Averaged Configurations from Molecular Dynamics Simulations

Peter G. Kusalik

Department of Chemistry, University of Calgary,
Calgary, Alberta, T2N 1N4, Canada
peter.kusalik@ucalgary.ca

Abstract. One of the challenges in the large scale simulations required for biomolecular system is the recording, monitoring and visualization of configurational information from molecular dynamics trajectories. A detailed record of instantaneous configuration along the full trajectory can quickly become unmanageable. In this paper we will describe an alternative approach where configurations averaged over trajectory segments are used to follow the detailed molecular behaviour of a system over multiple-nanosecond simulations. We will then discuss the successful application of this approach to molecular dynamics simulations of crystal growth.

1 Introduction

Analysis of system configurations can be a key component in the simulation of molecular systems. These system configurations, which provide full molecular detail of the (relative) positions of the particles that make up the system, are frequently used to provide a means of exploring the microscopic behaviour captured by a molecular simulation. It is typical for these configurations to consist simply of the positions of all particles at a particular instant in time along the dynamical trajectory of the system. A set of such instantaneous configurations from a simulation can be saved and later analyzed further or visualized.

Complications arise in simulations, such as those required for biomolecular systems, where large length and time scales are necessary to capture the behaviour of interest. In these cases, the storage required to maintain this detailed record of the systems evolution (for example, as instantaneous configurations recorded every 20 fs) can be prohibitive. Moreover, the visualization of a system at this resolution over a multiple nanosecond trajectory becomes quite impractical. One remedy to these problems is to select and record instantaneous configurations on a far coarser grid in time (perhaps every 20 ps, for example). However, while such an approach may retain some basic aspects of the dynamics exhibited by the system, considerable information is lost nonetheless. For instance, it may be unclear if the instantaneous configuration recorded at the end of a relatively long trajectory segment is representative of the behaviour exhibited by the system during that segment. Below we will demonstrate that an alternative approach, that employs averaged configurations, captures considerably more detailed molecular information.

2 Generation of Averaged Configurations

An averaged configuration can be produced for any particular trajectory segment by averaging molecular coordinates,

$$\bar{x} = \frac{1}{N_s} \sum_{t=1}^{N_s} x_t \; , \tag{1}$$

over the N_s time steps of the segment. The choice of segment length can be an important consideration; it typically is made as long as possible while still providing a reasonable vantage point (i.e., frequency of sampling) from which to observe the complex processes characterizing the molecular behaviour of interest. With an appropriate selection of trajectory segment length over which to average, extraneous particle motion (e.g. thermal motion) can be effectively removed so that any net (more gross) motion can be more easily observed.

For each trajectory segment one has in principle a distribution of values for each degree of freedom of the system. To provide additional information into the nature of these distributions (beyond their means), we also find it advantageous to monitor their second-moments, or widths. The root mean-squared (RMS) deviations,

$$\sigma = (\overline{x^2} - \bar{x}^2)^{1/2} \; , \tag{2}$$

are also clearly measures of the diffusive motion exhibited by a molecule during the trajectory segment.

The treatment of molecular orientations is worth specific attention. At least for small molecules (like water), we find it advantageous to separate and average their degrees of freedom as positions and orientations. This approach allows for the conservation of molecular geometry during the averaging process. An averaging procedure for orientations has been developed recently[1] in terms of an average quaterion, or orientational centroid. In this procedure the orientational centroid, qc, minimizes the function

$$G(q_c) = \sum_{t=1}^{N_s} \Gamma^2(q_c, q_t) \; , \tag{3}$$

where $\Gamma(q_c, q_t)$ is the arc length between the centroid (average) orientation and the orientation qt. A simple Monte Carlo search algorithm can be used [1] to determine a value of qc for a set of Ns orientations. Analysis of system configurations can be a key component in the simulation of molecular systems. These system configurations, which provide full molecular detail of the (relative) positions of the particles that make up the system, are frequently used to provide a means of exploring the microscopic behaviour captured by a molecular simulation. It is typical for these configurations to consist simply of the positions of all particles at a particular instant in time along the dynamical trajectory of the system. A set of such instantaneous configurations from a simulation can be saved and later analyzed further or visualized.

Complications arise in simulations, such as those required for biomolecular systems, where large length and time scales are necessary to capture the behaviour of interest. In these cases, the storage required to maintain this detailed record of the systems evolution (for example, as instantaneous configurations recorded every 20 fs) can be prohibitive. Moreover, the visualization of a system at this resolution over a multiple nanosecond trajectory becomes quite impractical. One remedy to these problems is to select and record instantaneous configurations on a far coarser grid in time (perhaps every 20 ps, for example). However, while such an approach may retain some basic aspects of the dynamics exhibited by the system, considerable information is lost nonetheless. For instance, it may be unclear if the instantaneous configuration recorded at the end of a relatively long trajectory segment is representative of the behaviour exhibited by the system during that segment. Below we will demonstrate that an alternative approach, that employs averaged configurations, captures considerably more detailed molecular information.

3 Visualization of Averaged Configurations

We have employed averaged configurations extensively in our studies of heterogeneous crystal growth, where the detailed analysis of multiple nanosecond trajectories is required to uncover the underlying processes associated with crystal growth. Figure 1 compares an averaged configuration (from a 20 ps trajectory segment) with the instantaneous configuration from the end of this trajectory segment. Perhaps the most striking aspect of the averaged configuration is the clarity of its crystalline structure and the distinctiveness of the interfacial layer.

To enhance further our ability to extract visual information from averaged

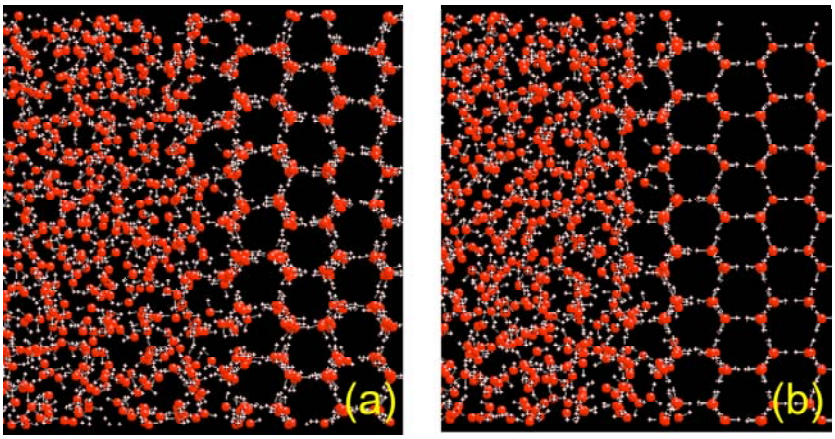

Fig. 1. (a) Instantaneous configuration and (b) averaged configuration of the same interfacial region of an ice/water system

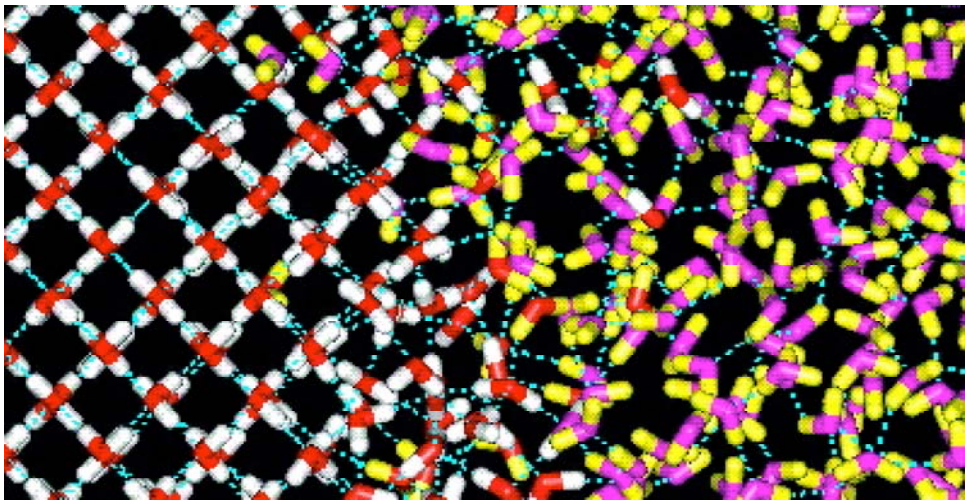

Fig. 2. Averaged configuration of the interfacial region of an ice/water system where the molecules have been colored to reflect their solid or liquid character, as discussed in the text.

configurations, we have labeled water molecules within each averaged configuration as being translationally solid or liquid, as well as rotationally solid or liquid. This was done by identifying appropriate thresholds for the RMS deviations (i.e., diffusive behaviour) for both positions and orientations that are consistent with values found in the bulk crystal. The molecules could be then coloured according it their solid/liquid labels. Specifically, if a molecular was labeled translationally solid its oxygen was coloured red, otherwise the oxygen was coloured magenta. If a molecular was labeled rotationally solid its hydrogens were coloured white, otherwise the hydrogens were coloured yellow. We can see from Figure 2, where the interfacial region of an averaged configuration for another ice/water system is shown, that these labels provide considerable insight it the molecular behaviour occurring at the interface.

References

1. L. Hernandez de la Pea and P. G. Kusalik,
Mol. Phys. 102, 927 (2004).