Manuscript version: Author's Accepted Manuscript

The version presented in WRAP is the author's accepted manuscript and may differ from the published version or Version of Record.

Persistent WRAP URL:

http://wrap.warwick.ac.uk/111463

How to cite:

Please refer to published version for the most recent bibliographic citation information.
If a published version is known of, the repository item page linked to above, will contain details on accessing it.

Copyright and reuse:

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher's statement:

Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk.

Systematic derivation of hybrid coarse-grained models

Nicodemo Di Pasquale [a]

Department of Mathematics, University of Leicester, University Rd, Leicester LE1 7RH, UK[†]

Thomas Hudson

Warwick Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK[‡]

Matteo Icardi

School of Mathematical Sciences, University of Nottingham, Nottingham NG7 2RD, UK[§]
(Dated: December 3, 2018)

Molecular dynamics represents a key enabling technology for applications ranging from biology to the development of new materials. However, many real-world applications remain inaccessible to fully-resolved simulations due their unsustainable computational costs and must therefore rely on semi-empirical coarse-grained models. Significant efforts have been devoted in the last decade towards improving the predictivity of these coarse-grained models and providing a rigorous justification of their use, through a combination of theoretical studies and data-driven approaches. One of the most promising research effort is the (re)discovery of the Mori-Zwanzig projection as a generic, yet systematic, theoretical tool for deriving coarse-grained models. Despite its clean mathematical formulation and generality, there are still many open questions about its applicability and assumptions. In this work, we propose a detailed derivation of a hybrid multi-scale system, generalising and further investigating the approach developed in [Español, P., EPL, 88, 40008 (2009)]. Issues such as the general coexistence of atoms (fully-resolved degrees of freedom) and beads (larger coarse-grained units), the role of the fine-to-coarse mapping chosen, and the approximation of effective potentials are discussed. The theoretical discussion is supported by numerical simulations of a monodimensional nonlinear periodic benchmark system with an open-source parallel Julia code, easily extensible to arbitrary potential models and fine-to-coarse mapping functions. The results presented highlight the importance of introducing, in the macroscopic model, non-constant fluctuating and dissipative terms, given by the Mori-Zwanzig approach, to correctly reproduce the reference fine-grained results, without requiring *ad-hoc* calibration of interaction potentials and thermostats.

---

[a] Alphabetical order
[†] Email: ndp8@leicester.ac.uk
[‡] Email: t.hudson.1@warwick.ac.uk
[§] Email: matteo.icardi@nottingham.ac.uk

I. INTRODUCTION

Molecular Dynamics simulations (MD) have become a standard tool in many applied research areas, such as the study of biological molecules [1], soft matter [2] and, more generally, in condensed matter physics [3, 4]. Despite their widespread use, fully atomistic MD simulations are heavily constrained in terms of system size and simulation time, due to the computational power required to generate trajectories which are long enough to exhibit characteristic system behaviour. In fact, while solving the equations of motion for few atoms represents a simple task, calculating solutions for interesting systems such as proteins, involving hundreds of thousands of atoms and billions of time steps, rapidly becomes intractable. The key numerical bottleneck in large-scale simulations is the evaluation of forces, which are generally both numerically stiff and expensive to compute. Moreover, the precise evolution of the full system often contains irrelevant information. In the case of a protein molecule in water, a large part of the overall computation time is used to compute the evolution of water molecules, while the conformation of the protein is usually the most interesting aspect of the simulation, and may be described by relatively few variables.

Another problem often encountered in the simulation of long chain-like molecules is that as the size of the system increases, the relaxation time required to approach equilibrium becomes comparable with the maximum trajectory length achievable in simulations. In order to overcome these limitations, Coarse-Grained (CG) models have been developed alongside atomistic ones. Several 'coarse-graining' techniques have been established (both systematically and empirically), allowing faster and larger Molecular Dynamics (MD) simulations than those possible with fully atomistic simulations. These typically reduce the number of degrees of freedom by grouping a number of atoms into a single particle, referred to as a 'bead'. The size of such beads can range from a few atoms to entire molecules.

In this work, we focus on the Mori-Zwanzig (MZ) formalism, which is a mathematical framework which allows for systematic treatment of coarse-graining approaches. Coarse-graining an MD system requires choosing a mapping operator which describes the reduced degrees of freedom. This choice leads to an effective potential, which governs the motion of the coarse-grained variables. While this effective potential preserves various equilibrium properties of the system, it fails to preserve dynamical properties, which require a proper treatment of the fluctuations in the system. The MZ formalism provides a framework which allows us to do exactly that, and can therefore inform a choice of effective dynamics which more closely replicate the dynamical properties of the coarse-grained system.

The main objective and novelty of this work is to propose a general framework for hybrid models, explore the implications of many of the various approximations required to use the MZ formalism in practice, and provide a practical open-source implementation as a testing platform to understand and quantify the accuracy of various modelling assumptions. The methodologies and the results, obtained here for a simple one-dimensional test-case, can form the guidelines to apply the coarse-graining to more realistic chemical systems. With this in mind, below and in the conclusions we provide a summary of the specific issues and lessons learnt, while extensive computational studies on multidimensional systems are left for future works.

The the paper is organised as follows: after a literature review of the existing methods (section II), and the definition of the notation used throughout the paper (section III), we present the general formal equations which arise upon coarse-graining using the MZ formalism in section IV. We provide a detailed derivation to make this approach more accessible to non-specialists, including a discussion of all relevant derivation steps in the main text and additional details in the Supplemental Material, SM. In section V, we then present an analysis of the equations in their final form, while in section VI some of the approximations usually considered in literature are discussed. Our numerical results are presented in section VII, where these equations are implemented for a simple test case, which allows us to discuss practical algorithmic details. In section VIII we draw before drawing some conclusions and present the outlook for future studies. Before proceeding with this programme, we outline below the most important features of the MZ formalism, with a particular focus on the consequences for practical coarse-grained MD simulation.

A. Summary of results and practical consequences

The MZ formalism treats the chosen coarse-grained variables as observables of the underlying 'true' dynamics, which leads to the derivation of the MZ equations eq. (6). These govern the coarse-grained dynamics for the variables of choice, without approximations, resulting in a system that is no simpler to solve than the full dynamics. The value of this formalism, however, lies in separating out different contributions to the evolution of the coarse-grained variables, each of which can be assigned heuristic meaning, and can inform an approximation strategy. In particular, the terms in this equation fall into three important groups, which we discuss in turn.

The first term in eq. (6), after appropriate manipulations, corresponds in eq. (26) to a derivative of the *effective potential* (defined rigorously in eq. (17)), representing the *mean force* between beads and between atoms and beads. In section VI A, we argue that this quantity can be approximated by many methods, including various well-known techniques available in the literature such as Iterative Boltzmann Inversion, Force Matching or the Relative Entropy method. In all cases, the force experienced by a bead in a hybrid simulation is related to the effective potential via eq. (19), and each of these methods seek to approximate the resulting coarse-grained forces, i.e. derivatives of V^{eff}. In the numerical experiments carried out here, we use a direct way to approximate the effective potential that allows us to concurrently approximate the other terms in the MZ equation. In its current form, the direct method chosen is easily applied to a system of this size, but unlike other methods available in the literature, might not be appropriate from a computational point of view for more complex systems. Approximating the effective potential necessarily relies on some assumptions to simplify the calculation of this highly complex multidimensional integral eq. (17). These are stated explicitly and discussed in section VII B.

Regardless of the method used to approximate the effective potential, the MZ formalism provides a framework that paves the way towards a more quantitative estimation of the errors committed.

The remaining terms in eq. (6) (and its more explicit form eq. (26)), usually neglected in applied coarse-graining approaches, correspond to *memory effects* (represented by time convolution of the friction matrix eq. (9)), and a *fluctuating force*. These terms naturally appear as a way to implicitly reintroduce the effects of the degrees of freedom that have been averaged out, similar to an external 'bath'. Their overall effect is to slow beads down to retrieve the correct dynamical properties of system. In particular, the fluctuating force term represents those forces which cause deviations from the 'mean force'. Sampling fluctuating forces compatible with given coarse-grained degrees of freedom requires the resolution of a constrained (or orthogonal) dynamics. These also appear in the memory term which is computed as their time covariance, eq. (9), and represents the effects of the past history of the system on its evolution. When this covariance decays sufficiently fast in time, the memory term can be conveniently approximated as a simple friction term (see eq. (43)), and the fluctuating force can be approximated as a white noise. This desirable property makes the system Markovian, section VI C, and this approximation is generally appropriate when there is a separation of timescales between the coarse-grained variables and those neglected.

In summary, from the theoretical derivation and the numerical results, we note the following practical consequences of our results:

- So far, most coarse-grained approaches, have focused on the approximation and computation of the effective potential. While its central role is recognised here, we argue that a proper understanding of the additional fluctuations and memory terms can avoid ad-hoc fitting and modifications of the potential, keeping it uniquely defined for a given mapping.

- We observe that, to retrieve satisfactory equilibrium and dynamical properties, a simple Langevin dynamics, with constant friction and diffusion coefficients, is not appropriate even in the particularly simple numerical example we consider.

- A better description of static properties are obtained when the friction and diffusion coefficients are parametrised (similarly to the effective potential) with respect to the system configuration (e.g., positions).

- The preservation of more complex dynamical properties instead requires the strict validity of the timescale separation assumption or else a proper parametrisation and implementation of the memory term.

II. STATE OF THE ART

Any coarse-graining procedure must prescribe a means by which to compute interactions between beads, and numerous variants exist in the literature [5–7], roughly divided between structural and thermodynamic models. The former class includes models where interactions among beads are obtained by molecular structures coming from atomistic simulations. In the Iterative Boltzmann Inversion [8], an iterative procedure progressively optimises a tabulated potential in order to reproduce a known observed quantity, such as the Radial Distribution Function (RDF); a similar approach, that aims at obtaining interactions among beads using the RDF, through Monte Carlo simulations is the Inverse Monte Carlo [9]. The force-matching model [10–12] uses a least squares minimisation over forces sampled from finer scale models. In the Relative Entropy framework [13] the CG interactions are obtained by minimising an entropy function which represents the overlap between two molecular ensembles, namely the CG and the underlying atomistic one. The second class of thermodynamic models are those where thermodynamic data, either from atomistic simulations or experiments, are used to obtain interactions among beads. The approach is to decide a fixed functional form and then choosing parameters to reproduce empirical observables. The parameters thus obtained can be optimised to be valid for a generic class of chemical species as in MARTINI force field [14] or be specific for each type of molecules [15]. In all the cases these calculated interactions represent an appropriate average of the interactions felt by the group of atoms in the full atomistic model that constitute the bead. In the last few years new data-driven models have been developed which take advantage of the increased computational power available. In this particular case, machine learning algorithms are trained by using information extracted from the more detailed atomistic models [16].

Most CG approaches, however, suffer from a major flaw: even if static equilibrium properties are preserved in the passage from atomistic to CG systems, the same is not true of the dynamical properties. This effect arises from the fact that neglecting degrees of freedom during the coarse-graining results in the elimination of high-frequency fluctuating components of the force [17]. Without these, the CG system exhibits an artificially 'accelerated' dynamics which can, however, be corrected to achieve more accurate recovery of dynamical observables by adding appropriate friction terms. An example of a method to derive these terms is given in Izvekov and Voth [18], where friction coefficients are derived in the framework of the multi-scale coarse-graining method [11]. In the framework of nonlinear equations, neglecting the friction terms mentioned above has been referred to as the *first optimal description* [19, 20], and in effect, this choice evolves the mean value of the CG variables only. It has been shown that the first optimal description is accurate only for short times [21] and is therefore clear that for sampling applications, this is insufficient to appropriately predict many observables of interest.

A second necessary limitation of coarse-grained models is the inevitable price paid for the speed-up of simulations: a loss of information about the model caused by the procedure of grouping atoms into beads. If coarse-graining is carried out incoherently, the limited level of detail offered by CG models may not be enough to correctly predict the system evolution, for example due to phenomena observed only at atomistic level ruled out by the coarse-graining procedure. An example of one such phenomenon is hydrogen bonds, which require special treatment to be considered [22] in a CG system. Another example arises in the simulation of coarse-grained polymer melts, where the entanglement regime may not be properly described because of the "chain-crossability" resulting from the soft CG potentials [23]. These limitations have led to the development of hybrid atomistic/CG models which in principle should combine both a lower computational cost than fully atomistic simulations, along with improved accuracy over other CG schemes. Examples of such hybrid approaches include the recent works of Rzepiela *et al.* [24], where massless virtual sites which mediate interaction between atoms and beads were used for butane and dilanine in water; and Di Pasquale *et al.* [25, 26] for polystyrene and polyethylene melt, in which the atomistic and CG descriptions were seamlessly embedded in the same molecule. This represents an important step towards a computable and yet accurate multi-scale framework. However, it was shown that when the dynamic in hybrid models of complex systems such as polymers is considered, other effects arise. In melt, poly-ethylene chains seem to move according either to Rouse or entanglement regime depending on their resolution (i.e. on the ratio atom/beads) [27]. Therefore, a rigorous treatment of these models becomes essential.

As mentioned above, the CG approach studied here is the Mori-Zwanzig projection formalism. This is a promising approach to better understand and derive generic coarse-grained models, and due to its generality,

can be applied to a very wide class of dynamical models than many standard CG techniques, including multi-scale (or hybrid) CG/atomistic models. In 2009, Español [28] proposed a derivation of a hybrid atomistic/CG model based on the Mori-Zwanzig operator formalism, and the equations of motion were derived for a generic polymer chain represented by its centre of mass and a single, fully atomistic, molecule. This work was based on previous studies [29, 30] which proposed a derivation of a CG model in terms of equations of motion for beads starting from atomistic systems by using the Mori-Zwanzig (MZ) projection [31, 32] operator. Guenza [33] proposed the use of projection operator technique for simulations in dense melts, and this approach was later derived from first principles [34]. At least formally, the MZ projection operator allows the reduction of the dimensionality of the system under analysis in a mathematically optimal way. One of the first applications of the MZ projection operator to CG dynamics is reported in Givon *et al.* [35] where the problem of the scale separation is also addressed.

The Mori-Zwanzig formalism, as it will become clear later, has two intrinsic advantages: First, given an arbitrary and generic mapping between atomistic and CG variables, it allows a rigorous derivation of equations for the exact evolution of CG variables, including terms which may be identified as fluctuating forces and a dissipative memory (consistent with the fluctuation-dissipation theorem). Secondly, under the hypotheses used to derive the equations, it provides an explicit, rigorous means by which to compute the CG interactions, and a framework in which to quantify and progressively improve the accuracy if desired, which is the ultimate goal of any coarse-graining methodology. Despite being given in an explicit form, in practice the different terms in the CG equations of motion derived via the MZ formalism are not easy to compute in a realistic scenario, and to date, it appears no systematic way of selecting an optimal CG strategy has been developed. In recent years, there has been a significant effort towards addressing these issues and thus to render simulations using this theoretical tool more practicable. It is worth mentioning here that the MZ formalism leads to equations with the same structure [29] as those of the DPD equations [36]. With regards to the particularly cumbersome estimation of the memory terms, Li *et al.* [17, 37, 38] proposed a way to introduce them into a CG simulation by using non-Markovian DPD.

Notwithstanding these difficulties, the MZ formalism represents an appropriate formal framework for systematic coarse-graining and for better understanding the various hypotheses behind the derivation of other CG models. In this work we present a full bottom-up derivation of a general CG model appropriate for application to MD simulations, and generalise the work of Español [28] by considering the presence of an arbitrary number of beads and atoms in the same molecule, rather than assuming that each molecule is represented as a single bead. Atomistic/CG models Rzepiela *et al.* [24], Di Pasquale *et al.* [25], Español [28] are usually referred to as "hybrid". More generally, we can view them as a broad class of models in which the fully coarse-grained model and the fully atomistic models are limiting cases, with all the possible choices and combinations of atoms/beads in between. This clearly has many advantages but induces the very challenging task of choosing the appropriate (optimal) description. We believe that, a better theoretical and practical understanding of the MZ formalism represents the starting point to quantify the approximations used in CG simulations. A few other details of the derivation and modelling assumptions, partly overlooked in previous literature, are also discussed in the following. These, together with the unavoidable further approximations introduced in the practical implementations, are discussed for a simple test case (Lennard-Jones periodic chain), solved and coarse-grained through an efficient parallel open-source implementation in `Julia v0.6` [39, 40].

III. DEFINITIONS AND NOTATION

Our starting point is a full dynamical system composed by M^{FG} atoms, to which we refer as the *fine-grained* (FG) system. The term *particle* is used to indicate either an atom or larger group of atoms, called beads. We will assume that the chemical structure of the system is fixed (i.e. no reactions occur) and the connectivity of the atoms is specified at the outset, allowing us to fix a global index for the particles in the FG system. Throughout this paper, FG particles are referred to with lower-case indices, e.g. i, j. Correspondingly, phase space variables characterising the degrees of freedom of particles in the FG system will be denoted with lower case bold letters, e.g. $\mathbf{r}$ represents their positions, with $\mathbf{p}$ representing their momenta In the same spirit, we suppose that the coarse-grained (CG) description of the same system is composed of M^{CG} particles. Each FG particle is mapped to a fixed (unique) particle in the CG system. CG
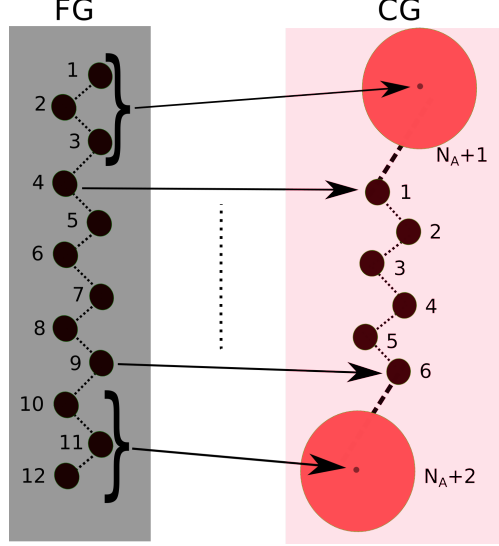
FIG. 1: Sketch of the mapping of a molecule from the FG system to the CG system where the index mapping functions between the different resolutions are highlighted. In the picture, $N^A = 6$, $M^{CG} = 8$, therefore, $S_1 = \{4\}$, $S_2 = \{5\}$, ..., $S_6 = \{9\}$, $S_7 = \{1, 2, 3\}$, $S_8 = \{10, 11, 12\}$

particles will be indexed with capital letters, e.g. I, J. Phase space variables characterising the degrees of freedom of particles in the CG system will be denoted with capital letters, e.g. $\mathbf{R}$ represents the position of the CG particles while $\mathbf{P}$ represents their momenta.

Following [12] we define S_J, the set of atoms included within the bead of index J:

$$S_J = \left\{ i \in \{1, \ldots, M^{FG}\} \text{ such that atom } i \text{ forms part of CG particle } J \right\}.$$

Having assumed that each atom belongs to a single bead, the number of atoms corresponding to CG particle J will be referred to as $s_J = \#S_J$. Given an atom with index k in the FG system, we denote the index of the CG particle containing atom k as J_k. A sketch of a possible mapping between a molecule in FG system and the corresponding molecule in CG resolution is shown in fig. 1.

A general CG (hybrid) system may include particles corresponding to single atoms (i.e. J for which $s_J = 1$), and into proper 'beads' (i.e. J for which $s_J > 1$). The number of atoms in the CG system will be denoted $N^A := \#\{J \text{ such that } s_J = 1\}$, and the number of beads is denoted $N^B := M^{CG} - N^A$. For notational convenience, we will assume that the indices are ordered in such a way that indices $J \in \{1, \ldots, N^A\}$ correspond to atoms, and $J \in \{N^A + 1, \ldots, M^{CG}\}$ correspond to beads. The above can be summarised in the following rules that apply throughout our derivation:

- Lower-case letters (e.g. i, j) refer to particles in the FG system.

- Capital letters (e.g. I, J) label the particles (atoms or beads) in the CG system.

- FG phase space variables are denoted with lower case letters:

$$\boldsymbol{\gamma} = (\boldsymbol{\gamma_r}, \boldsymbol{\gamma_p}) = (\mathbf{r}_1, \ldots, \mathbf{r}_{M^{FG}}, \mathbf{p}_1, \ldots, \mathbf{p}_{M^{FG}}).$$

- CG phase space variables (also referred to as field or effective/averaged variables [30, 33, 41]) are denoted using capital letters:

$$\boldsymbol{\Gamma} = (\boldsymbol{\Gamma_R}, \boldsymbol{\Gamma_P}) = (\mathbf{R}_1, \ldots, \mathbf{R}_{M^{CG}}, \mathbf{P}_1, \ldots, \mathbf{P}_{M^{CG}}).$$

- The mass of the jth atom in FG system will be indicated with a lower case m as m_j. The mass of the Jth particle in CG system will be indicated as M_J.

6

IV. THE MORI-ZWANZIG FORMALISM FOR HAMILTONIAN SYSTEMS

The Mori-Zwanzig (MZ) formalism is a reformulation of a dynamical system using projection operators. For application to MD, the dynamical system considered is Hamiltonian, and in this case there are two equivalent MZ approaches, akin to the Schrödinger and Heisenberg 'pictures' in the Copenhagen interpretation of Quantum Mechanics. The former defines a projection operator acting on the Hilbert space of densities of states, and has been used in Kinjo and Hyodo [30], resulting in a generalised Fokker-Planck equation for the CG variables. The latter considers a projection operator defined in the Hilbert space of all possible choices of CG variables, which may be considered observables of the FG system [42]. This has been used in Hijòn *et al.* [29], resulting in a Generalised Langevin Equation (GLE). The duality between spaces of observables and spaces of densities of states indicate the mathematical equivalence between these approaches. In what follows, we choose to follow the second approach, i.e. we write the evolution equations for the CG variables as functions of the underlying FG system, since this results in a more physically-intuitive derivation.

A. The Mori-Zwanzig equations

As stated in section III, the state of the FG system is characterised by the variables $\boldsymbol{\gamma}$ in $6M^{FG}$-dimensional phase space. The evolution of the FG system is obtained by solving Hamilton's equations:

$$\frac{\mathrm{d}\boldsymbol{\gamma}(t)}{\mathrm{d}t} = \boldsymbol{\Lambda}\frac{\partial\mathcal{H}(\boldsymbol{\gamma}(t))}{\partial\boldsymbol{\gamma}} \tag{1}$$

with initial condition $\boldsymbol{\gamma}(0) = \boldsymbol{\gamma}_0$, where $\mathcal{H}(\boldsymbol{\gamma}(t))$ is the Hamiltonian of the system and $\boldsymbol{\Lambda}$ is the $6M^{FG}\times 6M^{FG}$ symplectic matrix:

$$\boldsymbol{\Lambda} = \begin{pmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{I} & \mathbf{0} \end{pmatrix}.$$

Here, $\mathbf{I}$ is the $3M^{FG} \times 3M^{FG}$ identity matrix and $\mathbf{0}$ is a $3M^{FG} \times 3M^{FG}$ matrix of zeroes.

To derive a reduced description of this system, we consider a vector of *observables*, $\boldsymbol{\Gamma}(\boldsymbol{\gamma})$. Throughout the present section, these observables could be quite general, but we will later apply the derivation made here to the particular choice $\boldsymbol{\Gamma}(\boldsymbol{\gamma}) = (\mathbf{R}_1(\boldsymbol{\gamma}), \ldots, \mathbf{R}_{M^{CG}}(\boldsymbol{\gamma}), \mathbf{P}_1(\boldsymbol{\gamma}), \ldots, \mathbf{P}_{M^{CG}}(\boldsymbol{\gamma}))$ corresponding to the positions and momenta of particles in the CG system. It is well-known [43] that the evolution of such a function of phase variables is given by

$$\frac{\mathrm{d}\boldsymbol{\Gamma}(\boldsymbol{\gamma}(t))}{\mathrm{d}t} = \mathfrak{L}\boldsymbol{\Gamma}(\boldsymbol{\gamma}(t)), \quad \boldsymbol{\Gamma}(\boldsymbol{\gamma}(0)) = \boldsymbol{\Gamma}(\boldsymbol{\gamma}_0) \qquad \text{where} \qquad \mathfrak{L} := \sum_{i=1}^{M^{FG}} \left(\frac{\partial\mathcal{H}}{\partial\mathbf{p}_i} \cdot \frac{\partial}{\partial\mathbf{r}_i} - \frac{\partial\mathcal{H}}{\partial\mathbf{r}_i} \cdot \frac{\partial}{\partial\mathbf{p}_i} \right). \tag{2}$$

The operator $\mathfrak{L}$ is called the *Liouvillean*, and in Poisson bracket notation, may be written $\mathfrak{L} = \{\mathcal{H}, \cdot\}$. We note that each of the terms in the definition of $\mathfrak{L}$ is a dot-product. Using semigroup notation, eq. (2) has the formal solution:

$$\boldsymbol{\Gamma}(\boldsymbol{\gamma}(t)) = \mathbf{e}^{t\mathfrak{L}}\boldsymbol{\Gamma}(\boldsymbol{\gamma}_0). \tag{3}$$

Equation (3) is the exact dynamics of the observables. It is clear that this equation does not provide a self-consistent evolution equation for the observables, since in general, the evolution implicitly depends upon the full evolution of the FG system.

The Zwanzig projection [31, 42] projects the equation for the observables $\boldsymbol{\Gamma}$ onto the space of functions of the CG variables, which may be regarded as a subspace of the FG observable space [44]. The Zwanzig projection is equivalent to taking the conditional expectation [45] $\langle F \rangle_{\boldsymbol{\Gamma}} = \mathbb{E}\left[F|\boldsymbol{\Gamma}\right]$ with respect to the equilibrium distribution, for a generic function $F = F(\boldsymbol{\gamma})$ acting on the phase space and for a specific value of the coarse-grained variables $\boldsymbol{\Gamma}(\boldsymbol{\gamma}) = \boldsymbol{\Gamma}$ [46]. An alternative approach, the so-called Mori projection [32],

can be seen as a finite-dimensional linear approximation of the Zwanzig projection [47]; in certain cases it can be shown that the Mori projection and the Zwanzig projection are approximately equivalent [47].

We denote the Zwanzig projection as $\mathcal{P}_{\mathbf{\Gamma}}$:

$$\mathcal{P}_{\mathbf{\Gamma}} F(\boldsymbol{\gamma}) = \langle F \rangle_{\mathbf{\Gamma}} = \frac{1}{\Omega(\mathbf{\Gamma})} \int F(\boldsymbol{\gamma}') \, \delta\left(\mathbf{\Gamma}(\boldsymbol{\gamma}') - \mathbf{\Gamma}\right) \rho^{eq}(\boldsymbol{\gamma}') \mathrm{d}\boldsymbol{\gamma}' \tag{4}$$

where the normalisation factor, also called the *structure function* [31], represents the number of micro-states compatible with the macro-state $\mathbf{\Gamma}(\boldsymbol{\gamma}) = \mathbf{\Gamma}$:

$$\Omega(\mathbf{\Gamma}) = \int \delta\left(\mathbf{\Gamma}(\boldsymbol{\gamma}') - \mathbf{\Gamma}\right) \mu(\mathrm{d}\boldsymbol{\gamma}'). \tag{5}$$

Here, $\mu(\mathrm{d}\boldsymbol{\gamma}) = \rho^{eq}(\boldsymbol{\gamma})\mathrm{d}\boldsymbol{\gamma}$ is an equilibrium probability measure which must be specified. The conditional expectation just defined depends crucially on the choice of the underlying measure with respect to which expectations are taken. Different choices are possible and the modeller must jointly choose both the dynamics and the ensemble with respect to which they condition, as each ensemble will give slightly different forms of evolution equations for the coarse-grained system. The most common choices are usually either $\rho^{eq}(\boldsymbol{\gamma}) = \frac{1}{\mathcal{Z}} \mathrm{e}^{-\beta\mathcal{H}(\boldsymbol{\gamma})}$, i.e. the canonical ensemble, or $\rho^{eq}(\boldsymbol{\gamma}) = \delta(\mathcal{H}(\boldsymbol{\gamma}) - E)$, i.e. the micro-canonical ensemble. The rest of the derivation will be obtained by using the canonical measure as probability equilibrium measure, but it is also perfectly possible to carry out the same derivation using the micro-canonical equilibrium measure.

Let $\mathcal{Q}_{\mathbf{\Gamma}} = \mathcal{P}_{\mathbf{\Gamma}} - \mathcal{I}$ be the projection orthogonal to $\mathcal{P}_{\mathbf{\Gamma}}$ with $\mathcal{I}$ being the identity operator. Splitting the identity by writing $\mathcal{I} = \mathcal{P}_{\mathbf{\Gamma}} + \mathcal{Q}_{\mathbf{\Gamma}}$ on the right-hand side of eq. (2), and performing some manipulations [48], a Generalised Langevin Equation (GLE) [29, 31, 49] can be obtained:

$$\frac{\mathrm{d}\mathbf{\Gamma}(t)}{\mathrm{d}t} = \mathfrak{L}_{\mathcal{P}} \mathbf{\Gamma}(t) - \int_0^t \boldsymbol{\mathcal{M}}(\mathbf{\Gamma}(t-s), s) \frac{\partial}{\partial \mathbf{\Gamma}} \mathcal{S}(\mathbf{\Gamma}(t-s)) \mathrm{d}s$$
$$+ \beta^{-1} \int_0^t \frac{\partial}{\partial \mathbf{\Gamma}} \boldsymbol{\mathcal{M}}\left(\mathbf{\Gamma}(t-s), s\right) \mathrm{d}s + \boldsymbol{\mathcal{F}}_{\mathbf{\Gamma}}(t, \boldsymbol{\gamma}_0) \tag{6}$$

where $\mathfrak{L}_{\mathcal{P}} = \mathcal{P}_{\mathbf{\Gamma}} \mathfrak{L}$ is the projected Liouvillean, the application of which yields functions depending only on $\mathbf{\Gamma}$. Recalling that the underlying equilibrium measure is the canonical ensemble, the function $\mathcal{S}$ is defined to be

$$\mathcal{S} = -\frac{1}{\beta} \ln \Omega(\mathbf{\Gamma}), \tag{7}$$

and, since it has units of energy, $\mathcal{S}$ is interpreted as a contribution to the free energy of the system, in view of the definition of $\Omega(\mathbf{\Gamma})$ in eq. (5).[50] We note that in [29], the alternative quantity $\mathcal{S}(\mathbf{\Gamma}) = k_B \ln \Omega(\mathbf{\Gamma})$ is defined, which is an entropy. Both definitions are identical up to a multiplicative factor, and depend upon on the assumed temperature of the system.

The last term $\boldsymbol{\mathcal{F}}_{\mathbf{\Gamma}}(t, \boldsymbol{\gamma}_0)$ is defined as:

$$\boldsymbol{\mathcal{F}}_{\mathbf{\Gamma}}(t, \boldsymbol{\gamma}_0) = \mathrm{e}^{t\mathfrak{L}_{\mathcal{Q}}} \mathfrak{L}_{\mathcal{Q}} \mathbf{\Gamma}(\boldsymbol{\gamma}_0) \tag{8}$$

where $\mathfrak{L}_{\mathcal{Q}} = \mathcal{Q}_{\mathbf{\Gamma}} \mathfrak{L}$ is the projected dynamics operator that still depends on the FG variables $\boldsymbol{\gamma}$. The quantity $\boldsymbol{\mathcal{F}}_{\mathbf{\Gamma}}(t, \boldsymbol{\gamma}_0)$ is given by the solution of an auxiliary set of equations called the orthogonal dynamics equations [35, 51], which we briefly discuss in the following sections. It is important to highlight here that the complexity of the eq. (6) is not reduced with respect the starting point (eq. (2)), it has simply been shifted into the calculation of the orthogonal dynamics. However, it may be argued that close to equilibrium, this term is uncorrelated with $\mathbf{\Gamma}$ and therefore approximated as a random noise [52] [53].

The friction matrix $\boldsymbol{\mathcal{M}}$ is a positive definite matrix whose components can be written in Green-Kubo form. In the case where the underlying measure is the canonical ensemble, it is defined to be [54]:

$$\boldsymbol{\mathcal{M}}(\mathbf{\Gamma}, t) = \beta \, \mathcal{P}_{\mathbf{\Gamma}} \left([\mathfrak{L}_{\mathcal{Q}}\mathbf{\Gamma}] \otimes \left[\mathrm{e}^{t\mathfrak{L}_{\mathcal{Q}}} \mathfrak{L}_{\mathcal{Q}}\mathbf{\Gamma}\right]\right) = \beta \, \mathcal{P}_{\mathbf{\Gamma}} \left(\boldsymbol{\mathcal{F}}_{\mathbf{\Gamma}}(0, \cdot) \otimes \boldsymbol{\mathcal{F}}_{\mathbf{\Gamma}}(t, \cdot)\right). \tag{9}$$

where $\otimes$ represents a tensor product.

The terms involving the friction matrix turn out to be the most difficult to compute in a real calculation because of the memory effect. In this work we will focus on the practical derivation of the relevant equations for a CG atomic system including both atoms and beads. We will briefly discuss the friction matrix and the memory effects but a more thorough analysis of these terms is left for future publications. A detailed derivation of eq. (6) and eq. (9), following Hijòn *et al.* [29], is reported in the SM. We remark that this derivation is quite general, requiring solely that $\boldsymbol{\Gamma}$ is a sufficiently smooth function of $\boldsymbol{\gamma}$.

In section V we will derive more explicit forms of the terms appearing in eq. (6) for a CG system, for specific choices of the CG variables which arises naturally from physical considerations, namely canonical mappings (i.e. transformations). We will show in section VII how these formal expressions can be computed from MD simulations for a simple system represented by a one dimensional chain composed by atoms with different mass connected through a Lennard-Jones potential. The scale separation hypothesis will be discussed for this simple system.

V.   CANONICAL MAPPINGS AND COARSE-GRAINING

In the previous section, the CG variables considered were quite general. We now consider a more concrete set of choices defining a coarse-graining procedure for an atomistic system. We suppose that the Hamiltonian for the FG system takes the form

$$\mathcal{H}(\mathbf{r}, \mathbf{p}) = \sum_{i=1}^{M^{FG}} \frac{\mathbf{p}_i^2}{m_i} + U(\mathbf{r}),\tag{10}$$

where $U(\mathbf{r})$ is the interaction potential between atoms and m_i is the mass of the ith atom. As is usual for molecular models, U is typically written as a sum of 2, 3 and 4-body potentials, chosen to accurately capture interatomic interactions.

A.   A particular class of mapping operators

In the terminology of section III, the mapping operators which fix the CG variables will be observables of the FG system. These are functions $\Psi, \Phi : \mathbb{R}^{6M^{FG}} \to \mathbb{R}^{6M^{CG}}$ such that

$$\boldsymbol{\Gamma_R} = \Psi(\boldsymbol{\gamma}) \quad \text{and} \quad \boldsymbol{\Gamma_P} = \Phi(\boldsymbol{\gamma}).$$

The simplest choice for a CG mapping is to simply select the position of a representative atom from each bead, along with its conjugate momentum. This particular mapping is used in CG models where the bead is centred on one of the atoms belonging to the bead itself (e.g. see the coarse-grained model for poly-ethylene presented in [25]):

$$\mathbf{R}_J = \Psi_J(\boldsymbol{\gamma}) = \mathbf{r}_i, \quad \text{for some } i \in S_J,\tag{11}$$

$$\mathbf{P}_J = \Phi_J(\boldsymbol{\gamma}) = \mathbf{p}_i, \quad \text{for some } i \in S_J.\tag{12}$$

With this definition, the construction of the projection operator is particular straightforward: we need simply to 'integrate out' all other positions and momenta.

The class of mapping operators can be extended if, before selecting a single degree of freedom for each CG particle, we first perform a *canonical transformation*. Canonical transformations have the attractive property that they preserve the Hamiltonian structure of the model.[55] If the pair of functions $(\boldsymbol{\zeta}, \boldsymbol{\theta}) : \mathbb{R}^{6M^{FG}} \to \mathbb{R}^{6M^{FG}}$ constitute a change of variable on the FG phase space, a simple criterion for the mapping to be a canonical transformation is that it preserves the Poisson bracket, which requires that

$$\sum_{i=1}^{M^{FG}} \frac{\partial \boldsymbol{\zeta}_k(\boldsymbol{\gamma})}{\partial \mathbf{r}_i} \cdot \frac{\partial \boldsymbol{\theta}_l(\boldsymbol{\gamma})}{\partial \mathbf{p}_i} = \delta_{kl} \quad \text{and} \quad \sum_{i=1}^{M^{FG}} \frac{\partial \boldsymbol{\zeta}_k(\boldsymbol{\gamma})}{\partial \mathbf{p}_i} \cdot \frac{\partial \boldsymbol{\theta}_l(\boldsymbol{\gamma})}{\partial \mathbf{r}_i} = -\delta_{kl}$$

where δ_{kl} is the Kronecker delta. In general, we will define a class of *canonical coarse-graining operators*, being the composition of the index selection operator with a canonical transformation, i.e. for each $J = 1, \ldots, M^{FG}$, setting $\Psi_J = \zeta_i$ and $\Phi_J = \theta_i$ for some i, depending on J. If the canonical transformation involved is linear, then the entire CG mapping operator is linear, and we will call mappings in this class the *linear canonical coarse-graining operators*.

One of the most often considered choices within this class is to define the mapping operator to select the centre of mass of a CG particle as a positional coordinate, and the relevant conjugate momentum, which is simply the sum of momenta of all atoms within the bead, i.e.

$$\mathbf{R}_I = \Psi_I(\boldsymbol{\gamma}) = \frac{1}{M_I} \sum_{k \in S_I} m_k \mathbf{r}_k = \sum_{k \in S_I} c_k \mathbf{r}_k \tag{13}$$

$$\mathbf{P}_I = \Phi_I(\boldsymbol{\gamma}) = \sum_{k \in S_I} \mathbf{p}_k. \tag{14}$$

where $M_I = \sum_{k \in S_I} m_i$ and for each k, $c_k = m_k/M_I > 0$ where I is such that $k \in S_I$. This is a linear canonical transformation, and has the convenient feature that it is local: each transformed degree of freedom depends only on a small number of FG variables, all of which are close to one another in space. For a hybrid description, the CG variables which correspond to single atoms are unchanged, since $c_k = 1$ for k for which $S_I = \{k\}$.

The particular linear canonical mapping operators defined above are convenient due to their immediate physical meaning; however, it is legitimate to ask if there are other ways in which to better characterise the behaviour of beads, particularly in view of the further practical approximations we will make below. In some cases, the centre of mass of a group of atoms could preserve too little information about the system: instead, the choice of the centre of mass could be complemented with additional degrees of freedom which are nonlinear functions of the atomic positions within a bead, in order to capture rotational degrees of freedom when beads are expected to have an anisotropic structure. The price to pay for the additional detail is in more complex implementation and less intuition about the resulting equations; nevertheless, this is a direction we aim to explore further in subsequent work.

The linear canonical coarse-graining operators defined in eqs. (13) and (14) are not injective functions of the fine-grained phase space. This property means that the same centre of mass and bead momentum may be given by different combinations of atomistic positions and momenta. This can be easily shown by considering a two atoms $\{\mathbf{r}_1, \mathbf{r}_2\}$ bead, with masses $\{m_1, m_2\}$ and $m_1 \neq m_2$. The position of the centre of mass will be on the line connecting the two atoms. If the two atoms are swapped, the centre of mass will be on the same exact position, i.e. $\Psi(\mathbf{r}_1, \mathbf{r}_2) = \Psi(\mathbf{r}_2, \mathbf{r}_1)$. In vacuum the swapping will not have any effect, but in a system with finite density (i.e. in a system where atoms 1 and 2 are surrounded by other atoms) the two physical configurations will be different because, in general, the energies of the two systems will be different. The non injectivity of the mappings in eqs. (13) and (14) reflects the loss of information due to the reduction of the degrees of freedom in the coarse-graining procedure[56].

Related to this loss of information is the so-called *back-mapping problem* [57, 58], which is the problem of obtaining an inverse of the position mapping operator (eq. (13)), given some additional constraints, which can be either geometrical [59] or may involve energy minimisation [60]. We will give a brief discussion about the use of Mori-Zwanzig projection operator in the back-mapping problem when we introduce the orthogonal dynamics (see section V F).

B. Separability and effective interactions

Typically the Hamiltonian of a molecular system is expressed additively as sum of the kinetic energy and a potential energy expressed as a sum of 2, 3, and 4-body interactions. When considering a hybrid CG system, this additive decomposition allows us to rewrite U in the form

$$U(\mathbf{r}) = U^{AA}(\mathbf{r}) + U^{AB}(\mathbf{r}) + U^{BB}(\mathbf{r}) \tag{15}$$

where $U^{AA}(\mathbf{r})$ is the sum of interaction potentials between the particles in the CG system that are simply atoms, $U^{BB}(\mathbf{r})$ is the sum of interactions between atoms that are grouped into beads in the CG system, and

$U^{AB}(\mathbf{r})$ is the sum of all other interactions, i.e. interactions involving both atoms which are part of beads and atoms which are not coarse-grained.

Following Español [28], it is convenient to start by writing down explicitly the free energy of a generic atomistic/coarse-grained system. When considering a separable Hamiltonian and a CG mapping of the form given in eqs. (13) and (14), the integral can be factorised into contributions arising from the momenta of CG particles and their positions, which results in the decomposition

$$\mathcal{S}(\boldsymbol{\Gamma}) = -\frac{1}{\beta}\log\left(\int\frac{1}{\mathcal{Z}}\mathbf{e}^{-\beta\mathcal{H}(\boldsymbol{\gamma})}\delta\left(\Psi\left(\boldsymbol{\gamma}\right) - \boldsymbol{\Gamma}_{\mathbf{R}}\right)\delta\left(\Phi\left(\boldsymbol{\gamma}\right) - \boldsymbol{\Gamma}_{\mathbf{P}}\right)\mathrm{d}\boldsymbol{\gamma}\right) = \mathcal{S}_{\mathbf{P}}(\boldsymbol{\Gamma}_{\mathbf{P}}) + \mathcal{S}_{\mathbf{R}}(\boldsymbol{\Gamma}_{\mathbf{R}}).$$

Computing, we find the two terms in the latter expression then correspond to:

a. *the Kinetic Energy of the CG particles:*

$$\mathcal{S}_{\mathbf{P}}(\boldsymbol{\Gamma}_{\mathbf{P}}) = -\frac{1}{\beta}\log\left(\frac{1}{\mathcal{Z}_{\mathbf{P}}}\int\mathrm{d}\boldsymbol{\gamma}'_{\mathbf{p}}\exp\left(-\beta\sum_{i=1}^{M^{FG}}\frac{|\mathbf{p}_k|^2}{2m_k}\right)\prod_{J=1}^{M^{CG}}\delta\left(\Phi_J(\boldsymbol{\gamma}'_{\mathbf{p}}) - \mathbf{P}_J\right)\right)$$

$$= \sum_{J=1}^{M^{CG}}\frac{|\mathbf{P}_J|^2}{2M_J} + const, \tag{16}$$

b. *the Potential Energy of the CG particles:*

$$\mathcal{S}_{\mathbf{R}}(\boldsymbol{\Gamma}_{\mathbf{R}}) = -\frac{1}{\beta}\log\left(\int\mathrm{d}\boldsymbol{\gamma}'_{\mathbf{r}}\frac{1}{\mathcal{Z}_{\mathbf{R}}}\mathbf{e}^{-\beta\left[U^{AA}(\mathbf{r}') + U^{AB}(\mathbf{r}) + U^{BB}(\mathbf{r}')\right]}\prod_{J=1}^{M^{CG}}\delta\left(\Psi_J(\mathbf{r}') - \mathbf{R}_J\right)\right)$$

$$= U^{AA}\left(\{\mathbf{R}_I\}_{I=1}^{N^A}\right)$$

$$- \frac{1}{\beta}\log\left(\int\mathrm{d}\boldsymbol{\gamma}'_{\mathbf{r}}\frac{1}{\mathcal{Z}_{\mathbf{R}}}\mathbf{e}^{\left[-\beta\left(U^{AB}(\{\mathbf{R}_I\}_{I=1}^{N^A},\mathbf{r}') + U^{BB}(\mathbf{r}')\right)\right]}\prod_{J=N^A+1}^{N^B}\delta\left(\Psi_J(\mathbf{r}') - \mathbf{R}_J\right)\right)$$

$$= U^{AA}\left(\{\mathbf{R}_I\}_{I=1}^{N^A}\right) + V^{\mathrm{eff}}(\boldsymbol{\Gamma}_{\mathbf{R}}). \tag{17}$$

In the latter derivation, we have used the fact that the δ function over the atoms can be trivially applied. The integral in the second line must be intended as performed over all the coordinates of atoms belonging to some bead. The integral in the second line is also defined to be the latter term on the third line, V^{eff}. V^{eff} can be interpreted as an effective potential due to the property that if I satisfies $S_I = \{k\}$, then

$$-\frac{\partial V^{\mathrm{eff}}}{\partial\mathbf{R}_I} = \left\langle -\frac{\partial U^{AB}}{\partial\mathbf{r}_k}\right\rangle_{\boldsymbol{\Gamma}} = \langle\mathbf{F}_k\rangle_{\boldsymbol{\Gamma}} \tag{18}$$

where $\mathbf{F}_k$ is the force acting on atom k in FG system, which corresponds to particle I in the CG system. A similar result can be obtained for beads:

$$-\frac{\partial V^{\mathrm{eff}}}{\partial\mathbf{R}_J} = \left\langle -\sum_{k\in S_J}\frac{1}{c_k s_J}\frac{\partial}{\partial\mathbf{r}_k}\left(U^{AB} + U^{BB}\right)\right\rangle_{\boldsymbol{\Gamma}} = \langle\Phi_J(\mathbf{F})\rangle_{\boldsymbol{\Gamma}} \tag{19}$$

where $\Phi_J(\mathbf{F})$ is the sum of the forces acting on the atoms in bead J. An explicit derivation of these properties involves the rewriting of the integral above as a Mori-Zwanzig projection [61].

It is important to remember that this definition of V^{eff} is still formal and, in general, cannot be computed analytically as it depends on all coarse-grained variables. To make such a definition useful in practice, further approximation and assumptions will be needed, as discussed in section VI.

Following Español [28], we note that eq. (19) represents the average force induced by all the atoms in the system not included in the bead J on the CG coordinate $\mathbf{R}_J$. The former interpretation is equivalent to the

one used in the Force Matching method [12] which derives the CG interactions from the forces calculated in an atomistic simulation. In Noid *et al.* [12] the effective potential was obtained by invoking the consistency condition, which asks that the equilibrium probability density of the CG system for a certain choice of the mapping operator is equal to the equilibrium probability density of the atomistic system constrained by the mapping operator (see Eq. 20 in [12]). In our case we showed that the consistency condition is automatically included in the MZ projection operator and solving the problem for the determination of the CG potential with Force Matching means finding a suitable approximation of the V^{eff} potential. The latter identification is even clearer in [62] where a more general framework for force matching was derived. In particular, the force matching method is obtained as a projection from the FG space onto the CG space. Further discussion of possible approximations of the effective potential is provided at the end of section VI A.

C.   Projected dynamics

To explicitly write eq. (6) for the CG variables we need to first compute the dynamics of $\mathbf{\Gamma}(\boldsymbol{\gamma})$. For the position and momentum of a particle $\{\mathbf{R}_J, \mathbf{P}_J\}$ in the CG system:

$$\mathfrak{L}\mathbf{R}_J = \sum_{k \in S_J} \frac{\mathbf{p}_k}{m_k} \frac{\partial \Psi_J(\boldsymbol{\gamma_r})}{\partial \mathbf{r}_k} = \sum_{k \in S_J} \frac{\mathbf{p}_k}{m_k} \frac{m_k}{M_J} = \frac{\mathbf{P}_J}{M_J} \tag{20}$$

$$\mathfrak{L}\mathbf{P}_J = -\sum_{k \in S_J} \frac{\partial U(\boldsymbol{\gamma_r})}{\partial \mathbf{r}_k} \frac{\partial \Phi_J(\boldsymbol{\gamma_p})}{\partial \mathbf{p}_k} = -\sum_{k \in S_J} \frac{\partial U(\boldsymbol{\gamma_r})}{\partial \mathbf{r}_k} \tag{21}$$

where we have simply used the definition of Φ, Ψ as chosen in eqs. (13) and (14). We note that by making this particular choice, we obtain that $\mathfrak{L}\mathbf{R}_J$ is written in terms of the respective CG momentum, and therefore $\mathfrak{L}_{\mathcal{P}}\mathbf{R}_J = \frac{1}{M_J}\mathbf{P}_J$. For the same reason, if $\mathcal{Q}_{\mathbf{\Gamma}}$ is applied to eq. (20), the result vanishes, i.e. $\mathfrak{L}_{\mathcal{Q}}\mathbf{R}_I = 0$.

Considering the CG momenta, we have

$$\mathfrak{L}_{\mathcal{P}}\mathbf{P}_I = -\frac{\partial U^{AA}}{\partial \mathbf{R}_I} - \frac{\partial V^{\text{eff}}}{\partial \mathbf{R}_I} \qquad \text{for } I = 1, \ldots, N^A, \text{ with } S_I = \{k\},$$

$$\mathfrak{L}_{\mathcal{P}}\mathbf{P}_J = -\frac{\partial V^{\text{eff}}}{\partial \mathbf{R}_J} \qquad \text{for } I = N^A + 1, \ldots, M^{CG},$$

and therefore the orthogonal projections are

$$\mathfrak{L}_{\mathcal{Q}}\mathbf{P}_I = -\frac{\partial U^{AB}}{\partial \mathbf{r}_k} + \frac{\partial V^{\text{eff}}}{\partial \mathbf{R}_I} \qquad \text{for } I = 1, \ldots, N^A, \text{ with } S_I = \{k\},$$

$$\mathfrak{L}_{\mathcal{Q}}\mathbf{P}_J = -\sum_{k \in S_J} \frac{\partial}{\partial \mathbf{r}_k} \left(U^{AB} + U^{BB} \right) + \frac{\partial V^{\text{eff}}}{\partial \mathbf{R}_J} \qquad \text{for } I = N^A + 1, \ldots, M^{CG}. \tag{22}$$

These functions are often referred to as the *fluctuating forces*, and we therefore define for a generic time t:

$$\mathcal{F}_I(t, \cdot) = \mathbf{e}^{t\mathfrak{L}_{\mathcal{Q}}} \mathfrak{L}_{\mathcal{Q}}\mathbf{P}_I(\cdot) \qquad \text{for } I = 1, \ldots, M^{CG}. \tag{23}$$

D.   Invariants and properties of the Friction Matrix

The Friction Matrix, defined in eq. (9), represents the non-Markovian contribution to the dynamics of the CG variables: the fact that the system 'remembers' earlier states can be seen from the time convolution with all the previous states along the trajectory. Nevertheless, we can deduce some important properties of $\mathcal{M}$, which will inform the choices we make in section VI to approximate it.

As observed in section V C, when choosing a linear canonical CG operator, $\mathfrak{L}_{\mathcal{Q}}\mathbf{R} = 0$. It follows that the only non-zero terms in the friction matrix arise from the correlations between fluctuating forces $\mathcal{F}$, defined

in eq. (23). Furthermore, since $\mathfrak{L}_\mathcal{Q} \mathbf{P}_I$ is expressed in terms of a function of $\mathbf{r}$ alone (see eq. (22)), it follows that $\mathcal{M}$ must depend only on $\mathbf{R}$, and not on $\mathbf{P}$.

By using the conservation of linear momentum, which states that $\mathfrak{L}\left(\sum_{J=1}^{M^{CG}} \mathbf{P}_J\right) = 0$, we see that $\mathfrak{L}_\mathcal{Q}\left(\sum_{J=1}^{M^{CG}} \mathbf{P}_J\right) = 0$, and hence

$$\mathbf{e}^{t\mathfrak{L}_\mathcal{Q}} \mathfrak{L}_\mathcal{Q} \sum_{J=1}^{M^{CG}} \mathbf{P}_J = 0. \tag{24}$$

Using this observation and the definition of $\mathcal{M}$ (eq. (9)), we find that

$$\mathcal{M}_{II} = \beta \, \mathcal{P}_\mathbf{\Gamma} \left([\mathfrak{L}_\mathcal{Q}\mathbf{P}_I]\left[\mathbf{e}^{t\mathfrak{L}_\mathcal{Q}} \mathfrak{L}_\mathcal{Q}\mathbf{P}_I\right]\right) = -\beta \, \mathcal{P}_\mathbf{\Gamma} \left(\mathfrak{L}_\mathcal{Q}\mathbf{P}_I \, \mathbf{e}^{t\mathfrak{L}_\mathcal{Q}} \mathfrak{L}_\mathcal{Q} \sum_{K \neq I} \mathbf{P}_K\right) = -\sum_{K \neq I} \mathcal{M}_{IK}. \tag{25}$$

E.   Final system of equations

Now that every term of eq. (6) is explicitly written, the evolution for the CG system is given by solving the following system of equations:

$$\frac{\partial \mathbf{R}_K}{\partial t} = \frac{\mathbf{P}_K}{m_K}$$

$$\frac{\partial \mathbf{P}_K}{\partial t} = -\omega_{N^A K} \frac{\partial U^{AA}}{\partial \mathbf{R}_K} - \frac{\partial V^{\text{eff}}}{\partial \mathbf{R}_K} - \sum_{J=1}^{M^{CG}} \int_0^t \mathcal{M}_{KJ}(t-s,s) \frac{1}{M_J} \mathbf{P}_J(t-s) \, \mathrm{d}s$$

$$+ \beta^{-1} \sum_{j=1}^{M^{CG}} \int_0^t \mathrm{d}s \frac{\partial}{\partial \mathbf{R}_J} \mathcal{M}_{KJ}(s) + \boldsymbol{\mathcal{F}}_K(t) \tag{26}$$

where $\omega_{N^A K}$ is equal to zero whenever the difference between the first and second subscript is negative and one otherwise.

F.   Orthogonal Dynamics

The last term of eq. (26), identified in first approximation as noise terms, is function of the evolution of the orthogonal components of the projection, i.e. all the terms where the orthogonal projection $\mathcal{Q}_\mathbf{\Gamma}$ is present. In general, the noise function $v(\boldsymbol{\gamma}, t) = e^{t\mathfrak{L}_\mathcal{Q}} v(\boldsymbol{\gamma}, 0)$, is defined as the solution of the orthogonal system [35, 51]:

$$\frac{\partial}{\partial t} v(\boldsymbol{\gamma}, t) = \mathfrak{L}_\mathcal{Q} v(\boldsymbol{\gamma}, t)$$

$$v(\boldsymbol{\gamma}, 0) = \mathcal{Q}_\mathbf{\Gamma} \alpha(\boldsymbol{\gamma}) \tag{27}$$

where α is a generic function of the FG phase-space $\boldsymbol{\gamma}$, and with $\mathcal{P}_\mathbf{\Gamma} \mathcal{Q}_\mathbf{\Gamma} \alpha(\boldsymbol{\gamma}) = 0$, i.e. $\mathcal{Q}_\mathbf{\Gamma} \alpha(\boldsymbol{\gamma})$ is a function in the null space of $\mathcal{P}$. The existence of the solutions for the orthogonal dynamics equations was discussed by Givon *et al.* [51]. Since the application of $\mathcal{P}_\mathbf{\Gamma}$ and differentiation with respect to time commute, it follows that v remains in the null space of $\mathcal{P}_\mathbf{\Gamma}$ for all time. In the particular case of the CG mapping described in section V A, we find that $v_\mathbf{R}(\boldsymbol{\gamma}, 0) = 0$, whereas for momenta, $v_\mathbf{P}(\boldsymbol{\gamma}, 0) = \boldsymbol{\mathcal{F}}_I(0) = \mathfrak{L}_\mathcal{Q}\mathbf{\Gamma_P}$. By using the fact that trajectories generated by orthogonal dynamics must be in the null space of $\mathcal{P}$ at all times, we can describe the orthogonal dynamics by using constrained Hamiltonian dynamics, where the constraints are represented by the mapping functions for positions and momenta defined earlier (eqs. (13) and (14)). In particular, the mapping defined in eqs. (13) and (14) permits us to write the orthogonal equations explicitly as constrained equations [63]. The equivalence of orthogonal and constrained dynamics

13

is shown by rewriting the FG Hamiltonian with centre of mass and internal coordinates [64]. The orthogonal projection operator in this case discards the centre of mass coordinates, and the operator $\mathcal{L}_Q$ describes the evolution of the orthogonal coordinates, which can be chosen to correspond to internal coordinates in each bead. In the framework of the constrained dynamics, if we write the constrained Lagrangian for this system, we can derive a Hamiltonian, and a Liouville operator for the constrained system equivalent to $\mathcal{L}_Q$ [63]. The constraints, eqs. (13) and (14), can be written as:

$$\Psi_I(\mathbf{r}) - \mathbf{R}_I = 0$$
$$\Phi_I(\mathbf{p}) - \mathbf{P}_I = 0$$

The equation of motion for the constrained dynamics are given by:

$$\frac{\mathrm{d}\mathbf{r}_k}{\mathrm{d}t} = \mathbf{p}_k - \frac{\mathbf{P}_I}{m_I}, \qquad \frac{\mathrm{d}\mathbf{p}_k}{\mathrm{d}t} = \frac{\partial U(\mathbf{r})}{\partial \mathbf{r}_k} - \frac{m_k}{M_I}\mathbf{F}_I \qquad (28)$$

where I is the index such that $k \in S_I$, and $\mathbf{F}_I$ is given by the expressions in eq. (22) [65]. Simulating them provides a tractable approach to computing the fluctuating forces, and therefore to describe the memory kernel through the Green-Kubo relation given in eq. (9): indeed, this is exactly the approach we take in the numerical example presented in section VII.

In section V A we gave a brief introduction to the back-mapping problem, which involves finding atomistic positions which are consistent with CG variables. While a comprehensive discussion of this topic is beyond the scope of this paper, we note that the orthogonal dynamics presented here provide a way to sample the space of configurations which are consistent with the coarse-grained variables. That means that given a CG configuration, solutions of the back-mapping problem can be obtained by evolving the system under equations eq. (28), since solutions will automatically fulfil all the constraints (i.e. bonds, atoms, dihedrals), preserving the CG variables. This approach can naturally explore the whole (compatible) phase-space and find its probability density, clearly showing the intrinsic uncertainty in the back-mapping problem, instead of arbitrarily introducing ad-hoc minimisation of the constraints or the energy.

VI. DISCUSSION

In this section, we discuss the relationship between the derivation and modelling choices presented above, and other approaches in the literature, as well as the Markovian approximation, an approximation which renders the MZ equations eq. (26) more tractable in practice.

A. Approaches to the Effective Potential

The Mori-Zwanzig projection of the equations of motion for coarse-grained variables leads to the definition of the effective potential V^{eff} in eq. (19). This is, in general, a complex function on a high-dimensional space since it depends on all coarse-grained variables, and is therefore difficult to compute. In section VII B we will discuss about approximations often (sometimes implicitly) used to simplify it and make it practically computable. While it is not our intention to propose a new technique to compute the effective potential, here we wish to highlight the fact that the effective potential derived in the MZ framework, and those used in other CG approaches are all closely related.

The effective potential eq. (19) can be connected with the widely used concept of the Potential of Mean Force (PMF) [66], $w^{(n)}$, which represents the potential that returns the average force acting on an atom j integrated over all the configurations of the atoms $n+1, \ldots, M^{FG}$ while atoms $1, \ldots, n$ are kept fixed [67]:

$$-\frac{\partial w^{(n)}}{\partial \mathbf{r}_j} = \frac{\int \mathrm{d}\mathbf{r}'_{n+1} \cdots \mathrm{d}\mathbf{r}'_{M^{FG}} \left(-\frac{\partial U(\boldsymbol{\gamma}_\mathbf{r})}{\partial \mathbf{r}_j}\right) e^{-\beta U(\mathbf{r})}}{\int \mathrm{d}\mathbf{r}'_{n+1} \cdots \mathrm{d}\mathbf{r}'_{M^{FG}} e^{-\beta U(\mathbf{r})}} = \frac{\int \mathrm{d}\boldsymbol{\gamma}'_\mathbf{r} \left(-\frac{\partial U(\boldsymbol{\gamma}'_\mathbf{r})}{\partial \mathbf{r}_j}\right) e^{-\beta U(\boldsymbol{\gamma}'_\mathbf{r})} \prod_{i=1}^{n} \delta\left(\mathbf{r}'_i - \mathbf{r}_i\right)}{\int \mathrm{d}\boldsymbol{\gamma}'_\mathbf{r} e^{-\beta U(\boldsymbol{\gamma}'_\mathbf{r})} \prod_{i=1}^{n} \delta\left(\mathbf{r}'_i - \mathbf{r}_i\right)}. \qquad (29)$$

14

This can be easily identified with the effective potential defined in eq. (19) when using the mapping functions defined in eqs. (11) and (12) (i.e., CG positions identified by selected atoms):

$$-\frac{\partial V^{\text{eff}}}{\partial \mathbf{r}_j} = \frac{\int d\boldsymbol{\gamma}'_{\mathbf{r}} \left(-\frac{\partial U(\boldsymbol{\gamma}'_{\mathbf{r}})}{\partial \mathbf{r}_j}\right) e^{-\beta U(\boldsymbol{\gamma}'_{\mathbf{r}})} \prod_{i=1}^n \delta\left(\mathbf{r}'_i - \mathbf{r}_i\right)}{\int d\boldsymbol{\gamma}'_{\mathbf{r}} e^{-\beta U(\boldsymbol{\gamma}'_{\mathbf{r}})} \prod_{i=1}^n \delta\left(\mathbf{r}'_i - \mathbf{r}_i\right)} = -\frac{\partial w^{(n)}}{\partial \mathbf{r}_j} \tag{30}$$

However, this equivalence is only valid for a particular choice of the mapping function in CG system. We can therefore define a *Generalised Potential of Mean Force*, $W^{(M^{CG})}$ (GPFM), averaging forces on atoms, while using a more general arbitrary mapping function as a constraint. If we consider the force acting on the atom J_k averaged over the configuration of all the other atoms, keeping fixed the centres of mass of M^{CG} beads, we will obtain the following expression for the GPFM

$$-\frac{\partial W^{(M^{CG})}}{\partial \mathbf{r}_{J_k}}\Bigg|_{\mathbf{R}_1,\dots,\mathbf{R}_{M^{CG}}} = \frac{\int d\boldsymbol{\gamma}_{\mathbf{r}} \left(-\frac{\partial U(\mathbf{r})}{\partial \mathbf{r}_{J_k}}\right) e^{-\beta U(\mathbf{r})} \prod_{I=1}^{M^{CG}} \delta\left(\Psi_I - \mathbf{R}_I\right)}{\int d\boldsymbol{\gamma}_{\mathbf{r}} e^{-\beta U(\mathbf{r})} \prod_{I=1}^n \delta\left(\Psi_I - \mathbf{R}_I\right)}, \tag{31}$$

where J_k maps the general index in the FG space to the index in the CG space. By using the same symbols defined in section III for the number of atoms in the group J, s_J, we divide both side of eq. (31) by $s_J c_{J_k}$, where c_{J_k} is defined in eq. (13), and summing over s_J we obtain the relation between the GPMF and the effective potential:

$$\frac{\partial W^{(M^{CG})}}{\partial \mathbf{R}_J} = \sum_{k=1}^{s_J} \frac{1}{s_J c_{J_k}} \frac{\partial W^{(M^{CG})}}{\partial \mathbf{r}_{J_k}} = -\frac{\int d\boldsymbol{\gamma}_{\mathbf{r}} \sum_{k=1}^{s_J} \left(-\frac{1}{s_J c_{J_k}} \frac{\partial U(\hat{r})}{\partial \mathbf{r}_{J_k}}\right) e^{-\beta U(\mathbf{r})} \prod_{I=1}^{M^{CG}} \delta\left(\Psi_I - \mathbf{R}_I\right)}{\int d\boldsymbol{\gamma}_{\mathbf{r}} e^{-\beta U(\mathbf{r})} \prod_{I=1}^n \delta\left(\Psi_I - \mathbf{R}_I\right)} = \frac{\partial V^{\text{eff}}}{\partial \mathbf{R}_J} \tag{32}$$

where we have used the fact that $\frac{\partial \mathbf{R}}{\partial \mathbf{r}} = \frac{\partial \Psi}{\partial \mathbf{r}} = c$, and, with a slight abuse of notation, W can be therefore thought as a function of $\mathbf{r}$ (through the map Ψ) or $\mathbf{R}$. The definition of GPMF given above can be easily extended to the case where only a subset, $\{1, \dots, n\}$, say, of the M^{CG} centres of mass are fixed, and a further averaging is performed over all configurations of the remaining $\{n+1, \dots, M^{CG}\}$ centres of mass. In this case we obtain:

$$-\frac{\partial W^{(n)}}{\partial \mathbf{r}_{J_k}}\Bigg|_{\mathbf{R}_1,\dots,\mathbf{R}_n} = \frac{\int d\mathbf{R}_{n+1} \cdots d\mathbf{R}_{M^{CG}} \int d\boldsymbol{\gamma}_{\mathbf{r}} \left(-\frac{\partial U(\hat{r})}{\partial \mathbf{r}_{J_k}}\right) e^{-\beta U(\mathbf{r})} \prod_{I=1}^{M^{CG}} \delta\left(\Psi_I - \mathbf{R}_I\right)}{\int d\mathbf{R}_{n+1} \cdots d\mathbf{R}_{M^{CG}} \int d\boldsymbol{\gamma}_{\mathbf{r}} e^{-\beta U(\mathbf{r})} \prod_{I=1}^n \delta\left(\Psi_I - \mathbf{R}_I\right)} \tag{33}$$

It is important to highlight here that eq. (33) is no more directly related to the effective potential which requires that all CG variables are constrained. The GPMF equals the effective potential only in the case where $n = M^{CG}$. In view of the physical interpretation of the GPMF along with eq. (32) we argue that the GPMF with $n < M^{CG}$ is the n-body approximation to the full GPMF and therefore to the full effective potential. Being the effective potential V^{eff}, and similarly the PMF, generally a function of all CG variables, is usually compute it in an approximate form [68], with the most common approach being to assume a multi-body expansion. In practical coarse-graining applications this typically reduces to writing the effective potential as a sum of pairwise and three-body potentials.

One way to achieve this is by employing the quantity $W^{(2)}$, that leads to an extension of the Theorem of Reversible Work [67]. Following eq. (33), define the 2-body distribution function

$$\rho^{(2)}(\mathbf{R}_1, \mathbf{R}_2) = M^{CG}(M^{CG} - 1)\frac{\int d\mathbf{R}_3 \cdots d\mathbf{R}_{M^{CG}} \int d\boldsymbol{\gamma}_{\mathbf{r}} e^{-\beta U(\mathbf{r})} \prod_{I=1}^{M^{CG}} \delta\left(\Psi_I - \mathbf{R}_I\right)}{\int d\mathbf{R}_1 \cdots d\mathbf{R}_{M^{CG}} \int d\boldsymbol{\gamma}_{\mathbf{r}} e^{-\beta U(\mathbf{r})} \prod_{I=1}^{M^{CG}} \delta\left(\Psi_I - \mathbf{R}_I\right)} \tag{34}$$

which represents a joint distribution probability for finding a bead in the position $\mathbf{R}_1$ and any other bead at position $\mathbf{R}_2$. From eq. (34), the Radial Distribution Function (RDF) for a homogeneous and isotropic coarse-grained system can be defined as in [67]:

$$g(|\mathbf{R}|) = \frac{\rho^{(2)}(0, \mathbf{R})}{\rho^2}, \qquad \text{where} \qquad \rho := \rho^{(1)}(\mathbf{R}_1) = \frac{M^{CG}}{V}, \tag{35}$$

where V is the volume of the box containing the beads, so in other words, $\rho^{(1)}$ is the density of beads.

If we sum the forces on the atoms in the bead 1 averaged over all configurations of the beads $\{3, \ldots, M^{CG}\}$ keeping fixed the positions of the beads 1 and 2 we obtain:

$$\sum_{k=1}^{s_1} -\frac{1}{s_1 c_{1_k}} \frac{\partial W^{(2)}}{\partial \mathbf{r}_{1_k}}\bigg|_{\mathbf{R}_1, \mathbf{R}_2} = -\frac{\partial W^{(2)}}{\partial \mathbf{R}_1}\bigg|_{\mathbf{R}_1, \mathbf{R}_2}$$

$$= \frac{\int d\mathbf{R}_3 \cdots d\mathbf{R}_{MCG} \int d\boldsymbol{\gamma} \sum_{k=1}^{s_k} \left(-\frac{1}{s_1 c_{1_k}} \frac{\partial U(\mathbf{r})}{\partial \mathbf{r}_k}\right) e^{-\beta U(\mathbf{r})} \prod_{I=1}^{MCG} \delta\left(\Psi_I - \mathbf{R}_I\right)}{\int d\mathbf{R}_3 \cdots d\mathbf{R}_{MCG} \int d\boldsymbol{\gamma} e^{-\beta U(\mathbf{r})} \prod_{I=1}^{MCG} \delta\left(\Psi_I - \mathbf{R}_I\right)}$$

$$= -\beta^{-1} \frac{\int d\mathbf{R}_3 \cdots d\mathbf{R}_{MCG} \int d\boldsymbol{\gamma} e^{-\beta U(\mathbf{r})} \prod_{I=2}^{MCG} \delta\left(\Psi_I - \mathbf{R}_I\right) \sum_{k=1}^{s_1} \left(\frac{1}{s_1 c_{1_k}} \frac{\partial \Psi_1}{\partial \mathbf{r}_k}\right) \frac{\partial}{\partial \Psi_1} \delta\left(\Psi_1 - \mathbf{R}_1\right)}{\int d\mathbf{R}_3 \cdots d\mathbf{R}_{MCG} \int d\boldsymbol{\gamma} e^{-\beta U(\mathbf{r})} \prod_{I=1}^{MCG} \delta\left(\Psi_I - \mathbf{R}_I\right)}$$

$$= \beta^{-1} \frac{\partial}{\partial \mathbf{R}_1} \ln \int d\mathbf{R}_3 \cdots d\mathbf{R}_{MCG} \int d\boldsymbol{\gamma} e^{-\beta U(\mathbf{r})} \prod_{I=1}^{M^{CG}} \delta\left(\Psi_I - \mathbf{R}_I\right)$$

$$= \beta^{-1} \frac{\partial}{\partial \mathbf{R}_1} \ln \left(\frac{M^{CG}(M^{CG}-1) \int d\mathbf{R}_3 \cdots d\mathbf{R}_{MCG} \int d\boldsymbol{\gamma} e^{-\beta U(\mathbf{r})} \prod_{I=1}^{MCG} \delta\left(\Psi_I - \mathbf{R}_I\right)}{\int d\mathbf{R}_1 \cdots d\mathbf{R}_{MCG} \int d\boldsymbol{\gamma} e^{-\beta U(\mathbf{r})} \prod_{I=1}^{MCG} \delta\left(\Psi_I - \mathbf{R}_I\right)}\right) = \beta^{-1} \frac{\partial}{\partial \mathbf{R}_1} \ln g(|\mathbf{R}|)$$

where we used the properties of the Dirac delta and the fact that $\sum_{i=1}^{s_I} \frac{1}{s_I c_{I_k}} \frac{\partial \Psi_I}{\partial \mathbf{r}_i} = 1$ (which follows from eq. (13)).

The expression obtained above shows interesting connection with some CG techniques. The Iterative Boltzmann Inversion (IBI) algorithm [69] is based on the Henderson theorem [70] that relates the radial distribution function with the pair potential that reproduces this RDF. The IBI algorithm calculates coarse-grained potential by refining a potential function obtained as an inversion of the expression of the potential of mean force for two beads. The iterations usually starts with the following initial guess:

$$u_0 = -\frac{1}{\beta} \ln g \quad \text{and} \quad u_{k+1} = u_k + \frac{1}{\beta} \ln \left(\frac{g_k}{g}\right) \tag{36}$$

where g_k is the radial distribution function obtained at the k-th iteration and u_0 is the potential at the 0-th iteration, which in our language represents the two-body approximation of the effective potential V^{eff}. It was shown [71] in the context of the fixed point iteration theory, that iteration in eq. (36) is well defined. Hanke [71] proved that $\|u_{k+1} - u\|_{\mathcal{V}_u} \leq C \|u_k - u\|_{\mathcal{V}_u}$ if u_k is sufficiently close to u, where $C > 0$, u is the true pair potential and the norm $\|\cdot\|_{\mathcal{V}_u}$ is defined over the Banach space $\mathcal{V}_u$ of the perturbations of the potential u. The initial iteration of the IBI technique therefore uses the quantity $W^{(2)}(\mathbf{r})$ as first guess. As we have previously identified, this quantity is an approximation of the effective potential. It can be shown that this potential of mean force is a Lennard-Jones type potential and for this reason the first iteration in IBI technique is well defined. We conclude this section by considering another method used to derive effective CG interaction, the Relative Entropy (RE) framework introduced by Scott Shell [13] and reformulated as a stochastic optimisation problem by Bilionis and Zabaras [72]. This is based on using the Kullback-Leibler divergence [73] to define a distance metric to estimate how "close" two probability distributions are. Given a parameter-dependent CG configurational distribution, the RE method defines the effective CG interactions by minimising the distance, i.e. the entropy, between the CG distribution and the real FG distribution. Relative entropy has the intuitive property that it is zero when the two distributions represent the same one. The CG interaction that minimises the relative entropy is exactly the effective potential V^{eff} [74], which arises naturally from the MZ projection techniques, so that the RE framework is another way to obtain the V^{eff} potential. It was shown [74] that the IBI and Force Matching methods (which we briefly discussed in section V B) are equivalent within the Relative Entropy framework. As we have argued, this result is expected, since the main objective of all these methods is to find a suitable approximation to the effective potential. In particular, the latter line of reasoning entails that the CG techniques just described construct approximations of the many body potential of mean force [75]. Our identification of the effective potential as a potential of mean force within the GPMF allows us to merge the two concepts of potential of mean force and effective potential, identifying the latter as the only quantity of interest.

We want to finally highlight here that, in the systematic bottom-up CG with MZ, the best CG approximation is given by the effective potential V^{eff} and a reliable coarse-grained model crucially requires a good approximation of this quantity, together with the additional fluctuating and dissipative/memory terms. The importance of V^{eff} lies in the fact that it formally represents the exact projected dynamics derived from *first principles* without any calibration. As we have seen above, several methods have been proposed to compute alternative forms of the effective potentials, mostly relying on *a-posteriori* efficient calibration against reference FG simulations. Therefore, all the other approximations (e.g., neglecting memory and fluctuating terms, sampling errors, time-stepping errors, etc.) are implicitly affecting this procedure, and the resulting CG models are guaranteed to recover only the quantities and scenarios used for calibration. For other observables, instead, a quantitative estimation of the errors is not generally available. This is only partially overcome by the MZ approach which however relies on computing additional terms.

B. Hybrid systems and the virtual sites approximation

To simplify the implementation of atom-bead interactions, we now review a method recently used by Di Pasquale *et al.* [25]. The expression in eq. (18) represents the force acting on each atom in the hybrid system due to all the other atoms that are being coarse-grained. In hybrid systems, it may be that atoms which are not coarse-grained into beads can nevertheless be assigned a *virtual bead*.

Supposing this procedure is carried out, then in accordance with the definitions given in section III, we introduce the total number of the virtual sites in the system M^{VS}, which is the total number of CG and virtual beads, and in analogy with the procedure laid out in section III, set

$$S_\mu^{VS} = \{I \in \{1, \dots, M^{CG}\} \text{ such that } I \text{ is and atom and } I \text{ is in the } \mu\text{th virtual site}\},$$
$$\text{and } \mu \in \{1, \dots, M^{VS}\}. \qquad (37)$$

The virtual site notation uses Greek letters to highlight the fact that virtual site index is not the same as the CG index. We assume that all real and virtual beads are disjoint, and the number of elements in each set of specific atoms S_μ^{VS} will be indicated as $\sigma_\mu = \#S_\mu^{VS}$.

Defining virtual sites allows us to obtain a simpler form for the mixed interaction between atoms and beads. Summing all forces acting on all atoms belonging to a virtual bead eq. (18), divided by $c_{\mu_k}\sigma_\mu$, we obtain:

$$\sum_{i=1}^{\sigma_\mu} -\frac{\partial V^{\text{eff}}}{\partial \mathbf{R}_{\mu_i}} \frac{1}{c_{\mu_i}\sigma_\mu} = \left\langle \sum_{i=1}^{\sigma_K} -\frac{\partial U^{AB}}{\partial \mathbf{r}_{\mu_i}} \frac{1}{c_{\mu_i}\sigma_\mu} \right\rangle_{\mathbf{\Gamma}} \qquad (38)$$

where in the above equation we used the linearity of the projection $\mathcal{P}$. Comparing with eq. (19), the forces arising from other beads acts on the group of atoms as if they were a proper bead. A virtual site represents the expected location of the atom inside a virtual site had they been coarse-grained. In such a system we will write the interaction between beads and virtual sites as:

$$-\frac{\partial V^{\text{eff}}}{\partial \mathbf{R}_\mu^{VS}} = \left\langle \sum_{i=1}^{\sigma_\mu} \left(-\frac{\partial U^{AB}}{\partial \mathbf{r}_{\mu_i}} \frac{1}{c_{\mu_i}\sigma_\mu} \right) \right\rangle_{\mathbf{\Gamma}} = \left\langle \mathbf{F}_\mu^{VS} \right\rangle_{\mathbf{\Gamma}} \qquad (39)$$

In the above equation we wrote a way to average the force on a group of atoms in the same way used for beads. The only difference with eq. (19) is the presence of the contribution of the potential term involving the interactions among beads and atoms U^{AB} only. The importance of this identification lies in the fact that, if the virtual beads are identical to a subset of the true beads, there is no need to perform additional calculation. Once the interaction between beads is obtained, we need a way to map the forces between beads back to atoms forming virtual beads. In this case we can use again the mapping operator for the virtual site $\Psi_\mu^{VS}(\mathbf{r})$ obtaining for the kth atom inside the virtual bead

$$-\frac{\partial V^{\text{eff}}}{\partial \mathbf{R}_{\mu_k}} = -\frac{\partial V^{\text{eff}}}{\partial \mathbf{R}_\mu^{VS}} \frac{\partial \mathbf{R}_\mu^{VS}}{\partial \mathbf{R}_{\mu_k}} = -\frac{\partial V^{\text{eff}}}{\partial \mathbf{R}_\mu^{VS}} \frac{\partial \Psi_\mu^{VS}(\mathbf{r})}{\partial \mathbf{r}_{\mu_k}} = c_{\mu_k} \left\langle \mathbf{F}_\mu^{VS} \right\rangle_{\mathbf{\Gamma}}. \qquad (40)$$

17

The force acting on the kth atom belonging to the μth virtual site is the force acting on the virtual site weighted by the coefficient c_{μ_k} which depends on the mapping operator. The result shown in eq. (40) was used in [25] without clear justification other that the *a posteriori* correctness of final results. Here, it is shown how this assumption can be justified within the MZ framework. The virtual sites approach avoids the computation of an extra interaction between beads and atoms by using the interaction between beads already computed, thus reducing the computational overhead.

C. Markovian approximation of the Memory Kernel

A good description of the memory term represents probably the biggest challenge in the derivation of the correct dynamic for a coarse-grained model. The problems faced in order to give a proper description of this term are twofold: the calculation of the friction matrix, which in its formal derivation depends from all the fast degrees of freedom coarse-grained away (i.e. the orthogonal dynamics, see eq. (9) and section V F) and the time integral of which the friction matrix is one of the arguments, that needs to be evaluated at each time step (see second term on the RHS eq. (6)). Different ways to consider the friction matrix were reported [19, 47] as well as to include memory effects in CG simulations, [37, 76–78].

One way to reduce the complexity of the equations of motion given in eq. (26) is to approximate the memory kernel by assuming the system is 'memoryless', i.e. the fluctuating forces at a given time have no correlation with those at previous times. This assumption is typically valid if the timescale over which the CG variables change is much longer than the timescale over which the fluctuating forces vary [79]. If the characteristic timescale for the evolution of the CG variables is τ_R, and the timescale over which the components of the friction matrix decay is τ_D, then if $\tau_R \gg \tau_D$, then $\boldsymbol{\mathcal{M}}(t)$ can be considered as 'memoryless'. This is the Markovian approximation, which assumes that

$$\boldsymbol{\mathcal{M}}(\boldsymbol{\Gamma}, t) \approx \tilde{\boldsymbol{\mathcal{M}}}(\boldsymbol{\Gamma})\delta(t). \tag{41}$$

If $\boldsymbol{\mathcal{M}}(\boldsymbol{\Gamma}, t)$ exhibits exponential decay in t at a sufficiently large rate uniformly on the phase space, i.e. $\boldsymbol{\mathcal{M}}$ takes the form

$$\boldsymbol{\mathcal{M}}(\boldsymbol{\Gamma}, t) = \lambda \mathbf{e}^{-\lambda t} f(\boldsymbol{\Gamma}, t)$$

where f is bounded in time, then the integrals involving the memory kernel may be approximated using Laplace's method as

$$\int_0^t \boldsymbol{\mathcal{M}}(\boldsymbol{\Gamma}(t-s), s)\frac{\partial}{\partial \boldsymbol{\Gamma}}\mathcal{S}(\boldsymbol{\Gamma}(t-s))\,\mathrm{d}s + \beta^{-1}\int_0^t \frac{\partial}{\partial \boldsymbol{\Gamma}}\boldsymbol{\mathcal{M}}\left(\boldsymbol{\Gamma}(t-s), s\right)\mathrm{d}s$$
$$= \boldsymbol{\mathcal{M}}(\boldsymbol{\Gamma}(t), 0)\frac{\partial}{\partial \boldsymbol{\Gamma}}\mathcal{S}(\boldsymbol{\Gamma}(t)) + \frac{1}{\beta}\frac{\partial}{\partial \boldsymbol{\Gamma}}\boldsymbol{\mathcal{M}}\left(\boldsymbol{\Gamma}(t), 0\right) + O(\lambda^{-1}).$$

In the above approximation, λ should be viewed as a measure of the timescale separation, i.e. $\lambda = \tau_R/\tau_D \gg 1$. In this case, $\tilde{\boldsymbol{\mathcal{M}}}$ is the time independent friction matrix defined as:

$$\tilde{\boldsymbol{\mathcal{M}}}(\boldsymbol{\Gamma}) := \boldsymbol{\mathcal{M}}(\boldsymbol{\Gamma}, 0) = \beta \left\langle \boldsymbol{\mathcal{F}}_{\boldsymbol{\Gamma}}(0, \cdot) \otimes \boldsymbol{\mathcal{F}}_{\boldsymbol{\Gamma}}(0, \cdot) \right\rangle_{\boldsymbol{\Gamma}}. \tag{42}$$

We note that

$$\langle \boldsymbol{\mathcal{F}}_I(0, \cdot)\boldsymbol{\mathcal{F}}_J(0, \cdot)\rangle_{\boldsymbol{\Gamma}} = \langle \mathfrak{L}_{\mathcal{Q}}\mathbf{P}_I \otimes \mathfrak{L}_{\mathcal{Q}}\mathbf{P}_J\rangle_{\boldsymbol{\Gamma}} = \left\langle \left(\frac{\partial V^{\mathrm{eff}}}{\partial \mathbf{R}_I} - \sum_{i \in S_I}\frac{\partial U}{\partial \mathbf{r}_i}\right)\left(\frac{\partial V^{\mathrm{eff}}}{\partial \mathbf{R}_J} - \sum_{j \in S_J}\frac{\partial U}{\partial \mathbf{r}_j}\right)\right\rangle_{\boldsymbol{\Gamma}}, \tag{43}$$

which is a function of $\boldsymbol{\Gamma_R}$ alone, and therefore derivatives of $\tilde{\boldsymbol{\mathcal{M}}}$ with respect to $\mathbf{P}_J$ vanish in eq. (44). It is also clear that the approximate friction matrix is symmetric, i.e. $\tilde{\mathcal{M}}_{IJ} = \tilde{\mathcal{M}}_{JI}$.

The validity of the Markovian approximation has been repeatedly questioned in literature, in particular for system where the timescale of the sound propagation is comparable with the timescale of the evolution

of the CG variables [80–82]. In [83, 84] it was argued that the Markovian approximation may fail to be valid for a chain of oscillators interacting with nonlinear potential, such as a chain of Lennard-Jones particles, but in more complicated systems, such as those treated in more realistic MD simulations, it may be considered valid.

Assuming the approximation eq. (41) is valid, the equations of motion for the atomistic/CG system (eq. (26)) become:

$$\frac{\partial \mathbf{P}_K}{\partial t} = -\omega_{N^A K} \frac{\partial U^{AA}}{\partial \mathbf{r}_K} - \frac{\partial V^{\text{eff}}}{\partial \mathbf{R}_K} - \sum_{J=1}^{M^{CG}} \tilde{\mathcal{M}}_{IJ} \frac{\mathbf{P}_J}{M_J} + \beta^{-1} \sum_{J=1}^{M^{CG}} \frac{\partial}{\partial \mathbf{R}_J} \tilde{\mathcal{M}}_{KJ}(s) + \boldsymbol{\mathcal{F}}_K(t) \tag{44}$$

We will make use of the last equation in the following section when we will apply the theory to a simple system.

If, instead of $\tilde{\mathcal{M}}_{IJ}$, we further average over the positions $\boldsymbol{\Gamma}_\mathbf{R}$ of CG particles, $\langle \cdots \rangle$, we could simplify eq. (44) again, by noticing that $\langle \tilde{\mathcal{M}}_{KJ} \rangle$ does not depend on $\mathbf{R}$ anymore:

$$\frac{\partial \mathbf{P}_K}{\partial t} = -\omega_{N^A K} \frac{\partial U^{AA}}{\partial \mathbf{r}_K} - \frac{\partial V^{\text{eff}}}{\partial \mathbf{R}_K} - \sum_{J=1}^{M^{CG}} \langle \tilde{\mathcal{M}}_{KJ} \rangle \frac{\mathbf{P}_J}{M_J} + \boldsymbol{\mathcal{F}}_K(t). \tag{45}$$

In this case, we have obtained the usual Langevin dynamics for the momentum, where a friction constant which is uniform in particle positions is usually assumed.

We want to stress here that we use the Markovian approximation only as a means by which to simplify the calculation; more refined methods to calculate the friction matrix, and we will consider them and consider memory effect exist, and will be considered in future work.

VII. NUMERICAL EXPERIMENTS

Having derived the general structure of equations of motion for a generic CG system, we now numerically investigate the effects of some of the choices and assumptions made when a FG system is coarse-grained, along with their conceptual and practical implications. With this aim in mind, our choice is to consider a simple one-dimensional system with only one type of bead: by studying such a simple system, we hope to better understand which of our assumptions are valid, and which might fail in a more complex system. Moreover, we focus on the differences between a fully coarse-grained system and its fine-grained version, i.e. we do not consider a hybrid system. Although similar systems have been studied before [83], our aim is to carefully consider and test the errors made in the various approximations required and their parametric dependence in order to better understand their validity.

The particular system we consider is a periodic chain of M^{FG} atoms of two different species. The two species are assumed to have different masses and stiffnesses, in a repeating pattern: such a system might be viewed as a toy model for a binary alloy. The system is coarse-grained by combining single repeating units into beads. Following the choices made in section V, the CG variables are taken to be the centres of mass of each bead and the corresponding momenta. The number of beads, M^{CG}, will therefore be equal to the number of times the pattern is repeated in the chain. An open-source parallel code, written in Julia 0.6.2, has been developed to solve the full, constrained and coarse-grained dynamics in the model described and, more importantly, to compute all of the components needed for the CG system. The implementation we have developed allows for a relatively straightforward recalibration of the test system via Julia 'types' which are initialised by the user to fix all parameters of the simulation, including the repeating pattern of the beads, the number of beads, and the mass and stiffnesses of the interatomic potentials. Dynamics are implemented using symplectic or pseudo-symplectic schemes to maximise accuracy at low computational cost, and the sampling algorithms employed take advantage of parallelisation to maximise the efficiency of the code.

The particular test cases under study here involve atoms with two different masses, $m = 1$ and $M = 10$, and the repeating pattern has been chosen to consist of three atoms in two different configurations. We distinguish four cases, denoted as A, B, C and D. In the first two cases, each bead contains two atoms of mass M and a single atom of mass m. In the second two cases, each bead is composed of two atoms of

mass m and a single bead of mass M. The one-dimensional chain in all the configurations considered, along with their coarse-grained version is sketched in fig. 2. The difference between systems A and B, C and D is given by the different interactions among the atoms, as will be explained in detail when we will introduce the specific form of the inter-atomic potential. Throughout, the number of beads will generally be set to $M^{CG} = 10$; clearly, by fixing the number of beads we fix the number of atoms also, as $M^{FG} = 3M^{CG} = 30$.
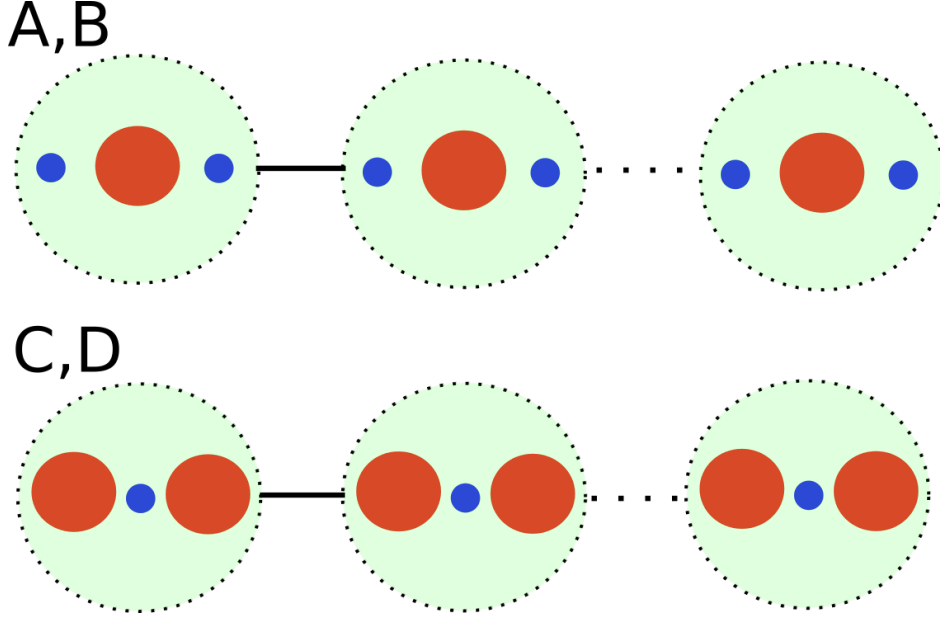
FIG. 2: Sketch of the one-dimensional chain and the two chosen configurations. The two species are indicated as red and blue particles. Small blue spheres represent atoms of mass m and big, red spheres represent atoms of mass M. The beads are indicated as a dotted line surrounding three atoms.

The interaction potential between atoms is chosen to be a simple 12-6 Lennard-Jones potential:

$$U_{i,i+1}(r) = 4\epsilon_{i,i+1}\left(\left(\frac{\sigma_{i,i+1}}{r}\right)^{12} - 2\left(\frac{\sigma_{i,i+1}}{r}\right)^{6}\right). \tag{46}$$

We choose $\sigma_{i,i+1} = 1$ so that the minimum of the potential is attained at $r^* = 1$, and the well-depth $\epsilon_{i,i+1}$ will be varied to analyse the effect of the stiffness of the interactions. We choose three values for this parameter $\{0.1, 1, 100\}$; the potentials $U_{i,i+1}(r)$ describing the mixed interaction among two different atoms are fixed by choosing the parameters $\epsilon_{i,i+1}$ which obey the Lorentz-Berthelot mixing rule. The sequence of well-depth values for all the system is then given by

$$\epsilon_{i,i+1} = \begin{cases} \sqrt{\epsilon^a \epsilon^b} & i = 1, 2, 4, 5, \ldots, M^{FG} - 2, M^{FG} - 1, \\ \epsilon^a \ (\text{or } \epsilon^b) & i = 3, 6, \ldots, M^{FG}. \end{cases} \tag{47}$$

where the superscripts a, b indicate two different values for ϵ taken from the set $\{0.1, 1, 100\}$. These choices are paired with the different mass patterns to give 4 distinct systems. The pairs A and C, and B and D have the same 'internal' interactions (i.e. interactions among atoms belonging to the same bead) and same inter-bead interactions (i.e. interactions among atoms in different beads): a summary of the four cases is reported in table I. In this way, we seek to isolate the effects of the different beads and mapping on the properties investigated.

Letting $\mathbf{p}_i$ and $\mathbf{r}_i$ respectively be the momentum and position of the ith atom, the Hamiltonian for system

	Interactions ($\epsilon_{i,i+1}$)			Bead structure
	$m-M$	$M-M$	$m-m$	
A	10	100	1	m-M-m
B	1	.1	10	m-M-m
C	10	1	100	M-m-M
D	1	10	.1	M-m-M

TABLE I: Summary of parameters in the 4 configurations considered.

$X = A, B, C, D$ reads

$$\mathcal{H}^X(\mathbf{r}, \mathbf{p}) = \frac{1}{2}\mathbf{p}^T(M^X)^{-1}\mathbf{p} + \sum_{i=1}^{M^{FG}-1} U_{i,i+1}^X\left(\mathbf{r}_{i+1} - \mathbf{r}_i\right) + U_{M^{FG},1}^X(\mathbf{r}_1 - \mathbf{r}_{M^{FG}} + r^* M^{FG}), \qquad (48)$$

where the mass matrix M^X and pair potentials $U_{i,i+1}^X$ have been chosen with the relevant parameters. The final term reflects the coupling of the right-hand end of the chain to the left-hand end, where we have chosen a total chain length of $r^* M^{FG} = 30$; other fixed volumes could be selected by making other choices.

As mentioned above, in all cases, the dynamics were implemented using symplectic schemes, and the time step chosen was $\Delta t = 10^{-3}$; we note that $\Delta t \leq \min \sqrt{M/U_{i,i+1}''(1)} \approx 5.89 \times 10^{-3}$, the shortest timescale in the system. Below, we present the results of simulating different dynamics: for convenience, we term these dynamics:

(FGD) The *Fine-Grained Dynamics* when solving eq. (1) with the Hamiltonian defined in (48);

(OD) the *Orthogonal Dynamics* when solving eq. (28);

(DCGD) the *Deterministic Coarse-Grained Dynamics* when solving eq. (44) neglecting the dissipative and fluctuating terms; and

(MMZD) the *Markovian Mori-Zwanzig Dynamics* when solving the full eq. (44), with the memory and fluctuating terms replaced by a position-dependent[85] Langevin dynamics.

To numerically integrate (MMZD), we have used a version of the BAOAB integrator [86, 87] adapted to the case of non-constant diffusion.

A. Sampling algorithm

In order to compare the predictions of an approximation of the MZ projection to the true dynamics, we must calculate an approximation of the effective potential, given in eq. (17), and the friction matrix, as given in eq. (9). Here we report the algorithm we used to generate the results that will be used for the discussion. The basic idea is to sample using the OD.

When the mapping operator onto CG variables is chosen to be a canonical transformation, the OD has the property that it remains Hamiltonian, and therefore preserves energy, sampling an appropriate marginal of the NVE ensemble. Moreover, in the case of a *linear* canonical transformation, the OD takes the explicit form given in eq. (28); these equations may be integrated using a leapfrog scheme, just as for the FGD.

We first fix an initial condition for the simulation, along with the corresponding value for the Hamiltonian $\mathcal{H}(\mathbf{p}, \mathbf{r}) = E$. A trajectory of the FGD is then computed using a leapfrog scheme, and samples of fine-grained positions are recorded. For each sample, CG positions and momenta are extracted, and a longer trajectory of the OD is then run with the sample positions and momenta as initial condition. Along trajectories of the OD, time averages of the first and second moments of the effective forces between beads are stored. The first quantity gives the value of the force acting between adjacent beads given the positions of the centre of

mass, while the latter is used to compute the friction matrix. Here, we implicitly assume that (28) generates an ergodic dynamics on the relevant marginal of the NVE ensemble; practically speaking, this assumption cannot be verified, but is commonly assumed in MD simulations.

Since the OD sampling step can be run concurrently with no communication between processes, it is ideal for parallelisation: each time average provides a value for the effective force between beads given a distance between centres of mass of adjacent beads.

The method we used to calculate the effective potential is very similar to the force-matching method [10]. However, as mentioned above, here we *directly* derive the effective potential from the sampling performed during OD simulations, without prescribing a functional form for the CG potential, and therefore have no need of minimisation techniques. Using the constrained dynamics eq. (22), we sample coefficients of the macroscopic dynamics, using the following algorithm:

1: **procedure** CONSTRAINED SAMPLE(Configuration of atomistic system, target$_1$, target$_2$)
2:     Generate initial condition for fine-grained system
3:     **while** # of samples $<$ target$_1$ **do**
4:         Run fine-grained dynamics to generate initial condition for constrained dynamics
5:         **while** length of constrained dynamics trajectory $<$ target$_2$ **do**
6:             Compute timestep using integrator for constrained dynamics
7:             Update statistical estimators
8:         **end while**
9:     **end while**
10: **end procedure**

B.   Data-driven approximation of the effective potential and memory kernel

Additional assumptions must be made in order to enable practical calculation; we highlight each such assumption in order to understand whether we may consistently extrapolate from the conclusions we draw for this simple system to a more realistic situation.

Firstly, we make an *equivalence assumption* for the interaction between beads, i.e., using the symmetry of the system, we sample interactions between similar beads concurrently. In our case, this assumption is valid as every bead is identical and, since they lie in a periodic arrangement, are equivalent by translation. We can therefore calculate the interaction between any combination of beads in the same arrangement relative to one another and assume that the interaction computed is valid for any equivalent pair in the system.

Secondly, we make *parametric assumptions*, i.e. we choose the form of the coarse-grained variables, which leads to a choice of parametrisation of the effective potential. The choice of the variables is often intuitive and completely natural (inter-particle distances, angles, etc.), but currently there exist no systematic way (apart from purely data-driven approaches) to find an 'optimal' representation of the potential.

In many approaches, a third *functional assumption* is made, in which the functional form of the potential is chosen, and optimised within some parameter space. In the following test case, we make no functional form assumptions, leaving the functional form free, and we therefore avoid any intrinsic bias in our method.

The last and most significant simplification is a *sparsity assumption*: not only do we make a specific choice of coarse-grained variable, we also assume that the effective potential may be approximated as a sum of n-body interactions between beads, where n is of a lower dimension than the dimension of the coarse-grained system. In this test case, we compute only 2-body interactions based on the distance to the 'nearest-neighbour' beads only, and use this as an approximation to the full effective potential, neglecting all other variables. These assumptions mean that we drastically reduce the dimensionality of the space which we need to sample to compute the effective potential, from the total number of coarse-grained variables to simply one, i.e., the inter-bead distance. This inevitably simplifies the exploration of the entire constrained phase space, as well as the representation and interpolation of the approximate potential in the parameter space, and thus improves the computation time required to adequately sample the mean force. Clearly, it is possible, at the expense of further computational effort and more sophisticated interpolation methods, to improve the effective potential by accounting for multi-body interactions.
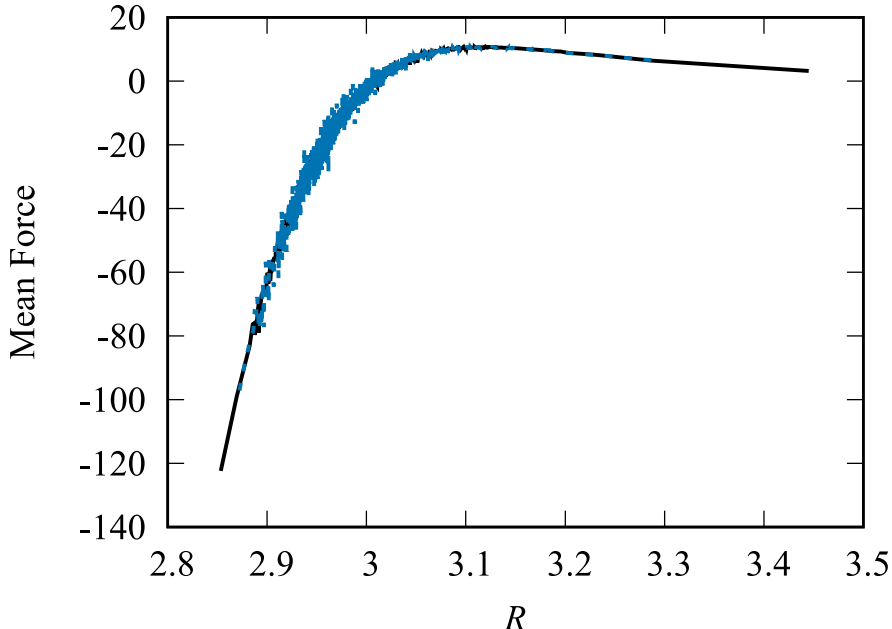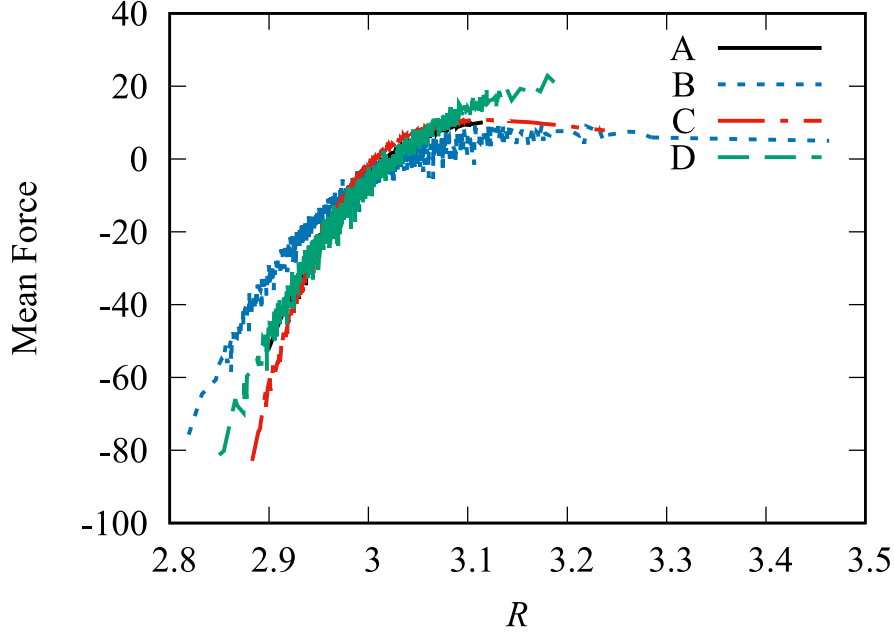
FIG. 3: Plot of the mean force for the system C. Mean force calculated with all the beads constrained (dotted blue line), and mean force calculated with a single bead constrained while the rest of the system evolve with the full dynamics (solid black line).

It is important to highlight here that, although in most previous studies these approximations were not explicitly identified, they are crucial to make any coarse-graining method applicable. In principle, any of the above assumptions could cause a significant deterioration in the accuracy and effectiveness of a systematic MZ projection approach. For example, in realistic systems, even chemically identical beads do not fully satisfy the equivalence assumption as they occupy different relative positions within a larger molecule. An overly sparse approximation of the effective potential could fail to appropriately penalise certain configurations, resulting in a molecule which is too flexible, or fluctuates too much. In spite of the caveats above, we are able to illustrate the validity of both the equivalence and sparsity assumptions in this test case. In fig. 3, the mean force for system C is shown, calculated by constraining a single pair of beads; superimposed are samples of the mean force between two particles calculated by constraining all beads in the system. The pair potential clearly accounts for the majority of the effective force, while significantly reducing sampling error, and in this case therefore appears a reasonable and statistically-stable approximation.

In fig. 4, we report the results of the computation of the mean force between beads computed via OD simulations in each of the four test cases A-D. The main behaviour of each curve depends from the parameters used in the description of the interactions (see table I), however we can find some expected similarities. In each case, we observe that the equilibrium bond length lies close to $R = 3$, which corresponds to both the minimum of the $U_{i,i+1}$ potential for three consecutive atoms, as well as the value expected due to the volume constraint: changing the density of the system would result in different minima for the effective potential.

While the mean force observed for the four systems are qualitatively similar, the fluctuating part of the force exhibits more significant differences, shown in fig. 5. Even in such a simple system, all the four chains A-D have a clear dependence of the fluctuations on the inter-bead distance R, demonstrating that a simple Langevin dynamics (with white noise fluctuations and dissipation having a magnitude independent of the state of the system) is not an adequate approximation. The fluctuations for cases B and D are noisier and of a much greater magnitude than that in systems A and C, which appears to be a reflection of the increased stiffness of the bonds between beads in the latter cases. This would seem to result in greater fluctuations in the force for similar interatomic displacements, and leads us to hypothesis that the mixing timescale for the OD may be determined by the bead mass divided by the stiffness of inter-bead interactions. It is interesting

23

FIG. 4: Plot of the mean force for the four systems considered: system A (black solid line), B (dotted blue line), C (red dashed-dotted line), D (dashed green line). The black line is partially covered by, and mostly overlapping with the blue ones.

to notice that, despite these differences, the corresponding effective potentials look instead very similar. This means that, in practical applications, even when the effective potential is easily and stably computed, significant dissipative effects might be present that are not visible in the deterministic effective potential. Another interpretation of the presence of larger fluctuations in the cases B and D is that we have made a 'worse' choice of mapping, and parametrisation of the potential; in view of eq. (6), a memory of larger magnitude means the effective potential accounts less of the force felt by the beads.

The integration of the mean force (fig. 4) with respect the distance R between the beads leads to the effective potential V^{eff}.

In fig. 6, the effective potential is plotted: we note that oscillations in the mean force are averaged through by the integration. In practice, we use a kriging-type interpolation algorithm [88] to obtain a smooth approximation of the mean force directly from the values shown in fig. 4. This approach allows us to compute forces at arbitrary values of the inter-bead distance, and simultaneously filters the noise resulting from sampling error [89].

If we assume that the forces on bead I can be decomposed as independent 'stress' contributions $\sigma_{I+1,I}$ and $-\sigma_{I,I-1}$, due to the interaction with bead $I-1$ and bead $I+1$, then

$$\mathcal{M}_{II}(\mathbf{R}) = \mathbb{E}[\mathcal{F}_I \otimes \mathcal{F}_I | \mathbf{R}] = \mathbb{E}[(\sigma_{I+1,I} - \sigma_{I,I-1}) \otimes (\sigma_{I+1,I} - \sigma_{I,I-1}) | \mathbf{R}]$$
$$= \mathbb{E}[\sigma_{I+1,I} \otimes \sigma_{I+1,I} | \mathbf{R}] + \mathbb{E}[\sigma_{I,I-1} \otimes \sigma_{I,I-1} | \mathbf{R}];$$

on the latter line we have assumed that the cross terms are negligible, i.e. [90]:

$$\mathbb{E}[\sigma_{I+1,I} \otimes \sigma_{I,I-1} | \mathbf{R}], \ \mathbb{E}[\sigma_{I,I-1} \otimes \sigma_{I+1,I} | \mathbf{R}] \approx 0.$$

This assumption is natural in view of the hypothesis that pair interactions are sufficient to capture the behaviour of the system, and thus fluctuating forces between adjacent beads are not correlated. Similarly, we obtain

$$\mathcal{M}_{I,I+1}(\mathbf{R}) = \mathbb{E}[\mathcal{F}_I \otimes \mathcal{F}_{I+1} | \mathbf{R}] = \mathbb{E}[(\sigma_{I+1,I} - \sigma_{I,I-1}) \otimes (\sigma_{I+2,I+1} - \sigma_{I+1,I}) | \mathbf{R}]$$
$$= -\mathbb{E}[\sigma_{I+1,I} \otimes \sigma_{I+1,I} | \mathbf{R}].$$

24
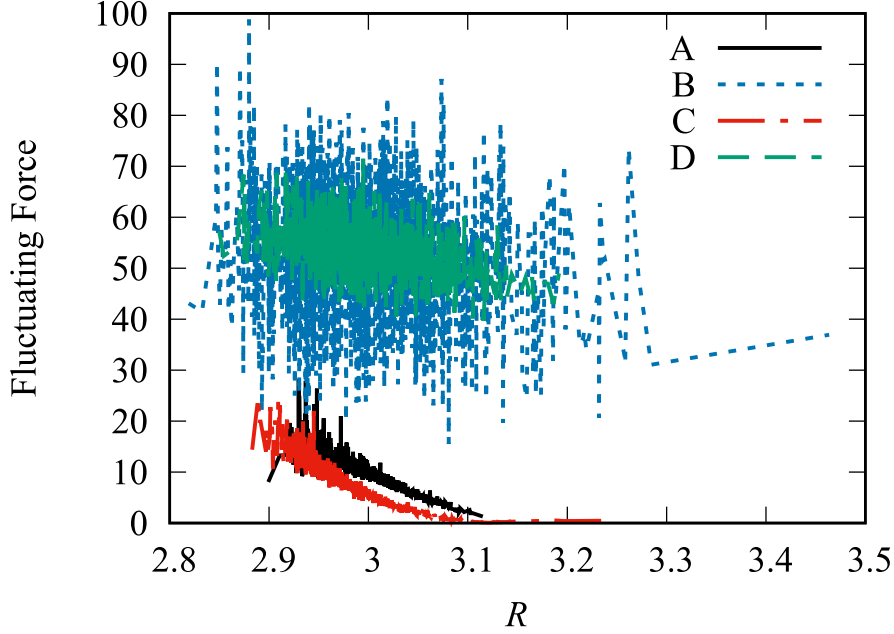
FIG. 5: Plot of the fluctuating force for the four systems considered: system A (black solid line), B (dotted blue line), C (red dashed-dotted line), D (dashed green line).

Given the interparticle distances $\mathbf{R}_{I+1} - \mathbf{R}_I$, we compute a smooth approximation of the variance of the fluctuating force between beads, i.e.

$$\gamma_{I+1,I} := \mathbb{E}[\sigma_{I+1,I} \otimes \sigma_{I+1,I} | \mathbf{R}]$$

again using the same kriging-type interpolation algorithm as for the effective potential. We then define the tridiagonal matrix

$$\Sigma_{IJ} = \begin{cases} -\gamma_{I+1,I} & I = J - 1 \mod M^{CG}, \\ \gamma_{I+1,I} + \gamma_{I,I-1} & I = J, \\ -\gamma_{I,I-1} & I = J + 1 \mod M^{CG}. \end{cases} \tag{49}$$

The fluctuating forces and memory terms may then be generated together by computing a realisation of the stochastic dynamics

$$-\tfrac{1}{2}\Sigma(\mathbf{R}_t)M^{-1}\mathbf{P}_t \, dt + \Sigma(\mathbf{R}_t)^{1/2} \, d\mathbf{W}_t, \tag{50}$$

where $\Sigma(\mathbf{R})^{1/2}$ is the positive definite square root of the matrix $\Sigma(\mathbf{R})$, and M is the matrix of bead masses, and $d\mathbf{W}_t$ denotes an increment of a Brownian motion in $\mathbb{R}^{M^{CG}}$. In practice, a realisation of this dynamics is computed in a similar way to the BAOAB scheme described in [86].
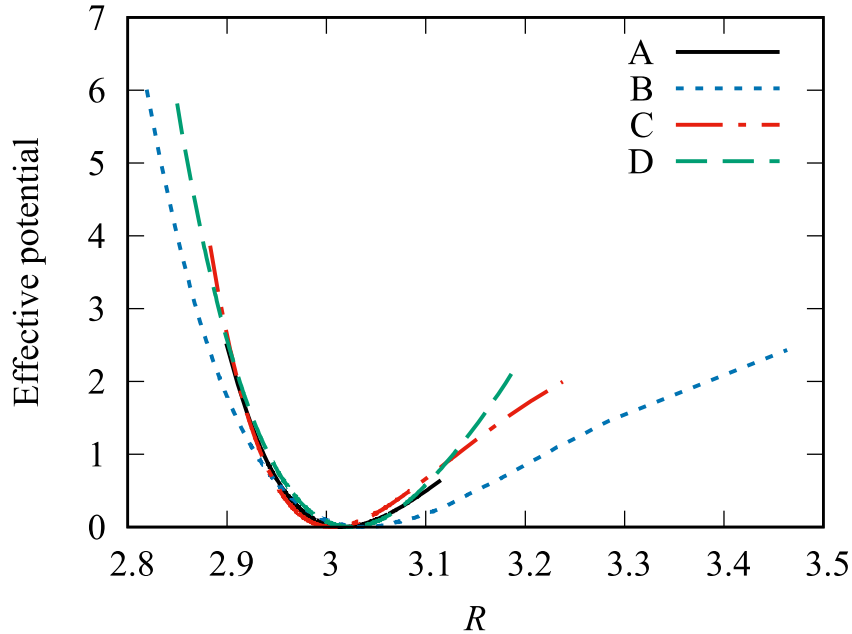
FIG. 6: Plot of the effective potential for the four systems considered: system *A* (black solid line), *B* (dotted blue line), *C* (red dashed-dotted line), *D* (dashed green line).

C. Timescale separation and the validity of the memory approximation

Sampling the fluctuating force allows us to calculate the components of the friction matrix by using eq. (42) that will be used to simulate the MMZD. The usual physical justification for eq. (42) relies on the assumption that the decay of the time of the correlation of the fluctuating force is much smaller that the relaxation time of the system [29]. These two quantities can be estimated by using the fluctuating force auto-correlation function (FFACF) and the velocity auto-correlation functions for the beads (VACF) calculated via a FGD simulation.

These autocorrelation functions for the model for the system *C* are plotted in fig. 7.

The results are distinct from those shown in [83] for a chain of Lennard-Jones particles, as the timescale separation is less distinct. According to the usual rule-of-thumb, which states that the velocity autocorrelation function should decay more slowly than the autocorrelation function of the fluctuating forces, it should follow that there is no clear scale-separation which fully justifies the Markovian approximation. Possible explanations for the lack of scale separation are that the system analysed here is much smaller than system shown in [83], and the bead size chosen here is also much smaller.

Nevertheless, we note that the autocorrelations of both forces and momenta in the MMZD (represented by blue dashed curves in fig. 7) lack oscillations, and indeed, this agrees with the significantly weaker oscillations observed in the autocorrelations computed from the FGD (shown in black). This suggests that while a distinct timescale separation is not evident from the autocorrelation function, there may be a 'homogenization' effect due to the rapid oscillation of the fluctuating forces, and hence the Markovian approximation remains an appropriate first estimation of the memory effect, even in the present case.
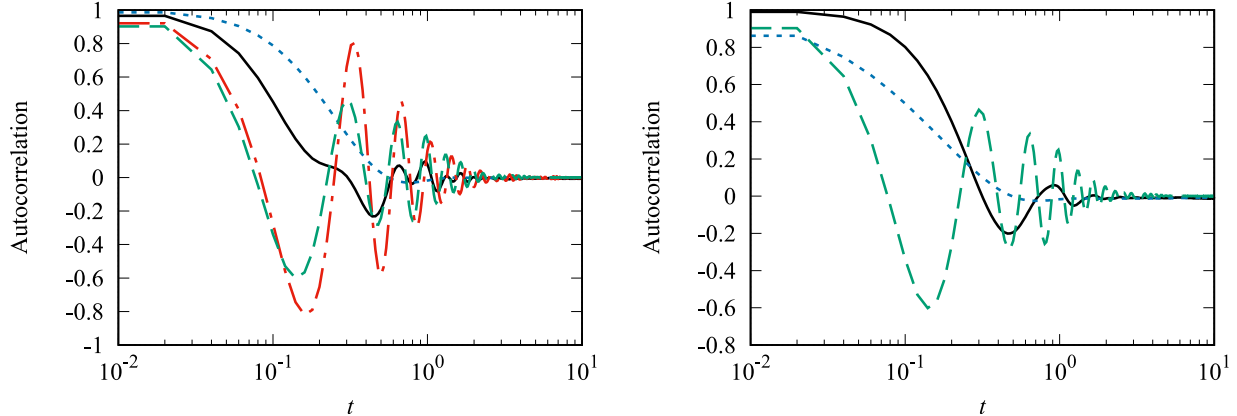
FIG. 7: Time autocorrelation function (ACF) of forces and momenta. In the left panel the ACF of the forces in the FGD (black solid line) is compared with that of the forces in the OD (red dot-dashed line), the MMZD (blue dashed line), and the 'orthogonal' component of the force in the FGD (green dashed line). In the right panel, the ACF of the momenta in the FGD (black solid line) is compared with the ACF of the momenta in the MMZD (blue dashed line), and the ACF of the momenta in the OD (green dashed line). All calculations refer to the system C (see table I) and are averaged in time and over all the beads.

D. The effect of the memory approximation on observed system properties

After obtaining an approximation of both the effective forces and the fluctuations, we compare the benefits of simulating the MMZD over the DCGD (i.e. dynamics evolving under the effective potential only, ignoring the memory terms), using as a target results obtained from the FGD. The observables chosen for comparison are the distribution of distances between centres of mass of adjacent beads, the momentum distribution of a single bead, and the evolution of the mean-squared displacement between the centres of mass of two adjacent beads over time. The two former observables are linear spatial statistics, and as such, it might be expected that either dynamics would capture them well. The latter statistic is a dynamical property of the system, and is an average of a nonlinear function, so we would expect to see non-trivial differences between the observables when computed during a simulation of the DCGD and the MMZD.

In fig. 8, we report the distributions of the momenta and positions obtained through simulation of the FGD (shown in black), as well as the DCGD using only the effective potential (shown in red), and the MMZD eq. (26) (shown in blue). The addition of the fluctuating force correction clearly improves the distributions, and indeed captures the tail behaviour of the distributions surprisingly accurately, in view of the relatively crude approximation of the memory term using eq. (42). In all cases, the mean distance between particles obtained is 3.00 to three significant figures. However, it is clear that even in this extremely simple toy system, the addition of the fluctuating terms does indeed improve the accuracy of the approximation in the case of these distributions.

The most important result shown in fig. 8 is the behaviour of the curve for the Markov-Mori-Zwanzig dynamics (dashed blue curve) relative to the curve for the Deterministic Coarse-Grained dynamics (red dot-dashed line). A CG potential obtained from some static distribution of the FG system (e.g. potentials obtained via IBI which are based on radial distribution functions) require an optimization, usually represented by a series of iterations to recover the correct distribution. In our case, we performed no optimisation of the potential: through sampling, we have simply approximated all of the terms in eq. (44). In particular, we can see that the main effect of the addition of the fluctuation-dissipation terms is to reduce the width of the distributions: while the average is satisfactorily captured by the DCGD (as shown by the overlapping peaks of the distributions), the tails are not correctly described. DCGD predicts a higher variance of bead velocities, pushing the beads further from equilibrium than they are in the real simulation. The presence of the additional terms provides improved variance without the need for optimisation of the effective potential. The only significant difference between the distributions of the positions of the MMZD and the FGD is the

27
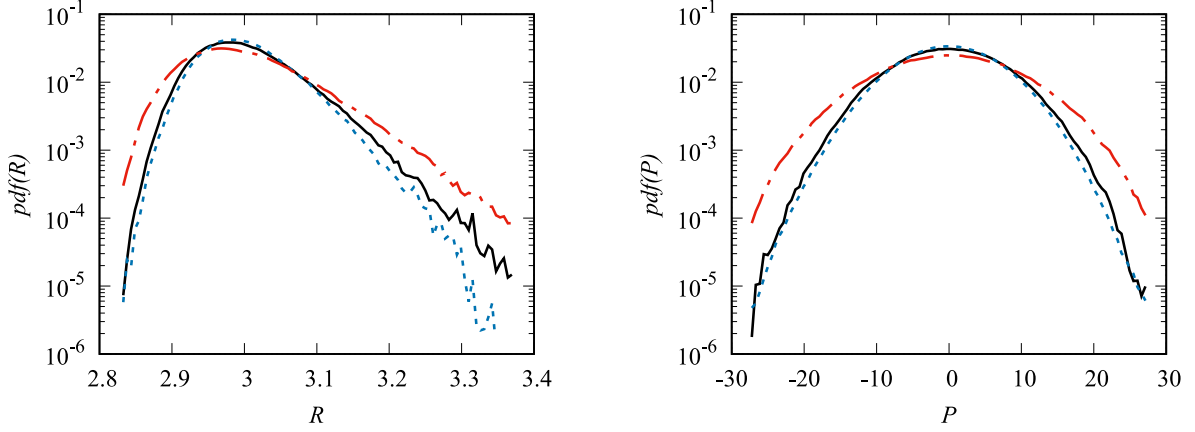
FIG. 8: PDFs of inter-bead distances and bead momenta. Comparison between the Fine-Grained dynamics (black solid line), the Deterministic Coarse-Grained dynamics (red dot-dashed line), and the Markov-Mori-Zwanzig dynamics (blue dashed line). All calculations refer to system C (see table I).

right tail of the position distribution; the region around the equilibrium position ($R \in [2.85, 3.2]$) is relatively accurate, but away from the equilibrium position the two distributions start to diverge; this difference may be either due to insufficient sampling, or simply because the Markovian approximation is insufficient in these regions: improving the approximation would inevitably improve the description of the system.
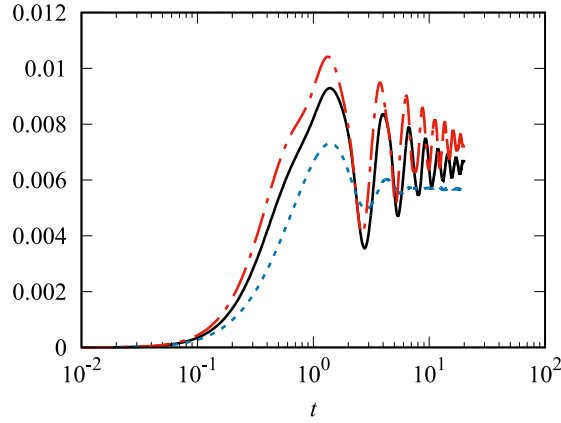
FIG. 9: Mean-squared displacement (MSD) of the beads. Comparison between the Fine-Grained dynamics (black solid line), the Deterministic Coarse-Grained dynamics (red dot-dashed line), and the Markov-Mori-Zwanzig dynamics (blue dashed line). All calculations refer to system C (see table I).

In fig. 9, we plot the mean-squared displacement (MSD) of the inter-bead distance in time for the MMZD, DCGD and FGD, showing the effect of the different approximations on the dynamics of the beads. As expected, the particles simulated under the DCGD move faster than in the FGD, whereas the opposite is true for the MMZD. In this case we do not observe as good an agreement between the FGD and the MMZD. This is likely to be due to the crudeness of the Markovian approximation, as discussed above. In particular, the results seem to indicate that the value of the friction matrix is too large: each peak of the blue curve lies beneath the black one. The same seems true in view of the fact that the autocorrelation of the momenta appears to decay too rapidly, as seen in the right-hand side of fig. 7.

The usual way to deal with the difference in the dynamical quantities between CG and FG simulations (e.g. diffusion coefficient, MSD) is to include into the CG model a way to map the time scales span by the

FG system into the CG one [91]. The time mapping consists in a scaling of the dynamic quantities obtained by CG simulations by a suitable parameter that measure how faster the CG dynamics is with respect the underlying FG system one [92–94]. In the framework of the MZ theory the same results can be achieved by using better defined quantities which can be formally derived from first principles. A more refined description of those is therefore essential for a better description of CG models.

VIII.   CONCLUSION

While the development of the Mori-Zwanzig (MZ) theory in the context of non-equilibrium thermodynamics dates back to 1961, only in recent years has it been developed as a theoretical framework in which to systematically derive Coarse-Grained models that might be applicable to Molecular Dynamics. The first purpose of this work has been to give an accessible, rigorous derivation of a class of CG models through the MZ formalism. We recalled the theory, giving a full derivation in the Supplemental Material, and provided a sketch of the most crucial points in the derivation of the various equations within the MZ framework.

We then discussed the relationship between the MZ approach and other coarse-graining approaches used in practice, including various methods of computing effective potentials, and the Virtual Sites approximation [25]. In the latter case, the MZ formalism provides a justification of the approach, previously used on an *ad-hoc* basis and justified retroactively. More generally, it can provide ways to assess *a-posteriori* the validity of coarse-graining approaches, and indeed provides a framework for quantifying their approximation errors.

Finally, we numerically studied a toy model, with which we illustrated some features of the MZ derivation. An open source code to efficiently compute the effective potential, friction matrix and fluctuating terms was written in Julia. We consider four systems with different parameter, and in all cases, no significant scale separation between the dynamics of the beads and individual atoms was observed. Nevertheless, a comparison was made between the true and approximate dynamics, and as expected, systems with 'more rigid' beads, having greater internal stiffness, showed much weaker memory effects. In all cases, we observed that a simple Langevin dynamics (with a friction coefficient independent of position) does not appear to be appropriate for generic CG systems.

Despite the lack of a clear scale separation, the Markovian approximation we used, was able to represent very accurately the equilibrium distribution of positions and momenta, while a deterministic CG fails to be as accurate, although neither approach succeeded in accurately capturing the dynamical observable we chose to test, suggesting that oversimplifying the memory means that fluctuations in the autocorrelation of forces are not accurately reproduced.

The projection operator we considered in this work is applied uniformly in the space of the FG variables to derive the CG model. However in multiscale simulation schemes based on spatially adaptive molecular resolution [95], the projection operator must be applied to an open subspace of the FG domain [96]. The latter also demonstrates the power of the MZ theory in the framework of the coarse-grained simulations and the needs to be more thoroughly investigated in future.

---

[1] M. Karplus and J. A. McCammon, Nature Structural Technology **9**, 646 (2002).
[2] J. Barrat, J. Banschnagel, and A. Lyulin, Soft Matter **6**, 3430 (2010).
[3] P. Li and K. M. J. Merz, Chemical Reviews **117**, 1564 (2017).
[4] M. S. Daw, S. M. Foiles, and M. I. Baskes, Materials Science Reports **9**, 251 (1993).
[5] P. Carbone and C. Avendaño, WIREs Comput Mol Sci **4**, 62 (2013).
[6] E. Brini, A. E. Algaer, P. Ganguly, C. Li, F. Rodríguez-Ropero, and N. F. A. van der Vegt, Soft Matter **9**, 2108 (2013).
[7] J. McCarty, A. Clark, J. Copperman, and M. Guenza, The Journal of chemical physics **140**, 204913 (2014).
[8] F. Müller-Plathe, Chem. Phys. Chem. **3**, 754 (2002).
[9] A. P. Lyubartsev and A. Laaksonen, Physical Review E **52**, 3730 (1995).
[10] S. Izvekov, M. Parrinello, C. J. Burnham, and G. A. Voth, The Journal of Chemical Physics **120**, 10896 (2004).
[11] S. Izvekov and G. A. Voth, Journal of Chemical Physics **123**, 134105 (2005).

[12] W. G. Noid, J. Chu, G. A. Ayton, V. Krishna, S. Izvekov, G. A. Voth, A. Das, and H. C. Andersen, Journal of Chemical Physics **128**, 244114 (2008).

[13] M. Scott Shell, Journal of Chemical Physics **129**, 144108 (2008).

[14] S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman, and A. H. Vries, Journal of Physical Chemistry **111**, 7812 (2007).

[15] W. Shinoda, W. Devane, and M. L. Klein, Molecular Simulation **33**, 27 (2007).

[16] V. Harmandaris, E. Kalligiannaki, M. Katsoulakis, and P. Plecháč, Journal of Computational Physics **314**, 355 (2016).

[17] Z. Li, X. Bian, B. Caswell, and G. E. Karniadakis, Soft Matter **10**, 8659 (2014).

[18] S. Izvekov and G. Voth, Journal of Chemical Physics **125**, 151101 (2006).

[19] A. J. Chorin, O. H. Hald, and R. Kupferman, Physica D **166**, 239 (2002).

[20] A. J. Chorin, A. P. Kast, and R. Kupferman, Proceedings National Academy of Science USA **95**, 4094 (1998).

[21] O. H. Hald and R. Kupferman, SIAM Journal of Numerical Analysis **39**, 983 (2001).

[22] H. A. Karimi-Varzaneh, P. Carbone, and F. Múller-Plathe, Jurnal of Chemical Physics **129**, 154904 (10 pages) (2008).

[23] Y. Masubuchi, Annual Review of Chemical and Biomolecular Engineering **5**, 11 (2014).

[24] A. Rzepiela, M. Louhivori, C. Peter, and S. J. Marrink, Phys. Chem. Chem. Phys. **13**, 10437 (2011).

[25] N. Di Pasquale, D. Marchisio, and P. Carbone, Journal of Chemical Physics **137**, 164111 (2012).

[26] N. Di Pasquale, R. J. Gowers, and P. Carbone, Journal of Computational Chemistry **35**, 1199 (2014).

[27] N. Di Pasquale and P. Carbone, Journal of Chemical Physics **146**, 084905 (2017).

[28] P. Español, EPL **88**, 40008 (2009).

[29] C. Hijòn, P. Español, E. Vanden-Eijnden, and D.-B. R., Faraday Discussions **144**, 301 (2010).

[30] T. Kinjo and S. A. Hyodo, Physical Review E: Statistical, Nonlinear, and Soft Matter Physics **75**, 051109 (2007).

[31] R. Zwanzig, Physical Review **124**, 983 (1961).

[32] H. Mori, Progress of Theoretical Physics **33**, 423 (1965).

[33] M. G. Guenza, Journal of Chemical Physics **110**, 7574 (1999).

[34] I. Lyubimov and M. G. Guenza, Physical Review E: Statistical, Nonlinear, and Soft Matter Physics **84**, 031801 (2011).

[35] D. Givon, R. Kupferman, and A. Stuart, Nonlinearity **17**, R55 (2004).

[36] R. D. Groot and P. B. Warren, Journal of Chemical Physics **107**, 4423 (1997).

[37] Z. Li, B. X., X. Li, and G. E. Karniadakis, Journal of Chemical Physics **143**, 243128 (2015).

[38] Z. Li, H. S. Lee, E. Darve, and G. E. Karniadakis, Journal of Chemical Physics **146**, 014104 (2017).

[39] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah, SIAM Review **59**, 65 (2017).

[40] T. Hudson, M. Icardi, and N. D. Pasquale, "Hybridzwanz/hybridzwanzjulia v1.0," (2018).

[41] K. Schweizer, Journal of Chemical Physics **91**, 5802 (1989).

[42] S. Nordholm and R. Zwanzig, Journal of Statistical Physics **13**, 340 (1975).

[43] H. Goldstein, *Classical Mechanics* (Addison-Wesley Pubclishing Comapany Inc., 1980).

[44] This is true as long as some very broad measurability assumptions on the CG variables chosen are met; in practice, choosing CG variables which are continuous and differentiable in the FG variables is already sufficient, and this particular choice is implicitly assumed in the derivation presented here. In theory however, much more general choices can be made.

[45] Since identifying the projection with a conditional expectation will be useful for the following sections, we provide a more detailed discussion in section I of the SM.

[46] A. J. Chorin, O. H. Hald, and R. Kupferman, PNAS **97**, 2968 (2000).

[47] D. Kauzlarić, P. Español, A. Greiner, and S. Succi, Macromolecular Theory and Simulations **20**, 526 (2011).

[48] For a more detailed derivation, see section II of the SM.

[49] E. Darve, J. Solomon, and A. Kia, Proceedings National Academy of Science USA **106**, 10884 (2009).

[50] In the micro-canonical ensemble, i.e. with microcanonical equilibrium measure, the free energy instead becomes $\mathcal{S}(\mathbf{\Gamma}) = -E \log \Omega(\mathbf{\Gamma})$.

[51] D. Givon, O. H. Hald, and R. Kupferman, Israel Journal of Mathematics **145**, 221 (2005).

[52] D. Evans and G. Morris, *Statistical Mechanics of Nonequilibrium Liquids* (London: Academic, 1990).

[53] See section III SM for a sketch of the proof using the Mori projector.

[54] In the microcanonical ensemble this definition becomes $\boldsymbol{\mathcal{M}}(\mathbf{\Gamma}, t) = \frac{1}{E}\mathcal{P}_{\mathbf{\Gamma}}(\boldsymbol{\mathcal{F}}_{\mathbf{\Gamma}}(0, \cdot) \otimes \boldsymbol{\mathcal{F}}_{\mathbf{\Gamma}}(t, \cdot))$.

[55] In the language used in the mathematical literature, canonical transformations belong to the wider class of 'symplectomorphisms'.

[56] In the context of the relative entropy framework [13], which will be further discussed in a later section (see section VI A), the non injectivity of the mapping operator is the reason why the mapping entropy term arises (see Eq. 4, [13]), as a measure of the degeneracy of the mapping. Higher mapping entropy values correspond to more atomistic configurations that can be mapped into the same CG variables.

[57] C. Peter and K. Kremer, Faraday Discussion **5**, 4357 (2009).

[58] C. Peter and K. Kremer, Faraday Discussion **144**, 9 (2010).

[59] G. Santangelo, A. Di Matteo, F. Müller-Plathe, and G. Milano, J. Phys. Chem. B **111**, 2765 (2007).

[60] X. Chen, P. Carbone, G. Santangelo, A. Di Matteo, G. Milano, and F. Müller-Plathe, Physical Chemistry Chemical Physics **11**, 1977 (2009).

[61] See section V of the SM for a detailed derivation of eqs. (18) and (19).

[62] E. Kalligiannaki, V. Harmandaris, M. A. Katsoulakis, and P. Plecháč, The Journal of chemical physics **143**, 08B613_1 (2015).

[63] R. Akkermans and W. J. Briels, Journal of Chemical Physics **113**, 6409 (2000).

[64] Similar to the computation shown in SM eq. S19.

[65] The derivation for previous equations is reported in section VI of the SM.

[66] J. W. Kirkwood, Journal of Chemical Physics **3**, 300 (1935).

[67] D. Chandler, *Introduction to Modern Statistical Mechanics* (Oxford University Press, 1987).

[68] L. Mones, N. Bernstein, and G. Csányi, Journal of Chemical Theory and Computation **12**, 5100 (2016).

[69] D. Reith, M. Putz, and F. Müller-Plathe, J. Comput. Chem. **24**, 1624 (2003).

[70] R. L. Henderson, Physics Letters **49A**, 197 (1974).

[71] M. Hanke, Journal of Statistical Physics **170**, 536 (2017).

[72] I. Bilionis and N. Zabaras, The Journal of chemical physics **138**, 044313 (2013).

[73] S. Kullback and R. A. Leibler, Journal of Chemical Physics **134**, 094112 (1951).

[74] A. Chaimovic and M. Scott Shell, Journal of Chemical Physics **134**, 094112 (2011).

[75] E. Kalligiannaki, A. Chazirakis, A. Tsourtis, M. Katsoulakis, P. Plecháč, and V. Harmandaris, The European Physical Journal Special Topics **225**, 1347 (2016).

[76] H. Lei, N. A. Baker, and X. Li, PNAS **113**, 14183 (2016).

[77] Y. Yoshimoto, Z. Li, I. Kinefuchi, and G. E. Karniadakis, Journal of Chemical Physics **147**, 244110 (2017).

[78] G. Jung, M. Hanke, and F. Schmid, Journal of Chemical Theory and Computation **13**, 2481 (2017).

[79] M. S. Green, Journal of Chemical Physics **22**, 398 (1954).

[80] D. Cubero and S. N. Yaliraki, Physical Review E **72**, 032101 (2005).

[81] D. Cubero and S. N. Yarilaki, Journal of Chemical Physics **122**, 034108 (2005).

[82] D. Cubero, Journal of Chemical Physics **128**, 147101 (2008).

[83] C. Hijòn, M. Serrano, and P. Español, Journal of Chemical Physics **125**, 204101 (2006).

[84] C. Hijòn, M. Serrano, and P. Español, Journal of Chemical Physics **128**, 147102 (2008).

[85] (i.e., fluctuation and dissipation terms that depends on the inter-particle distance).

[86] B. Leimkuhler and C. Matthews, The Journal of Chemical Physics **138**, 174102 (2013), http://dx.doi.org/10.1063/1.4802990.

[87] B. Leimkuhler and C. Matthews, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **472**, 20160138 (2016), http://rspa.royalsocietypublishing.org/content/472/2189/20160138.full.pdf.

[88] D. R. Jones, Journal of Global Optimization **21**, 345 (2001).

[89] Details of this algorithm are discussed further in the section VII of the M.

[90] The equivalence between the projection and an expected value is discussed in the section I of the SM.

[91] D. Fritz, K. Koschke, V. Harmandaris, and van der Vegt N.D.F, Phys. Chem. Chem. Phys. **13**, 10412 (2011).

[92] V. Harmandaris and K. Kremer, Macromolecules **42**, 791 (2009).

[93] D. Fritz, C. R. Herbers, K. Kremer, and van der Vegt N.D.F, Soft Matter **5**, 4556 (2009).

[94] P. K. Depa and J. K. Maranas, Journal of Chemical Physics **126**, 054903 (2007).

[95] R. Potestio, S. Fritsch, P. Español, R. Delgado-Buscalioni, K. Kremer, R. Everaers, and D. Donadio, Physical Review Letters **110**, 108301 (2013).

[96] L. Delle Site and M. Praprotnik, Physics Reports **639**, 1 (2017).