Commun. Comput. Phys.
doi: 10.4208/cicp.OA-2017-0213

Vol. **23**, No. 3, pp. 629-639
March 2018

Deep Potential: A General Representation of a Many-Body Potential Energy Surface

Jiequn Han[1,†], Linfeng Zhang[1,†], Roberto Car[1,2] and Weinan E[1,3,4,*]

[1] *Program in Applied and Computational Mathematics, Princeton University, Princeton, NJ 08544, USA.*
[2] *Department of Chemistry, Department of Physics, and Princeton Institute for the Science and Technology of Materials, Princeton University, Princeton, NJ 08544, USA.*
[3] *Department of Mathematics, Princeton University, Princeton, NJ 08544, USA.*
[4] *Center for Data Science and Beijing International Center for Mathematical Research, Peking University, and Beijing Institute of Big Data Research, Beijing, 100871, China.*

Received 5 October 2017; Accepted 17 October 2017

Abstract. We present a simple, yet general, deep neural network representation of the potential energy surface for atomic and molecular systems. It is "first-principle" based, in the sense that no *ad hoc* approximations or empirical fitting functions are required. When tested on a wide variety of examples, it reproduces the original model within chemical accuracy. This brings us one step closer to carrying out molecular simulations with quantum mechanics accuracy at empirical potential computational cost.

AMS subject classifications: 81V55, 92E10, 68T99

Key words: Potential energy surface, deep learning, molecular simulation.

1 Introduction

A representation of the potential energy surface (PES) for general systems of atoms or molecules is the basic building block for molecular dynamics (MD) and/or Monte Carlo (MC) simulations, which are common tools in many disciplines, including chemistry, physics, biology, and materials science. Until now, this problem has been addressed using two very different approaches. At one extreme, empirical potentials have been constructed by fitting limited experimental and/or numerical data from accurate quantum

†These authors contributed equally to this work.
*Corresponding author. *Email addresses:* weinan@math.princeton.edu (W. E), jiequnh@princeton.edu (J. Han), linfengz@princeton.edu (L. Zhang), rcar@princeton.edu (R. Car)

mechanical calculations. Well-known examples include the Lennard-Jones potential [1], the Stillinger-Weber potential [2], the embedded-atom method (EAM) potential [3], the CHARMM [4]/AMBER [5] force fields, the reactive force fields [6], etc. These potentials are computationally efficient, allowing large-scale simulations (up to millions of atoms), but their construction is very much an art and their accuracy and transferability are limited. At the other extreme, methods based on first-principle quantum theory such as density functional theory (DFT) [7] have been proposed, the most well-known example being the *ab initio* molecular dynamics (AIMD) [8] scheme. These methods promise to be much more accurate but they are also computationally expensive, limiting our ability to handling systems of hundreds to thousands of atoms only. Until recently, the drastic disparity between these two approaches in terms of accuracy and computational cost has been a major dilemma to be confronted with in molecular simulation.

Recent advances in machine learning, particularly deep learning, have ushered some new hope in addressing this dilemma [9–20]. Several promising new ideas have been suggested, in which deep neural networks are used to represent the potential energy surface. Of particular interest are the Behler-Parrinello neural network (BPNN) [10] and the deep tensor neural network (DTNN) [9]. BPNN uses the so-called symmetry functions as input and a standard neural network as the fitting function; DTNN, on the other hand, uses as input a vector of nuclear charges and an inter-atomic distance matrix, and introduces a sequence of interaction passes where "the atom representations influence each other in a pair-wise fashion" [9]. Both methods are able to predict with great accuracy the potential energy surface of materials in condensed phase, in the case of BPNN, and of small organic molecules, in the case of DTNN. However, the construction of the local symmetry functions for BPNN contains an *ad hoc* and often tedious component where hand-crafted fitting functions and human intervention are required. On the other hand, methods like DTNN may face difficulties when one tries to transfer them directly to large systems.

In this work, we develop a new method, called Deep Potential (DP), that successfully addresses the inadequacies of the existing models. Deep Potential is a simple, yet general, end-to-end deep neural network representation of a many-atom potential energy surface. The network uses as input the raw coordinates of the atoms in a proper frame of reference, and naturally respects the symmetries of the system. Promising results are obtained in a variety of test cases, including small molecular isomers and condensed-phase systems. The Deep Potential method brings us closer to performing molecular modeling with the secure accuracy of first-principle based methods at a computational cost comparable to that of empirical potentials.

2 Results

2.1 Deep Potential framework

Our goal is to formulate a general and direct end-to-end representation of the many-body potential energy surface that uses the atomic configurations directly as the only