Catalysis Today xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Catalysis Today

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

Modeling palladium surfaces with density functional theory, neural networks and molecular dynamics

Tianyu Gao, John R. Kitchin*

Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, United States

ARTICLE INFO

Keywords:
Neural network
Potential energy surface
Molecular dynamics
Machine learning

ABSTRACT

In this work, we have constructed a high dimensional neural network (NN) potential energy function for simulating palladium surface properties. The NN potential was trained with 3035 density functional theory (DFT) calculations, and was shown to be nearly as accurate as DFT in molecular simulations. Important properties including lattice constants, elastic properties and surface energies as well as transition state energies and adatom diffusion barriers were predicted by the NN and were found to be in excellent agreement with DFT results. The computational time to run the NN was compared to DFT calculation time, and we found this implementation of the NN is roughly four orders of magnitude faster than DFT. This approach is general and applicable to other systems and may have applications in modeling catalytic processes at surfaces.

1. Introduction

With density functional theory atomic structure and electronic structure calculations have now been implemented extensively and used to obtain an atomic-level understanding of chemical phenomena and processes. *Ab initio* based simulations such as molecular dynamics (MD) [1] and Monte Carlo (MC) [2] simulation have become possible with progress in high-performance computing. But even with the aid of the most advanced computers, density-functional theory (DFT)-MD/MC simulations are still limited to both short time scales (hundreds of picoseconds) and a small scale of a few thousand atoms. It is still necessary to find methods to accelerate these kinds of simulations to extend them to larger length and time scales without sacrificing accuracy.

A reliable simulation requires an accurate description of the potential-energy surface (PES), which describes the potential energy as well as the forces as a function of the atomic coordinates. Progress has been made in introducing and modeling approximated PESs. There are basically two different approaches: physical potentials and mathematical potentials. In the physical approach, the PES is simplified by physical assumptions and approximations, which generally derive from two and more body-interactions. Parameters in the physical models are typically fitted to experimental or computational data to reproduce desired properties. Typically the parameters in physical models have physical interpretations. There are many physical potentials including the Lennard-Jones potentials [3], classical force fields [4–8], and reactive force fields [9,10]. The accuracy of a physical PES is heavily dependent on the underlying physical model, and they are not always systematically improvable.

The other approach to PES construction is the use of mathematical potentials, which employ mathematical fitting models. Generally, these methods including genetic programming [11], interpolating moving least squares (IMLS) [12,13], Gaussian approximation potentials [14,15] and many other improved mathematical forms. Neural networks (NN) have become increasingly promising with the introduction of "high-dimensional" NN systems by Behler and Parrinello [16–18].

A lot of progress has been made recently in creating and using NN potentials for a variety of systems including Si [19], Cu surfaces [17], Au bulk, surfaces and clusters [10], water clusters [14], metal oxides (ZnO [20] and TiO_2 [21]), and recently for the dissociation of N_2 at Ru (0001) surfaces [22]. For different systems ranging from single to multiple compositions, periodic bulk to clusters containing thousands of atoms, the NN mathematical model has been shown to have an arbitrary precision given a sufficiently large amount of training data. In the work described above, electronic structure calculations such as DFT calculations were used to train the NN to the desired level of accuracy.

We have been investigating the use of NN potentials for modeling surface properties of Au [10], Au-Pd segregation [23], and palladium-oxygen interactions [24]. That work has systematically increased the complexity of the composition and structure of the surfaces being modeled, as well as the use of the potentials in extending density functional theory calculations. The goals of this work are to methodically increase our understanding of how to construct these potentials, their advantages and shortcomings, and eventually to be able to reliably build these potentials for applications in catalysis.

---

* Corresponding author.
E-mail address: jkitchin@andrew.cmu.edu (J.R. Kitchin).

https://doi.org/10.1016/j.cattod.2018.03.045
Received 15 November 2017; Received in revised form 16 March 2018; Accepted 23 March 2018

In this paper, we constructed and studied a Behler-Parrinello NN for bulk palladium and palladium surfaces. Specifically, we examine diffusion of Pd adatoms at the surface. While not a catalytic reaction, we view it as a step towards modeling activated processes on surfaces. The training process was implemented with the Atomistic Machine-learning Potentials (AMP) code developed by Peterson and Khorshidi [25]. We show that the NN models a variety of bulk structures, surface energies, surface vacancy formation and diffusion, and adatom diffusion with accuracy that is comparable to DFT in most cases. We also show that the NN can be used in molecular dynamic simulations to estimate adatom diffusion constants, and that is considerably more cost effective than running DFT calculations for that purpose.

2. Computational details

2.1. DFT calculations

The reference DFT calculations for NN training were performed with the Vienna *ab initio* simulation package (VASP) [26–28] code with the Perdew-Burke-Ernzerhof [29] generalized gradient approximation (GGA-PBE) as the exchange-correlation functional. The projector augmented wave function (PAW) [30,31] was chosen to describe the core electrons. Monkhorst-Pack [32] grids of dense $14 \times 14 \times 14$ *k*-point grids for each primitive cell along with a 350 eV plane wave cutoff were utilized. These parameters achieved ≈ 1 meV convergence in total energies. The slab models consisted of five layers with the bottom two layers fixed, with 6 Å of vacuum on both sides. We used a similar density of *k*-points for the slabs in the non-vacuum directions. The geometries were optimized using a conjugate-gradient algorithm and stopped at an energetic criteria of 10^{-4} eV/atom. Transition states and energy barriers calculations were performed using standard nudged elastic band (NEB) method [33]. The training set contained a total of 3035 configurations of different bulk and slab configurations. The DFT calculations provided potential energies and atomic forces for each system, and both of them were utilized to train NN potentials.

2.2. Neural network

The NN potentials were constructed in an iterative way using the AMP code [25], which provides both atomic environment descriptors and regression models (neural networks). By introducing a cutoff radius and symmetry functions [34] the BPNN can be utilized to predict energies for systems that contain different numbers of atoms from the training set. The potential energy of the system is computed as the sum of atomic energy contributions, which are determined by the atomic environment within a cutoff radius. The cutoff radius used was 6.5 Å, which is typical for metals [17]. The neural network architecture reported in this work had two hidden layers with 18 nodes each. The default four radial and four angular Gaussian symmetry functions in AMP were used as input.

The training is performed by minimizing the energy and force root mean squared error (RMSE) between known values in a database of calculations and predictions from the neural network. The energy goal was set to 1 meV, and the force goal was set to 0.01 eV/Å. It has been shown that high dimensional NNs are capable of achieving a very high level of accuracy, but over-fitting needs to be avoided. To assess over-fitting, we reserved 10% of the training data for validation, and these configurations are not included in the training process. By using the trained NN to predict the validation set and comparing the RMSE of validation to that of training set, we can assess over-fitting. If the error of validation set is significantly larger than the error of training data, over-fitting has likely occurred and either more training data is required to eliminate it or a smaller NN can be used.

In this work, the training data set was generated iteratively. The initial set started from several ideal or structures with defects including ideal bulk equations of state and slab structures. The slabs and defect structures were allowed to relax, and each of the geometry relaxation steps from this were also included as training data. With this data a pair of NNs with different architectures was trained, which are typically of moderate quality and accuracy. We then utilized one of the NNs to carry out MD or some other simulation to generate a large quantity of new configurations, and used the second NN to validate the predictions on the same configurations. Some number of the "new" ones are usually chosen from the regions where two NNs of different architecture predict very distinct energies, e.g. the shaded regions shown in Fig. 1. These structures, reasonable or not, can be computed by DFT calculations and added into training set. Through this process, the NN can be systematically and self-consistently improved. Further details on the training process are described in the subsequent sections that describe each type of result.

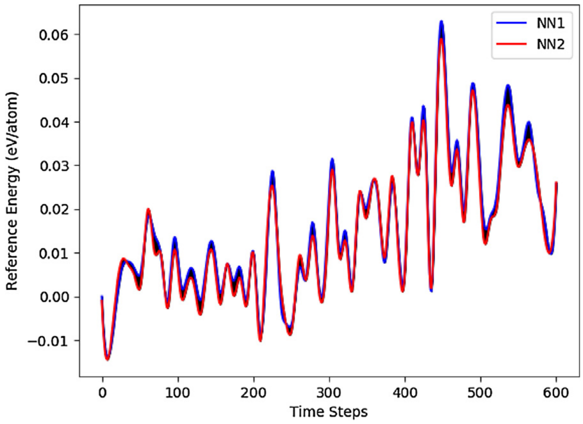

Fig. 1. New configurations generated by MD simulations and determined by two different NN. The shaded black region between the curves represents structures where the two NNs disagree. These regions contain candidates that can be added to the training set to improve its accuracy.

3. Results and discussion

3.1. NN potential and RSME

Among the total of 3035 calculations, 10% were divided into a validation set to monitor over-fitting. Gaussian functions are chosen to describe the local chemical environment of each function. We were able to use only 8 symmetry functions (four radial, and four angular) in this work. Our trained NN achieved a reasonable convergence of 5.51 meV/atom residual error for the training set and 1.92 meV/atom for the validation set, so the NN generally reproduced the PES with no evidence for over-fitting. The trained and validation NN RMSE and error distribution are shown in Fig. 2. The distribution mean value μ and standard deviation σ are shown, along with the best fit normal distribution. The errors are approximately described by a normal distribution with a 0.006 eV/atom standard deviation. We conclude that this NN is accurate for use in the kind of molecular simulations used in the training data, e.g. molecular dynamics.

3.2. Bulk properties

3.2.1. Equations of state and cohesive energy

We first show that the NN can describe bulk properties with great accuracy. We used the NN to predict the equations of state for different palladium bulk structures, including face-centered cubic (fcc), body-centered cubic (bcc), hexagonal close-packed (hcp), simple cubic (sc) and diamond (diam) lattice. These structures represent a range of coordination environments from four to twelve atoms and are the starting point to look at under-coordinated environments including bulk defects and surfaces. The equations of state for these bulk structures are shown