Complexity competes with transferability in machine-learning force fields: The example of Gold-Iron interactions

Magali Benoit,1 Jonathan Amodeo,2 Ségolène Combettes,1 Ibrahim Khaled,3 Aurélien Roux,3 and Julien Lam1,3

1CEMES, Université de Toulouse, CNRS
2Université de Lyon, INSA-Lyon, MATEIS, UMR 5510 CNRS, 69621 Villeurbanne, France
3Center for Nonlinear Phenomena and Complex Systems, Université Libre de Bruxelles, Code Postal 231, Boulevard du Triomphe, 1050 Brussels, Belgium

Machine-learning force fields have been increasingly employed in order to extend the possibility of current first-principle calculations. While the main approach consists in using a force field with the most universal and most complete analytical formulation, here we propose an alternative where we purposely design the simplest potential possible. In particular, we use a linear combination of traditional force fields and optimize the linear coefficients using a penalizing regression scheme that reduces the complexity of the resulting potential. After validating the method in the case of gold-iron structures, we study the competition between the complexity and the transferability of the potential.

Atomistic modeling is often divided in two different types of calculations. On the one hand, quantum methods including Hartree-Fock and DFT approaches are considered the most accurate and are employed for virtually any types of chemical species.2 On the other hand, classical force fields are used to perform large-scale and long-time simulations.3 However, it is still difficult to connect both approaches and one can hardly perform a simulation involving millions of atoms for nanoseconds while retaining the accuracy of quantum methods.

For that purpose, machine-learning interaction potentials (MLIP) have been proposed in the recent years and have shown great potentials.4,5 Numerous approaches are currently considered including Artificial Neural Networks6, Gaussian approximation methods7, Spectral Neighbor Analysis Potential8, Symmetric Gradient Domain Machine learning9 and Moment Tensor Potentials10. The success of the these techniques is recognized by the large variety of materials that were tackled including pure metals11–13, organic molecules14–16, oxydes17,18, water19–23, amorphous materials24–26 and hybrid perovskites27.

For all of these techniques, the main procedure consists in using a very universal analytical formulation for the force field which is then parameterized to match a database made of DFT calculations including total energy, forces and stress tensors. However, machine-learning potentials sometimes show poor transferability and can not always be predictive for systems that are not included in the learning database. In the worst scenario, the machine-learning potential is so-well fitted to its learning database that outside of it, non-physical behaviors are obtained when dealing with non-learned configurations. Altogether, such overfitting behavior originates from the underlying complexity of the machine-learning analytical formulation which is particularly enhanced when non-linear regression schemes are involved.

In order to fix such issue, the main proposal is to regularly check the accuracy of the potential as the machine-learning molecular dynamics simulations are carried out and improve the MLIP “on the fly.”28,29 Here, instead of increasing the complexity of the mathematical formulation to reach more accurate MLIP, we propose the simplest potential possible while remaining accurate enough for some practical applications. The choices that are made participate in the overall wish for a more physically informed usage of machine-learning30,31.

Regarding the studied materials, we choose bi-metallic gold-iron systems for which the aim is to combine properties of both of its chemical components. In particular, recent experiments have shown that the synthesized Au-Fe nanoparticles are made of an iron core wrapped in a gold shell and that the shape of the iron core depends strongly on the amount of surrounding gold.32–34 These nanoparticles have potential applications in systems with biological interests as iron is known for its intrinsic ferromagnetism and gold capping can protect the iron core from oxidation. However, rationalizing the results of the synthesis along with predicting the material properties would require numerical simulations which are sparse for gold-iron nanoparticles.35,36 Indeed, while full quantum calculations can not be employed to study clusters of more than tens of atoms, the difficulties in obtaining a reliable force field for gold-iron nanoparticles originate from both its multi-species nature and the preponderance of surface effects at the nanoscale.37

In the following, we will first describe the methodology for the MLIP and the DFT database generation. Then, we will employ it to obtain an accurate force field by learning on structures of gold-iron nanoclusters. Finally, we will demonstrate how the specifics of our method support a control over the complexity and the accuracy of the potential. Then, we will show that non-monotonic behavior with respect the complexity of the potential can emerge when testing the obtained potential on untrained...
Regarding the analytical formulation, the total potential energy of a configuration made of $N$ atomic positions is first given by $E_{\text{tot}} = \sum_{i=0}^{N} E_i$, where $E_i$ is its atomic counterpart. For $E_i$, we employ a linear combination of numerous descriptors indexed by $n$: \[
abla = \sum_n \omega_n X_n^{(i)} \tag{1}
\]
where $\omega_n$ is the linear coefficient associated with the descriptor $X_n^{(i)}$. At this stage, moment tensors [14], group-theoretical high-order rotational invariants [54] and bispectrum components [10, 55, 59] were previously proposed as descriptors for such linearized potentials. In our case, we will purposely consider a simpler formulation which consists in developing the descriptor space in explicit two-body, three-body and n-body interactions:

\[
\begin{align*}
\text{2B}^{(i)}_{n,ij} &= \sum_j f_n(R_{ij}) \times f_c(R_{ij}) \\
\text{3B}^{(i)}_{n,ijk} &= \sum_j f_n(R_{ij}) f_c(R_{ij}) f_n(R_{ik}) f_c(R_{ik}) \cos(\theta_{ijk}) \\
\text{NB}^{(i)}_{n,ijkl} &= \left[ \sum_j f_n(R_{ij}) \times f_c(R_{ij}) \right]^n
\end{align*}
\tag{2}
\]

where $R_{ij}$ is the distance between atoms $i$ and $j$, $\theta_{ijk}$ is the angle centered around the atom $i$ and $l$ is a positive integer which is not necessarily the same in the three-body and the n-body descriptors. We chose for the cutoff function $f_n$ to mimic traditional interatomic potentials as e.g., Morse, Lennard-Jones, Buckingham, ZBL or Yukawa potentials. Likewise, for the three-body interactions, the current formulation is very similar to what is done in the Stillinger-Webber potential [57]. Finally, the n-body interaction is a generalized form of the EAM potential where the embedding function is a polynomial of the atomic density [58]. Our potential formulation is analogous to what was developed by Seko et al. [15, 16, 54, 59, 60]. Yet, in the n-body interactions, we did not include any explicit angular dependence and did not mix different forms of the atomic density. As such, we purposely reduce the complexity of the potential formulation in order to reach the most simplified MLIP possible. Altogether, it allows us to retrieve some physical intuition but it also decreases the flexibility of the potential and possibly its accuracy.

For the fitting procedure of such a linear model, previous studies have proposed the use of genetic algorithm [10, 55], weighted ordinary least squares [56], Bayesian linear regression [27], ridge regression [10, 54, 60] and Lasso regression [15, 59]. In order to continue in our wish to reach the simplest MLIP possible, we employ the Lasso-Lars regression scheme for the fitting procedure [61]. In practice, along with the ordinary least square objective function, $\alpha^2$, the Lasso scheme adds a penalty on the sum over the absolute value of the coefficients $\omega_n$ and the employed error function is therefore given by:

\[
\chi^2 = \alpha^2 + \sum_n \left| \omega_n \right|
\tag{5}
\]

where $\alpha$ is a parameter that controls the degree of penalty. The so-called $L_1$ penalty enforces lots of the linear coefficients to be exactly 0. On top of that, using the Lars algorithm allows us to select the most relevant descriptors by measuring their correlation to the target. Choosing LassoLars as a regression scheme is at the expense of accuracy and flexibility for the MLIP but it allows for a considerable reduction in the complexity of the potential.

In the following, we will compare LassoLars with two commonly employed linear regression schemes naming Ridge and the coordinate descent Lasso. For this test, the database is generated with an equimolar mixture of binary Lennard-Jones particles thus allowing to directly verify the obtained MLIP. With such a binary system, an additional challenge for the fitting algorithm is to distinguish self-species and cross-species interactions. In practice, positions and forces were measured for 50 configurations of 64 atoms in the liquid regime. Regarding the basis of descriptors, only two-body interactions were considered and we used 17 Lennard-Jones functions with different distance parameters including those in the original simulations. All of the four employed methods manage to retrieve a linear combination of Lennard-Jones functions that matches the original interactions. However, it appears that only LassoLars can find the correct coefficients $\omega_n$ setting all coefficients to 0 except those of the original interactions [See Fig. 1]. Such result shows the advantage of using LassoLars instead of the commonly employed linear regression methods.

Now, we will apply our method to the gold-iron system for which there is a lack of interatomic potentials in the literature. This is largely due to the fact that, because these two metals are immiscible, there is little or no experimental data on which adjusting the parameters of traditional interatomic potentials. Ultimately, the aim of this work is study issues associated with transferability in machine-learning potentials. Therefore, we purposely did not design the most complete database possible and for instance did not include neither any bulk structures nor gold-iron interface. Instead, we constructed three types of nanoparticles: (1) Alloys with almost equimolar compositions, (2) Pure iron in the body-centered cubic (bcc) phase and (3) Pure gold in the face centered cubic (fcc) phase [See Fig. 2a]. By using nanoparticles instead of bulk structures, we enforce a change in the coordination number at the surface that should help us capture n-body interactions. Then, molecular dynam-
ics (MD) simulations were carried out to melt the constructed nanoparticles. We used simple pair-wise potentials made of Lennard-Jones and Morse interactions for gold and iron respectively and of Lennard-Jones interaction for the gold-iron cross-interaction. The Morse potential for iron was found in the literature[62] while the two Lennard-Jones potentials for gold-gold and iron-gold were simply parameterized in order to match the bulk lattice parameters and the cohesive energy. Along the melting path, we extracted configurations that are representative of the solid to liquid transition. For each of these configurations, forces were finally computed at the DFT-level using single-point calculations. Spin-polarized DFT calculations were performed with the VASP code [63], using PAW type pseudopotentials for iron and gold[64], a plane wave cutoff of 650 eV and a Methfessel-Paxton smearing parameter $\sigma$ of 0.01 eV. All calculations were done at the $\Gamma$-point of the Brillouin zone. Altogether, the database is made of 167262 atomic configurations with an almost equal proportion in the three types of nanoparticles (34% of alloy, 34% of pure gold and 32% of pure iron). We note that the choice of the initial interaction potential employed for these MD simulations is not critical as forces are computed at the DFT-level afterwards. This approach simply allows us to quickly sample configurations that are physically sound. In addition, by employing configurations that span from the crystalline to the liquid regime, we assure a large variety of atomic neighborhoods and the forces goes from $10^{-4}$ to 5 eV/Å.

Regarding the fitting procedure, we used only Gaussian functions $f_n = \exp(-a_n(R_{ij} - b_n)^2)$ for which $a_n$ and $b_n$ are respectively sampled over the lists $[0.5, 1.0, 1.5]$ and $[1.2, 3, 4.5, 6]$. For the three-body and the $n$-body interactions, we used respectively $l = [1, 2, 3, 4, 5]$ and $l = [4, 5, 6, 7]$. Altogether, we have 990 descriptors which include Au-Au, Fe-Fe and Fe-Au interactions. The training set is made of 95% of the whole database and the remaining serves as a test set. The root mean square error (RMSE) on the forces is computed to measure the accuracy of the obtained potential. In our case, by varying the cut-off from 3 Å to 8 Å, we found the best RMSE with a cutoff equal to 6 Å and kept that number for the rest of the study.

In Fig. 2.b, it appears that the fitting error for the train and the test sets are similar for all the values of $\alpha$. This shows that the MLIP does not lead to overfitting issues during the learning process. Then, the main advantage of using the LassoLars algorithm consists in having a parameter that enables for a direct control over both the accuracy and the complexity of the obtained potential. According to Fig. 2, by increasing $\alpha$, the number of non-zero coefficients can be reduced at the expense of increasing the RMSE. As such, with the LassoLars algorithm, one can simply chose which degree of accuracy or complexity is required for their usage. Furthermore, we note that regardless of the desired accuracy, at most only 184 descriptors were selected while we started with 990 ones. Finally, the presence of a plateau for the smallest values of $\alpha$ shows that the LassoLars regression manages to reduce the potential complexity by only selecting the relevant descriptors.

The remaining of the article is dedicated to testing the obtained potential on untrained systems. In particular, we wish to study how having this control over the complexity of the MLIP results in a control over its transferability as well.

In this first test, we wanted to check the ability of
our potential to retrieve bulk properties of the material while only being trained on nanostructures. In particular, we measured eight different lattice constants corresponding to the following cases: pure iron bcc, pure gold facefcc, alloys with 25%, 50% and 75% of gold in both bcc and fcc phases. For all of these cases, we computed the relative error on the lattice parameters when compared to DFT calculations obtained previously\[52\].

In addition, we also measured elastic constants for the pure bcc iron and fcc gold and compared them with experimental results from Ref. \[65, 66\]. Simulations were carried out using the large-scale molecular dynamics software LAMMPS \[67\]. Tri-periodic simulations cells were used and the different minimization runs were performed down to a net force of $10^{-6}$ eV/Å for low numbers of non-zero coefficients ($N_{\text{coeff}}<15$) which correspond to high degree of penalty ($\alpha > 5 \times 10^{-4}$). A large error is found for both the lattice and the elastic constants \[See Fig.3\]. This corresponds to very poorly fitting potentials with force RMSE that are too large. Then, the relative errors on the lattice parameters and the elastic constants are small for all of the obtained potentials once $N_{\text{coeff}} > 15$ \[See Fig.3\]. In our training set, we employed alloying proportions that are much closer to 50% and only used structures from nanoparticles. Therefore, being able to reach such a small relative error even for those extreme alloying proportions (25% and 75%) and in addition with bulk structures is an encouraging result for the transferability of our MLIP. Surprisingly, the relative error behaves non-monotonically with the number of non-zero coefficients. Such results mean that while having more complex formulation for the MLIP contributes to a better RMSE on the training set, it does not necessarily lead to a more transferable potential.

For the second test, three additional morphologies of gold-iron nanoparticles were designed: two Janus and one core-shell \[See Fig.4a\]. \[Fig.4b\] shows the corresponding RMSE on the forces without having trained the potential on these structures. The RMSE behavior is non-monotonic which again demonstrates that precautions should be made and over fitting can easily be found when using machine-learning approaches. The minimum is located for the three structures around 90 coefficients which corresponds to $\alpha = 10^{-5}$. There, the configuration that is the least well-fitted is the core-shell structure for which the RMSE remains smaller than 0.2 eV/Å. In practice, it shows that our methodology leads to a MLIP that can be applied from the alloy and pure metal structures to the Janus and core-shell morphologies. Such results could originate from the simplicity of the employed potential formulation. The use of a traditional framework for our machine-learning potential may indeed lead to its higher stability with respect to unknown structures.

To conclude, we proposed a novel method to derive accurate force fields based on DFT calculations and machine learning. We made three specific choices for our method: 1) The potential is developed as a linear combination, 2) Explicit two-body, three-body and n-body interactions are employed as descriptors, 3) The linear regression scheme is the LassoLars for which we presented the advantages on a practical example based on Lennard-Jones interactions. Then, we tested the method on the gold-iron system. First, the method allows one to obtain an RMSE on the forces around 0.1 eV/Å which is on par with most of the currently employed MLIP methods. Then, we show that the penalizing parameter $\alpha$ controls the complexity of the potential as it allows one to turn off most of the initially proposed descriptors. Finally, instead of trying to obtain a general purpose interatomic potential for the gold-iron system, we studied the intricate relationship between complexity and transferability.

![FIG. 3. Influence of the number of non-zero coefficient over the percentage of error on (a) the lattice constant for pure and alloy structures and (b) the elastic constants for pure structure.](image)

![FIG. 4. Images of untrained configurations with morphology Janus (a,b) and Core-shell (c). (d) Fitting error as a function of the number of non-zero coefficients for each untrained configurations and for the training set.](image)
within machine-learning approaches. Two very challenging tests were carried out: (1) Retrieve bulk properties while only training on nanoparticles, (2) Compute forces on nanoparticles with untrained morphologies. For both cases, we showed that non-monotonic behavior can be observed and that increasing the accuracy on the trained configurations and the overall complexity of the potential is not necessarily accompanied with an improvement on its transferability on untrained systems.

Acknowledgments

JL acknowledges financial support of the Fonds de la Recherche Scientifique - FNRS. Computational resources have been provided by the Consortium des Equipements de Calcul Intensif (CECI) and by the Fédération Lyonnaise de Modélisation et Sciences Numériques (FLMSN). JL thanks David Tew for having introduced him to the machine-learning potential problem and C. Patrick Royall for supervising the early method development. JL is also grateful to Francesco Turci, Joshua F. Robinson, Abdul-Rahman Allouche and Atsuto Seko for fruitful discussions.

[1] D. Marx, *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods* (Cambridge University Press, Leiden, 2009).
[2] R. Martin, *Electronic structure: basic theory and practical methods* (Cambridge University Press, Cambridge, UK New York, 2008).
[3] D. Frenkel and B. Smit, *Understanding Molecular Simulation* (Elsevier, Academic Press, 2002).
[4] A. J. Stone, *The theory of intermolecular forces* (Oxford University Press, Oxford, 2016).
[5] J. Behler, *Int. J. Quantum Chem.* 115, 1032 (2015).
[6] A. P. Bartók and G. Csányi, *Int. J. Quantum Chem.* 115, 1051 (2015).
[7] J. Behler, *J. Chem. Phys.* 145, 170901 (2016).
[8] J. Behler and M. Parrinello, *Phys. Rev. Lett.* 98, 146401 (2007).
[9] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, *Phys. Rev. Lett.* 104, 136403 (2010).
[10] A. P. Thompson, L. P. Swiler, S. M. Foiles, and G. J. Tucker, *J. Comput. Phys.* 285, 316 (2015).
[11] S. Chmiela, A. Tkatchenko, H. E. Sauceda, I. Poltavsky, K. T. Schütt, and K.-R. Müller, *Sci. Adv.* 3, e1603015 (2017).
[12] S. Chmiela, H. E. Sauceda, K.-R. Müller, and A. Tkatchenko, *Nat. Commun.* 9, 1 (2018).
[13] A. V. Shapoval, *Multiscale Model. Simul.* (2016).
[14] I. I. Novoselov, A. V. Yanilkin, A. V. Shapoval, and E. V. Podryabinkin, *Comput. Mater. Sci.* 164, 46 (2019).
[15] A. Seko, A. Takahashi, and I. Tanaka, *Phys. Rev. B* 92, 054113 (2015).
[16] A. Takahashi, A. Seko, and I. Tanaka, *J. Chem. Phys.* 148, 234106 (2018).
[17] C. Zeni, K. Rossi, A. Glielmo, Á. Fekete, N. Gaston, F. Baletto, and A. De Vita, *J. Chem. Phys.* 148, 214139 (2018).
[18] V. Bóto, R. Batra, J. Chapman, and R. Ramprasad, *J. Phys. Chem. C* 121, 511 (2017).
[19] T. Bereau, R. A. DiStasio, A. Tkatchenko, and O. A. van Lilienfeld, *J. Chem. Phys.* DETC2018, 241706 (2018).
[20] H. E. Sauceda, S. Chmiela, I. Poltavsky, K.-R. Müller, and A. Tkatchenko, *J. Chem. Phys.* 150, 114102 (2019).
[21] A. P. Bartók, S. De, C. Poelking, N. Bernstein, J. R. KerMODE, G. Csányi, and M. Cerviotti, *Sci. Adv.* 3, e1701816 (2017).
[22] M. Veit, S. K. Jain, S. Bonakala, I. Rudra, D. Hohl, and G. Csányi, *J. Chem. Theory Comput.* 15, 2574 (2019).
[23] N. Arthith, T. Morawietz, and J. Behler, *Phys. Rev. B* 83, 153101 (2011).
[24] V. Quaranta, M. Hellström, and J. Behler, *J. Phys. Chem. Lett.* 8, 1476 (2017).
[25] F. T. Nguyen, E. Székely, G. Imbalzano, J. Behler, G. Csányi, M. Ceriotti, A. W. Götz, and F. Paesani, *J. Chem. Phys.* DETC2018, 241725 (2018).
[26] A. P. Bartók, M. J. Gillan, F. R. Manby, and G. Csányi, *Phys. Rev. B* 88, 054104 (2013).
[27] T. Morawietz, V. Sharma, and J. Behler, *J. Chem. Phys.* 136, 064103 (2012).
[28] S. K. Natarajan and J. Behler, *Phys. Chem. Chem. Phys.* 18, 28704 (2016).
[29] T. Morawietz and J. Behler, *J. Phys. Chem. A* 117, 7356 (2013).
[30] V. L. Deringer and G. Csányi, *Phys. Rev. B* 95, 094203 (2017).
[31] A. P. Bartók, J. Kermode, N. Bernstein, and G. Csányi, *Phys. Rev. X* 8, 041048 (2018).
[32] M. A. Caro, V. L. Deringer, J. Koskinen, T. Laurila, and G. Csányi, *Phys. Rev. Lett.* 120, 166101 (2018).
[33] V. L. Deringer, N. Bernstein, A. P. Bartók, M. J. Cliffe, R. N. Kerber, L. E. Marbella, C. P. Grey, S. R. Elliott, and G. Csányi, *J. Phys. Chem. Lett.* 9, 2879 (2018).
[34] V. L. Deringer, M. A. Caro, R. Jana, A. Arava, S. R. Elliott, T. Laurila, G. Csányi, and L. Pastewka, *Chem. Mater.* 30, 7438 (2018).
[35] G. C. Sosso, V. L. Deringer, S. R. Elliott, and G. Csányi, *Mol. Simul.* 44, 866 (2018).
[36] R. Jimmouchi, J. Lahmsteiner, F. Karsai, G. Kresse, and M. Bokdam, *Phys. Rev. Lett.* 122, 225701 (2019).
[37] R. Jimmouchi, F. Karsai, and G. Kresse, *Phys. Rev. B* 100, 014105 (2019).
[38] Z. Li, J. R. Kermode, and A. De Vita, *Phys. Rev. Lett.* 114, 096405 (2015).
[39] L. J. Nelson, G. L. W. Hart, F. Zhou, and V. Ozoliņš, *Phys. Rev. B* 87, 035125 (2013).
[40] T. Barnard, H. Hagan, S. Tseng, and G. C. Sosso, *Mol. Syst. Des. Eng.* (2019), 10.1039/C9ME00109C.
[41] R. Ouyang, S. Curtarolo, E. Ahmetci, M. Scheffler, and
L. M. Ghiringhelli, Phys. Rev. Mater. 2, 083802 (2018).
[42] W. Huang, A. Urban, Z. Rong, Z. Ding, C. Luo, and G. Ceder, npj Comput. Mater. 3, 1 (2017).
[43] L. M. Ghiringhelli, J. Vybiral, E. Ahmetcik, R. Ouyang, S. V. Levchenko, C. Draxl, and M. Scheffler, New J. Phys. 19, 023017 (2017).
[44] M. Ceriotti, J. Chem. Phys. 150, 150901 (2019).
[45] L. M. Ghiringhelli, J. Vybiral, S. V. Levchenko, C. Draxl, and M. Scheffler, Phys. Rev. Lett. 114, 105503 (2015).
[46] K. T. Schütt, M. Gastegger, A. Tkatchenko, and K.-R. Müller, SpringerLink, 311 (2019).
[47] C. Langlois, P. Benzo, R. Arenal, M. Benoit, J. Nicolai, N. Combe, A. Ponchet, and M. J. Casanove, Nano Lett. 15, 5075 (2015).
[48] P. Benzo, S. Combettes, B. Pecassou, N. Combe, M. Benoit, M. Respaud, and M. J. Casanove, Phys. Rev. Mater. 3, 096001 (2019).
[49] A. Tymoczko, M. Kamp, C. Rehbock, L. Kienle, E. Catataruzza, S. Barcikowski, and V. Amendola, Nanoscale Horiz. 4, 1326 (2019).
[50] A. Tymoczko, M. Kamp, O. Prymak, C. Rehbock, J. Jakobi, U. Schürmann, L. Kienle, and S. Barcikowski, Nanoscale 10, 16434 (2018).
[51] J. Vernieres, S. Steinhauser, J. Zhao, P. Grammatikopoulos, R. Ferrando, K. Nordlund, F. Djurabekova, and M. Sowwan, Adv. Sci. 6, 1900447 (2019).
[52] F. Calvo, N. Combe, J. Morillo, and M. Benoit, J. Phys. Chem. C 121, 4680 (2017).
[53] S. Hong and T. S. Rahman, Phys. Chem. Chem. Phys. 17, 28177 (2015).
[54] A. Seko, A. Togo, and I. Tanaka, Phys. Rev. B 99, 214108 (2019).
[55] M. A. Wood and A. P. Thompson, J. Chem. Phys. DETC2018, 241721 (2018).
[56] A. M. Goryaeva, J.-B. Maillet, and M.-C. Marinica, Comput. Mater. Sci. 166, 200 (2019).
[57] F. H. Stillinger and T. A. Weber, Phys. Rev. B 31, 5262 (1985).
[58] M. S. Daw and M. I. Baskes, Phys. Rev. B 29, 6443 (1984).
[59] A. Seko, A. Takahashi, and I. Tanaka, Phys. Rev. B 90, 024101 (2014).
[60] A. Takahashi, A. Seko, and I. Tanaka, Phys. Rev. Materials 1, 063801 (2017).
[61] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, Annals of Statistics 32, 407 (2004).
[62] N. Hung, T. Hue, and N. Duc, VNU Journal of Science: Mathematics - Physics 31 (2015).
[63] G. Kresse and J. Furthmüller, Computational Materials Science 6, 15 (1996).
[64] P. E. Blöchl, Phys. Rev. B 50, 17953 (1994).
[65] T. Tsuchiya and K. Kawamura, J. Chem. Phys. 116, 2121 (2002).
[66] P. Söderlind, O. Eriksson, J. M. Wills, and A. M. Boring, Phys. Rev. B 48, 5844 (1993).
[67] S. Plimpton, J. Comput. Phys. 117, 1 (1995).