Making free-energy calculations routine: Combining first principles with machine learning

Ryosuke Jinnouchi,^{1,2,*} Ferenc Karsai,³ and Georg Kresse^{1,3}

¹University of Vienna, Faculty of Physics and Center for Computational Materials Sciences, Sensengasse 8/12, 1090, Vienna

²Toyota Central R&D Labs, Inc., 41-1, Yokomichi, Nagakute, Aichi 480-1192, Japan

³VASP Software GmbH, Sensengasse 8, 1090 Vienna, Austria



(Received 24 September 2019; accepted 3 February 2020; published 18 February 2020)

The chemical potentials of atoms and molecules in condensed matter are fundamental properties that allow one to predict a wide variety of thermodynamic properties. However, predictions using first principles are challenging. Here, an efficient and accurate method using machine-learned force fields is presented. A key point is that it requires training only at the end points of the thermodynamic pathway, rendering the training simple and efficient. Applications to liquid Si, and Li and F ions hydrated by water show that the method can predict accurate chemical potentials at low computational cost.

DOI: 10.1103/PhysRevB.101.060201

The temperature-dependent chemical potentials of atoms are exceptionally important fundamental properties. Their knowledge enables the quantitative evaluation of a wide variety of thermodynamic properties, such as coexistence points of different phases, concentration of minority species, or solubility in solvents to name only few. However, the accurate calculation of the chemical potential from first principles (FP) remains a very challenging task. To understand the problem imagine that we want to calculate the chemical potential of one or many atoms in a liquid using FP methods. Since the free energy is not an observable, one needs to calculate the free-energy difference between a known system and the desired real system using thermodynamic perturbation theory (TPT) [1] or thermodynamic integration (TI) [1,2] employing extensive molecular dynamics (MD) or Monte Carlo simulations [3–13]. The simplest approach is to perform a coupling constant integration from the solvent and noninteracting gaseous atoms to the atoms fully interacting with the solvent and each other. To make such computations practically feasible, cleverly constructed intermediate steps are often used; for instance, insertion of hard spheres into the liquid, charging the hard spheres, and then switching to the full FP description [11]. Each step then involves either TPT or TI. Brute force methods, where the coupling between the added atoms and the solvent is gradually switched on, are also feasible but suffer from bad statistics [10]. In fact, the initial step, where an infinitesimal small interaction between the inserted atom and the liquid needs to be used, is the most problematic step, since the not yet interacting atom can come very close to an atom in the solvent and as a result experiences a huge repulsive potential. Hence, the integrand in TI becomes infinite at zero coupling. This issue can be partly circumvented by variable transformations [12,13]. However, the remaining problem is that most FP codes become unstable when two atoms are very close. This means that the introduction of suitably constructed intermediate "surrogate" models becomes a necessity. Struc-

*e1262@mosk.tytlabs.co.jp

tures generated by classical force fields composed by physicsbased model functions, such as Lennard-Jones, Coulomb, and/or modified embedded atom models, have often been used as the surrogate models [14–17]. However, these models are not flexible enough to describe atomic interactions in a wide variety of materials, and construction of force fields always needs significant human intervention. The fact that only a handful of free-energy FP calculations have been reported to date for disordered materials shows convincingly, how challenging such calculations still are [7,8,10–13,18,19].

Machine learning (ML) is a rapidly evolving powerful approach that allows one to systematically construct force fields that can be used as surrogate models [20,21]. It has already been shown that MD simulations with machine-learned force fields (MLFFs) can provide quantitative predictions on atomic structures and dynamical properties of a wide variety of materials [22-32]. A recent study by Grabowski and coworkers showed that MLFFs can be used as surrogate models to efficiently compute anharmonicity in multicomponent alloys [33]. However, a method to compute the chemical potential of atoms and molecules in liquid phases relative to a well-defined reference state has not been suggested. For disordered systems, such a method is essential, because simple reference statistical models, like the quasiharmonic model, are not available for them. Integration from the ideal gas, however, always leads to large integrands even for say Lennard-Jones potentials. This causes large statistical errors related to the problematic singularity described before. In addition, the disordered structures of liquids need more statistics. Here we propose and test an efficient and accurate thermodynamic integration scheme using MLFFs. This scheme is computationally several orders of magnitude faster than the usual FP calculations, retains the accuracy of FP calculations, requires training only for the end points of the desired integration pathway, and entirely avoids the numerical instabilities observed in other calculations. This will pave the way towards routine calculations of thermodynamic potentials from FP for a huge variety of disordered materials, for instance, aqueous materials.

The MLFFs are generated following the learn-on-the-fly methodologies in our recent publications [31,32] [see also Supplemental Material (SM) [34]]. All computations are carried out using the Vienna Ab initio Simulation Package (VASP) [35,36]. For the present calculations, crucial modifications were made to allow for thermodynamic integration using these potentials. These modifications are discussed in detail in the next paragraph, and we call the corresponding method λ -MLFF. A key point of our λ -MLFF is that the training is necessary only for the two thermodynamic states located at the end points of the coupling constant pathway. Training on the intermediate points is unnecessary, and, in fact, would be counterproductive: the ML potential is not supposed to reproduce the hard well-like potential at small λ and short interatomic distances. The crucial insight is that TI requires only that the end points of the coupling constant integral are described accurately. Points at intermediate coupling rely on the surrogate model, which yields a smooth integrand very different from the first-principles calculations. However, the integral remains exact within the error bars of the ML potential. Our tests below show convincingly that this conjecture is correct, and that the remaining errors related to the MLFF can be easily removed by a final TPT step. This is not a trivial insight, but rather an important shift in our understanding of TI.

In the application to liquid Si, the interactions between all Si atoms are gradually switched on, i.e., we integrate from the ideal gas to the ML Si potential. In this case, training is necessary only for liquid Si, because the noninteracting case is automatically recovered by the λ -MLFF at zero coupling. Thus, we adopt the MLFF trained on solid, liquid, and interfacial Si that was used to calculate the melting point of Si in our previous study [32]. In the application to hydrated Li and F ions in water, the interactions between the water molecules are always accounted for exactly, and the integration is performed from gaseous ideal Li and F atoms to Li and F ions fully interacting with the water molecules (and each other). Accordingly, in this case, training was performed on the liquid phase without and with Li and F ions. Details of the trainings are shown in the SM [34].

The accuracy of our MLFFs is examined by comparing their structural properties to the ones obtained from FP calculations. In Figs. 1 and 2 the pair-correlation functions of liquid Si and LiF in H_2O are shown, respectively. The pair-correlation functions were sampled on a trajectory obtained from 30 ps molecular dynamics calculations. In both cases, we obtain excellent agreement between the force fields and FP calculations.

Thermodynamic properties of liquid Si are also accurately reproduced by the MLFF as reported in our previous study [32], but we emphasize that the method used before does *not* allow us to determine chemical potentials. The density of water at 300 K is calculated as 0.877 ± 0.009 g cm⁻³. This is in good agreement with the density of 0.864 ± 0.005 g cm⁻³ calculated by our FP calculation as well as a reported value of 0.89 g cm⁻³ obtained by a neural network potential [23]. All these values slightly deviate from the experimental value [37] of 0.997 g cm⁻³. The deviation is well understood [23,38–43] and mainly related to deficiencies of the approximate electronic exchange-correlation functional, here the revised



FIG. 1. Pair-correlation function of liquid Si at 1800 K. Results from FP calculations and MLFF are shown by black (continuous) and red (dashed) lines, respectively. The bonds in the structure are shown up to a neighbor distance of 2.6 Å.

Perdew-Burke-Ernzerhof (RPBE) [44] functional with the D3 dispersion corrections (+D3) [45,46], that we use in the FP calculations.

In thermodynamic integration a coupling parameter λ is introduced in the Hamiltonian to smoothly switch between



FIG. 2. Pair-correlation functions of LiF in H_2O at 300 K. Results from FP calculations and MLFF are shown by black (continuous) and red (dashed) lines, respectively.

two thermodynamic reference states. The change of the free energy and chemical potential along this pathway is then given by the integral of the expectation value of the derivative of the Hamiltonian $dH(\lambda)/d\lambda$:

$$\Delta \mu = \int_0^1 \left\langle \frac{dH(\lambda)}{d\lambda} \right\rangle_{\lambda} d\lambda, \tag{1}$$

where $\langle \cdot \rangle_{\lambda}$ means evaluation of the expectation value using an ensemble created by the Hamiltonian at coupling λ . As usual in MLFF, the Hamiltonian $H(\lambda)$ is given by kinetic and potential energy terms. The latter depends on the atomic environment around each atom *i*:

$$H(\lambda) = \sum_{i=1}^{N_a} \frac{|\mathbf{p}_i|^2}{2m_i} + \sum_{i \notin \mathbf{M}} U_i(\lambda) + \lambda \sum_{i \in \mathbf{M}} U_i(\lambda) + \sum_i^{N_a} U_{i,\text{atom}},$$
(2)

where N_a specifies the number of atoms, $U_{i,\text{atom}}$ denotes a reference potential energy of a single atom in the noninteracting case, and $U_i(\lambda)$ denotes the atomic potential energy relative to $U_{i,\text{atom}}$ of the interacting system. M is the subset of atoms whose interactions are controlled by the coupling parameter λ . At $\lambda = 0$, clearly the atoms in M do not directly contribute to the energy. However, the atoms in M are still visible to all other atoms, hence, the atomic potential energy $U_i(\lambda)$ also needs to be modified, by gradually blending out the atoms in M. As usual in kernel-based methods, we assume that the local energies can be approximated by a linear combination of functions $K(\mathbf{X}_i(\lambda), \mathbf{X}_{i_B})$ [47]:

$$U_i(\lambda) = \sum_{i_B=1}^{N_B} w_{i_B} K(\mathbf{X}_i(\lambda), \mathbf{X}_{i_B}).$$
(3)

Here $K(\mathbf{X}_i(\lambda), \mathbf{X}_{i_B})$ measures the similarity in the local structure surrounding atom *i* and atom *i_B* in the reference structures selected from the training set. The descriptors or fingerprints $\mathbf{X}_i(\lambda) = \mathbf{X}(\rho_i[(\mathbf{r}, \lambda)])$ are rotationally and translationally invariant functionals of the density distribution around atom *i*:

1

$$\rho_i(\mathbf{r}, \lambda) = \sum_{j \notin \mathbf{M}} f_{\text{cut}}(|\mathbf{r}_j - \mathbf{r}_i|)g[\mathbf{r} - (\mathbf{r}_j - \mathbf{r}_i)] \\
+ \lambda \sum_{j \in \mathbf{M}} f_{\text{cut}}(|\mathbf{r}_j - \mathbf{r}_i|)g[\mathbf{r} - (\mathbf{r}_j - \mathbf{r}_i)], \quad (4)$$

where f_{cut} is a cut-off function that smoothly eliminates the distribution of atoms outside a given cut-off radius, and $g(\mathbf{r})$ is a smoothed δ function. To blend out the atoms in M, we scale their contributions to the density distribution by the coupling parameter λ as well. Obviously, they become invisible to the MLFF at $\lambda = 0$, whereas they are fully accounted for at $\lambda = 1$. Importantly, the coupling parameter λ is introduced only in production runs *after* the force fields have been trained. The equations used to calculate the derivative $\frac{dH(\lambda)}{d\lambda}$ in Eq. (1) and the parameters for the descriptors and the similarity measure are summarized in the SM.

To test the approach for a simple well-studied system, we applied it to liquid Si. Here, all atoms are included in the set M, implying that we perform an integration from the ideal gas to the fully interacting liquid. The integration in Eq. (1) is carried out using the Gauss-Lobatto quadrature, and



FIG. 3. $\langle \frac{dH(\lambda)}{d\lambda} \rangle$ as a function of λ evaluated for different numbers of integration points N_{int} for (a) liquid silicon and (b) LiF in H₂O. Green solid lines are supposed to guide the eyes and shows a seventh-order polynomial fitted to all available data points.

a variable transformation is applied to λ (see SM) to cope with the somewhat stronger variations of the integrand at small λ values. The integrand of Eq. (1) $\frac{dH(\lambda)}{d\lambda}$ is shown in Fig. 3(a), and we observe that it completely lacks the problematic singularity at $\lambda = 0$ observed in FP calculations [12]. Note that all the data provided by the ML potential at $\lambda \neq 1$ or 0 are extrapolations, without deep physical meaning. Regardless, the predicted chemical potential is in excellent agreement with the previous full FP results as shown in Table I, although convergence with respect to the number of quadrature points is somewhat slow. We relate this to difficulties in integrating high-order contributions pertinent to the parametrization of the MLFF. However, adding more integration points is very cheap for MLFFs (four orders of magnitude faster for Si than the FP calculation; see detailed data in SM). Finally, to correct the results provided by the MLFF, TPT was used similar to the one applied in Ref. [12]. The most accurate value using 12 integration points and thermodynamic integration is within 1.5 meV to the previous reference calculations, but is obtained at a tiny fraction (approximately 1/1000) of the previous compute times. Also the previous calculations required multiple thermodynamic integration steps with tedious error control, whereas here we obtain the results in a single concise calculation.

TABLE I. Chemical potential of Si. N_{int} : Number of points used in the TI. MLFF: Machine-learning results. MLFF-TPT: Machinelearning results with corrections from thermodynamic perturbation theory. Standard deviations in parentheses are estimated by block averaging method. The theoretical literature result is taken from Ref. [12]. All units are in eV.

Nint	MLFF	MLFF-TPT	Ref.
8	-10.6702(0.0008)	-10.6733(0.0009)	
10	-10.6732(0.0008)	-10.6763(0.0009)	-10.6795(0.0011)
12	-10.6750(0.0006)	-10.6782(0.0008)	



FIG. 4. Snapshot of the structures for LiF in H_2O for (a) $\lambda = 1.0$ and (b) 0.002.

The solvation energy of LiF in water is a more relevant and more difficult property to compute. It is a measure of how much energy is gained when a LiF crystal dissolves in water. The computations are also significantly more challenging, as we now need to insert only two atoms (Li and F) into a simulation cell with many atoms. This means that the statistical error bars need to be made much smaller to obtain statistically converged results. Using FP methods this is hardly possible with present compute resources. Literally all previous calculations used some (uncontrolled) approximation along the chosen pathway or had significant error bars. Figure 4(a) shows a snapshot of the structure for a fully interacting system ($\lambda =$ 1.0) after equilibration. As reported in literature [48,49], the Li cation is hydrated by four neighboring O atoms. This can be clearly seen from the highlighted coordination tetrahedron in the figure. The F anion is hydrated by approximately five H atoms as reported previously in a FPMD [49]. With the decrease in the coupling parameter λ , interactions between H₂O and LiF and between Li and F are gradually turned off, and a physically meaningful hydration structure disappears. Figure 4(b) shows a snapshot of the structure at small interactions $(\lambda = 0.002)$, where the Li and F atoms behave essentially like gaseous noninteracting particles. Because of the lack of interactions, in this snapshot the F anion approaches the Li cation, and the ions are nonphysically close to some H₂O molecules (<1 Å). Nevertheless, all water molecules show an almost perfect coordination structure, with one hydrogen bond extending from each H atom, and every oxygen atom receiving two hydrogen bonds. Only in the direct vicinity of the LiF complex are slight disruptions of the network noticeable. Structures at intermediate λ values can be to some extent nonphysical in the vicinity of the Li and F atoms (short bonds and disrupted hydrogen bonds), but nevertheless our result for the chemical potential of -10.037 ± 0.040 eV per LiF (Table II) is in superb agreement with the observed

TABLE II. Chemical potential of LiF in H_2O . N_{int} : Number of points used in integration. MLFF: Machine-learning results. MLFF-TPT: Machine-learning results with corrections from thermodynamic perturbation theory. Standard deviations are given in parentheses. The literature results for theory and experiment are taken from Refs. [11] and [50], respectively. All units are in eV.

N _{int}	MLFF	MLFF-TPT	Ref.
8 10 Expt.	-10.084(0.030) -10.048(0.030)	-10.073(0.040) -10.037(0.040) -10.098	-10.417(0.041)

experimental value of -10.098 eV per LiF from Ref. [50] and in good agreement with previous FP calculations (using another type of revised PBE functional [51] plus D3 dispersion correction and a different FP code than applied here). We note that these calculations involved many intricate integration steps and that even the authors admit that "longer trajectories would be required for a precise determination of the error bar" [11].

In summary, we have developed an efficient and accurate thermodynamic integration method using MLFF to compute the temperature-dependent chemical potential. In this scheme the density distribution of those atoms that are supposed to be activated in the course of the thermodynamic integration are scaled by the coupling constant variable λ to continuously change the system from the case where some atoms are noninteracting to the fully interacting material. We call this scheme, λ -MLFF, as it allows a seamless description of the coupling constant integral along the integration pathway parametrized by λ . A remarkable point of the present approach is that training of MLFFs needs to be done only for ensembles at the two end points of the constructed thermodynamic pathway: the fully interacting and the noninteracting systems. Although the force fields provide only extrapolated interactions at intermediate points along the thermodynamic pathway, since the appearing structures are outside of the training data, the thermodynamic integral reproduces the freeenergy differences obtained by FP calculations exactly. The present method is widely applicable to a wide variety of materials, such as solvation energies, electrochemistry, and free energies of adsorbates on surfaces relevant in catalysis and chemistry.

The authors gratefully thank Ryoji Asahi for his many suggestions on the application and use of machine-learning methods to materials sciences.

- [1] R. W. Zwanzig, J. Chem. Phys. 22, 1420 (1954).
- [2] J. G. Kirkwood, J. Chem. Phys. 3, 300 (1935).
- [3] W. L. Jorgensen and C. Ravimohan, J. Chem. Phys. 83, 3050 (1985).
- [4] P. Bash, U. Singh, F. Brown, R. Langridge, and P. Kollman, Science 235, 574 (1987).
- [5] T. P. Straatsma and H. J. C. Berendsen, J. Chem. Phys. 89, 5876 (1988).
- [6] P. Kollman, Chem. Rev. **93**, 2395 (1993).
- [7] O. Sugino and R. Car, Phys. Rev. Lett. 74, 1823 (1995).
- [8] G. A. de Wijs, G. Kresse, and M. J. Gillan, Phys. Rev. B 57, 8223 (1998).
- [9] M. Ferrario, G. Ciccotti, E. Spohr, T. Cartailler, and P. Turq, J. Chem. Phys. 117, 4947 (2002).
- [10] K. Leung, S. B. Rempe, and O. A. von Lilienfeld, J. Chem. Phys. 130, 204507 (2009).

- [11] T. T. Duignan, M. D. Baer, G. K. Schenter, and C. J. Mundy, Chem. Sci. 8, 6131 (2017).
- [12] F. Dorner, Z. Sukurma, C. Dellago, and G. Kresse, Phys. Rev. Lett. 121, 195701 (2018).
- [13] M. Rang and G. Kresse, Phys. Rev. B 99, 184103 (2019).
- [14] J. Gao and X. Xia, Science 258, 631 (1992).
- [15] A. I. Duff, T. Davey, D. Korbmacher, A. Glensk, B. Grabowski, J. Neugebauer, and M. W. Finnis, Phys. Rev. B 91, 214311 (2015).
- [16] P. S. Hudson, S. Boresch, D. M. Rogers, and H. L. Woodcock, J. Chem. Theory Comput. 14, 6327 (2018).
- [17] T. J. Giese and D. M. York, J. Chem. Theory Comput. 15, 5543 (2019).
- [18] B. Grabowski, L. Ismer, T. Hickel, and J. Neugebauer, Phys. Rev. B 79, 134106 (2009).
- [19] C. Freysoldt, B. Grabowski, T. Hickel, J. Neugebauer, G. Kresse, A. Janotti, and C. G. Van de Walle, Rev. Mod. Phys. 86, 253 (2014).
- [20] J. Behler and M. Parrinello, Phys. Rev. Lett. 98, 146401 (2007).
- [21] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, Phys. Rev. Lett. 104, 136403 (2010).
- [22] J. Behler, J. Chem. Phys. 145, 170901 (2016).
- [23] T. Morawietz, A. Singraber, C. Dellago, and J. Behler, Proc. Natl. Acad. Sci. USA 113, 8368 (2016).
- [24] M. Hellström and J. Behler, Phys. Chem. Chem. Phys. 19, 82 (2017).
- [25] V. Botu, J. Chapman, and R. Ramprasad, J. Phys. Chem. C 121, 511 (2017).
- [26] K. Miwa and H. Ohno, Phys. Rev. Mater. 1, 053801 (2017).
- [27] K. Gubaev, E. V. Podryabinkin, G. L. Hart, and A. V. Shapeev, Comput. Mater. Sci. 156, 148 (2019).
- [28] S. Chmiela, H. E. Sauceda, K.-R. Müller, and A. Tkatchenko, Nat. Commun. 9, 3887 (2018).
- [29] A. P. Bartók, J. Kermode, N. Bernstein, and G. Csányi, Phys. Rev. X 8, 041048 (2018).
- [30] T. L. Jacobsen, M. S. Jørgensen, and B. Hammer, Phys. Rev. Lett. 120, 026102 (2018).
- [31] R. Jinnouchi, J. Lahnsteiner, F. Karsai, G. Kresse, and M. Bokdam, Phys. Rev. Lett. **122**, 225701 (2019).
- [32] R. Jinnouchi, F. Karsai, and G. Kresse, Phys. Rev. B 100, 014105 (2019).
- [33] B. Grabowski, Y. Ikeda, P. Srinivasan, F. Ktekarmann, C. Freysoldt, A. I. Duff, A. Shapeev, and J. Neugebauer, npj Comput. Mater. 5, 80 (2019).

- PHYSICAL REVIEW B 101, 060201(R) (2020)
- [34] See Supplemental Material http://link.aps.org/supplemental/ 10.1103/PhysRevB.101.060201 for a detailed description of the applied computational methods and settings, which includes Refs. [52–55].
- [35] G. Kresse and J. Furthmüller, Phys. Rev. B 54, 11169 (1996).
- [36] G. Kresse and J. Furthmüller, Comput. Mater. Sci. 6, 15 (1996).
- [37] P. J. Linstrom and W. Mallard, NIST Chemistry Webbook; NIST Standard Reference Database No. 69 (National Institute of Standards and Technology, Gaithersburg, MD, 2001).
- [38] K. Laasonen, F. Csajka, and M. Parrinello, Chem. Phys. Lett. 194, 172 (1992).
- [39] K. Laasonen, M. Parrinello, R. Car, C. Lee, and D. Vanderbilt, Chem. Phys. Lett. 207, 208 (1993).
- [40] M. Sprik, J. Hutter, and M. Parrinello, J. Chem. Phys. 105, 1142 (1996).
- [41] B. Santra, J. Klimeš, D. Alfè, A. Tkatchenko, B. Slater, A. Michaelides, R. Car, and M. Scheffler, Phys. Rev. Lett. 107, 185701 (2011).
- [42] M. J. Gillan, D. Alfè, and A. Michaelides, J. Chem. Phys. 144, 130901 (2016).
- [43] M. Chen, H.-Y. Ko, R. C. Remsing, M. F. Calegari Andrade, B. Santra, Z. Sun, A. Selloni, R. Car, M. L. Klein, J. P. Perdew, and X. Wu, Proc. Natl. Acad. Sci. USA 114, 10846 (2017).
- [44] B. Hammer, L. B. Hansen, and J. K. Nørskov, Phys. Rev. B 59, 7413 (1999).
- [45] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, J. Chem. Phys. 132, 154104 (2010).
- [46] S. Grimme, S. Ehrlich, and L. Goerigk, J. Comput. Chem. 32, 1456 (2011).
- [47] A. P. Bartók, R. Kondor, and G. Csányi, Phys. Rev. B 87, 184115 (2013).
- [48] H. H. Loeffler and B. M. Rode, J. Chem. Phys. 117, 110 (2002).
- [49] E. Pluharová, O. Marsalek, B. Schmidt, and P. Jungwirth, J. Phys. Chem. Lett. 4, 4177 (2013).
- [50] P. Hünenberger and M. Reif, Single-Ion Solvation: Experimental and Theoretical Approaches to Elusive Thermodynamic Quantities (The Royal Society of Chemistry, London, 2011).
- [51] Y. Zhang and W. Yang, Phys. Rev. Lett. 80, 890 (1998).
- [52] J. P. Boyd, *Chebyshev and Fourier Spectral Methods* (Dover, New York, 2001).
- [53] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, 2nd ed. (Oxford University Press, New York, 2017).
- [54] P. E. Blöchl, Phys. Rev. B 50, 17953 (1994).
- [55] J. Sun, A. Ruzsinszky, and J. P. Perdew, Phys. Rev. Lett. 115, 036402 (2015).