Fast and efficient optimization of Molecular Dynamics force fields for microporous materials: Bonded interactions via force matching

Andrea Gabrieli, Marco Sant *, Pierfranco Demontis, Giuseppe B. Suffritti

Dipartimento di Chimica e Farmacia, Università degli Studi di Sassari, Via Vienna 2, 07100 Sassari, Italy

ARTICLE INFO

ABSTRACT

A fast and efficient implementation of the *force matching* technique is presented and effectively applied to two important members of the microporous family. It enables the development of classical Molecular Dynamics force fields, tailored to each specific structure, starting from detailed *ab-initio* data. Reliable bonded interaction constants have been obtained for systems of increasing complexity, namely: CH_4, CO_2, Silicalite, and ZIF-8. The quality of the resulting parametrizations has been assessed, comparing model and reference vibrational spectra.

© 2014 Elsevier Inc. All rights reserved.

1. Introduction

One of the most widespread tools to investigate the dynamics of sorbed molecules within microporous materials, keeping into account also the flexibility of the framework [1], is the classical Molecular Dynamics (MD) technique [2,3]. This kind of simulations can follow the time evolution of a million atom system up to the microseconds scale. The drawback is the need to feed the program with a *force field* (FF) ruling the atoms interactions, on which will ultimately depend the quality of the results.

At the same time, the field of *ab-initio* Molecular Dynamics computations is rapidly growing. These give accurate results without need of an explicit FF (requiring as input only atoms types and initial positions). The drawback of this technique, on the other hand, is the large computational cost which becomes prohibitive at the time and space scales accessible to classical MD.

To exploit the advantages of both techniques, we could use short but detailed *ab-initio* computations to develop accurate FFs for classical MD, by means of the *force matching* method [4–14].

The aim of this work is to investigate the potential of the force matching technique and efficiently apply it to selected porous materials to obtain the FF constants that more closely reproduce the reference system dynamics.

Our group of research has recently published two papers dealing with the refinement of FF parameters for classical MD

simulations in both Silicalite [15] and Zeolitic Imidazolate Framework 8 (ZIF-8) [16], see Fig. 1. In those papers, the parameters optimization has been done with a trial and error procedure which is tedious and time consuming (more than one month for each structure). Moreover, this approach becomes practically unfeasible as the crystals complexity grows (e.g., ZIF-8 has already 46 bonded interaction terms to be tuned). The study of such systems requires a great amount of time even via the force matching technique, for this reason we made a big effort in improving the overall implementation performance.

The final force fields will be based on the CHARMM [17] functional form, tuned via an automated optimization procedure. In this work we will focus on the bonded part of the force field, which is responsible for the vibrational spectrum of modeled molecules and crystals, and can thus be accurately validated on the basis of this macroscopic property. In the CHARMM formalism, in fact, the exclusion policy has the effect of zeroing the weight of the non-bonded interactions with respect to the molecular frequencies of vibration. The correct reproduction of the vibrational frequencies is essential, in particular, for assuring a good thermal bath in simulating molecules sorbed in tight confinements [1].

The CO_2 and CH_4 FFs here presented are primarily intended as test cases to assess the feasibility of the procedure, while the Silicalite and ZIF-8 ones may be employed for production runs coupling the bonded parametrization with partial charges and vdW parameters taken from the literature.

In upcoming work, we will focus on developing complementary but independent techniques to obtain the partial charges and the

* Corresponding author.
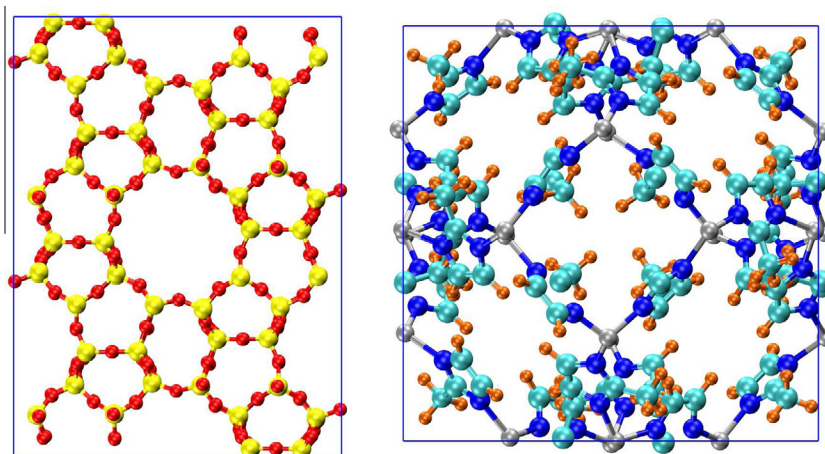   *E-mail address:* msant@uniss.it (M. Sant).

Fig. 1. Crystallographic structures: Silicalite (left) and ZIF-8 (right).

vdW parameters, using additional data taken from the same *ab-initio* run (e.g., matching of the electric field for the charges). The rationale behind this approach is to prevent competition among partial charges and vdW parameters during a unique force matching optimization; where the value of some parameters may be spoiled in the attempt to compensate for the lack of the others. This further development will make even more valuable the approach presented here.

In the first part of this paper we will illustrate the theory behind the force matching technique and our implementation of the method. In the second part we will apply the procedure to systems of increasing complexity, starting with CH_4, then CO_2, Silicalite and finally ZIF-8.

2. Theoretical background

2.1. Computational methods

2.1.1. Force matching

With this procedure the interaction parameters of a model system (e.g., a molecule or a crystal) are adjusted until they reproduce, within the wanted degree of accuracy, the forces of a given reference system [4]. In general, the reference is an highly detailed and thus accurate set of data, still very expensive from a computational point of view (here, *ab-initio*). The model system is in general a coarse-graining of the reference one, where some details are averaged out to attain high computational speed (here, classical MD).

The core of the whole procedure is the minimization of the sum of the squared residuals (merit function) between reference (F) *ab-initio* forces and model (f) MD forces:

$$\chi = \sum_{j=1}^{S} \sum_{i=1}^{3N} (F_{i,j} - f_{i,j})^2, \tag{1}$$

where N is the number of atoms in the system, clearly each atomic force has 3 components (x, y, z), and S is the number of snapshots (system configuration frames) taken with an arbitrary stride.

In accordance with the original work of Ercolessi and Adams [4], we evaluate the quality of the match looking at $h = \sqrt{\chi/(3NS)}$, the root mean square deviation per atomic force component, and compare it, h/g, against $g = \sqrt{\frac{1}{3NS}\sum_{j=1}^{S}\sum_{i=1}^{3N} F_{i,j}^2}$, the root mean square of the reference forces, representing their magnitude.

2.1.2. DFT computations

Ab-initio Molecular Dynamics computations are more accurate than classical MD simulations, yet they run about 1000 times

slower. For this reason they are not suitable to follow the time evolution of some macroscopic properties like self-diffusion. They have the big advantage of not requiring a structure dependent force field, but just knowledge of atomic species and positions. From this comes the idea of the force matching, where the system forces are stored over a short accurate trajectory, trying subsequently to obtain the same forces during a classical MD simulation.

In this work, the reference data are obtained performing Born–Oppenheimer Molecular Dynamics (BOMD) simulations using the CP2K open source code [18–21]. The energy of the system is evaluated via Density Functional Theory (DFT) [22,23] computations in the framework of the Gaussian and Plane Waves (GPW) [24] method.

The accuracy of DFT computations is continuously improving, thanks to the refinement of theoretical models and the growing of computational power [25]. This fact makes reasonable the expectation that DFT results will approach more and more the experimental limit. It becomes clear, then, that the force matching technique bridging detailed but expensive DFT computations and fast (but often based on too approximated FF) MD simulations will become more and more valuable.

2.1.3. Classical MD force fields

The FFs developed in this work rely on the CHARMM functional form:

$$\begin{aligned}
E_{\text{pot}} &= \sum_{\text{bonds}} k_{\text{b}}(b - b_0)^2 \\
&+ \sum_{\text{angles}} k_{\theta}(\theta - \theta_0)^2 + \sum_{\text{UB}} k_{\text{u}}(u - u_0)^2 \\
&+ \sum_{\text{dihedrals}} k_{\psi}(1 + \cos(n\psi - \delta)) \\
&+ \sum_{\text{impropers}} k_{\omega}(\omega - \omega_0)^2 \\
&+ \sum_{\text{vdW}} \epsilon \left[\left(\frac{R_{\min_{ij}}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{\min_{ij}}}{r_{ij}} \right)^6 \right] \\
&+ \sum_{i<j} \frac{q_i q_j}{\epsilon r_{ij}}.
\end{aligned} \tag{2}$$

where the first five terms refer to the bonded interactions, namely bonds (being k_{b} the force constant and b_0 the equilibrium distance), angles (with force constant k_{θ} and equilibrium angle θ_0), Urey–Bradley (UB, fictitious bond between two atoms, 1 and 3, connected to a common atom 2, k_{u} is the force constant and u_0 the equilibrium distance), dihedrals (torsion angles, with force constant k_{ψ}, function multiplicity n, and phase shift δ), and impropers (out of plane bending, with force constant k_{ω} and equilibrium angle ω_0). The last two