Journal of Computational Physics 300 (2015) 592–604

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

Wavelets as basis functions to represent the coarse-graining potential in multiscale coarse graining approach

M. Maiolo [a,c], A. Vancheri [a], R. Krause [b], A. Danani [a,*]

[a] SUPSI, Department of Innovative Technology, Galleria 2, 6928 Manno, Switzerland
[b] USI, Institute of Computational Science, Via Buffi 13, 6906 Lugano, Switzerland
[c] ZHAW, Institut für Angewandte Simulation, Grüental, CH-8820 Wädenswil, Switzerland

ARTICLE INFO

Article history:
Received 17 February 2014
Received in revised form 13 May 2015
Accepted 24 July 2015
Available online 30 July 2015

Keywords:
Molecular dynamics
Coarse graining
MSCG
Multiresolution analysis
Wavelets
Sparsity

ABSTRACT

In this paper, we apply Multiresolution Analysis (MRA) to develop sparse but accurate representations for the Multiscale Coarse-Graining (MSCG) approximation to the many-body potential of mean force. We rigorously framed the MSCG method into MRA so that all the instruments of this theory become available together with a multitude of new basis functions, namely the wavelets. The coarse-grained (CG) force field is hierarchically decomposed at different resolution levels enabling to choose the most appropriate wavelet family for each physical interaction without requiring an a priori knowledge of the details localization. The representation of the CG potential in this new efficient orthonormal basis leads to a compression of the signal information in few large expansion coefficients. The multiresolution property of the wavelet transform allows to isolate and remove the noise from the CG force-field reconstruction by thresholding the basis function coefficients from each frequency band independently. We discuss the implementation of our wavelet-based MSCG approach and demonstrate its accuracy using two different condensed-phase systems, i.e. liquid water and methanol. Simulations of liquid argon have also been performed using a one-to-one mapping between atomistic and CG sites. The latter model allows to verify the accuracy of the method and to test different choices of wavelet families. Furthermore, the results of the computer simulations show that the efficiency and sparsity of the representation of the CG force field can be traced back to the mathematical properties of the chosen family of wavelets. This result is in agreement with what is known from the theory of multiresolution analysis of signals.

1. Introduction

Despite the tremendous increase in computer power and algorithms efficiency, an adequate investigation of many complex systems remains well beyond the present capabilities of conventional atomistic simulation methods. To overcome this problem, in the recent years, coarse-grained (CG) models have become very popular. In these models, groups of atoms are clustered into new CG sites and averaging over less important details allows to focus on essential features, consequently obtaining a significant jump in the accessible spatial/temporal scales.

* Corresponding author.
E-mail addresses: massimo.maiolo@zhaw.ch (M. Maiolo), alberto.vancheri@supsi.ch (A. Vancheri), rolf.krause@usi.ch (R. Krause), andrea.danani@supsi.ch (A. Danani).

http://dx.doi.org/10.1016/j.jcp.2015.07.039
0021-9991/

In the CG approach, a mapping operator (typically the center of mass or the center of geometry of a set of atoms) is used to define the location of each CG site in an atomistic system as a function of the positions of the atoms. Then the key issue is the choice of the properties used to link the models of different resolutions. Several CG techniques are based on a "bottom-up" approach which employs information from the atomistic model to construct the potential for a CG model of the corresponding system.

A straightforward and popular approach is the Boltzmann inversion [1], together with its extensions like iterative Boltzmann inversion and inverse Monte Carlo method [2,3]. Another class of CG methods uses variational approaches to determine in a systematic way the CG potential within a rigorous statistical mechanical framework, such as the Multiscale Coarse-Graining (MSCG) introduced by Izvekov and Voth [4,5].

The basic assumption of the MSCG method is that the atomistic and the coarse-grained canonical distribution functions are compatible. The compatibility condition between CG and all-atom (AA) models is obtained by requiring that the probability to find the CG system in a given configuration is equal to the probability to find the AA system in the subset of AA configurations corresponding to the CG reduction defined by the mapping operator [6–8].

The MSCG force field is determined by minimizing a defined variational residual which generates a linear least-squares problem. The resulting inverse problem is solved by choosing an appropriate basis set of functions. The numerical implementation of the variational principle in MSCG can become a hard demanding task. There are two main challenges that arise by solving the least-squares problem. First, with a uniform grid, the huge number of equations requires an enormous amount of storage memory and computational power. Secondly, especially for complex systems, a poor sampling can induce statistical noises in the construction of the CG potential or in the worst case a degeneracy of the solution. Lu et al. have discussed in great details different new algorithms aiming at enhancing the efficiency and the stability of the MSCG computations [9]. In particular, they proposed various dense and sparse matrix algorithms to speed up the calculations and to regularize ill-posed MSCG problems.

In the present paper, we introduce a rigorous framework for the use of the wavelet transform in the context of the multiresolution analysis (MRA) for implementing the MSCG variational principle. This approach represents an innovative way to address the problems related to the sparsity of the CG potential representation. In a slightly different context, Ismail et al. introduced a coarse-graining methodology based on the wavelet transform, as a method for sampling polymer chains [10].

Thank to the MRA, the CG force field is decomposed at multiple space-scales simultaneously through a sequence of nested vector spaces to reflect the actual physical processes underlying the observed behavior as closely as possible. Moreover, the wavelet transform produces excellent results in signal denoising [11,12]. By shrinking or suppressing the wavelet coefficients under a predetermined threshold we are able to reduce the noise strongly.

The paper is organized in the following manner. Section 2 is devoted to the MSCG theory. In Section 3, wavelets as new basis functions are introduced. In Section 4, the implementation of the wavelet-based MSCG approach and the demonstration of its accuracy is discussed using two different condensed-phase systems, namely liquid water and methanol. Simulations of liquid argon were performed to test different wavelets properties. Finally, the last two Sections 5 and 6 contain discussion and conclusions, respectively.

2. The MSCG method: background

The first step in the construction of a CG force field is the reduction of the degrees of freedom of the atomistic system by means of a map that groups atoms into CG sites generally called beads. Let $\mathbf{r} = \{\mathbf{r}_1, \ldots, \mathbf{r}_n\}$ be the Cartesian coordinates of an instantaneous configuration of our atomistic model, composed by n atoms, and let $\mathbf{R} = \{\mathbf{R}_1, \ldots, \mathbf{R}_N\}$ be the corresponding state of our CG model composed by N CG sites. The configuration of the CG system is specified by a linear mapping operator $\mathbf{M}(\mathbf{r}) = \{\mathbf{M}_1(\mathbf{r}), \ldots, \mathbf{M}_N(\mathbf{r})\}$ of the form

$$\mathbf{M}_I(\mathbf{r}) = \sum_{i=1}^{n} c_{Ii}\mathbf{r}_i = \mathbf{R}_I \quad \text{for } I = 1, \ldots, N. \tag{1}$$

The CG site I is composed by all the atoms i such that $c_{Ii} \neq 0$. The numerical value of the coefficients is typically chosen so that $\mathbf{R}_I$ is either the center of mass or the geometrical center of the bead I. A basic assumption of the CG approach is that the atomistic and the coarse-grained canonical distribution functions are compatible. Equating AA and CG canonical distribution functions brings to the following effective potential U^{CG} of the CG system:

$$U^{\mathrm{CG}}(\mathbf{R}) = -\beta^{-1} \ln \int e^{-\beta U(\mathbf{r})} \delta\big(\mathbf{R} - \mathbf{M}(\mathbf{r})\big) \mathrm{d}\mathbf{r}\,. \tag{2}$$

It is also called potential of mean force (PMF) and represents the free energy of the atomistic system conditioned to the surface $\mathbf{M}(\mathbf{r}) = \mathbf{R}$. It is defined only up to an arbitrary additive constant. The PMF defines the CG force field through the relation:

$$\mathbf{F}_I(\mathbf{R}) = -\frac{\partial U^{\mathrm{CG}}}{\partial \mathbf{R}_I} \tag{3}$$

where $\mathbf{F}_I(\mathbf{R})$ is the force acting on the CG site I.