Computational Materials Science 127 (2017) 141–160

Contents lists available at ScienceDirect

Computational Materials Science

journal homepage: www.elsevier.com/locate/commatsci

Reliable Molecular Dynamics: Uncertainty quantification using interval analysis in molecular dynamics simulation

Anh V. Tran, Yan Wang *

Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA

ARTICLE INFO

Article history:
Received 10 May 2016
Received in revised form 16 September 2016
Accepted 14 October 2016

Keywords:
Molecular dynamics
Interatomic potential
Uncertainty quantification
Interval analysis

ABSTRACT

In molecular dynamics (MD) simulation, atomic interaction is characterized by the interatomic potential as the input of simulation models. The interatomic potentials are derived experimentally or from first principles calculations. Therefore they are inherently imprecise because of the measurement error or model-form error. In this work, a Reliable Molecular Dynamics (R-MD) mechanism is developed to extend the predictive capability of MD given the input uncertainty. In R-MD, the locations and velocities of particles are not assumed to be precisely known as in traditional MD. Instead, they are represented as intervals in order to capture the input uncertainty associated with the atomistic model. The advantage of the new mechanism is the significant reduction of computational cost from traditional sensitivity analysis when assessing the effects of input uncertainty. A formalism of generalized interval is incorporated in R-MD, as an intrusive uncertainty quantification method, to model the propagation of uncertainty during the simulation. Error generating functions associated with embedded atomic method (EAM) interatomic potentials are developed to capture the bounds of input variations to demonstrate interval interatomic potentials. Four different uncertainty propagation schemes are proposed to capture the uncertainty of the output. An example of uniaxial tensile loading of single-crystal aluminum is used to demonstrate the R-MD mechanism.

1. Introduction

Modeling and simulation tools are crucial for engineers to design and develop new materials efficiently. Uncertainty is always involved in model selection, calibration, and validation processes. Reliable simulation predictions require us to quantify input uncertainty of models. There are two elements of uncertainty in modeling and simulation: aleatory uncertainty and epistemic uncertainty. Aleatory uncertainty is the inherent randomness in the phenomenon being observed, and the impossibility of exhausting all descriptions deterministically. Epistemic uncertainty can be generally related to the lack of perfect knowledge about the involved physical processes [1].

Molecular dynamics (MD) is one of the most widely used atomistic simulation tools. In MD simulation, the aleatory uncertainty corresponds to any fluctuation of the simulated system, e.g. the natural thermal fluctuation that can be described by Boltzmann distribution at an equilibrium microscopic state. The epistemic uncertainty includes, but is not limited to, the imprecise interatomic potentials, the finite size effect, the boundary condition imposed on the simulation cell, and the cutoff radius of the interatomic potentials. The aleatory uncertainty associated with the thermal fluctuation is generally inseparable from MD simulation, and sometimes is induced by the ensemble integrator. For example, in Langevin thermostat, this thermal fluctuation is accounted by the friction-noise in the stochastic differential equations [2]. The epistemic uncertainty in MD simulations is mostly caused by the imperfection of the interatomic potential. These interatomic potentials are typically derived from first principles calculations or approximated based on experimental data. These results are contaminated by both systematic and random errors. The systematic errors of first principles calculations come from different approximations and assumptions in the models, such as Born-Oppenheimer approximation, Hartree-Fock approximation, and the assumed finite linear combination of the variational solution based on the set of basis functions [3]. On the other hand, the systematic errors of experimental results involve measurement bias and calibration errors. Based on the results, an interatomic potential model is formulated with a set of parameters to minimize a measurable error, which usually in turn is converted to a least-square error problem. Because of the non-negative residual in curve fitting and approximation error techniques used in deriving

* Corresponding author.
E-mail addresses: yan.wang@me.gatech.edu, prof.yan.wang@gmail.com (Y. Wang).

http://dx.doi.org/10.1016/j.commatsci.2016.10.021
0927-0256/

the interatomic potentials, MD simulations include both model-form uncertainty and parameter uncertainty. Furthermore, the epistemic uncertainty from interatomic potential in MD simulations is amplified because the number of interacting pairs, which scales at least as N^2, where N is the number of atoms. Therefore, quantifying uncertainty in MD simulations is a critical problem, in both assessing the accuracy and reliability of the simulation prediction.

Uncertainty quantification (UQ) problems are divided into two main paradigms, intrusive and non-intrusive methods on probabilistic and non-probabilistic frameworks. In non-intrusive UQ techniques, the simulation is viewed as a black box, and the simulator is modeled as a one-to-one non-linear function that maps from the input domains to the output or quantities of interests. Popular techniques, including stochastic collocation, Monte Carlo, and global sensitivity analysis, rely on statistical techniques to build comprehensive output distributions based on the assumed input distributions. Generalized polynomial chaos expansion is a widely used technique, and can be utilized either intrusively or non-intrusively. As an intrusive technique, it has been applied to solve stochastic differential equations and partial differential equations with random inputs. As a non-intrusive technique, it is typically used together with Smolyak sparse grid and nested sets in stochastic collocation methods.

Other intrusive UQ techniques, such as local sensitivity analysis and interval-based approaches, aim to provide the output probability density function or its bounded support for expensive simulation by incorporating and propagating the uncertainty *internally* using minimal number of runs. In interval-based approaches, the uncertainty is coupled into the input and represented by intervals. The simulator is thus extended to handle the interval inputs and propagate the uncertainty throughout the simulation. The output uncertainty, which is also represented as intervals, is computed at every time step at a relatively cheap computational cost.

Various UQ methods have been applied to multi-scale simulation for materials. Comprehensive literature reviews are available in [4,5]. Frederiksen and Jacobsen [6] applied Bayesian update to train the interatomic potentials parameters with experimental data sets by minimizing the square error between experimental data and simulation results. Jacobson et al. [7] constructed response surfaces with Lagrange interpolation to study the sensitivity of macroscopic properties with respect to interatomic potential parameters. Cailliez and Pernot [8] calibrated Lennard-Jones potential for Argon based on Bayesian calibration/prediction framework. Rizzi et al. [9,10] assumed uniform distribution for the four-site, TIP4P, water model parameters and constructed the generalized polynomial chaos representation by non-intrusive spectral projection and Bayesian inference approaches, then later on, calibrated these force-field parameters based on Bayesian inference. Angelikopoulos et al. [11] applied the Bayesian calibration to calibrate the water-carbon interactions based on water contact angles in water wetting of graphene, the aggregation of fullerenes in aqueous solution, and the water transport across carbon nanotubes. Rizzi et al. [9] applied polynomial chaos expansion to study the effect of input uncertainty in MD. Cailliezf et al. [12] applied the efficient global optimization algorithms in parameter space to calibrate the potential parameters for TIP4P model, based on probabilistic kriging metamodels. Wen et al. [13] studied the effect of different spline interpolations on the potential predictions by calculating the quasi-harmonic thermal expansion and finite-temperature elastic constant of a one-dimensional chain in tabulated interatomic potentials. Hunt et al. [14] developed a software package for non-intrusive propagation of uncertainties in input parameters, using surrogate models and adaptive sampling methods, such as Monte Carlo, Latin Hypercube, and Smolyak sparse grids, based on generalized polynomial chaos expansion. Li et al. [15] discussed the cut- and random sample-high dimensional model representation to quantify the uncertainty induced by potential surfaces.

As an intrusive approach on non-probabilistic framework, we recently proposed an interval-based reliable MD (R-MD) mechanism [16,17] that incorporates Kaucher interval arithmetic [18] into classical MD to quantify output uncertainty. Classical interval arithmetic provides a complete solution by capturing all possibilities for simple algebraic operations, such as addition, subtraction, multiplication, and division. Kaucher interval arithmetic generalizes and extends [19] classical interval arithmetic with better topology and algebraic properties. Compared to classic interval arithmetic, Kaucher interval arithmetic is preferred for three reasons. Firstly, the over-estimation problem is significantly reduced. Secondly, the self-dependency problem, which also results in an over-estimation of a function, where dependent variables are repeated more than once, is mitigated. Thirdly, the negation and reciprocal operations with respect to addition and multiplication exist. In contrast to the Kaucher interval space, the classical interval space only forms a semi-group algebraic structure because of the lack of invertibility. In R-MD, the input uncertainty associated with interatomic potentials is captured in interval forms, either as intervals or as interval functions. Consequently, the atomistic positions, velocities, and forces are also interval-valued. Fig. 1 plots a schematic sketch of simple 2D R-MD simulation cell, where the atomistic positions are interval-valued. The exact atomistic positions and velocities are unknown, but bounded by intervals. In this paper, the details of how Kaucher interval arithmetic is applied in simulation including interval potential, interval force computation, and interval statistical ensemble are described. In Section 2, we review the algebraic operations of Kaucher interval arithmetic. In Section 3, the formulation of interval potential and interval force are discussed, and four R-MD uncertainty propagation schemes are implemented in the framework of Large-scale Atomic/Molecular Massively Parallel Simulator, also known as LAMMPS [20]. An application to tensile uniaxial deformation of aluminum single crystal is demonstrated in Section 4, including UQ results, comparisons between different schemes, finite-size effects, and comparison with sensitivity analysis results as a part of verification process. Following are the discussion in Section 5 and conclusion in Section 6.

2. Kaucher interval arithmetic

The classical interval space, denoted as $\mathbb{IR}$, is a collection of classical interval, where the upper bound is strictly greater than or

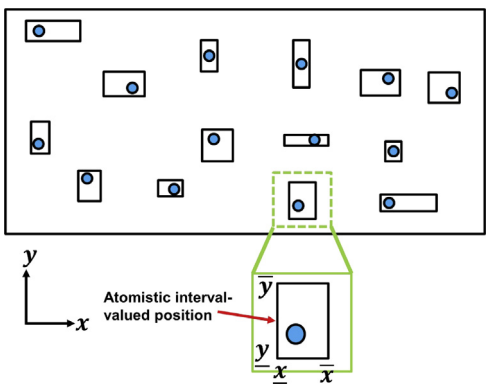

Fig. 1. Schematic illustration of R-MD in 2D.