Chemical Physics Letters 652 (2016) 130–135

Contents lists available at ScienceDirect

Chemical Physics Letters

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

Research paper

Modeling of DFT quality neural network potential for sodium clusters: Application to melting of sodium clusters (Na_{20} to Na_{40})

Siva Chiriki, Satya S. Bulusu *

Department of Chemistry, Indian Institute of Technology Indore, India

ARTICLE INFO

Article history:
Received 9 February 2016
Revised 6 April 2016
In final form 8 April 2016
Available online 13 April 2016

Keywords:
High dimensional neural networks
Monte Carlo simulations
Melting point
Sodium clusters

ABSTRACT

The present work demonstrates the use of computationally inexpensive neural network (NN) potential for studying global optimizations and phase transitions in small to medium sized sodium clusters with DFT accuracy. Accuracy of NN potential has been tested by performing global optimizations in the size range of 16–40 atoms. We performed Monte Carlo (MC) simulations using NN potential to study the melting behaviour. Melting study in the size range of 20–40 atoms shows a characteristic premelting peak and a main melting peak. Our results using NN potentials support the idea of stepwise melting in small Na clusters (Aguado, 2011).

1. Introduction

Computer simulation techniques such as molecular dynamics (MD) or Monte Carlo (MC) methods are widely used to obtain temperature dependent and time dependent properties for many chemical systems. These techniques generally use empirical, semi-empirical or ab-initio potentials to evaluate the potential energy at every time step. Ab initio potentials can be used to get accurate description of any chemical system but are prohibitive as the system size increases. For example, performing a global optimization on gold clusters with 50 atoms or more using ab initio potentials become computationally very expensive task because of exponential growth in local minima with increase in number of atoms in a cluster. Empirical potentials, on the other hand are good alternative to ab initio potentials in terms of computational speed. But the computational speed usually comes at the expense of considerable loss in accuracy in predicting chemical properties. This happens because functional form and fitting parameters used to construct an empirical potential are non-transferable. For example, empirical potentials that are adequate for bulk metals cannot describe structure and dynamical properties of small metal clusters. Therefore, constructing a potential that is good both in accuracy and computational efficiency is an active field of research [1–4].

Herein, we report the construction of NN potential to study the structure and dynamical properties of small to medium sized sodium (Na) clusters. We used the high dimensional feed forward neural networks to fit the DFT energies and gradients of sodium clusters. The details about how to construct such a neural network can be found from the recent literature [5–7]. The NN potentials combine the advantages of the speed of empirical potentials and the accuracy of ab initio methods. A number of recent papers have already shown the success of NN potentials in global optimizations (GO) [8] and molecular dynamics of bulk systems [9–12]. We tested the NN potential by performing GO on size selected Na clusters and by performing finite temperature calculations.

For Na clusters, a comprehensive GO has been performed to probe the structures of neutral and anion Na_n clusters for $n \leqslant 80$. Lowest energy structures obtained from GO were confirmed experimentally using photo electron spectra [13]. It was shown that Na clusters are based on fused icosahedral (Ih) packing. In case of small sized Na clusters ($n \leqslant 55$), theoretical studies on melting transitions were carried using second moment approximation (SMA) potential [14,15], distance dependent tight binding (DDTB) model [14], orbital free (OF) approach [16] and DFT methods [17–19]. It is well known that melting transition in small clusters depends on the ground state structure and therefore melting transition predicted is different for different levels of theory. Irrespective of the levels of theory melting transition in small clusters has a characteristic premelting peak along with a main melting peak. Other dominant factor in this size regime include electronic shell closing effects on melting point. Detailed investigations on melting properties based on electronic effects were done using ab initio molecular dynamic simulations [19]. It is confirmed from these studies that clusters which are highly symmetrical show a sharp

* Corresponding author.
E-mail address: sbulusu@iiti.ac.in (S.S. Bulusu).

http://dx.doi.org/10.1016/j.cplett.2016.04.013
0009-2614/

melting peak which can be explained due to electronic effects. Using OF and DFT methods a step wise melting in Na_{30} [16] and Na_{41}^{+} [17] was demonstrated to explain the mechanisms involved in melting process as the temperature increased. Apart from theoretical studies several experimental studies [20–25] exist on melting of Na clusters.

Section 2 describes the methodology and fitting procedures, in Section 3 we discuss GO results and MC simulation results obtained using our methodology.

2. Methodology and computational details

To train NN, we have to compute DFT energies and gradients on various Na clusters. All DFT calculations are done using TURBOMOLE software [26]. The Perdew, Burke, and Ernzerhof (PBE) functional [27] has been used for treating exchange–correlation of electrons. To reduce the computational effort we used resolution of identity (RI) approximation in DFT [28] along with def2-TZVP orbital basis set [29].

2.1. Initial data generation

To construct NN potential the first step is to generate initial data. For initial data generation, we made many random initial structures with clusters, ranging from 3 to 24 atoms, with various shapes, and bond lengths ranging from 2.5 Å to 4.0 Å. Starting from these structures we performed MC simulations at T = 300 K to generate small data set comprising of 10,000 data points using the SMA potential [30,31]. We then refined the data by removing correlations in structures and dangling bonds. We used the refined data, of about 3000 clusters, to construct preliminary NN potential. Once we got a preliminary potential (E_{RMSE} = 20 meV/atom), it was then used to run MC simulations at different temperatures. Approximately 30,000 data points with energy span of 0.21 eV per atom from the lowest energy structure, were used to train the NN.

2.2. Weight optimization of NN

We employed global extended kalman filter [32–34] for optimizing network weights. In the global extended kalman filter, the weight update was done at every kth cluster in the training set. The propagation error (PE), comprises of error in energy for kth cluster and sum of all the errors of the force components of each atom for a kth cluster as given in Eq. (1). The control parameters, w_e and w_f will balance the error due to energy and force. The force error depends on the number of atoms present in the cluster. During training process, clusters with different number of atoms are involved and hence the w_f parameter varies with number of atoms in the cluster. After running many tests, we found $w_e = 1$ and $w_f = e^{0.03N}$. Where, N is number of atoms per cluster. The weight penalty term w_α is used to avoid sudden changes in the weights during training process.

$$PE = w_e \Delta E + w_f \Delta F + w_\alpha W \tag{1}$$

$$\Delta E = \left(E_k^{DFT} - E_k^{NN}\right) \tag{2}$$

$$\Delta F = \frac{1}{3N_{atom}^k} \sum_{i=1}^{N_{atom}^k} \sum_{\alpha} \left(F_{k,i,\alpha}^{DFT} - F_{k,i,\alpha}^{NN}\right) \tag{3}$$

$$W = \sum_{n=1}^{N_w} (w_n)^2 \tag{4}$$

E_k^{DFT} and E_k^{NN} are energies calculated using DFT and NN potential respectively for the kth cluster in the data set of N_c structures. $F_{k,i,\alpha}^{DFT}$ and $F_{k,i,\alpha}^{NN}$ are the αth $= (x, y, z)$ force component on atom i of kth cluster in the data set of N_c structures and w_n is nth weight in the NN of N_w weights.

2.3. Functional form of NN and preparation of input for NN

The functional form and input vector used in present study is exactly as described in Ref. [7]. The functional form of NN for each atom i in the cluster is given in Eq. (5). The total energy (6) of the cluster is the sum of all atomic energy contributions.

$$E_i = b_1^3 + \sum_{l=1}^{30} a_{l1}^{23} \cdot f_l^2 \left[b_l^2 + \sum_{k=1}^{30} a_{kl}^{12} \cdot f_k^1 \left(b_k^1 + \sum_{j=1}^{51} a_{jk}^{01} \cdot G_{ij} \right) \right] \tag{5}$$

$$E = \sum_i E_i \tag{6}$$

We prepared the input vector for NN as 9 symmetrized coordinates from radial symmetric function and 42 symmetrized coordinates from angular symmetric function for each atom in the cluster. The cutoff function used is a cosine function with cutoff radius, R_c = 9 Å.

$$G_i^{rad} = \sum_j e^{-\eta R_{ij}^2} f_c(R_{ij}) \tag{7}$$

$$G_i^{ang} = 2^{1-\zeta} \sum_{j,k \neq i} (1 + \lambda \cos\theta_{ijk})^{\zeta} e^{-\eta(R_{ij}^2 + R_{ik}^2 + R_{jk}^2)} f_c(R_{ij}) f_c(R_{ik}) f_c(R_{jk}) \tag{8}$$

where $f_c(R_{ij}) = 0.5\left[\cos\left(\frac{\pi R_{ij}}{R_c}\right) + 1\right]$ for $R_{ij} \leqslant R_c$ and $f_c(R_{ij}) = 0$ for $R_{ij} > R_c$.

2.4. Training and testing of NN potential

The total data set was divided into a training set containing 85% of the data and a testing set containing 15% of remaining data. We found that the NN architecture of 51-30-30-1 (input layer of 51 nodes – hidden layer one of 30 nodes – hidden layer two of 30 nodes – output layer of 1 node) is adequate for the present work. The training process was done iteratively by minimizing the error function. We are validating the trained weights at every iteration by calculating the average root mean square error (RMSE) of energy and average RMSE of force for a test set using Eqs. (9) and (10).

$$\mathrm{RMSE}(E) = \sqrt{\frac{1}{N_{ts}} \sum_{k=1}^{N_{ts}} \left(E_k^{DFT} - E_k^{NN}\right)^2} \tag{9}$$

$$\mathrm{RMSE}(F) = \sqrt{\frac{1}{N_{ts}} \sum_{k=1}^{N_{ts}} \left[\frac{1}{3N_{atom}^k} \sum_{i=1}^{N_{atom}^k} \sum_{\alpha} \left(F_{k,i,\alpha}^{DFT} - F_{k,i,\alpha}^{NN}\right)^2 \right]} \tag{10}$$

where N_{ts} is the number of clusters in the test set. We found that the average RMSE of energy is 9 meV per atom and average RMSE of force is about 0.14 eV/Å per atom for the test set. Fig. 1 below shows the energies calculated for various cluster sizes in the test set using DFT and NN potentials. From the correlation coefficient ($R^2 = 0.98$) it is evident that NN predicts energies in good agreement with DFT.

3. Results and discussion

3.1. Global optimization of size selected Na clusters using NN potential

For testing the accuracy of our potential, we performed global optimizations using basin hopping (BH) method [35] and