a non-linear exactly solvable many-body particle system, here we show that the numerical results agree with the analytical predictions. We also apply our 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 non-linear memory friction contribution is finite and non-negligible. In order to estimate the importance of the non-linear memory friction, we have to compare the linear memory kernel $\Gamma^p(t)$ and the non-linear memory function $\Gamma^x(A, t)$. For this we multiply the linear-friction memory kernel at time zero, $\Gamma^p(0)$, by the root mean square velocity and obtain $\Gamma^p(0)\sqrt{\langle \dot{A}^2 \rangle} = 171 \text{ ps}^{-3}$, which can be directly compared with the maximal value of the non-linear memory friction function $\Gamma^x(A_{\max}, t_{\max}) = 97 \text{ ps}^{-3}$, which is obtained for $A_{\max} = 0.26$ and $t_{\max} = 0.043 \text{ ps}$. The value of $\Gamma^x(A_{\max}, t_{\max})$ thus turns out to be roughly half the value of $\Gamma^p(0)\sqrt{\langle \dot{A}^2 \rangle}$, which means that non-linear memory friction effects are not negligible. Interestingly, our results demonstrate that non-linear friction memory leads to oscillations in the memory function $\Gamma^{\text{app}}(t)$ of the approximate GLE, which are not present in $\Gamma^p(t)$, as shown in fig. 5C. Finally, we show that the random force in the GLE from our hybrid projection scheme exhibits small but detectable deviations from a Gaussian distribution. All these results lead us to conclude that the GLE derived from our hybrid projection scheme is practically useful and allows to detect and model non-linear friction effects that have been neglected in previous applications of the approximate GLE with linear memory friction.

VIII. METHODS AND MATERIALS

MD simulations are performed using the Gromacs MD package (version 2020-Modified) [54]. For the MD simulation of the butane molecule, we use the GROMOS53A6 force field [57] with the TIP4P/2005 rigid water model [58]. The simulation box has side lengths of 3.35 nm and contains 1250 water molecules. We constrain the

butane bond lengths and angles using the SHAKE algorithm [59]. For long-range electrostatic interactions, we use the particle-mesh Ewald [60], with a cut-off of 1 nm. The simulation time step is 1 fs, and the total simulation time is 100 ns. All simulations are performed in the NVT ensemble with a temperature of 300 K, controlled with a velocity rescaling thermostat [61]. Input files of the MD simulations are available upon request. The Langevin simulations are performed using the Leap Frog algorithm for numerical integration. Our Python codes for extracting the GLE parameters and running Langevin simulations are also available upon request. When computing non-linear memory contributions, the time resolution of the trajectory and the number of bins in reaction-coordinate space have to be chosen with care. In our analysis of Langevin and MD simulations, we use 200 bins to discretize the reaction-coordinate space. For the butane dihedral angle system in fig. 5, we exclude boundary regions in the trajectory for $|A| > 0.4$ in the computation of conditional correlations, that means we exclude observable values with a small fraction of realizations along the trajectory, since these would lead to significant noise in the extracted memory functions and thus destabilize the numerical extraction. Such noise effects are clearly visible in the effective mass profile in fig. 5F.

ACKNOWLEDGMENTS

We acknowledge support by Deutsche Forschungsgemeinschaft Grant CRC 1114 "Scaling Cascades in Complex System", Project 235221301, Project B03 and by the ERC Advanced Grant 835117 NoMaMemo. We gratefully acknowledge computing time on the HPC clusters at the physics department and ZEDAT, FU Berlin.

Appendix A: Derivation of Eq. (13)

In the following, we derive eq. (13) by using the definition of conditional correlation functions in eq. (9), the relations in eq. (8), eq. (12) and the definition of the PMF in eq. (11). We start with

$$\langle LB_{t'} \rangle_{A_t} = \frac{\langle \delta(A(\hat{\omega}_0) - A(\omega_t)), LB(\hat{\omega}_0, t') \rangle}{\langle \delta(A(\hat{\omega}_0) - A(\omega_t)) \rangle}, \quad (\text{A1a})$$

$$= \int_{-\infty}^{\infty} da \delta(A(\omega_t) - a) \frac{\langle \delta(A_0 - a), LB_{t'} \rangle}{\langle \delta(A_0 - a) \rangle}, \quad (\text{A1b})$$

where the average is over variables with a hat and consequently, the Liouville operator L only acts on variables with a hat. We consider the rightmost term in eq. (A1b)

$$\frac{\langle \delta(A_0 - a), LB_{t'} \rangle}{\langle \delta(A_0 - a) \rangle} = - \frac{\langle L\delta(A_0 - a), B_{t'} \rangle}{\langle \delta(A_0 - a) \rangle} \quad (\text{A1c})$$

$$= \frac{\langle \dot{A}_0 \frac{d}{da} [\delta(A_0 - a)], B_{t'} \rangle}{\langle \delta(A_0 - a) \rangle}, \quad (\text{A1d})$$

where we used eq. (8) and eq. (12). We next pull out the derivative w.r.t. a in eq. (A1d) from the inner product and use the product rule of differentiation, which gives

$$\frac{\langle \delta(A_0 - a), LB_{t'} \rangle}{\langle \delta(A_0 - a) \rangle} = \frac{d}{da} \frac{\langle \delta(A_0 - a) \dot{A}_0, B_{t'} \rangle}{\langle \delta(A_0 - a) \rangle} + \frac{\langle \delta(A_0 - a) \dot{A}_0, B_{t'} \rangle}{\langle \delta(A_0 - a) \rangle} \frac{d}{da} \ln \langle \delta(A_0 - a) \rangle. \quad (\text{A1e})$$

Finally, we use the definition of the PMF in eq. (11) and insert eq. (A1e) into eq. (A1b) to obtain eq. (13).

Appendix B: Multi-Dimensional Hybrid GLE

Here, we derive eq. (41) for a multidimensional observable that is a function of particle positions only. We denote the set of observables using the vector $\mathbf{A}(\mathbf{R}_t) = (A_1(\mathbf{R}_t), A_2(\mathbf{R}_t), \dots, A_n(\mathbf{R}_t))$. As before, all observables implicitly depend on time 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 projection operator reads for general vectorial projection function \mathbf{B}_0

$$P_H A_{m,t} = (P_p + P_x) A_{m,t}, \\ = \sum_{k=1}^n \frac{\langle A_{m,t}, \dot{B}_{k,0} \rangle}{\langle \dot{B}_{k,0}^2 \rangle} \dot{B}_{k,0} + \langle A_{m,t} \rangle_{\mathbf{B}_0}. \quad (\text{B1})$$

Choosing $\mathbf{B}_0 = \mathbf{A}_0$, as in the main text, the projection in eq. (B1) leads to the following potential term

$$e^{tL} P_x L \dot{\mathbf{A}}_0 = k_B T (\nabla_A^T \cdot M^{-T}(\mathbf{A}_t))^T - M^{-1}(\mathbf{A}_t) \cdot \nabla_A U_{\text{PMF}}(\mathbf{A}_t). \quad (\text{B2})$$

where we introduced the inverse generalized mass matrix

$$M_{kl}^{-1}(\mathbf{A}) = \beta \langle \dot{A}_{k,0}, \dot{A}_{l,0} \rangle_{\mathbf{A}}. \quad (\text{B3})$$

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

$$\ddot{\mathbf{A}}_t = k_B T (\nabla_A^T \cdot M^{-T}(\mathbf{A}_t))^T - M^{-1}(\mathbf{A}_t) \cdot \nabla_A U_{\text{PMF}}(\mathbf{A}_t) - \int_0^t ds \Gamma^p(t-s) \cdot \dot{\mathbf{A}}_s + \int_0^t ds \Gamma^x(\mathbf{A}_{t-s}, s) + \mathbf{F}^R(t), \quad (\text{B4})$$

where the following relations hold

$$\langle F_k^R(t) \rangle = 0, \quad \langle F_k^R(t), F_l^R(0) \rangle = \langle \dot{A}_0^2 \rangle \Gamma_{kl}^p(t), \\ \langle F_k^R(t), A_{l,0} \rangle = 0, \quad \langle F_k^R(t), \dot{A}_{l,0} \rangle = 0, \quad (\text{B5})$$

for all $k, l = 1, 2, \dots, n$. The k -th component of the vectorial non linear memory memory friction function $\Gamma^x(\mathbf{A}, s)$ is given by

$$\Gamma_k^x(\mathbf{A}, s) = P_x L F_k^R(s) = [\nabla_A \cdot \mathbf{D}_k(\mathbf{A}, s) - \beta \mathbf{D}_k(\mathbf{A}, s) \cdot \nabla_A U_{\text{PMF}}(\mathbf{A})], \quad (\text{B6a})$$

$$\mathbf{D}_k(\mathbf{A}, s) = \langle \dot{\mathbf{A}}_0, F_k^R(s) \rangle_{\mathbf{A}}. \quad (\text{B6b})$$

Appendix C: Idempotency of the Hybrid Projection Operator

The linear operator P_H in eq. (29) is a projection, if it is idempotent, i.e., $P_H^2 = (P_p + P_x)^2 = P_H$. Clearly, we have $P_p^2 = P_p$ and $P_x^2 = P_x$. Therefore, one has to check that $P_p P_x A_t = P_x P_p A_t = 0$ for an arbitrary observable $A_t = A(\omega_t)$. This is true because of the following: 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

$$\dot{B}_0 = LB_0 = \sum_{n=1}^N \frac{\mathbf{p}_n}{m_n} \cdot \nabla_{\mathbf{r}_n} B_0. \quad (\text{C1})$$

The operator P_p 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

$$P_x P_p A_t \propto P_x \dot{B}_0 = 0, \quad (\text{C2})$$

since the operator P_x involves an integral over the particle momenta but adds no momentum dependence.

P_x maps any observable onto a function which depends on particle positions only. Since \dot{B}_0 is linearly proportional to the particle momenta, P_p applied on a function that depends on particle positions only gives zero. Therefore, it follows that

$$P_p P_x A_t = 0. \quad (\text{C3})$$

Appendix D: Self-Adjointedness and Orthogonality of Hybrid Projection

Here, we prove that the projection P_H in eq. (29) is self-adjoint w.r.t. the inner product in eq. (6), i.e., for any observables $A_t = A(\omega_t)$ and $C_{t'} = C(\omega_{t'})$, we have $\langle A_t, P_H C_{t'} \rangle = \langle P_H A_t, C_{t'} \rangle$. For this, we consider the projection operators P_p and P_x separately.

Using the definition in eq. (29c), we find

$$\langle A_t, P_p C_{t'} \rangle = \left\langle A_t, \frac{\langle \dot{B}_0, C_{t'} \rangle}{\langle \dot{B}_0^2 \rangle} \dot{B}_0 \right\rangle \quad (\text{D1a})$$

$$= \langle A_t, \dot{B}_0 \rangle \frac{\langle \dot{B}_0, C_{t'} \rangle}{\langle \dot{B}_0^2 \rangle} \quad (\text{D1b})$$

$$= \left\langle \frac{\langle A_t, \dot{B}_0 \rangle}{\langle \dot{B}_0^2 \rangle} \dot{B}_0, C_{t'} \right\rangle \quad (\text{D1c})$$

$$= \langle P_p A_t, C_{t'} \rangle. \quad (\text{D1d})$$

Using the definition in eq. (29b) and eq. (9), we find

$$\begin{aligned}
& \langle A_t, P_x C_{t'} \rangle \quad (D2a) \\
&= \left\langle A(\omega'_t), \int da \delta(B(\mathbf{R}'_0) - a) \frac{\langle \delta(B(\hat{\mathbf{R}}_0) - a), C(\hat{\omega}_{t'}) \rangle}{\mathbb{P}(a)} \right\rangle \\
&= \int da \langle A(\omega'_t), \delta(B(\mathbf{R}'_0) - a) \rangle \frac{\langle \delta(B(\hat{\mathbf{R}}_0) - a), C(\hat{\omega}_{t'}) \rangle}{\mathbb{P}(a)} \\
&= \left\langle \int da \frac{\langle A(\omega'_t), \delta(B(\mathbf{R}'_0) - a) \rangle}{\mathbb{P}(a)} \delta(B(\hat{\mathbf{R}}_0) - a), C(\hat{\omega}_{t'}) \right\rangle \\
&= \langle P_x A_t, C_{t'} \rangle. \quad (D2b)
\end{aligned}$$

This means that the hybrid projection P_H in eq. (29) is self-adjoint and thus is an orthogonal projection, i.e.,

$$\langle P_H A_t, Q_H C_{t'} \rangle = 0, \quad (D3)$$

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

Appendix E: Average of Complementary Observables Vanishes

In the following, we prove eq. (30), i.e., we show that the equilibrium average of any observable that lies in the complementary subspace at all times vanishes. For this, we must show for an arbitrary observable $A(\omega_t) = A_t$ that $\langle P_H A_t \rangle = \langle A_t \rangle$ holds. First, from the definition of P_p in eq. (29), it follows that

$$\langle P_p A_t \rangle \propto \langle \dot{B}_0 \rangle = 0, \quad (E1)$$

since our projection function $B_0 = B(\mathbf{R}_0)$ is a function of positions only and therefore, its velocity $\dot{B}_0 = LB_0$ is linear in the momenta (see eq. (C1)). For the P_x projection operator we find

$$\langle P_x A_t \rangle = \langle \langle A_t \rangle_{B_0} \rangle = \langle A(t) \rangle. \quad (E2a)$$

From this, it immediately follows that $\langle P_H A_t \rangle = \langle A(t) \rangle$ and thus all equilibrium averages in the complementary subspace vanish, i.e., $\langle Q_H A_t \rangle = \langle (1 - P_H) A_t \rangle = 0$. In particular, the equilibrium average of the random force vanishes at all times, i.e., $\langle F(t) \rangle = \langle Q_H F(t) \rangle = 0$.

Appendix F: Generalized Mass of Distance Observables

We demonstrate that the generalized mass $M(A)$ defined in eq. (33) is constant for an observable that corresponds to the scalar distance between particle positions, which is a non-linear function of particle positions. In this case, the force term dU_{eff}/dA in eq. (41) reduces to dU_{PMF}/dA . 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

$$A_0 = A(\mathbf{R}_0) = \sqrt{(\mathbf{r}_0^N - \mathbf{r}_0^O)^2}. \quad (F1)$$

Applying the Liouville operator on eq. (F1) gives the velocity of the observable

$$LA_0 = \dot{A}_0 = \left(\frac{\mathbf{p}_0^N}{m_N} - \frac{\mathbf{p}_0^O}{m_O} \right) \cdot \frac{(\mathbf{r}_0^N - \mathbf{r}_0^O)}{A_0}. \quad (F2)$$

As can be seen in eq. (F2), the velocity \dot{A}_0 is linear in the momenta \mathbf{p}_0^N and \mathbf{p}_0^O . Computing the effective mass according to the definition in eq. (33), i.e.,

$$\langle \dot{A}_0^2 \rangle_{A_0} = \frac{k_B T}{M(A_0)} = \frac{\langle \delta[A(\hat{\mathbf{R}}_0) - A_0], \dot{A}(\hat{\mathbf{R}}_0)^2 \rangle}{\langle \delta[A(\hat{\mathbf{R}}_0) - A_0] \rangle}, \quad (F3)$$

requires the computation of the numerator on the r.h.s. of eq. (F3). Given an Hamiltonian of the form in eq. (1), factorization of the phase-space integral leads to

$$\begin{aligned}
& \langle \delta[A(\hat{\mathbf{R}}_0) - A_0], \dot{A}(\hat{\mathbf{R}}_0)^2 \rangle \\
&= k_B T \left(\frac{1}{m_N} + \frac{1}{m_O} \right) \langle \delta[A(\hat{\mathbf{R}}_0) - A_0] \rangle. \quad (F4)
\end{aligned}$$

Inserting eq. (F4) for the numerator on the r.h.s. of eq. (F3), we find

$$M(A_0) = M = \frac{m_N m_O}{m_N + m_O}, \quad (F5)$$

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

$$A_0 = \frac{1}{N_R} \sum_{n=1}^{N_R} A_{n,0}, \quad (F6)$$

with $A_{n,0}$ being the initial value of the n -th distance. Eq. (F3) becomes

$$\langle \dot{A}_0^2 \rangle_{A_0} = \left\langle \left(\frac{1}{N_R} \sum_{n=1}^{N_R} \dot{A}_{n,0} \right)^2 \right\rangle_{A_0} = \frac{1}{N_R^2} \sum_{n=1}^{N_R} \langle \dot{A}_{n,0}^2 \rangle_{A_0}. \quad (F7)$$

As before, terms consisting of mixed momentum factors average to zero, only diagonal terms contribute. In analogy to eq. (F5), the effective mass is constant also for this case.

Appendix G: Alternative Discretization of Eq. (47)

Here, we present an alternative discretization of eq. (47). Similar to eq. (48), the equation derived here still has an overall error of the order $\mathcal{O}(\Delta t^2)$. The advantage over eq. (48) is that we use the trapezoidal rule for the integration involving the memory kernel $\Gamma^p(t)$; note that we keep the rectangular rule for the integration of the memory function $\Gamma^x(A, t)$. We discretize eq. (47) in the following way

$$\begin{aligned} F^R(i, j+1) &= F^R(i+1, j) \\ &+ \frac{\Delta t}{2} \Gamma^p(j) \dot{A}_{i+1} + \frac{\Delta t}{2} \Gamma^p(j+1) \dot{A}_i \\ &- \Delta t \Gamma^x(A_{i+1}, j) + \mathcal{O}(\Delta t^2). \end{aligned} \quad (\text{G1})$$

Note that now, the r.h.s. of eq. (G1) depends on $\Gamma^p(j+1)$. To compute $\Gamma^p(j+1)$, we need the yet unknown $F^R(i, j+1)$. In the absence of the memory function $\Gamma^x(A, t)$, it has been demonstrated how one can work around this problem [6, 36, 46]. The trick is to multiply eq. (G1) by $F^R(i, 0)$ and average according to eq. (6). This gives

$$\begin{aligned} \langle F^R(i, j+1), F^R(i, 0) \rangle &= \langle F^R(i+1, j), F^R(i, 0) \rangle \\ &+ \frac{\Delta t}{2} \Gamma^p(j) \langle \dot{A}_{i+1}, F^R(i, 0) \rangle + \frac{\Delta t}{2} \Gamma^p(j+1) \langle \dot{A}_i, F^R(i, 0) \rangle \\ &- \Delta t \langle \Gamma^x(A_{i+1}, j), F^R(i, 0) \rangle. \end{aligned} \quad (\text{G2})$$

By identifying the l.h.s. of eq. (G2) with $\langle \dot{A}_0^2 \rangle \Gamma^p(j+1)$ and solving for $\Gamma^p(j+1)$, we find

$$\Gamma^p(j+1) = \frac{\xi(j) + \frac{\Delta t}{2} \Gamma^p(j) \zeta - \Delta t \eta(j)}{1 - \frac{\Delta t}{2} \chi}, \quad (\text{G3a})$$

$$\xi(j) = \frac{\langle F^R(i+1, j), F^R(i, 0) \rangle}{\langle \dot{A}_0^2 \rangle}, \quad (\text{G3b})$$

$$\zeta = \frac{\langle \dot{A}_{i+1}, F^R(i, 0) \rangle}{\langle \dot{A}_0^2 \rangle}, \quad (\text{G3c})$$

$$\eta(j) = \frac{\langle \Gamma^x(A_{i+1}, j), F^R(i, 0) \rangle}{\langle \dot{A}_0^2 \rangle}, \quad (\text{G3d})$$

$$\chi = \frac{\langle \dot{A}_i, F^R(i, 0) \rangle}{\langle \dot{A}_0^2 \rangle}. \quad (\text{G3e})$$

The function $\eta(j)$ in eq. (G3d) appears due to the presence of the non-linear friction $\Gamma^x(A, t)$ and is computed using

$$\eta(j) = \int_{-\infty}^{\infty} da \frac{\langle \delta(A_{i+1} - a), F^R(i, 0) \rangle}{\langle \dot{A}_0^2 \rangle} \Gamma^x(a, j), \quad (\text{G4a})$$

$$\langle \delta(A_{i+1} - a), F^R(i, 0) \rangle = \frac{\sum_{\substack{0 \leq i \leq N_{\text{traj}}-1 \\ A_{i+1} \in I_a}} F^R(i, 0)}{\sum_{\substack{0 \leq i \leq N_{\text{traj}}-1 \\ A_{i+1} \in I_a}} 1}. \quad (\text{G4b})$$

The alternative discretization is then found by replacing $\Gamma^p(j+1)$ on the r.h.s. of eq. (G1) by eq. (G2).

Appendix H: Solving the Harmonic Hamiltonian Model Using Hybrid Projection

We derive the GLE for the harmonic Hamiltonian model eq. (49) using our hybrid projection in eq. (29). The Liouville operator defined in eq. (3) reads

$$L = L_x + L_q, \quad (\text{H1a})$$

$$L_x = \frac{p}{m} \frac{\partial}{\partial x} - \left(kx + \sum_{n=1}^N k_n (x - q_n) \right) \frac{\partial}{\partial p}, \quad (\text{H1b})$$

$$L_q = \sum_{n=1}^N \left(\frac{v_n}{m_n} \frac{\partial}{\partial q_n} - k_n (q_n - x) \frac{\partial}{\partial v_n} \right) \quad (\text{H1c})$$

and acts on the initial values $x_0, p_0, q_{n,0}, v_{n,0}$. The hybrid projection is given by

$$P_H = P_x + P_p, \quad (\text{H2a})$$

$$P_p A_t = \frac{\langle p_0, A_t \rangle}{\langle p_0^2 \rangle} p_0, \quad P_x A_t = \langle A_t \rangle_{x_0}, \quad (\text{H2b})$$

with the conditional average in eq. (H2b) being defined in eq. (9). Using eq. (H1) and eq. (H2), it follows that

$$-U'_{\text{PMF}}(x_t) = e^{tL} P_H L p_0 = -k x_t, \quad (\text{H3})$$

as confirmed in fig. 1A. To compute the random force $F^R(t)$, we use the operator expansion

$$F^R(t) = e^{tQ_H L} Q_H L p_0 = \sum_{j=0}^{\infty} \frac{t^j}{j!} (Q_H L)^j Q_H L p_0, \quad (\text{H4})$$

and repeatedly apply the operator $Q_H L$ on $Q_H L p_0 = \sum_{n=1}^N k_n (q_{n,0} - x_0)$. We find

$$\begin{aligned} F^R(t) &= \sum_{n=1}^N \left[\left(\mu_n t - \frac{(\mu_n t)^3}{3!} + \frac{(\mu_n t)^5}{5!} + \dots \right) \mu_n v_{n,0} \right. \\ &\quad \left. + k_n \left(1 - \frac{(\mu_n t)^2}{2!} + \frac{(\mu_n t)^4}{4!} + \dots \right) (q_{n,0} - x_0) \right], \end{aligned} \quad (\text{H5})$$

with $\mu_n = \sqrt{k_n/m_n}$. Identifying the sums in the parenthesis as the series expansions of sine and cosine, respectively, $F^R(t)$ follows as

$$F^R(t) = \sum_{n=1}^N (\mu_n \sin(\mu_n t) v_{n,0} + k_n \cos(\mu_n t) (q_{n,0} - x_0)). \quad (\text{H6})$$

The result in eq. (H6) equals the result given in eq. (I4b) for the same model, obtained by setting $\alpha(x) = x$, which follows by explicit solution of the equations of motion. Using $F^R(t)$ in eq. (H6) to compute $D(x, t)$, one obtains

$$D(x, t) = \left\langle \frac{p_0}{m}, F^R(t) \right\rangle_x = 0. \quad (\text{H7})$$

Hence, the memory function $\Gamma^x(x, t)$ in eq. (38) vanishes. The memory function $\Gamma^p(t)$ follows as

$$\Gamma^p(t) = \frac{\langle F^R(t), F^R(0) \rangle}{\langle p_0^2 \rangle} = \frac{1}{m} \sum_{n=1}^N k_n \cos(\mu_n t). \quad (\text{H8})$$

The friction integral in the GLE reads

$$\int_0^t ds \Gamma^p(s) p_{t-s} = \int_0^t ds \Gamma(s) \dot{x}_{t-s}, \quad (\text{H9})$$

with $\Gamma(s) = \sum_{n=1}^N k_n \cos(\mu_n s)$ being the result in eq. (I4c) obtained by explicit solution of the equations of motion for the special case $\alpha(x) = x$.

Appendix I: Derivation of a GLE for the Non-Linear Zwanzig Model

We demonstrate how to derive a GLE from the Hamiltonian version of the non-linear Zwanzig model defined in eq. (55) [53]. In the main text, we perform numerical simulations of the resulting GLE to produce trajectories on which we can test our numerical extraction techniques. The coordinates of the relevant particle are given by (x, p) , a non-linear function $\alpha(x)$ enters the coupling to the auxiliary variables $\{q_n, v_n\}$. 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 . The harmonic model defined by eq. (49) follows from the non-linear model eq. (55) in the special case $\alpha(x) = x$, so the solution of the model in eq. (49) is obtained by setting $\alpha(x) = x$ in the final results obtained in this section. As we will show here, a non-linear $\alpha(x)$ induces non-linear memory friction in the corresponding GLE.

The equations of motion for the Hamiltonian in eq. (55) read

$$m\ddot{x}_t = -U'(x_t) - \sum_n k_n \alpha'(x_t) (\alpha(x_t) - q_{n,t}), \quad (\text{I1a})$$

$$m_n \ddot{q}_{n,t} = -k_n (q_{n,t} - \alpha(x_t)), \quad (\text{I1b})$$

where the prime superscript denotes a derivative w.r.t. the argument, i.e., $U'(x) = dU/dx$. Eq. (I1b) can be solved to give

$$q_{n,t} = q_{n,0} \cos(\mu_n t) + \frac{v_{n,0}}{m_n \mu_n} \sin(\mu_n t) + \mu_n \int_0^t ds \sin(\mu_n(t-s)) \alpha[x_s], \quad (\text{I2})$$

where $\mu_n^2 = k_n/m_n$. By partial integration, the solution in eq. (I2) can be written in the form

$$q_{n,t} = (q_{n,0} - \alpha(x_0)) \cos(\mu_n t) + \frac{v_{n,0}}{m_n \mu_n} \sin(\mu_n t) - \int_0^t ds \cos(\mu_n(t-s)) \alpha'(x_s) \dot{x}_s + \alpha(x_t). \quad (\text{I3})$$

Inserting eq. (I3) into eq. (I1a) leads to a GLE for x , i.e.,

$$m\ddot{x}_t = -U'(x_t) + F_Z^R(t, x_t) - \int_0^t ds \Gamma[t-s, x_t, x_s] \dot{x}_s, \quad (\text{I4a})$$

$$F_Z^R(t, x_t) = \sum_n \alpha'(x_t) \left(\mu_n v_{n,0} \sin(\mu_n t) + k_n (q_{n,0} - \alpha(x_0)) \cos(\mu_n t) \right), \quad (\text{I4b})$$

with the memory function

$$\Gamma(t-s, x_t, x_s) = \sum_n k_n \alpha'(x_t) \alpha'(x_s) \cos(\mu_n(t-s)). \quad (\text{I4c})$$

Appendix J: Markovian Embedding of the Non-Linear Zwanzig Hamiltonian Model

Here, we show how to map the non-linear Zwanzig Hamiltonian system defined in eq. (55) onto a Markovian stochastic system of equations, for which numerical simulations can be efficiently performed. The results obtained here include the harmonic model in eq. (49) by setting $\alpha(x) = x$. Consider the memory function in eq. (I4c). It contains a sum over cosines, i.e., $\sum_{n=1}^N k_n \cos(\mu_n t)$. For $N \rightarrow \infty$, this represents a Fourier series of an even, periodic function in time with Fourier coefficients k_n . In the continuous limit, i.e., $k_n \rightarrow k(\mu) d\mu/2\pi$, this defines an arbitrary even function $f(t)$

$$\sum_{n=1}^{\infty} k_n \cos(\mu_n t) \rightarrow \int_{-\infty}^{\infty} \frac{d\mu}{2\pi} k(\mu) \cos(\mu t) = f(t) \quad (\text{J1})$$

$$= K e^{-|t|/\tau} \left(\cos\left(\frac{2\pi}{T} t\right) + c \sin\left(\frac{2\pi}{T} |t|\right) \right) \quad (\text{J2})$$

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

$$k(\mu) = \int_{-\infty}^{\infty} dt \cos(\mu t) f(t). \quad (\text{J3})$$

The memory function in eq. (I4c) becomes

$$\Gamma(t-s, x_t, x_s) = \alpha'(x_t) \alpha'(x_s) K e^{-|t-s|/\tau} \times \left(\cos\left(\frac{2\pi}{T}(t-s)\right) + c \sin\left(\frac{2\pi}{T}|t-s|\right) \right). \quad (\text{J4})$$

Now, consider the random force in eq. (I4b). For $N \rightarrow \infty$, it can be rewritten as

$$F_Z^R(t, x_t) = \alpha'(x_t) \tilde{F}_Z^R(\{q_{n,0}, v_{n,0}\}, t). \quad (\text{J5})$$

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. (55), the distribution is given by the Boltzmann distribution. Thus, the initial values $q_{n,0}, v_{n,0}$ are Gaussian distributed random variables with

$$\langle (q_{n,0} - \alpha(x_0)) \rangle = 0, \quad \langle v_{n,0} \rangle = 0, \quad (\text{J6a})$$

$$\langle \alpha(x_0), v_{n,0} \rangle = 0, \quad \langle q_{n,0}, v_{n,0} \rangle = 0, \quad (\text{J6b})$$

$$\langle v_{n,0}, v_{m,0} \rangle = \delta_{n,m} \frac{k_B T}{m_n}, \quad (\text{J6c})$$

$$\langle (q_{n,0} - \alpha(x_0)), (q_{m,0} - \alpha(x_0)) \rangle = \delta_{n,m} \frac{k_B T}{k_n}. \quad (\text{J6d})$$

From this, it follows that \tilde{F}_Z^R is a stationary Gaussian process fulfilling

$$\langle \tilde{F}_Z^R(t) \rangle = 0, \quad (\text{J7a})$$

$$\langle \tilde{F}_Z^R(t), \tilde{F}_Z^R(0) \rangle = k_B T \sum_n k_n \cos(\mu_n t) \rightarrow k_B T f(t), \quad (\text{J7b})$$

The equal sign in eq. (J7b) follows from the explicit form given in eq. (I4b) and from the relation in eq. (J6d), 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. (J4) reads

$$m\ddot{x}_t = -U'(x_t) - k\alpha'(x_t)(\alpha(x_t) - y_t), \quad (\text{J8a})$$

$$m_y\ddot{y}_t = -k(y_t - \alpha(x_t)) - \gamma\dot{y}_t + \sqrt{2k_B T\gamma}\eta(t). \quad (\text{J8b})$$

with $\langle \eta(t) \rangle = 0$, $\langle \eta(t), \eta(s) \rangle = \delta(t - s)$ being white noise. The relations between the parameters in eq. (J8) and the parameters in eq. (J1) are given by

$$\nu^2 = 4\tau_\Gamma^2 \mu^2 - 1, \quad (\text{J9a})$$

$$T = \frac{4\pi}{\nu} \tau_\Gamma, \quad \tau = 2\tau_\Gamma, \quad K = k, \quad (\text{J9b})$$

$$\tau_\Gamma = \frac{m_y}{\gamma}, \quad \mu^2 = \frac{k}{m_y}, \quad c = \frac{1}{\nu}. \quad (\text{J9c})$$

By solving eq. (J8b) and inserting the result into eq. (J8a), we find the random force

$$\begin{aligned} \tilde{F}^R(t) &= k e^{-|t|/2\tau} \left(\cos\left(\frac{2\pi}{T}t\right) + \frac{1}{\nu} \sin\left(\frac{2\pi}{T}t\right) \right) \\ &\times (y_0 - \alpha(x_0)) + \frac{2}{\gamma\nu} e^{-|t|/2\tau} \sin\left(\frac{2\pi}{T}t\right) \\ &\times p_{y,0} + \sqrt{2k_B T\gamma} \int_0^t ds 2e^{-(t-s)/2\tau} \frac{\tau}{\nu} \\ &\times \sin\left(\frac{2\pi}{T}(t-s)\right) \eta(s), \end{aligned} \quad (\text{J10})$$

where the variable y_0 has the same distribution as $q_{n,0}$, and $p_{y,0} = m_y \dot{y}_0$ has the same distribution as $v_{n,0}$. The equivalence of \tilde{F}^R in eq. (J5) and \tilde{F}_Z^R in eq. (J10) follows from the fact that their first and second moments are the same. Using this, we have mapped the non-linear Hamiltonian Zwanzig model defined by eq. (55) onto the set of coupled Markovian stochastic equations in eq. (J8), which can be used to perform numerical simulations.

Appendix K: Transformation Between Different GLEs

When applied to the non-linear Hamiltonian Zwanzig model defined by eq. (55), our hybrid projection operator $P_H = P_x + P_p$, given in eq. (H2), leads to a GLE of the form

$$\begin{aligned} \dot{p}_t &= -U'(x_t) + F^R(t) \\ &- \int_0^t ds \Gamma^p(t-s) p_s \\ &+ \int_0^t ds \Gamma^x(t-s, x_s), \end{aligned} \quad (\text{K1a})$$

$$\Gamma^p(t) = \frac{\langle F^R(t), F^R(0) \rangle}{\langle p_0^2 \rangle}, \quad (\text{K1b})$$

$$D(s, x_{t-s}) = \left\langle \frac{p_0}{m}, F^R(s) \right\rangle_{x_{t-s}}, \quad (\text{K1c})$$

where $\Gamma^x(x_{t-s}, s)$ follows from eq. (39). The two GLE's in eq. (I4a) and eq. (K1a) obviously have a different mathematical structure, but they describe the exact same dynamics. To see this, consider the random force $F_Z^R(t, x_t)$ in eq. (I4b)

$$F_Z^R(t, x_t) = \alpha'(x_t) \tilde{F}_Z^R(t), \quad (\text{K2a})$$

$$\begin{aligned} \tilde{F}_Z^R(t) &= \sum_n \left(k_n (q_{n,0} - \alpha(x_0)) \cos(\mu_n t) \right. \\ &\left. + \mu_n v_{n,0} \sin(\mu_n t) \right). \end{aligned} \quad (\text{K2b})$$

The time derivative of $F_Z^R(t, x_t)$ is given by

$$\dot{F}_Z^R(t, x_t) = \tilde{F}_Z^R(t) \frac{d}{dt} \alpha'(x_t) + \alpha'(x_t) \frac{d}{dt} \tilde{F}_Z^R(t). \quad (\text{K3})$$

Since the function $\alpha'(x_t)$ depends on time only via x_t , its time derivative can be written using the Liouville operator, i.e., $\frac{d}{dt} \alpha'(x_t) = L \alpha'(x_t)$. The same is not true for the function $\tilde{F}_Z^R(t)$. By applying the Liouville operator, we find $L \tilde{F}_Z^R(t) = \dot{\tilde{F}}_Z^R(t) - \frac{p_0}{m} \alpha'(x_0) \sum_n k_n \cos(\mu_n t)$. Hence, we can write

$$\begin{aligned} \dot{F}_Z^R(t, x_t) &= L \left[\alpha'(x_t) \tilde{F}_Z^R(t) \right] \\ &+ \sum_n k_n \alpha'(x_t) \alpha'(x_0) \frac{p_0}{m} \cos(\mu_n t) \end{aligned} \quad (\text{K4a})$$

$$\begin{aligned} &= L F_Z^R(t, x_t) \\ &+ \sum_n k_n \alpha'(x_t) \alpha'(x_0) \dot{x}_0 \cos(\mu_n t). \end{aligned} \quad (\text{K4b})$$

Eq. (K4b) is a linear, inhomogeneous first-order differential equation for $F_Z^R(t, x_t)$. The general solution reads

$$F_Z^R(t, x_t) = e^{tL} F_Z^R(0, x_0) + \int_0^t ds e^{(t-s)L} \sum_n k_n \alpha'(x_s) \alpha'(x_0) \times \dot{x}_0 \cos(\mu_n(s)) \quad (\text{K5a})$$

$$= e^{tL} F_Z^R(0, x_0) + \int_0^t ds \Gamma[t-s, x_t, x_s] \dot{x}_s, \quad (\text{K5b})$$

where we used Γ from eq. (I4c). By using the Dyson identity from eq. (19) for e^{tL} , we can write eq. (K5b) in terms of the general projection operators P and Q as

$$F_Z^R(t, x_t) = e^{tQL} F_Z^R(0, x_0) + \int_0^t ds \Gamma[t-s, x_t, x_s] \dot{x}_s + \int_0^t ds e^{(t-s)L} P L e^{sQL} F_Z^R(0, x_0). \quad (\text{K6})$$

From eq. (K2), it follows that

$$F_Z^R(0, x_0) = \alpha'(x_0) \sum_n k_n (q_{n,0} - \alpha(x_0)) \quad (\text{K7a})$$

$$= Q_H L p_0 = F^R(0), \quad (\text{K7b})$$

where we used the definition of the random force $F^R(t)$ in eq. (20b) and the equation of motion for the complementary variables. This means that $F_Z^R(0, x_0)$ coincides with the random force $F^R(t) = e^{tQL} Q L p_0$ at time $t = 0$. Therefore, by inserting the result in eq. (K6) into eq. (I4a), we obtain eq. (K1). Thus we have proven that the GLE obtained by explicitly solving the non-harmonic Hamiltonian Zwanzig model, eq. (I4a), is equivalent to the GLE obtained from our hybrid projection scheme, eq. (K1).

Appendix L: Global Data Smoothing

In the main text, we use Legendre polynomial expansions to smooth the data for the dihedral angle dynamics. For $D(A, t)$, the Legendre expansion reads

$$D(A, t) = \sum_n c_n(t) h_n(A), \quad (\text{L1})$$

where h_n denotes the Legendre polynomial of order n . The coefficients $c_n(t)$ follow from

$$c_n(t) = \frac{2n+1}{2} \int_{-1}^1 dA D(A, t) h_n(A). \quad (\text{L2})$$

In this way, the data is globally smoothed while spatial symmetries can be conserved. For example, if a function is even in A , only even orders of Legendre polynomials are used in the smoothing procedure.

-
- [1] N. G. van Kampen, Remarks on Non-Markov Processes, Brazilian Journal of Physics **28**, 90 (1998), publisher: Sociedade Brasileira de Física.
 - [2] P. Español and I. Zuniga, Force autocorrelation function in brownian motion theory, The Journal of chemical physics **98**, 574 (1993).
 - [3] L. Bocquet, J. Piasecki, and J.-P. Hansen, On the brownian motion of a massive sphere suspended in a hard-sphere fluid. i. multiple-time-scale analysis and microscopic expression for the friction coefficient, Journal of statistical physics **76**, 505 (1994).
 - [4] R. Zwanzig and M. Bixon, Hydrodynamic theory of the velocity correlation function, Physical Review A **2**, 2005 (1970).
 - [5] T. Franosch, M. Grimm, M. Belushkin, F. M. Mor, G. Foffi, L. Forró, and S. Jeney, Resonances arising from hydrodynamic memory in brownian motion, Nature **478**, 85 (2011).
 - [6] D. Lesnicki, R. Vuilleumier, A. Carof, and B. Rotenberg, Molecular Hydrodynamics from Memory Kernels, Physical Review Letters **116**, 147804 (2016), publisher: American Physical Society.
 - [7] J. O. Daldrop, B. G. Kowalik, and R. R. Netz, External Potential Modifies Friction of Molecular Solutes in Water, Physical Review X **7**, 041065 (2017), publisher: American Physical Society.
 - [8] J. E. Straub, M. Borkovec, and B. J. Berne, Calculation of dynamic friction on intramolecular degrees of freedom, Journal of Physical Chemistry **91**, 4995 (1987).
 - [9] B. Berne, M. Tuckerman, J. E. Straub, and A. Bug, Dynamic friction on rigid and flexible bonds, The Journal of chemical physics **93**, 5084 (1990).
 - [10] M. Tuckerman and B. Berne, Vibrational relaxation in simple fluids: Comparison of theory and simulation, The Journal of chemical physics **98**, 7301 (1993).
 - [11] F. Gottwald, S. D. Ivanov, and O. Kühn, Applicability of the caldeira-leggett model to vibrational spectroscopy in solution, The journal of physical chemistry letters **6**, 2722 (2015).
 - [12] S. Adelman, Generalized langevin theory for many-body problems in chemical dynamics: Reactions in liquids, The Journal of Chemical Physics **73**, 3145 (1980).
 - [13] G. Ciccotti and J.-P. Ryckaert, On the derivation of the generalized langevin equation for interacting brownian particles, Journal of Statistical Physics **26**, 73 (1981).
 - [14] E. Guàrdia and J. Padró, Generalized langevin dynamics simulation of interacting particles, The Journal of chemical physics **83**, 1917 (1985).
 - [15] L. Bocquet, J.-P. Hansen, and J. Piasecki, Friction tensor for a pair of brownian particles: Spurious finite-size effects and molecular dynamics estimates, Journal of statistical physics **89**, 321 (1997).

- [16] M. Canales and G. Sese, Generalized langevin dynamics simulations of nacl electrolyte solutions, *The Journal of chemical physics* **109**, 6004 (1998).
- [17] S. S. Plotkin and P. G. Wolynes, Non-Markovian Configurational Diffusion and Reaction Coordinates for Protein Folding, *Physical Review Letters* **80**, 5015 (1998), publisher: American Physical Society.
- [18] E. Medina, R. Satija, and D. E. Makarov, Transition Path Times in Non-Markovian Activated Rate Processes, *The Journal of Physical Chemistry B* **122**, 11400 (2018), publisher: American Chemical Society.
- [19] R. Satija and D. E. Makarov, Generalized Langevin Equation as a Model for Barrier Crossing Dynamics in Biomolecular Folding, *The Journal of Physical Chemistry B* **123**, 802 (2019), publisher: American Chemical Society.
- [20] C. Ayaz, L. Tepper, F. N. Brünig, J. Kappler, J. O. Daldrop, and R. R. Netz, Non-Markovian modeling of protein folding, *Proceedings of the National Academy of Sciences* **118**, 10.1073/pnas.2023856118 (2021), publisher: National Academy of Sciences Section: Physical Sciences.
- [21] K. Lindorff-Larsen, S. Piana, R. O. Dror, and D. E. Shaw, How fast-folding proteins fold, *Science* **334**, 517 (2011).
- [22] R. B. Best, G. Hummer, and W. A. Eaton, Native contacts determine protein folding mechanisms in atomistic simulations, *Proceedings of the National Academy of Sciences* **110**, 17874 (2013).
- [23] H. S. Chung, S. Piana-Agostinetti, D. E. Shaw, and W. A. Eaton, Structural origin of slow diffusion in protein folding, *Science* **349**, 1504 (2015).
- [24] R. Zwanzig, Memory Effects in Irreversible Thermodynamics, *Physical Review* **124**, 983 (1961).
- [25] H. Mori, Transport, Collective Motion, and Brownian Motion, *Progress of Theoretical Physics* **33**, 423 (1965).
- [26] S. Nakajima, On Quantum Theory of Transport Phenomena Steady Diffusion, *Progress of Theoretical Physics* **20**, 948 (1958).
- [27] H. Grabert, P. Hänggi, and P. Talkner, Microdynamics and nonlinear stochastic processes of gross variables, *Journal of Statistical Physics* **22**, 537 (1980).
- [28] A. J. Chorin, O. H. Hald, and R. Kupferman, Optimal prediction and the Mori–Zwanzig representation of irreversible processes, *Proceedings of the National Academy of Sciences* **97**, 2968 (2000), publisher: National Academy of Sciences Section: Physical Sciences.
- [29] T. Kinjo and S.-A. Hyodo, Equation of motion for coarse-grained simulation based on microscopic description, *Physical Review E* **75**, 051109 (2007), publisher: American Physical Society.
- [30] C. Hijón, P. Español, E. Vanden-Eijnden, and R. Delgado-Buscalioni, Mori–Zwanzig formalism as a practical computational tool, *Faraday Discussions* **144**, 301 (2010), publisher: Royal Society of Chemistry.
- [31] H. Meyer, T. Voigtmann, and T. Schilling, On the non-stationary generalized Langevin equation, *The Journal of Chemical Physics* **147**, 214110 (2017), publisher: American Institute of Physics.
- [32] E. Herrera-Delgado, J. Briscoe, and P. Sollich, Tractable nonlinear memory functions as a tool to capture and explain dynamical behaviors, *Physical Review Research* **2**, 043069 (2020).
- [33] E. Darve, Numerical Methods for Calculating the Potential of Mean Force, in *New Algorithms for Macromolecular Simulation*, Lecture Notes in Computational Science and Engineering, edited by B. Leimkuhler, C. Chipot, R. Elber, A. Laaksonen, A. Mark, T. Schlick, C. Schütte, and R. Skeel (Springer, Berlin, Heidelberg, 2006) pp. 213–249.
- [34] O. F. Lange and H. Grubmüller, Collective Langevin dynamics of conformational motions in proteins, *The Journal of Chemical Physics* **124**, 214903 (2006), publisher: American Institute of Physics.
- [35] A. Carof, R. Vuilleumier, and B. Rotenberg, Two algorithms to compute projected correlation functions in molecular dynamics simulations, *The Journal of Chemical Physics* **140**, 124103 (2014), publisher: American Institute of Physics.
- [36] G. Jung, M. Hanke, and F. Schmid, Iterative Reconstruction of Memory Kernels, *Journal of Chemical Theory and Computation* **13**, 2481 (2017), publisher: American Chemical Society.
- [37] J. O. Daldrop, J. Kappler, F. N. Brünig, and R. R. Netz, Butane dihedral angle dynamics in water is dominated by internal friction, *Proceedings of the National Academy of Sciences* **115**, 5169 (2018), publisher: National Academy of Sciences Section: Biological Sciences.
- [38] H. S. Lee, S.-H. Ahn, and E. F. Darve, The multi-dimensional generalized langevin equation for conformational motion of proteins, *The Journal of chemical physics* **150**, 174113 (2019).
- [39] V. Klippenstein, M. Tripathy, G. Jung, F. Schmid, and N. F. A. van der Vegt, Introducing Memory in Coarse-Grained Molecular Simulations, *The Journal of Physical Chemistry B* **125**, 4931 (2021), publisher: American Chemical Society.
- [40] H. Vroylandt, L. Goudenège, P. Monmarché, F. Pietrucci, and B. Rotenberg, Likelihood-based parametric estimator for memory kernel in molecular dynamics, *arXiv preprint arXiv:2110.04246* (2021).
- [41] R. F. Grote and J. T. Hynes, The stable states picture of chemical reactions. II. Rate constants for condensed and gas phase reaction models, *The Journal of Chemical Physics* **73**, 2715 (1980), publisher: American Institute of Physics.
- [42] P. Hänggi and F. Mojtabai, Thermally activated escape rate in presence of long-time memory, *Physical Review A* **26**, 1168 (1982), publisher: American Physical Society.
- [43] E. Pollak, H. Grabert, and P. Hänggi, Theory of activated rate processes for arbitrary frequency dependent friction: Solution of the turnover problem, *The Journal of Chemical Physics* **91**, 4073 (1989), publisher: American Institute of Physics.
- [44] J. Kappler, F. Noé, and R. R. Netz, Cyclization and Relaxation Dynamics of Finite-Length Collapsed Self-Avoiding Polymers, *Physical Review Letters* **122**, 067801 (2019), publisher: American Physical Society.
- [45] F. N. Brünig, P. Hillmann, W. K. Kim, J. O. Daldrop, and R. R. Netz, Proton-transfer spectroscopy beyond the normal-mode scenario, *arXiv preprint arXiv:2109.08514* (2021).
- [46] V. Klippenstein and N. F. A. van der Vegt, Cross-correlation corrected friction in (generalized) Langevin models, *The Journal of Chemical Physics* **154**, 191102 (2021), publisher: American Institute of Physics.
- [47] F. Glatzel and T. Schilling, The Interplay between Memory and Potentials of Mean Force: A Discussion on the Structure of Equations of Motion for Coarse Grained Ob-

- servables, arXiv:2107.01111 [cond-mat] (2021), arXiv: 2107.01111.
- [48] R. Zwanzig, *Nonequilibrium statistical mechanics* (Oxford UnivPress, Oxford [u.a.], 2001).
 - [49] F. J. Dyson, The Radiation Theories of Tomonaga, Schwinger, and Feynman, *Physical Review* **75**, 486 (1949), publisher: American Physical Society.
 - [50] R. P. Feynman, An Operator Calculus Having Applications in Quantum Electrodynamics, *Physical Review* **84**, 108 (1951), publisher: American Physical Society.
 - [51] D. J. Evans, *Stat. mechanics of nonequilibrium liquids*, second edition. ed. (University Press, Cambridge, 2008).
 - [52] E. Darve, J. Solomon, and A. Kia, Computing generalized Langevin equations and generalized Fokker–Planck equations, *Proceedings of the National Academy of Sciences* **106**, 10884 (2009).
 - [53] R. Zwanzig, Nonlinear generalized Langevin equations, *Journal of Statistical Physics* **9**, 215 (1973).
 - [54] S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, *et al.*, Gromacs 4.5: a high-throughput and highly parallel open source molecular simulation toolkit, *Bioinformatics* **29**, 845 (2013).
 - [55] D. Chandler, Statistical mechanics of isomerization dynamics in liquids and the transition state approximation, *The Journal of Chemical Physics* **68**, 2959 (1978), publisher: American Institute of Physics.
 - [56] J. Zheng, K. Kwak, J. Xie, and M. D. Fayer, Ultrafast Carbon–Carbon Single-Bond Rotational Isomerization in Room-Temperature Solution, *Science* **313**, 1951 (2006), publisher: American Association for the Advancement of Science.
 - [57] C. Oostenbrink, A. Villa, A. E. Mark, and W. 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, *Journal of Computational Chemistry* **25**, 1656 (2004).
 - [58] J. L. F. Abascal and C. Vega, A general purpose model for the condensed phases of water: TIP4P/2005, *The Journal of Chemical Physics* **123**, 234505 (2005).
 - [59] J.-P. Ryckaert, G. Ciccotti, and H. J. C. Berendsen, Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, *Journal of Computational Physics* **23**, 327 (1977).
 - [60] T. Darden, D. York, and L. Pedersen, Particle mesh Ewald: An $N \log(N)$ method for Ewald sums in large systems, *The Journal of Chemical Physics* **98**, 10089 (1993).
 - [61] G. Bussi, D. Donadio, and M. Parrinello, Canonical sampling through velocity rescaling, *The Journal of chemical physics* **126**, 014101 (2007).