Neural Network and ReaxFF comparison for Au properties

Jacob R. Boes[1], Mitchell C. Groenenboom[2], John A. Keith[*2], and John R. Kitchin[†1]

[1]Department of Chemical Engineering, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA 15213, USA
[2]Department of Chemical and Petroleum Engineering, University of Pittsburgh, Benedum Hall, 3700 O'Hara Street, Pittsburgh, PA 15261, USA

Abstract

We have studied how ReaxFF and Behler-Parrinello neural network atomistic potentials should be trained to be accurate and tractable across multiple structural regimes of Au as a representative example of a single component material. We trained these potentials using subsets of 9,972 Kohn-Sham density functional theory calculations and then validated their predictions against the untrained data. Our best ReaxFF potential was trained from 848 data points and could reliably predict surface and bulk data, but it was substantially less accurate for molecular clusters of 126 atoms or fewer. Training the ReaxFF potential to more data also resulted in overfitting and lower accuracy. In contrast, Behler-Parrinello neural networks could be fit to 9,734 calculations, and this potential performed comparably or better than ReaxFF across all regimes. However, the Behler-Parrinello neural network potential in this implementation brings significantly higher computational cost.

Keywords: Kohn-Sham density functional theory, neural networks, reactive force fields, potential energy surfaces, machine learning

1 Introduction

Growing interest in nanoscale phenomena has driven the need for accurate computational models to study chemical properties at an atomistic level. Advances in high performance

---

[*]jakeith@pitt.edu
[†]jkitchin@andrew.cmu.edu

computing now enable powerful simulation tools such as Born-Oppenheimer molecular dynamics (BO-MD) and extended Monte Carlo techniques that employ first principles quantum chemistry (QC) methods (e.g. Kohn-Sham density functional theory (KS-DFT)) on increasingly complex reactions. However, many genuine nanoscale systems of interest are currently too large to be studied with first principles QC and thus require the use of computationally efficient atomistic potentials.

Atomistic potentials approximate the potential energy surfaces (PES) for atomic systems by mapping potential energies and forces as functions of atomic positions. Physical potentials are parameterized to fit analytical expressions for the known physics of pairwise and many body interactions, e.g.: Lennard-Jones potentials,[1] classical force fields,[2–6] the modified and standard embedded atom methods (MEAM and EAM),[7,8] reactive force fields (ReaxFF),[9] and charge-optimized many-body (COMB) potentials.[10] Another intriguing category are mathematical potentials that "learn" the PES directly through a minimization of residuals with no *a priori* knowledge of the underlying physics, e.g. Gaussian regression functions[11] and artificial neural networks (NN).[12] We refer to these as mathematical potentials.

Physical potentials have been used for decades to capture the underlying physics of complex systems. It is substantially more challenging to develop so-called *reactive* potentials that can accurately model the underlying physics of systems where bonds are broken and formed. One approach is to develop potentials that use chemical bond-order dependent functional terms that are fit exclusively (or primarily) to QC datasets.[13] A challenge with these approaches is knowing what data should be used for training and how to parameterize potentials that best fit the data.

In contrast, mathematical potentials are becoming more popular in chemical applications, and recent work has shown reasons for excitement.[14,15] Specifically, recent descriptive models from Behler and Parrinello[16] have expanded the applications of neural networks to "high-dimensional" systems that can account for variable numbers of atoms, multiple compositions, and reactions involving thousands of atoms. Such networks have already been implemented on a large range of systems, including: Si bulk structures,[17] water clusters,[18] Cu surfaces,[19] ZnO,[20] and even a quaternary system of Cu/Au/H/O.[21] This opens the door for mathematical potentials to be developed that are accurate and transferable across diverse

bulk, surface, and cluster regimes.

In this work, we compare the performance of a widely used physical potential, ReaxFF, with the recently developed Behler-Parrinello neural network (BPNN) potential. We have trained both to subsets of a full dataset comprised of $\approx$ 10,000 KS-DFT calculations. We chose Au for this study due to its diversity of known nanoscale structures. The fact that long-range electronic interactions are screened in Au makes it an appropriate system to model with atomistic physical and mathematical potentials that are less suited for long-range interactions such as ReaxFF and BPNN. (We note that long range effects can be incorporated into BPNN potentials, e.g. as has been done for ZnO[20]).

We have benchmarked both potentials to our QC dataset that contains information from KS-DFT bulk equation of state (EOS) data, vacancy formation energies, surface energies, adatom diffusion profiles, slipping barriers, and cluster binding energies. Parameterization of ReaxFF potentials were automated using the Monte Carlo Force Field optimization (MCF-Fopt) tool in ADF.[22,23] Our BPNN was parameterized using the Atomistic Machine-learning Potentials (AMP) code from the Peterson group at Brown University.[24] This allows feed-forward neural networks to be developed inside the atomic simulation environment (ASE).[25] All details of the trained BPNNs are stored in a JSON file which can be found in the supporting information (SI) file.

2 Methods

2.1 Density Functional Theory

KS-DFT calculations for training set data were performed using the Vienna *ab initio* simulation package (VASP)[26–29] with the Perdew-Burke-Ernzerhof generalized gradient approximation (GGA-PBE)[30,31] exchange-correlation functional. Core electrons were described using the projector augmented wave function (PAW).[32,33] k-points were represented using Monkhorst-Pack grids[34] with a density of at least $14 \times 14 \times 14$ for a single atom of Au in the primitive ground state configuration. Kohn-Sham orbitals were expanded up to energy cutoffs of at least 300 eV to attain an energy convergence of at least 5 meV/atom. All

calculations involving relaxations were completed with relaxation criteria of < 0.05 eV/Å. Unless otherwise noted, transition states were determined using the climbing image nudged elastic band (NEB) method.[35] The details for all the KS-DFT calculations are included in an ASE database that is embedded in the SI file. Instructions on how to access this database and reproduce these calculations can also be found in the SI along with more details on the methods used in this work.

The full KS-DFT training set contained 9,972 calculations that included 905 bulk, 1,022 surface, and 8,045 cluster configurations. The majority of these calculations (9,076 calculations) were taken from coordinate relaxation steps performed by VASP. These structures are the incremental steps taken from its initially guessed positions to the ground state configurations predicted by GGA-PBE. Each of the structures in a particular relaxation are very similar from one relaxation step to the next. The remaining 896 calculations are either the local ground state configurations or images from optimized NEB calculations. Our bulk Au data were obtained by plotting EOS data for a variety of bulk structures. Vacancy formation and diffusion calculations were also included in the bulk dataset. Our surface dataset includes calculations on fcc(111) surfaces as well as a variety of fcc(100) surface diffusion pathways that were originally generated in previous work by Pötting et. al.[36] The training set used single-point energies on the latter coordinates (without geometry relaxations) calculated using the methods listed above. Our cluster data include various 3D ordered, planar, and disordered structures that contain up to 126 atoms. The SI file has further details about the data.

2.2 Reactive Force Field

Bond order based reactive force fields, such as Tersoff,[37] Brenner,[38] and ReaxFF[9,39] potentials, differ from classical force fields, such as UFF,[3,4] CHARMM,[2] or AMBER,[5,6] which require that defined bonds remain fixed over the course of a simulation. ReaxFF potentials developed for Au and other metals normally employ three separate energy terms as seen in Equation 1.[40–42]

$$E_{total} = E_{bond} + E_{over} + E_{vdw} \qquad (1)$$

E_{bond} is for bond energies of atom pairs, E_{over} is an energy penalty to prevent overcoordination, and E_{vdw} accounts for van der Waals interactions and interatomic repulsions when interatomic distances are too small. ReaxFF potentials can also be parameterized to include 3-body terms which provide energy contributions from valence angles between sets of three Au atoms. Backman et. al. developed a Tersoff potential for Au that involves 3-body terms,[43] but these terms are not always added to ReaxFF potentials for metals due to increased computational cost. Our 3-body terms have the same form as valence angle interactions in hydrocarbon ReaxFF potentials.[39] We report comparisons in equations of state and timings for ReaxFF using 3-body and 2-body terms in the SI file. Future work will discuss these topics in greater detail.

We parameterized our Au ReaxFF using the MCFFopt tool implemented in ADF.[22,23] MCFFopt seeks to minimize an objective function by randomly changing force field parameters within a predefined range. The Monte Carlo nature of this process allows some parameter changes that increase the objective function. This "annealing" allows the optimizer to sample a larger parameter space and potentially produce multiple distinct parameter sets. This approach can also find parameter sets with less total error than the traditional parabolic search parameter optimization.[23] Further information on running the MCFFopt procedure and optimized force field parameters are available in the SI file.

Au ReaxFF potentials appear to have an optimal training set size. Fitting to larger training sets does not always improve the quality of the ReaxFF potential, and this overfitting is found to bias predictions toward certain geometry types. As a result, the ReaxFF training set was constructed using the 848 ground state geometries from within the training set. Out of these geometries, the number of calculations classified as bulk, surface, and cluster structures are roughly equal. During ReaxFF parameterizations, each geometry in the training set is also assigned a weight depending on its relative importance in the overall fitting procedure. Our goal was to produce a ReaxFF potential with reasonable accuracy across these three different structure regimes, so most of the geometries were given a weight

of one (specific details are given in the SI file). In principle, one could increase weights to parts of the PES so that properties, such as desired lattice constants, bulk moduli, or barrier heights would be accurately reproduced. However, weighting a potential in this way will affect its ability to make accurate predictions in less-weighted regions of the PES.

Figure 1 shows the error distribution of residual error between the trained ReaxFF and KS-DFT training set data labeled by geometry type. Errors in bulk data greater than 0.2 eV stem from an unphysical convex region in the ReaxFF functional form which causes bulk EOS data to significantly deviate from the KS-DFT data at atomic volumes ranging from 60-200 Å^3/atom. Since these atomic volumes fall outside those found in most simulations involving bulk and surface structures of Au, these inaccuracies are not a cause for significant concern. However, large errors in bond energies for pairs of atoms at intermediate distances may be problematic for molecular clusters. Images of the entire EOS for each bulk structure can be found in the SI.
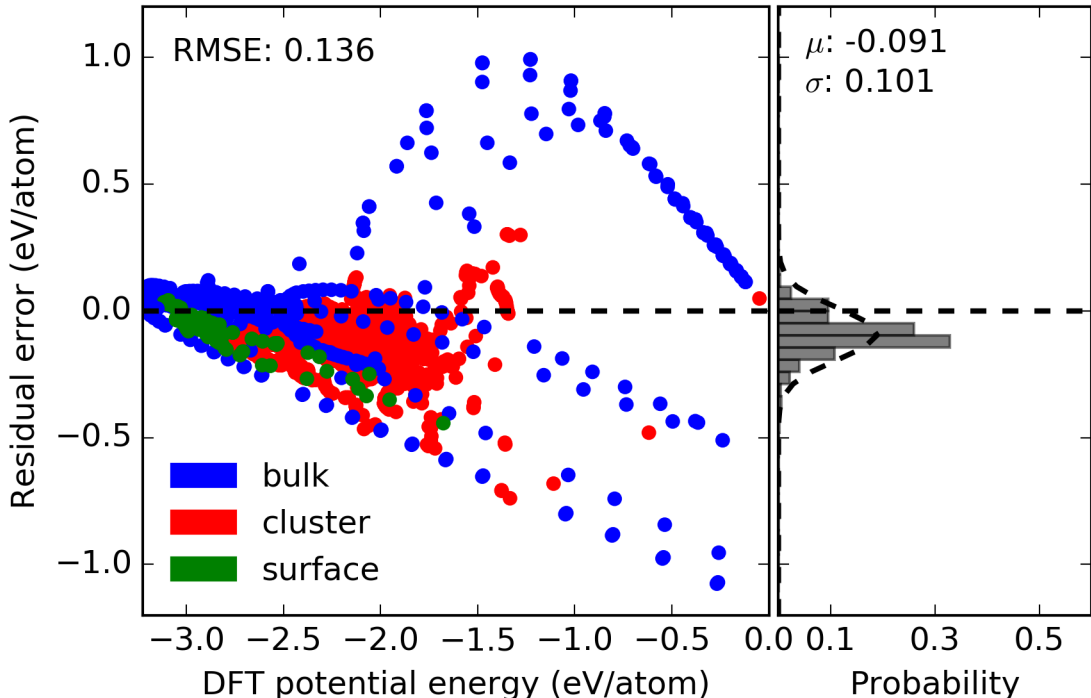

Figure 1: Energy residual error to the training set data broken down by bulk, surface, and cluster geometries for the ReaxFF potential.

A predefined validation set consisting of 238 calculations (out of the total 9,972 KS-DFT calculations) was set aside to test the transferability of predictions from our ReaxFF and BPNN potentials. This validation set was chosen to represent a variety of different Au structure types which are represented in the results section of this work. By reporting probability distributions for both the training and validation sets, we can determine the degree that our potentials show selection bias. For an ideal fitting procedure, the probability distributions for both the training and validation set would match, and any differences between the two would signify an over- or under-sampling. Figure 2 shows the residual error for the validation calculations labeled by geometry type. Significant deviations were found in bulk and cluster calculations from the validation and training set data.
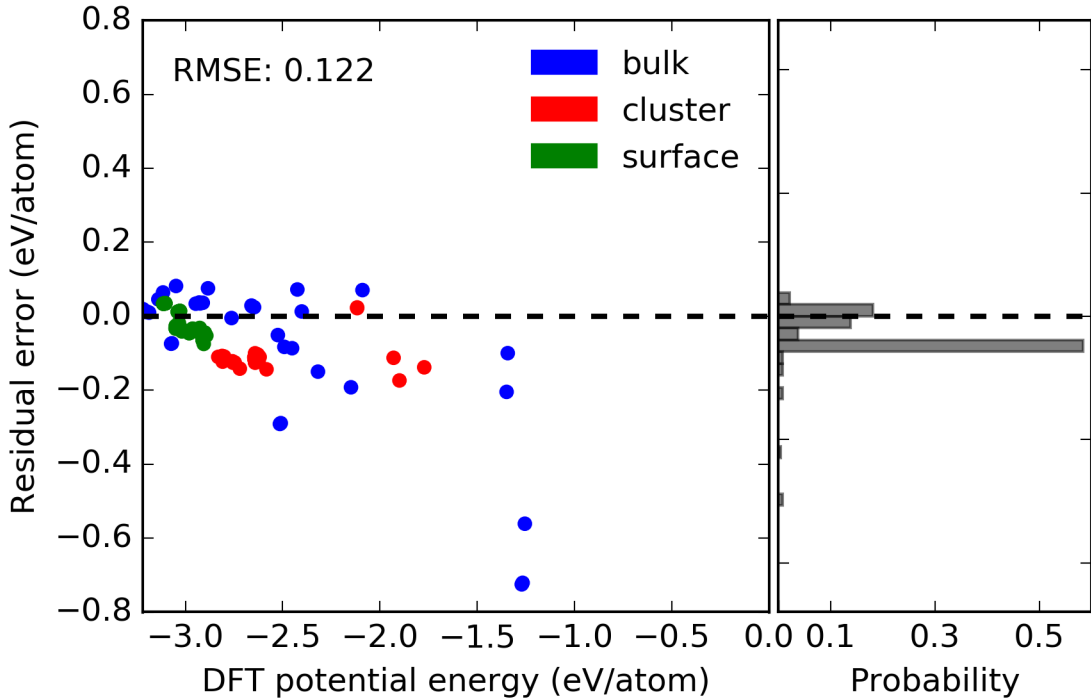

Figure 2: Energy residual error to validation set data broken down by bulk, surface, and clusters for the ReaxFF potential.

2.3 Neural Network

The NN is a machine learning algorithm. Unlike ReaxFF, potentials constructed from the NN have no physical basis, making them highly flexible, but also unsuitable for extrapolation. Nevertheless, NN potentials are growing in popularity due to their abilities to accurately characterize a PES from QC calculations. BPNN potentials were selected for this study because they can be trained to relatively large training sets.[16] BPNN potential fitting is facilitated through symmetry functions and utilization of multiple feed-forward NNs, one for each chemical species in the system. These modifications eliminate many of the shortcoming of traditional Cartesian NN potentials that are only applicable to systems with a fixed number of atoms.

Unlike other mathematical potentials, BPNN potentials[16,44] utilize a cutoff radius, R, that is applied to each atom in each image in the training set to keep the cost of the symmetry function small. Thus, the goal is to find a value of R that is large enough to capture meaningful atomic interactions but one that is not too large to result in high computational costs. These standard BPNN potentials are not suited for systems of atoms that have long-range interactions that extend outside the cutoff radii. For Au, we used $R = 6.5$ Å as long-range interactions are assumed to be negligible. (We find ≈ 2 meV/nearest-neighbor energy differences between gas phase Au and a primitive fcc unit cell with 6.5 Å nearest-neighbor distance). In the absence of this cutoff radius, it has been proven that NNs are capable of arbitrary levels of accuracy.[45]

The general structure of a feed-forward NN consists of nodes in an input layer, one or more hidden layers, and an output layer. In our BPNN potential, the nodes of the input layer are the Cartesian-coordinates of each atom in the unit cell. Each hidden layer is a linear combination of the values of the nodes from the previous layer. Each layer is also multiplied by an activation function (which often has a bounded non-linear form) to allow the NN potential to fit to arbitrary functions. The most accurate BPNN potential that we produced, and report in this work, utilizes four hidden layers with 40 nodes per layer and a hyperbolic tangent activation function. These specifications make our BPNN potential large compared to other BPNN potentials and more at risk for overfitting. However, the

root-mean square error (RMSE) of the validation set is similar to that of the training set, showing that overfitting has not occurred.[46] For further details on the theory behind BPNNs, we refer to previous work.[16,44]

We trained BPNN potentials using AMP, a code produced by the Peterson group at Brown University.[24] This software conveniently interfaces with the Atomic Simulation Environment (ASE) software package[25] for ease of reusability and reproducibility. The trained calculator parameters used by AMP are included in the SI file.

Of the 9,972 total calculations, 9,734 were used for training the BPNN potential. Figure 3 shows the error distribution from the training set. The mean, μ, and standard deviation, σ are given assuming a normal distribution fit. The RMSE is 0.017, similar to the standard deviation, indicating that the data is well approximated by a normal distribution overall.
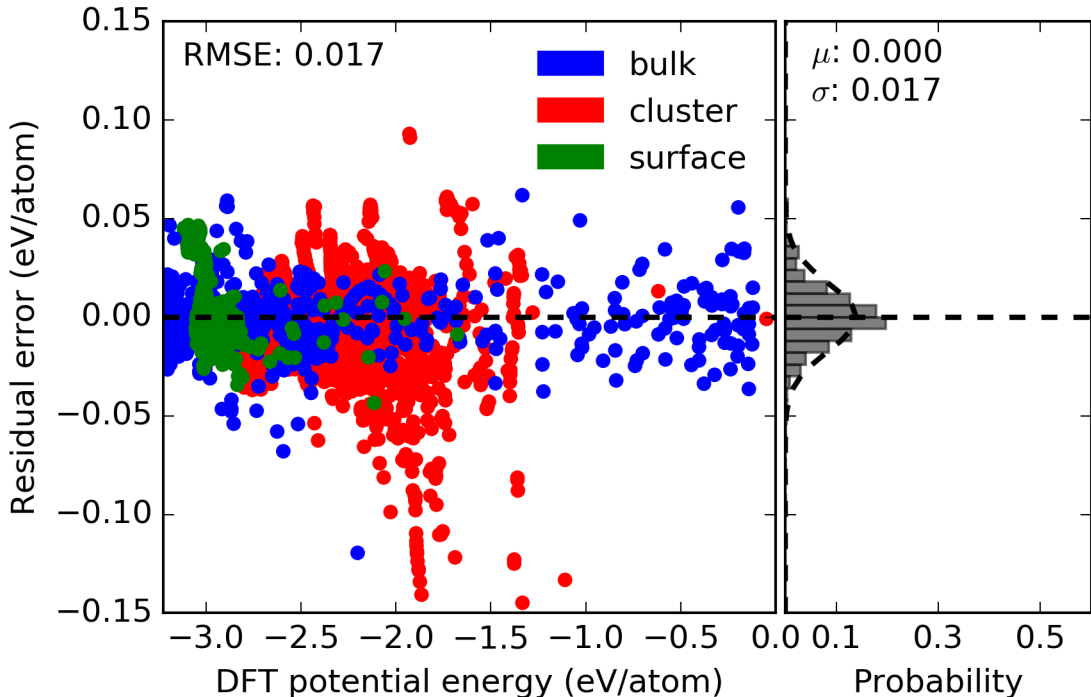

Figure 3: Energy residual error to the training dataset of the BPNN calculations. A RMSE of 0.017 eV/atom is calculated for the 9,734 structures included in the training set. The training set is also well described by a normal distribution.

Figure 4 shows the error distribution for the validation dataset. Overfitting can be

identified by a divergence between the RMSE of the training set and validation set data. In this case, the distribution is clearly not normal and arises from some underrepresented data in the training set, notably the fcc(100) terrace and dimer diffusion pathways (discussed below).
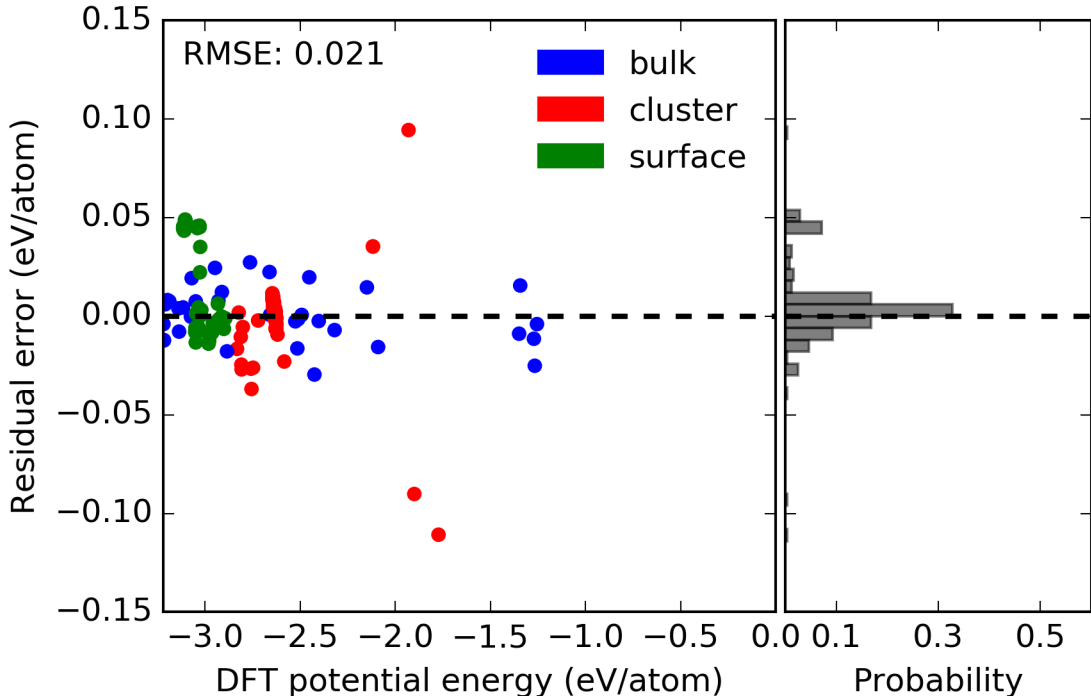

Figure 4: Energy residual error to the validation dataset of BPNN calculations. $\sigma = 0.21$, similar to the training set RMSE indicating little to no overfitting has occurred. The cluster of overpredicted surface calculations are from fcc(100) surface diffusion pathways, which are poorly represented in the training set.

3 Results and Discussion

We now benchmark the performance of the BPNN and ReaxFF potentials against KS-DFT energies across three different material regimes: bulk, surface, and molecular cluster structures. Both of our generated potentials can provide reasonably accurate descriptions of Au in the different material regimes. In general we find that ReaxFF potentials are more readily overfit, less transferrable to applications involving clusters of 126 atoms or fewer, and

overall less accurate than the BPNN. However, ReaxFF potentials demonstrate a notable strength by predicting barrier heights that resemble those found in their training sets. BPNN potentials in general are significantly more accurate than ReaxFF potentials, but they require significantly larger training sets to ensure well-balanced fitting. As explained below, they also currently bring substantially higher computational cost than ReaxFF potentials.

3.1 Bulk properties

3.1.1 Equations of state

EOS data for face centered cubic, simple cubic, and diamond structures are shown in Figure 5. All training and validation calculations are fit to a 3rd order inverse polynomial.[47] The metrics for each fit are included in Table 1. Results for the body centered cubic and hexagonal close packed EOS data are similar to the face centered cubic curve. Fits to all curves can be found in the SI file.
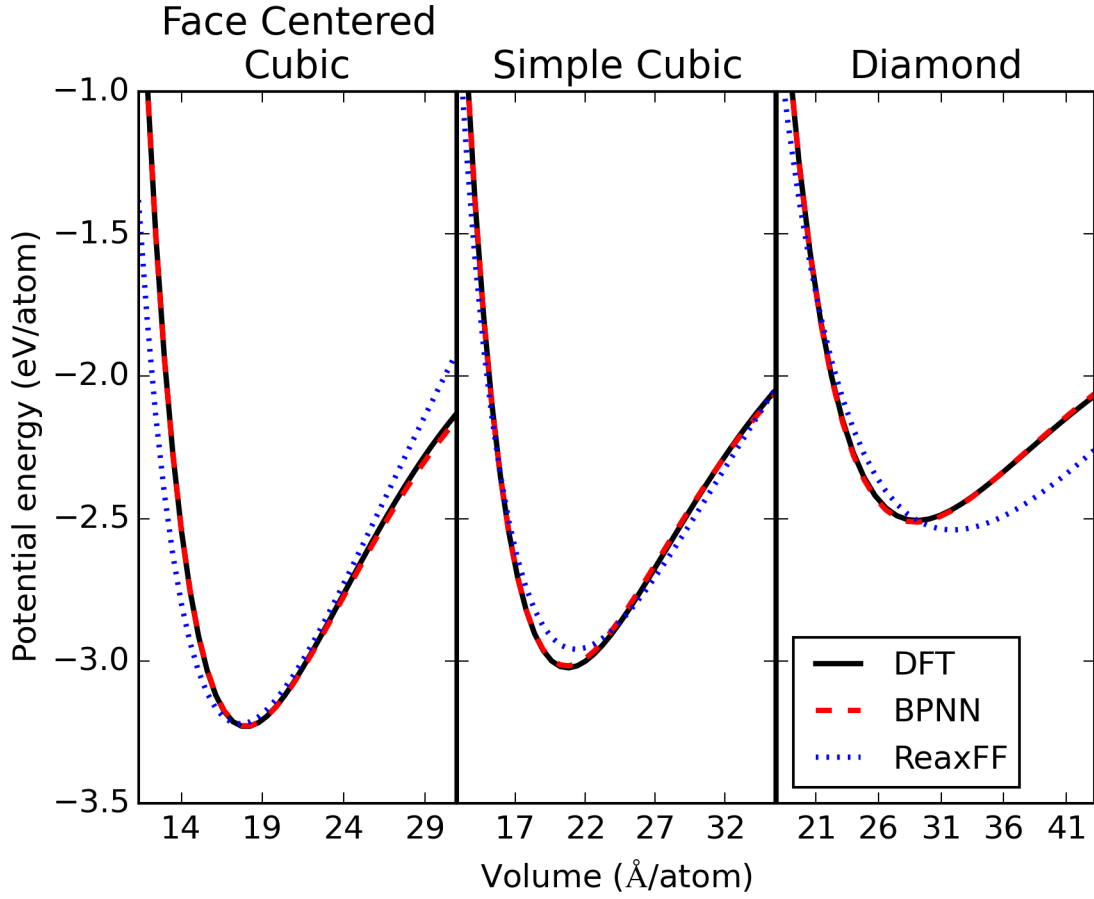
Figure 5: Comparison of EOS fits to KS-DFT, ReaxFF, and BPNN training and validation set data. Fits only include data within atomic volumes of ± 15 Å/atom as this is the region of interest for most applications.

Figure 5 shows that the EOSs are very well represented by our BPNN potential. Validation set data are also well behaved, indicating that overfitting has not occurred. Metric data shown in Table 1 shows excellent agreement in the minimum volume, minimum energy, and bulk modulus found using KS-DFT results. Data for the hcp and bcc structures shown in the SI file are reproduced similarly well.

We find that ReaxFF potentials with 3-body terms have substantially better fits compared to force fields which do not include 3-body interactions (see[41]). However, in both cases ReaxFF exhibits an unphysical convexity of the bond energy curve that creates problems manifested by large residual errors that can reach as high as ± 1 eV/atom at volumes

away from the minimum energy volume. Many simulations sample regions in the vicinity surrounding the minimum volume, so these deviations are not shown in Figure 5. Data from Table 1 shows reasonably good agreement for the equilibrium volume and minimum energy of the three structures. Bulk moduli are underpredicted by $\approx$ 20 GPa for each structure due to differences in the curvature of the EOS at the minimum. Again, one would likely improve the quality of predictions for individual properties by reweighting the parameterization to favor specific properties (e.g. bulk moduli), but this preferential fitting would also be expected to lower the quality of other predicted properties.

Table 1: Comparison of EOS metrics for KS-DFT, ReaxFF, and BPNN fits as shown in Figure 5.

Structure	Minimum volume ($\mathring{A}^3$)	Minimum energy (eV)	Bulk Mod. (GPa)
KS-DFT-fcc	17.97	-3.23	147
BPNN-fcc	17.99	-3.23	145
ReaxFF-fcc	17.60	-3.22	122
KS-DFT-sc	20.73	-3.02	110
BPNN-sc	20.66	-3.02	110
ReaxFF-sc	21.29	-2.96	84
KS-DFT-diam	29.04	-2.51	56
BPNN-diam	28.98	-2.51	57
ReaxFF-diam	31.92	-2.54	37

3.1.2 Bulk vacancy formation and diffusion barrier

Vacancy formation energies (E_v) are calculated using Equation 2. E_f, n_0, and E_i are the energies of the structure with vacancy, number of atoms in the structure before forming the vacancy, and energy of the structure before forming the vacancy, respectively. Our KS-DFT vacancy formation energies, shown in Figure 6, are in good agreement with other GGA-PBE calculations (0.42 eV), but both sets of data significantly underpredict experimental results (0.93 eV).[48] This is likely due to the well-known shortcoming of GGA-PBE in underpredicting atomization energies of Au.[?] In this work, vacancy formation is referenced to the energy

of a single atom in a primitive fcc unit cell. This may explain why the formation energies calculated here are slightly lower than those in the literature. The vacancies seem to reach the dilute concentration limit at ≈ 0.037 vacancies/atom. The anomalous increase in energy for the structure at ≈ 0.015 vacancies/atom is due to a minor structural perturbation into a different local minimum. More information about the nature of the reconfiguration can be found in the SI file.

$$E_v = E_f - \frac{n_0 - 1}{n_0} E_i \tag{2}$$

Our BPNN vacancy formation predictions are systematically overestimated by ≈ 0.4 eV while ReaxFF vacancy formation predictions are systematically underestimated by ≈ 0.3 eV. The preservation in trends indicates some error cancellation from the reference state for both fits. We find that neither method is sensitive enough to predict the subtle increase in energy for the reconfigured structure. Although the BPNN potential results are closer to experiment than the ReaxFF potential, this is simply a fortuitous error. BPNN potentials have no physical basis and therefore would reproduce the KS-DFT exactly with complete training.

The residual errors for structures with concentrations below 0.04 vacancies/atom are very low (less than 0.006 eV/atom, even for the point in the validation set having ≈ 0.037 vacancies/atom). Error cancellation between the vacancy structures and reference structure make it difficult to determine the level of precision needed to obtain accurate vacancy formation energies. A BPNN potential for Cu has been constructed with a higher level of accuracy (error < 0.11 eV), at the increased cost of a basis of calculations which is ≈ 3.5 times larger.[19] BPNN calculations were also performed using unit cells of the same size as the corresponding vacancy structure. The same trend was observed with slightly higher formation energies using the expanded reference super cell.
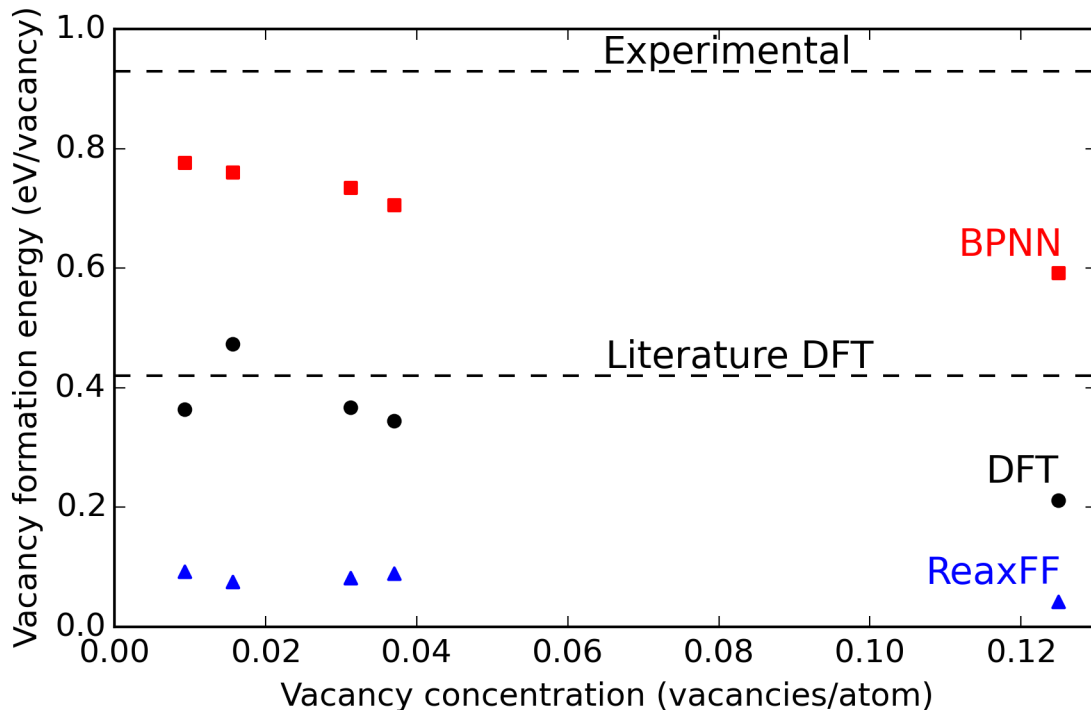
Figure 6: Bulk vacancy formation energies for fcc Au at various concentrations. BPNN fits to vacancy structures are systematically overpredicted by ≈ 0.4 eV, while ReaxFF fits are systematically underpredicted by ≈ 0.3 eV. Literature values are from Ref. 48.

Figure 7 shows the calculated bulk vacancy diffusion barrier using a vacancy concentration of ≈ 0.037 vacancies/atom (obtained from Figure 6). NEB calculations determined points along the minimum energy pathway that were then fit to a cubic spline. For diffusion calculations, the residual errors of both the BPNN and ReaxFF potentials are lower by about an order of magnitude as compared to the vacancy formation energy. This is due to error cancellation from the reference states that are similar to states along each reaction pathway. The BPNN potential overestimates this barrier by 0.04 eV while the ReaxFF potential underestimates the barrier by 0.05 eV.
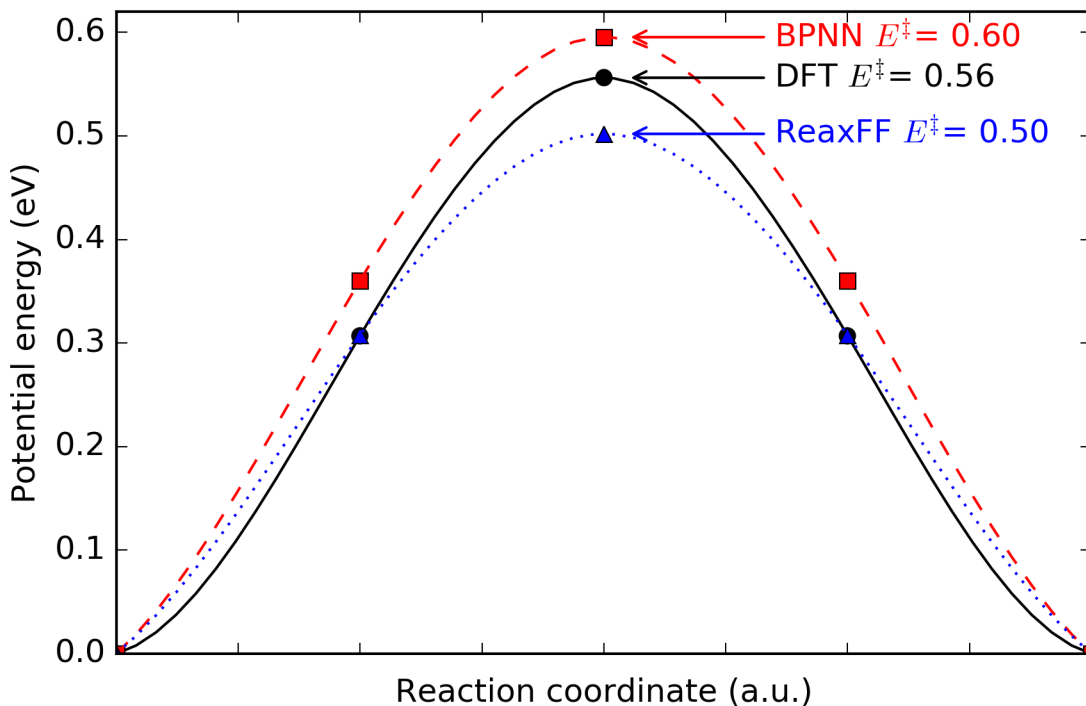
Figure 7: NEB predicted barrier for bulk vacancy diffusion through fcc Au. Transitions state energy (black, 0.56 eV) is overpredicted by the BPNN (red, 0.60 eV) and underpredicted by the ReaxFF (blue, 0.50).

3.2   Surface calculations

3.2.1   fcc(100) diffusion barriers

The training set for the ReaxFF potential in Reference 41 contains 166 surface diffusion barrier calculations from GGA-PBE using the SEQQUEST code.[49] NEB calculations with VASP were not used to recalculate the minimum energy pathways, but we recalculated single point energies on these structures using GGA-PBE in VASP to be consistent with the rest of our training set. Since NEB calculations were not done, there are significantly fewer points sampling the PES for these pathways compared to other pathways (8-10× fewer in most cases). Consequently, our BPNN fits to these pathways are expected to be less accurate compared to other pathways obtained from NEB calculations.

Figure 8 contains recreations from Figure 2 (a & b) in Ref. 41 using the BPNN potential

and ReaxFF potential. Note that the terrace and dimer diffusion pathways are not included in the training set for either potential, and they represent predictions by both potentials. For the terrace diffusion pathway, the ReaxFF potential performs quite well and shows that the ReaxFF potential can provide very accurate predictions of barrier heights when the training set contains similar pathways. The BPNN potential, which contains more than $10\times$ the training set data as the ReaxFF potential, can reasonably produce this adatom diffusion barrier but residuals fall between 0.2-0.3 eV. On the other hand, for a different adatom diffusion barrier, the BPNN potential predicts the dimer diffusion pathway quite well while the ReaxFF potential has higher residual errors between 0.1-0.2 eV. Larger training sets can be expected to reduce errors in both potentials, but reparameterization of these potentials with a larger training set will undoubtedly impact the accuracy when predicting other pathways.
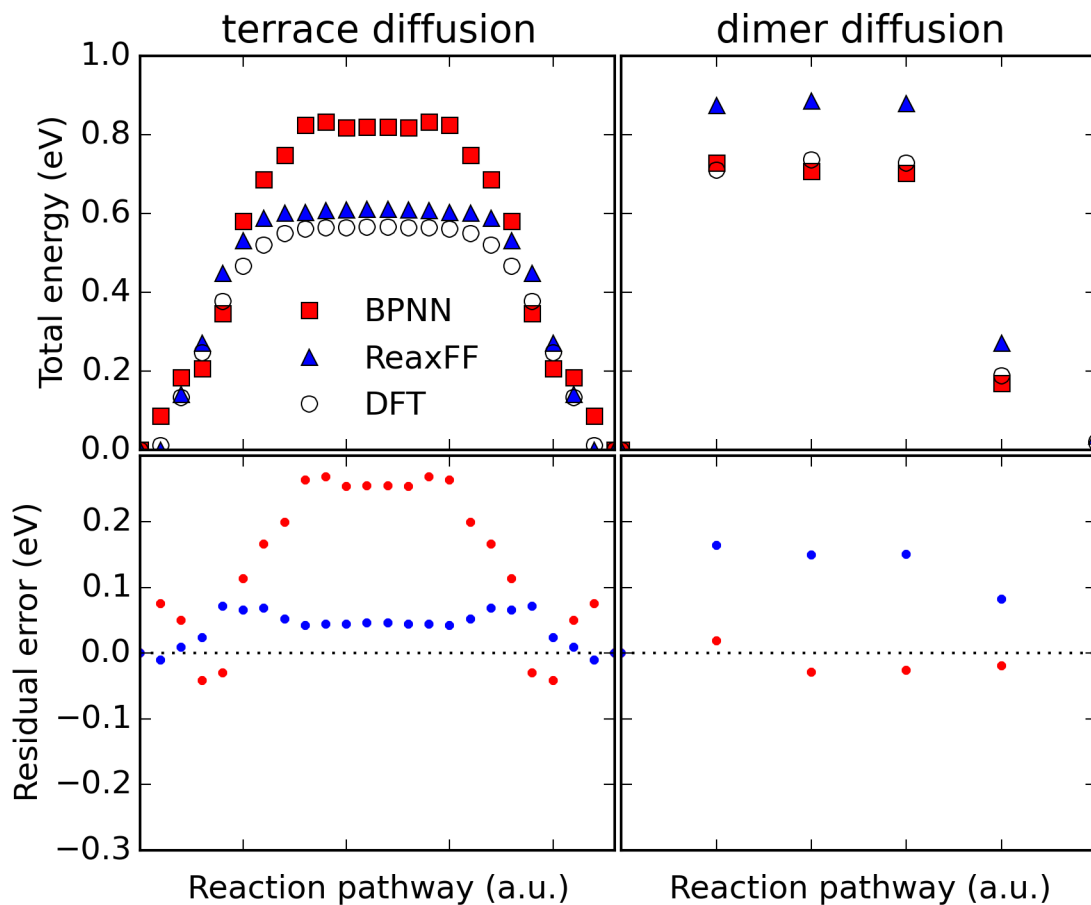
Figure 8: Residuals to diffusion pathways in the validation set. Structures are reproduced from those used in Ref. 41.

To assess the performance of these potentials under a wide range of adatom diffusions, Figure 9 shows the residuals for all 144 fcc(100) surface diffusion calculations. Solid shapes represent training set data and hollow shapes represent validation set data. Residuals are the same as those shown in Figure 8. Our ReaxFF potential (which has roughly 1/3 of its training set devoted to surface calculations) has 86.1% of these structures falling within a ± 0.1 eV tolerance of error. For the BPNN potential (with roughly 1/10 of its training set devoted to surface calculations), has 52.1% of these structures fall within a ± 0.1 eV tolerance of error.

Many of the calculations from the BPNN potential are underestimated compared to the reference KS-DFT data, signifying (as stated above) that these structures come from a poorly

sampled region of the PES and improvements could be attained with more training. For the
ReaxFF potential, errors appear to be less systematic, showing improved accuracy would
require more training to specific pathways. In practice, both ReaxFF and BPNN potentials
are normally trained with a specific application in mind, and so training sets, particularly
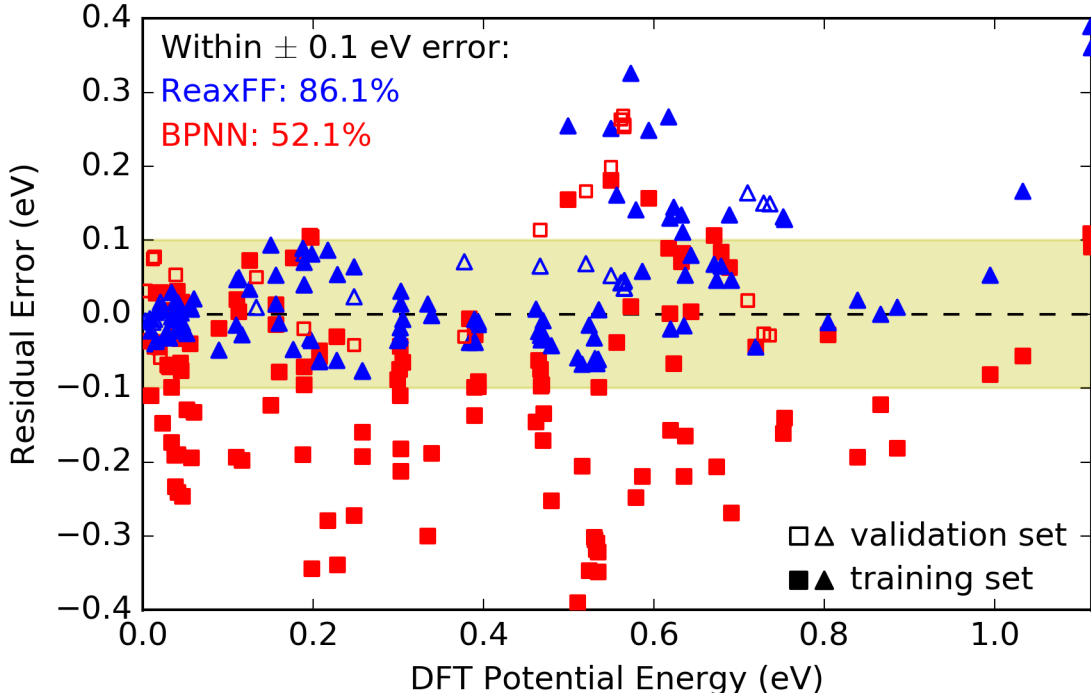those for ReaxFF potentials, can be smaller.

Figure 9: Residuals of 144 fcc(100) surface diffusion pathway calculations included from Ref.
41. Hollow markers represent residuals from the validation set which are shown in Figure 8.

3.2.2   fcc(111) surface slipping barrier

A slipping barrier is the minimum energy pathway required for a certain number of mono-
layers of atoms to move from their ground state site to the next most adjacent site of the
same kind. Slipping barriers were performed on fcc(100) and fcc(111) surfaces for one and
two layers in a five layer slab. Figure 10 shows the single-layer slipping barrier for the
fcc(111) surface. Both models find almost identical energies as KS-DFT (within 0.05 eV).
We can see that the ReaxFF potential finds a metastable intermediate instead of a single

barrier as found by KS-DFT and the BPNN potential. This ReaxFF potential also finds metastable intermediates when slipping in a different direction primarily over bridge sites (see the SI file), but residual errors are even lower. The very small difference in energies makes it difficult to assess if these are due to fitting errors or an unphysical component within the ReaxFF potential. Either way, both potentials can reproduce low energy slipping barriers within 0.05 eV with sufficient training.
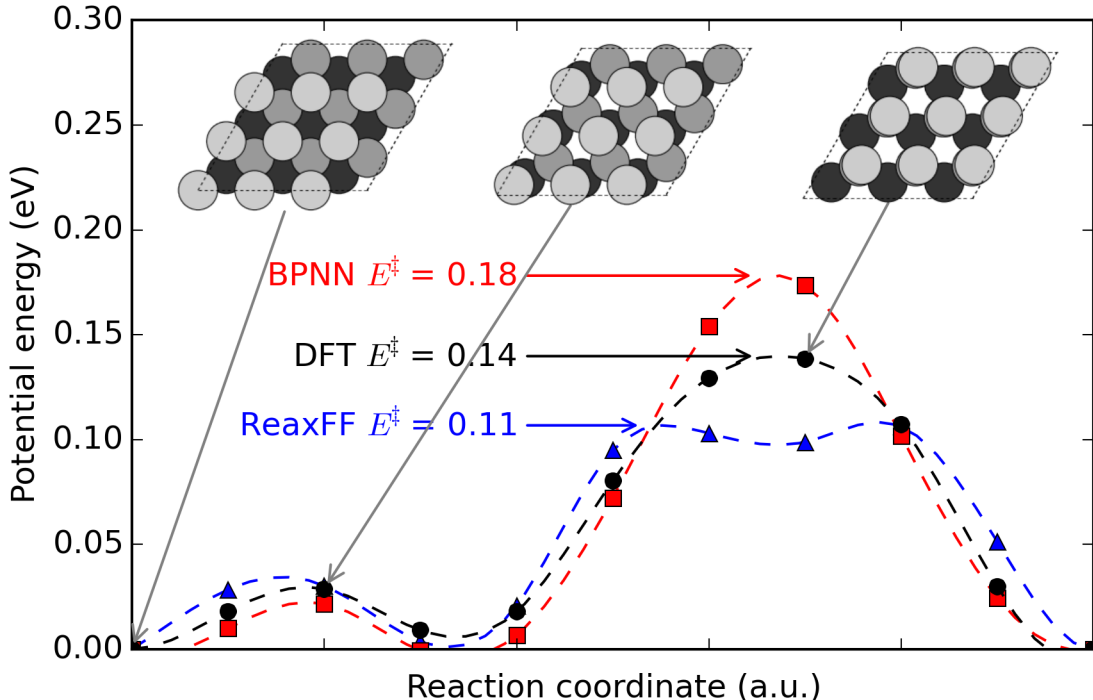

Figure 10: NEB predicted slipping barrier for a single layer of fcc(111). Initial, bridge, and top positions are shown for visual reference. The second local minima is representative of the hcp site. The darkest gray represents the deepest layer, while the lightest shade is the top layer.

3.3 Cluster predictions

3.3.1 6 atom clusters

Calculations on clusters up to 126 atoms make up $\approx 81\%$ of the entire database. To determine the robustness of the BPNN potential for determining the energetics of structures not

incorporated into the training set, several BO-MD simulations were performed on various clusters. For a six atom cluster, calculations were performed with NVT BO-MD without planar boundary conditions, where the temperature of the system was changed from 800 K to 300 K over the course of the simulation. The simulation using the BPNN potential started from a local minimum structure to determine if it would locate the known global minimum energy configuration. GGA-PBE found the global minimum to be planar and triangular (see Figure 11), which is also observed in the literature.[50] This structure was not included in the training set.

Figure 11 depicts the path taken by the BPNN BO-MD simulation (red) over the course of 2,000 time steps. Once every 100 steps we validated the energy using KS-DFT. The residuals are less than 0.05 eV/atoms for the BPNN potential, including the structure of the global minimum. We re-ran this simulation several times throughout development of the database. The first attempt at performing the described BO-MD simulation was with a dataset of $\approx 2,000$ cluster calculations with 20 atoms or fewer. In comparison with the full dataset, the residual error has been reduced dramatically, and the success rate of discovering the global minimum improved significantly. Further details of these initial attempts with the smaller database can be found in the SI file.

The 2,000 structures generated from the BO-MD run with our BPNN potential were then calculated using the ReaxFF potential. In this case, the ReaxFF potential did not identify the same minimum energy configuration of the six atom system. However, the cohesive energies of structures resembling the planar cluster are fairly consistent with KS-DFT data. Although the presented data shows situations where ReaxFF is not accurate, we note that this may signify an area where ReaxFF could be extended with additional functionality. For example, metal-metal bonds in small clusters could be treated with functional forms different than those used for bulk metal-metal bonds. This would likely correct systematic deviations, but such re-parameterizations may also adversely affect other structure types and/or increase computational cost. We note that Narayanan et. al. have reported a hybrid bond order potential that uses a screened Lennard-Jones term for bulk structures in combination with a highly trained Tersoff potential for smaller regimes.[51] This is a possible work-around to make other physical potentials accurate across different size regimes.
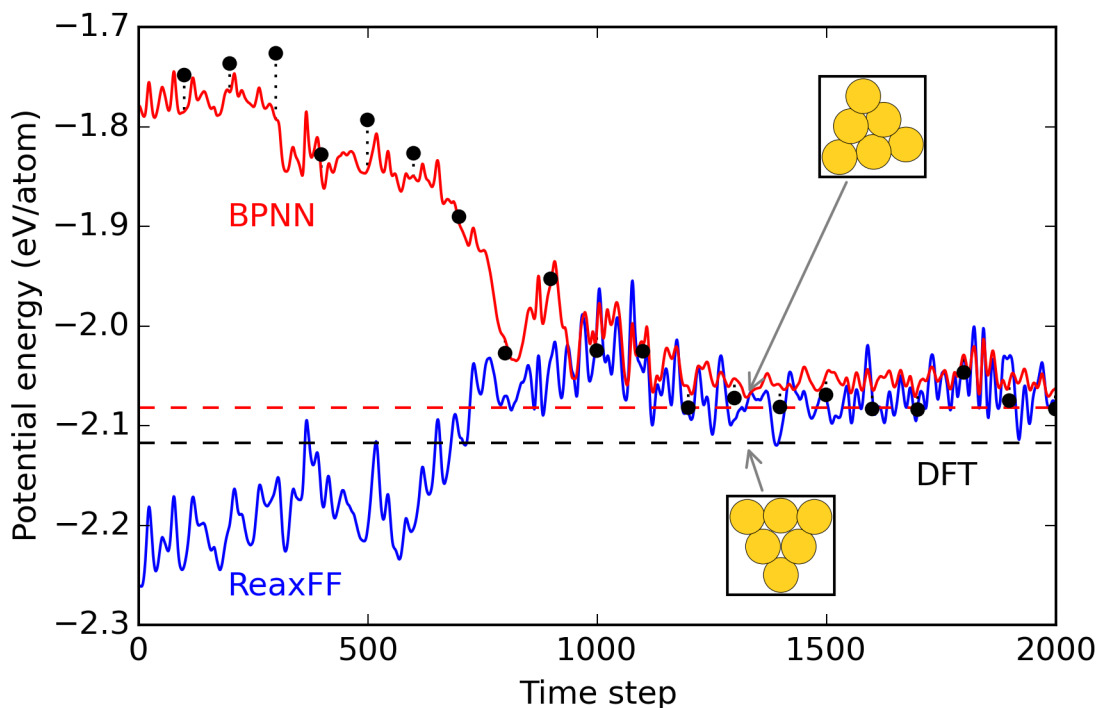
Figure 11: NVT BO-MD simulation of 6 atom cluster starting from local energy minimum and finding the global minimum. The temperature was reduced from 800 K to 300 K over the course of the simulation. Solid lines show BO-MD trajectories while dashed lines show energy predictions for the global minima from KS-DFT (black) and BPNN (red) and ReaxFF (blue).

3.3.2   38 atom clusters

A similar exploration for multiple local minima was implemented on a 38 atom cluster using minima hopping techniques.[52] This exploration of minimum energy structures works through a series of fixed temperature NVT BO-MD simulations followed by geometric optimization requiring a significant number of calculations between each iteration. After each iteration, the minimum geometry is stored and perturbed before restarting its search. The resulting minima predicted from 125 such iterations are shown in Figure 12.

Again, this approach located a lower energy minimum than the starting point geometry. The largest energy difference between minima occurred during the first iteration of the process. After this initial step, the energies do not change as dramatically. This can be

interpreted as a shift into a local minimum energy basin (a group of configurations with similar atomic positions and energies) which the BPNN potential proceeds to explore in the next 124 minima. A more complete analysis of the 38 atom Au cluster space would be time consuming and is beyond the scope of this work. Despite demonstrating low residual errors, the BPNN potential does not correctly predict the lowest energy structure determined by KS-DFT in this set of minima. Regardless, it is still capable of distinguishing between configurations in different basins, and thus could be a valuable tool for exploring minimum energy structures in conjunction with KS-DFT calculations.

Residual errors for the ReaxFF potential are consistently lower by -0.11 eV/atom compared to KS-DFT. If energetics are shifted by this amount (as show in the top of Figure 12) one finds that the trend in relative energies is in reasonable agreement with KS-DFT, although our ReaxFF potential does not correctly predict the lowest energy configuration either. The performance of the ReaxFF potential for clusters could always be improved by adding more cluster data to its training set, but we found that doing so rapidly deteriorates its ability to calculate bulk and surface properties. As a result, we do not recommend using ReaxFF in its standard formalism for applications involving clusters with fewer than 126 atoms.
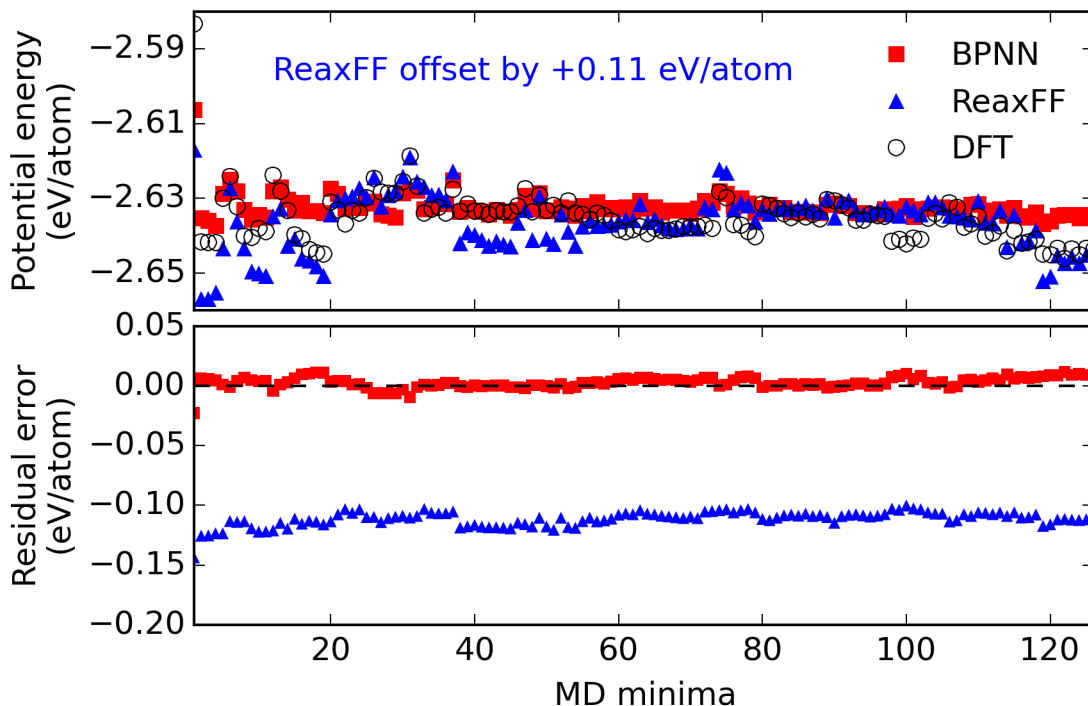
Figure 12: Local minima for 38 atom Au cluster predicted from the BPNN (red) and compared with KS-DFT (black) and ReaxFF (blue). The ReaxFF potential energies are offset by +0.11 eV/atom in the top figure to better depict the trend in energies.

3.4   Computational cost

An important aspect of these modeling approaches is their computational cost. This includes the time needed to produce the necessary QC training sets, train the potentials, and the time needed to run the calculations. Implementation and training of parameters for both the ReaxFF and BPNN potentials can be automated using instructions in our SI file, thus reducing the time needed to learn how to train potentials. The generation of meaningful QC data is also a significant bottleneck in time, particularly for BPNN potentials that require large training sets to be accurate. This is simplified in part by generating NEB data and geometry optimizations which contain many valuable calculations on which physical and mathematical potentials can be trained. One of the best ways to speed the progress of developing accurate and transferrable potentials is to make data and methods freely available

and easily accessible.

A fair comparison between calculation times between ReaxFF and BPNN potentials is not currently possible. The BPNN potential we developed used a Python code that is still in early stages of development.[24] In comparison, ReaxFFs and other force field codes have been implemented in the LAMMPS program which is already a high performance computing code.[53] Using available open source tools, BO-MD simulations on the 6 atom cluster using the C-compiled ReaxFF code performs $\approx$ 6,700 timesteps/second, while the Python BPNN in ASE performs $\approx$ 15 timesteps/second. Nevertheless, we consider BPNN potentials to be extremely promising for simulations requiring high accuracy, especially if they can be implemented into high performance codes that can dramatically accelerate their calculation times.

4   Conclusions

We have trained ReaxFF and BPNN potentials using subsets of $\approx$ 10,000 KS-DFT calculations. Our training sets consider Au in a variety of atomic configurations in bulk, surface, and cluster regimes that would be useful for practical atomistic modeling across all regimes. By virtue of being a mathematical potential, the BPNN potential can be trained to an arbitrary level of accuracy. Our most accurate BPNN potential was fitted to 9,734 calculations and yields an RMSE of 0.017 eV. Our ReaxFF potential (which contains 3-bond terms for higher accuracy) was fitted best to a significantly smaller training set consisting of 848 calculations (a value that is considerably larger than parameter sets in many other ReaxFF potentials). This potential provides an overall RMSE of 0.136 eV compared to the full KS-DFT dataset.

In applications on bulk structures, our BPNN almost exactly reproduces reference QC data of equations of state, while the ReaxFF potential is less accurate, particularly at atomic volumes that extend far beyond the equilibrium structures. When modeling surface structures and adatom diffusions, both the ReaxFF and BPNN potentials perform quite well with sufficient training, but obtaining a BPNN potential having comparable or higher accuracy than ReaxFF for adatom diffusions requires substantially larger training sets. For clusters, the BPNN potential exhibits essentially negligible residual errors compared to the KS-DFT

calculations it was trained to, while the ReaxFF potential exhibits sizable systematic errors of 0.11 eV. This highlights the challenge of developing a physical potential that is accurate across bulk, surface, and cluster data. Increasing the size of the training set for the ReaxFF potential to include more cluster data was found to be detrimental to the accuracy of bulk and surface data, thus showing an area needing improvement in terms of ReaxFF functionality.

Although BPNNs can be trained to the desired level of accuracy, the computational cost, both upfront in the form of training set data and during calculation time, are currently substantially higher than ReaxFF potentials. Nevertheless, BPNN potentials are very promising if trained for specific applications (hence requiring smaller training sets) and they will be highly intriguing as computational developments enable faster runtimes. Since accurate BPNN potentials contain substantially larger numbers of parameters than most physical potentials, it is unlikely that BPNN potentials will ever be as fast as ReaxFF potentials, but we have demonstrated that BPNN potentials can be trained to be substantially more accurate.

5   Acknowledgements

References

1. J. E. Jones, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences **106**, 463 (1924), URL http://dx.doi.org/10.1098/rspa.1924.0082.

2. B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and

M. Karplus, J. Comput. Chem. **4**, 187 (1983), URL http://dx.doi.org/10.1002/jcc.540040211.

3. A. K. Rappe, C. J. Casewit, K. S. Colwell, W. A. Goddard, and W. M. Skiff, J. Am. Chem. Soc. **114**, 10024 (1992), URL http://dx.doi.org/10.1021/ja00051a040.

4. C. J. Casewit, K. S. Colwell, and A. K. Rappe, J. Am. Chem. Soc. **114**, 10046 (1992), URL http://dx.doi.org/10.1021/ja00051a042.

5. W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, J. Am. Chem. Soc. **117**, 5179 (1995), URL http://dx.doi.org/10.1021/ja00124a002.

6. W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, J. Am. Chem. Soc. **118**, 2309 (1996), URL http://dx.doi.org/10.1021/ja955032e.

7. M. S. Daw and M. I. Baskes, Phys. Rev. Lett. **50**, 1285 (1983), URL http://dx.doi.org/10.1103/PhysRevLett.50.1285.

8. M. I. Baskes, Phys. Rev. B **46**, 2727 (1992), URL http://dx.doi.org/10.1103/PhysRevB.46.2727.

9. A. C. T. van Duin, S. Dasgupta, F. Lorant, and W. A. Goddard, J. Phys. Chem. A **105**, 9396 (2001), URL http://dx.doi.org/10.1021/jp004368u.

10. T. Liang, B. Devine, S. R. Phillpot, and S. B. Sinnott, J. Phys. Chem. A **116**, 7976 (2012), URL http://dx.doi.org/10.1021/jp212083t.

11. C. E. Rasmussen, *Advanced Lectures on Machine Learning: ML Summer Schools 2003, Canberra, Australia, February 2 - 14, 2003, Tübingen, Germany, August 4 - 16, 2003, Revised Lectures* (Springer Berlin Heidelberg, Berlin, Heidelberg, 2004), chap. Gaussian Processes in Machine Learning, pp. 63–71, ISBN 978-3-540-28650-9, URL http://dx.doi.org/10.1007/978-3-540-28650-9_4.

12. S. S. Haykin, *Neural Networks and Learning Machines* (Prentice Hall/Pearson, Upper Saddle River, New Jersey, 2008), 3rd ed.

13. T. Liang, Y. K. Shin, Y.-T. Cheng, D. E. Yilmaz, K. G. Vishnu, O. Verners, C. Zou, S. R. Phillpot, S. B. Sinnott, and A. C. van Duin, Annu. Rev. Mater. Res. **43**, 109 (2013), URL http://dx.doi.org/10.1146/annurev-matsci-071312-121610.

14. J. Behler, Phys. Chem. Chem. Phys. **13**, 17930 (2011), URL http://dx.doi.org/10.1039/c1cp21668f.

15. J. Behler, Journal of Physics: Condensed Matter **26**, 183001 (2014), URL http://dx.doi.org/10.1088/0953-8984/26/18/183001.

16. J. Behler and M. Parrinello, Phys. Rev. Lett. **98**, 146401 (2007), URL http://dx.doi.org/10.1103/PhysRevLett.98.146401.

17. J. Behler, R. Martoňák, D. Donadio, and M. Parrinello, physica status solidi (b) **245**, 2618 (2008), URL http://dx.doi.org/10.1002/pssb.200844219.

18. T. Morawietz and J. Behler, J. Phys. Chem. A **117**, 7356 (2013), URL http://dx.doi.org/10.1021/jp401225b.

19. N. Artrith and J. Behler, Phys. Rev. B **85**, 045439 (2012), URL http://dx.doi.org/10.1103/PhysRevB.85.045439.

20. N. Artrith, T. Morawietz, and J. Behler, Physical Review B **83**, 153101 (2011), URL http://dx.doi.org/10.1103/PhysRevB.83.153101.

21. N. Artrith and A. M. Kolpak, Computational Materials Science **110**, 20 (2015), URL http://dx.doi.org/10.1016/j.commatsci.2015.07.046.

22. G. te Velde, F. M. Bickelhaupt, E. J. Baerends, C. F. Guerra, S. J. A. van Gisbergen, J. G. Snijders, and T. Ziegler, J. Comput. Chem. **22**, 931 (2001), URL http://dx.doi.org/10.1002/jcc.1056.

23. E. Iype, M. Hütter, A. P. J. Jansen, S. V. Nedea, and C. C. M. Rindt, J. Comput. Chem. **34**, 1143 (2013), URL http://dx.doi.org/10.1002/jcc.23246.

24. A. Peterson and A. Khorshidi, *Amp: Atomistic machine-learning potentials*, https://bitbucket.org/andrewpeterson/amp, accessed: 2015-02-25.

25. S. R. Bahn and K. W. Jacobsen, Comput. Sci. Eng. **4**, 56 (2002), ISSN 1521-9615.

26. G. Kresse and J. Hafner, Physical Review B **47**, 558 (1993), URL http://dx.doi.org/10.1103/PhysRevB.47.558.

27. G. Kresse and J. Hafner, Physical Review B **49**, 14251 (1994), URL http://dx.doi.org/10.1103/PhysRevB.49.14251.

28. G. Kresse, Physical Review B **54**, 11169 (1996), URL http://dx.doi.org/10.1103/physrevb.54.11169.

29. G. Kresse and J. Furthmüller, Computational Materials Science **6**, 15 (1996), URL http://dx.doi.org/10.1016/0927-0256(96)00008-0.

30. J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996), URL http://dx.doi.org/10.1103/PhysRevLett.77.3865.

31. J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **78**, 1396 (1997), URL http://dx.doi.org/10.1103/PhysRevLett.78.1396.

32. P. E. Blöchl, Physical Review B **50**, 17953 (1994), URL http://dx.doi.org/10.1103/PhysRevB.50.17953.

33. G. Kresse, Physical Review B **59**, 1758 (1999), URL http://dx.doi.org/10.1103/PhysRevB.59.1758.

34. H. J. Monkhorst and J. D. Pack, Physical Review B **13**, 5188 (1976), URL http://dx.doi.org/10.1103/PhysRevB.13.5188.

35. G. Henkelman, B. P. Uberuaga, and H. Jónsson, The Journal of Chemical Physics **113**, 9901 (2000), URL http://dx.doi.org/10.1063/1.1329672.

36. K. Pötting, W. Schmickler, and T. Jacob, Chem. Eur. J. of Chem. Phys. **11**, 1395 (2010), URL http://dx.doi.org/10.1002/cphc.200900807.

37. J. Tersoff, Phys. Rev. B **37**, 6991 (1988), URL http://dx.doi.org/10.1103/PhysRevB.37.6991.

38. D. W. Brenner, Phys. Rev. B **42**, 9458 (1990), URL http://dx.doi.org/10.1103/PhysRevB.42.9458.

39. K. D. Nielson, A. C. T. van Duin, J. Oxgaard, W.-Q. Deng, and W. A. Goddard, J. Phys. Chem. A **109**, 493 (2005), URL http://dx.doi.org/10.1021/jp046244d.

40. T. T. Järvi, A. Kuronen, M. Hakala, K. Nordlund, A. C. van Duin, W. A. Goddard, and T. Jacob, The European Physical Journal B **66**, 75 (2008), URL http://dx.doi.org/10.1140/epjb/e2008-00378-3.

41. J. A. Keith, D. Fantauzzi, T. Jacob, and A. C. T. van Duin, Physical Review B **81**, 235404 (2010), URL http://dx.doi.org/10.1103/PhysRevB.81.235404.

42. J. M. Cabrera-Trujillo, J. M. Montejano-Carrizales, F. Aguilera-Granja, and Á. Posada-Amarillas, The European Physical Journal D **69**, 167 (2015), URL http://dx.doi.org/10.1140/epjd/e2015-60058-y.

43. M. Backman, N. Juslin, and K. Nordlund, The European Physical Journal B **85**, 317 (2012), URL http://dx.doi.org/10.1140/epjb/e2012-30429-y.

44. J. Behler, The Journal of Chemical Physics **134**, 074106 (2011), URL http://dx.doi.org/10.1063/1.3553717.

45. K. Hornik, M. Stinchcombe, and H. White, Neural Networks **2**, 359 (1989), URL http://dx.doi.org/10.1016/0893-6080(89)90020-8.

46. J. Behler, Int. J. Quantum Chem. **115**, 1032 (2015), URL http://dx.doi.org/10.1002/qua.24890.

47. A. B. Alchagirov, J. P. Perdew, J. C. Boettger, R. C. Albers, and C. Fiolhais, Phys. Rev. B **67**, 026103 (2003), URL http://dx.doi.org/10.1103/PhysRevB.67.026103.

48. W. Xing, P. Liu, X. Cheng, H. Niu, H. Ma, D. Li, Y. Li, and X.-Q. Chen, Physical Review B **90**, 144105 (2014), URL http://dx.doi.org/10.1103/PhysRevB.90.144105.

49. P. A. Schultz and P. Muller., *Seqquest electronic structure code*, http://dft.sandia. gov/Quest (2002), accessed: 2015-02-25.

50. N. Phaisangittisakul, K. Paiboon, T. Bovornratanaraks, and U. Pinsook, Journal of Nanoparticle Research **14** (2012), URL http://dx.doi.org/10.1007/ s11051-012-1020-4.

51. B. Narayanan, A. Kinaci, F. G. Sen, M. J. Davis, S. K. Gray, M. K. Y. Chan, and S. K. R. S. Sankaranarayanan, ArXiv e-prints (2015), accessed: 2015-02-25, 1511.03319.

52. S. Goedecker, The Journal of Chemical Physics **120**, 9911 (2004), URL http://dx.doi. org/10.1063/1.1724816.

53. S. Plimpton, Journal of Computational Physics **117**, 1 (1995), URL http://dx.doi. org/10.1006/jcph.1995.1039.