



Predicting hot-electron free energies from ground-state dataChiheb Ben Mahmoud ^{*}, Federico Grasselli ^{*}, and Michele Ceriotti [†]*Laboratory of Computational Science and Modeling, IMX, École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland*

(Received 11 May 2022; revised 28 August 2022; accepted 1 September 2022; published 27 September 2022)

Machine-learning potentials are usually trained on the ground-state, Born-Oppenheimer energy surface, which depends exclusively on the atomic positions and not on the simulation temperature. This disregards the effect of thermally excited electrons, that is important in metals, and essential to the description of warm dense matter. An accurate physical description of these effects requires that the nuclei move on a temperature-dependent electronic free energy. We propose a method to obtain machine-learning predictions of this free energy at an arbitrary electron temperature using exclusively training data from ground-state calculations, avoiding the need to train temperature-dependent potentials, and benchmark it on metallic liquid hydrogen at the conditions of the core of gas giants and brown dwarfs. This Letter demonstrates the advantages of hybrid schemes that use physical consideration to combine machine-learning predictions, providing a blueprint for the development of similar approaches that extend the reach of atomistic modeling by removing the barrier between physics and data-driven methodologies.

DOI: [10.1103/PhysRevB.106.L121116](https://doi.org/10.1103/PhysRevB.106.L121116)

In the past decade, machine-learning (ML) algorithms proved to be an efficient alternative to expensive first-principles (FP) calculations. The construction of ML interatomic potentials (MLIPs) trained on FP data has achieved a successful balance between computational cost and accuracy [1–4]. This greatly simplified sampling the finite-temperature properties of materials, and has been complemented by ML models that predict functional materials properties ranging from scalar quantities [5–8] to tensorial properties and fields [9–12]. Current ML strategies are usually designed to reproduce the ground-state, Born-Oppenheimer (BO) potential energy surface, and do not account for the temperature-dependent electronic excitations which may play a major role in metallic matter at planetary conditions, such as warm dense matter (WDM) [13–17], and that introduce subtle but important corrections in the thermophysical properties of ordinary metals [18,19]. The most common strategy to treat finite electron temperature is to replace the BO potential with a temperature-dependent electronic free energy $A(T^{\text{el}})$.

In traditional MLIP frameworks, that rely exclusively on nuclear coordinates as inputs, switching from the BO potential to $A(T^{\text{el}})$ would require training a separate model for every target electronic temperature T^{el} , recomputing also the training set—although the temperature can be included as an input of the model, which yields MLIPs that are explicitly temperature dependent, and interpolate between training data at different electron temperature [20]. One recent attempt to incorporate directly electronic excitations into ML simulations is to predict the single-particle density of states [21], and use it to evaluate *a posteriori* corrections to the

thermodynamic quantities, e.g., heat capacity or melting temperature, extracted from the molecular dynamics (MD) of ions whose MLIP is trained on ground-state data [22]. This approach is limited to condensed matter well below the Fermi temperature, where atomic forces are almost unaffected by the electronic excitations. Another recent method relies on Hamiltonian models based on the local density of states, and trained on finite-temperature data [23]. Despite its success in describing directly electron finite-temperature effects, this approach would still require generating data at specific target temperatures, limiting its transferability to conditions which span broad temperature ranges.

In this Letter we first show that, within a density functional theory (DFT) framework, the total free energy, atomic forces, and the stress tensor of the system can be rigorously approximated as the sum of a $T^{\text{el}} = 0$ K contribution and a finite- T^{el} correction depending exclusively on the *ground-state* electronic density of states (DOS). This general result underpins a framework that relies only on ground-state calculations to learn $A(T^{\text{el}})$ and its derivatives in the presence of thermally excited electrons. Thus, a consistent ground-state training set and model can be generated, and used to sample the finite-electron-temperature distributions, using T^{el} as an external parameter. We test our method on simple metals, where the atomic forces are evaluated at increasing electronic temperatures. We then validate our framework by constructing the equation of state (EOS) of hydrogen at conditions relevant in gas giants and brown dwarfs, for temperatures up to 50 000 K and pressures up to 1600 GPa, by means of MD simulations driven by our T^{el} -dependent MLIPs. We also compute the heat capacity of hydrogen at 400 GPa in the high-temperature regime.

Let us start by considering the standard representation of the DFT energy,

$$E = E_{\text{b}} - E_{\text{dc}} + E_{\text{ion}}, \quad (1)$$

^{*}These authors contributed equally to this work.[†]michele.ceriotti@epfl.ch

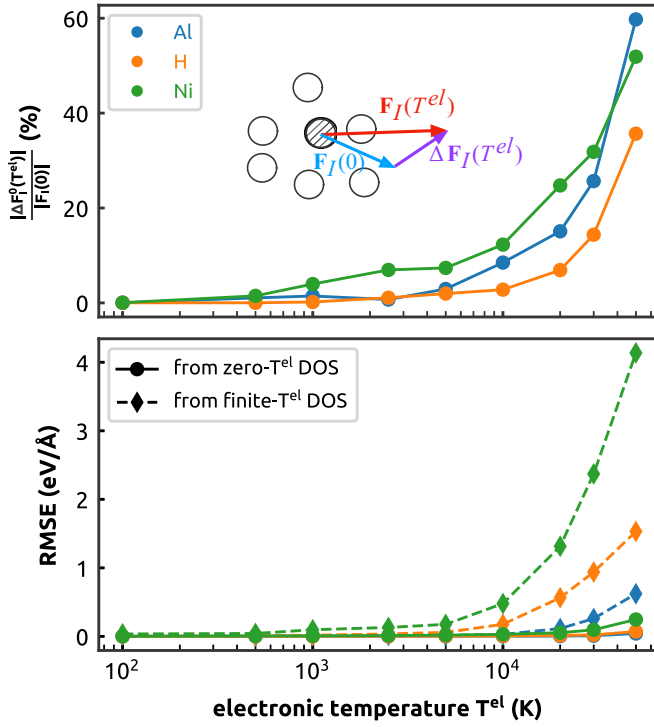


FIG. 1. Upper panel: Relative deviation of Hellmann-Feynman atomic force vs the electronic temperature with respect to the ground-state force for a given ion and a Cartesian direction. The sketch represents the decomposition of the finite- T^{el} atomic force component within our framework. Lower panel: Root-mean-square errors (RMSEs) of ten force components computed with Eq. (10) compared to their Hellmann-Feynman counterparts. Solid lines: Using the DOS from $T^{el} = 0$ K calculations. Dashed lines: Using the DOS from finite- T^{el} calculations. Blue: aluminum; orange: hydrogen; green: nickel.

as a sum of the electrostatic interactions between the ions E_{ion} , the band energy $E_b = \sum_i f_i \epsilon_i$, expressed in terms of the Kohn-Sham (KS) eigenvalues ϵ_i and level occupations f_i , and the “double-counting term”

$$E_{dc} = \frac{1}{2} \iint \frac{\rho(\mathbf{r}')\rho(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}' - E_{xc}[\rho] + \int V_{xc}[\rho](\mathbf{r})\rho(\mathbf{r})d\mathbf{r}. \quad (2)$$

Here, E_{xc} is the exchange-correlation (XC) functional, $V_{xc}[\rho](\mathbf{r}) = \delta E_{xc}/\delta\rho(\mathbf{r})$ is the XC potential, and $\rho(\mathbf{r}) = \sum_i f_i |\phi_i(\mathbf{r})|^2$ is the DFT density, expressed in terms of the KS eigenfunctions $\phi_i(\mathbf{r})$ and occupations f_i .

Whenever an electronic temperature T^{el} is introduced, f_i become fractional, and the correct energy functional becomes the Helmholtz free energy [24–26],

$$A(T^{el}) = E(0) + \Delta E(T^{el}) - T^{el}S(T^{el}), \quad (3)$$

where $\Delta E(T^{el})$ is the finite- T^{el} contribution to the energy, and $S(T^{el})$ is the KS electronic entropy. From Eq. (3) one can obtain the finite- T^{el} Hellmann-Feynman forces [27], whose relative deviation with respect to $T^{el} = 0$ K forces becomes significant at large T^{el} , as reported in the upper panel of Fig. 1. In principle, a T^{el} -dependent XC functional should be employed [28]. However, it is often possible to rely on the zero-temperature approximation (ZTA), where the XC

functional depends on T^{el} only through the T^{el} dependence of the density: $E_{xc}[\rho(T^{el})]$. The ZTA performs well at both low and high T^{el} and also satisfies exact conditions as discussed in Ref. [29], and we adopt it as the basis of our framework.

A change in the occupation of the levels, e.g., as a consequence of thermal excitations, determines a change in the density, and thus, self-consistently, in the KS eigenenergies and eigenfunctions. For instance, the functional derivative of E_b with respect to f_i is

$$\frac{\delta E_b}{\delta f_i} = \epsilon_i + \sum_j f_j \frac{\delta \epsilon_j}{\delta f_i}. \quad (4)$$

Nonetheless, it can be proved, following a reasoning similar to that used in Ref. [30], and in Ref. [31] for the energy variation due to infinitesimal atomic displacements (cf. Supplemental Material [32]), that the second term in Eq. (4) *cancels exactly* with the variation of the double-counting term $\delta E_{dc}/\delta f_i$. Therefore, the change in E due to a finite change in the occupations can be approximated by

$$\Delta E \approx \Delta E_b^0 \equiv \sum_i \epsilon_i^0 \Delta f_i, \quad (5)$$

where $\epsilon_i^0 \equiv \epsilon_i(\{\Delta f_k = 0\})$ are the eigenenergies computed at vanishing variation on all the f_k . The “0” superscript labels quantities obtained from unperturbed eigenenergies, computed at $T^{el} = 0$ K.

We now focus on the specific case where the set of f_i are Fermi-Dirac distributed, $f_i = f(\frac{\epsilon_i - \mu(T^{el})}{k_B T^{el}})$, $f(x) = 1/(1 + e^x)$ being the Fermi function, $\mu(T^{el})$ being the chemical potential of the electron system, and k_B the Boltzmann constant. From Eq. (5), the finite- T^{el} correction to the DFT energy is

$$\Delta E_b^0(T^{el}) = \int_{-\infty}^{+\infty} \epsilon g^0(\epsilon) \left[f\left(\frac{\epsilon - \mu(T^{el})}{k_B T^{el}}\right) - f\left(\frac{\epsilon - \mu(0)}{k_B 0^+}\right) \right] d\epsilon, \quad (6)$$

where $g^0(\epsilon) = \sum_i \delta(\epsilon - \epsilon_i^0)$ is the electronic DOS. $\mu(T^{el})$ is computed by enforcing charge-conservation:

$$N = \int_{-\infty}^{+\infty} g^0(\epsilon) f\left(\frac{\epsilon - \mu(T^{el})}{k_B T^{el}}\right) d\epsilon, \quad (7)$$

where N is the number of electrons. In the Supplemental Material [32], we justify the use of $g^0(\epsilon)$ in Eq. (7) and when evaluating the electronic entropy $S(T^{el})$,

$$S(T^{el}) \approx S^0(T^{el}) \equiv \int_{-\infty}^{+\infty} g^0(\epsilon) s\left(\frac{\epsilon - \mu(T^{el})}{k_B T^{el}}\right) d\epsilon, \quad (8)$$

where $s(x) = f \ln f + (1 - f) \ln(1 - f)$. Therefore, our approximation for the free energy reads

$$A(T^{el}) \approx E(0) + \Delta E_b^0(T^{el}) - T^{el}S^0(T^{el}). \quad (9)$$

The number of states above $\mu(T^{el})$ that must be included to reliably compute the finite-temperature contribution to the free energy depends on the temperature. This means that for training configurations we have to include a larger number of empty states than that usually needed for $T^{el} = 0$ K calculations. The finite- T^{el} correction terms in Eq. (9) are independent of the alignment of the DOS (cf. Supplemental Material [32]), as long as it is chosen consistently when

computing the chemical potential, the band energy, and the entropy terms. Our derivation justifies other approximations made in the literature such as the fixed-DOS approximation of Ref. [33], which assumes that the electronic DOS is approximately independent of T^{el} . In fact, the cancellations ensure the validity of Eq. (5), even if the self-consistent energy levels (and thus the DOS itself) changed substantially by changing T^{el} . If one wanted to go beyond this ground-state approximation, it would not be sufficient to obtain the finite- T^{el} DOS, and to use it in expressions similar to Eqs. (6)–(8). Without access to the self-consistent finite temperature E_{dc} , doing so would lead to *worse* results, as shown in the lower panel in Fig. 1. If one was prepared to perform self-consistent calculations at multiple temperatures, our perturbative expressions could also be applied to a reference temperature different from $T^{\text{el}} = 0$, and serve as the basis of more accurate temperature-interpolation schemes (cf. Supplemental Material [32]). Our derivation directly translates to the calculation of derivatives of the free energy, such as forces and stresses. For instance, according to the Born-Oppenheimer approximation, the force acting on the I th nucleus in the the DFT ensemble is $\mathbf{F}_I(T^{\text{el}}) = -\nabla_I A(T^{\text{el}}) \approx \mathbf{F}_I(0) + \Delta \mathbf{F}_I^0(T^{\text{el}})$, where

$$\begin{aligned} \mathbf{F}_I(0) &\equiv -\nabla_I E(0), \\ \Delta \mathbf{F}_I^0(T^{\text{el}}) &\equiv -\nabla_I [\Delta E_b^0(T^{\text{el}}) - T^{\text{el}} S^0(T^{\text{el}})]. \end{aligned} \quad (10)$$

In this decomposition, the electronic temperature T^{el} enters as an *external parameter*.

These equations would be of limited practical value if the end goal was to compute $A(T^{\text{el}})$ for a given structure and temperature by means of a self-consistent electronic structure calculation. However, they become very useful in the context of data-driven modeling, as they provide a rigorous basis for the development of an ML framework to learn finite- T^{el} interatomic force fields without the need to train on finite- T^{el} calculations. The $T^{\text{el}} = 0$ K quantities, $E(0)$ and $\mathbf{F}_I(0)$, can be modeled by any of the widely used MLIPs [2,4,34–36]. The hot-electron correction, Eq. (10), can be accessed by training an ML model for the DOS. In this Letter we use the Gaussian approximation potentials (GAPs) [34] and an atom-centered model for the DOS as detailed in Ref. [21]. Both models rely on a linear expansion of a target quantity y (be it the energy or the DOS) of a given structure A on a set of positive-definite functions, also called kernels, $k(A, M_j)$, measuring the similarity between the structure A and a structure M_j belonging to the set of M reference environments, also called the active set:

$$y(A) = \sum_{j \in M} w_j k(A, M_j). \quad (11)$$

Notice that, for the DOS, $y = g^0(\epsilon, A)$, the weights $w_j(\epsilon)$ are a function of a discretized energy grid. The weights do not depend on the atomic positions of A . Therefore, gradients needed to compute atomic forces, as in Eq. (10), do not act on the weights, but on kernels alone. In practice, since neither the kernels nor their gradients depend on ϵ , they can be collected out of integrals involving the DOS such as Eq. (6), which makes the implementation of this approach particularly simple in the case of kernel (or linear) models (cf. Supplemental Material [32]). We construct the kernels $k(A, M_j)$

TABLE I. Table of the validation root-mean-square error (RMSE) of the ML models on the energies and forces compared to the reference DFT data, at the same level of theory introduced in Eqs. (9) and (10). The electronic temperature is $T^{\text{el}} = 35\,000$ K. The training set consists of 28 000 structures and the errors are reported for a validation set of 2500 configurations.

	RMSE
$E(0)$	11.05 meV/atom
$\Delta E_b^0(T^{\text{el}}) - T^{\text{el}} S^0(T^{\text{el}})$	13.43 meV/atom
$A(T^{\text{el}})$	12.22 meV/atom
$\mathbf{F}_I(0)$	0.87 eV/Å
$\Delta \mathbf{F}_I^0(T^{\text{el}})$	0.66 eV/Å
$\mathbf{F}_I(T^{\text{el}})$	0.81 eV/Å

from the smooth overlap of atomic positions (SOAP) representation [37,38] with radial scaling [39] as implemented in LIBRASCAL [40], and we use its interface with I-PI [41] to run MD simulations with finite electron temperature.

We demonstrate the practicality of our theoretical framework in ML workflows incorporating the electronic finite-temperature effects in atomistic simulations by constructing the EOS of metallic liquid hydrogen at conditions similar to those found in the core of a young Jupiter [42], and we compare our ML approach to explicit first-principles molecular dynamics (FPMD) simulations results at finite- T^{el} . We build a training set made of $\sim 28\,000$ structures, each containing 128 atoms, and densities ranging between 0.6 and 1.77 g cm $^{-3}$. It consists of configurations from Ref. [43], complemented by snapshots obtained from MD simulations performed with preliminary versions of the MLIP. We employ QUANTUM ESPRESSO [44–46] (QE) for DFT calculations of the data set, using the optimized norm-conserving Vanderbilt pseudopotential [47] version 1.2, which is shown to perform well even at \sim TPa pressures [48]. Dispersion interactions are included via a van der Waals density functional [49–52]. Details of the reference calculations and of the GAP model construction are given in the Supplemental Material [32]. Table I shows the root-mean-square error (RMSE) of the different (free) energies and forces for $T^{\text{el}} = 0$ K and at $T^{\text{el}} = 35\,000$ K. The ML models are in good agreement with the corresponding DFT calculations and the RMSE of the total free energy is well below the typical thermal energy at the temperatures we consider in this study, and comparable to the values observed in previous simulations of liquid systems at high ionic temperatures [53].

In order to gauge the importance of finite- T^{el} effects, and to obtain accurate reference calculations consistent with our computational setup, we run two sets of FPMD trajectories targeting the pressures 400, 800, 1200, and 1600 GPa for each of the ionic temperatures $T^i = 10\,000$, 20 000, 35 000, and 50 000 K. The electronic temperature of the first set is $T^{\text{el}} = 0$ K, while $T^{\text{el}} = T^i$ in the second set. The DFT calculations are performed with QE and Γ -point sampling. We evolve the ion dynamics with I-PI for at least 8 ps, after an equilibration of 1 ps, with a time step of 0.1 fs. T^i is controlled by stochastic velocity rescaling [54] with a time constant $\tau = 5$ fs, and an isotropic barostat [55] with a time constant

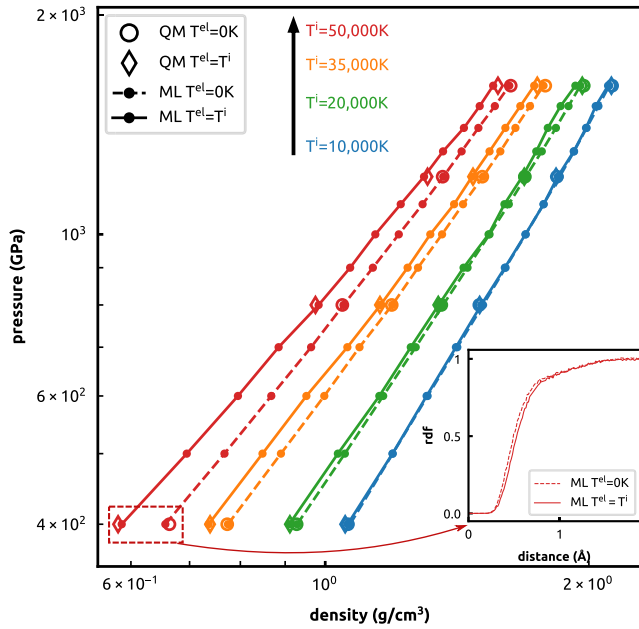


FIG. 2. Hydrogen isotherms of different equations of state (EOS). The open circles correspond to the EOS computed with “cold”-electron- Γ -point DFT. The open diamonds correspond to the EOS computed with finite-electron-temperature- Γ -point DFT. The dashed lines correspond to the EOS computed with ML trained on $T^{\text{el}} = 0$ K data. The solid lines correspond to the EOS computed with finite-temperature ML framework, where $T^{\text{el}} = T^i$. The temperatures range from 10 000 to 50 000 K denoted by the different colors as shown in the legend. The statistical error bars computed by block averages are smaller than the size of the markers. The small inset represents the radial distribution functions of hydrogen at $P = 400$ GPa and $T^i = 50$ 000 K computed from the ML trajectories. The dashed line corresponds to $T^{\text{el}} = 0$ K and the solid line to $T^{\text{el}} = T^i$.

$\tau = 20$ fs, thermalized with an optimal-sampling generalized Langevin thermostat [56]. Due to the high temperature and the fast intrinsic timescale of hydrogen, such relatively short simulations are sufficient to obtain converged results with a small statistical uncertainty. We report the results of these simulations in Fig. 2, by the open symbols. The differences due to the finite electron temperature grow steadily between 10 000 and 50 000 K, and at the highest temperature they range between 4% at 1600 GPa and 10% at 400 GPa, providing an indication of the impact of finite T^{el} in this range of pressure and density.

We then run two analogous sets of trajectories based on the finite- T^{el} MLIP, temperatures as for the FPMD, and pressures spanning the range between 400 and 1600 GPa in intervals of 100 GPa. As for the case of FPMD, the first set of simulations does not include any finite-temperature effects (dashed lines in Fig. 2), while the second incorporates them (solid lines). Our ML EOS are in excellent agreement with the reference curves obtained with explicit finite- T^{el} FPMD, up to the statistical uncertainties. We also observe a small shift in the radial distribution at the lower-pressure and higher-temperature range, corresponding to the difference in particle densities. As an additional demonstration of the importance of incorporating finite- T^{el} effects, we compute constant-pressure heat capac-

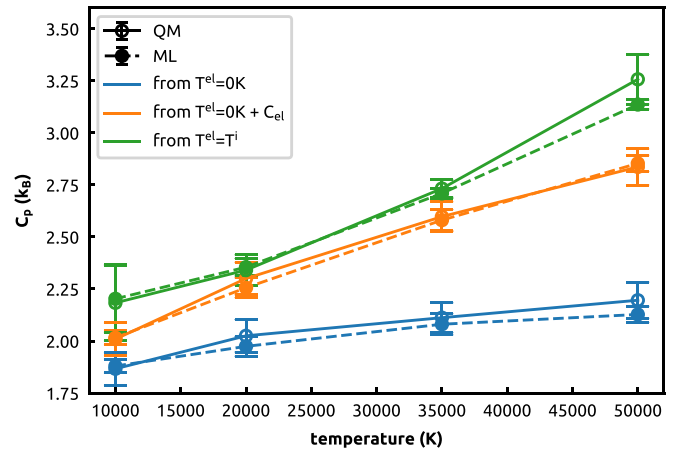


FIG. 3. Specific heat capacity C_p of hydrogen from NpT simulations at 400 GPa. The solid lines represent the DFT calculations and the dashed lines represent the ML calculations. Blue: C_p from the fluctuations of the ions’ enthalpy at $T^{\text{el}} = 0$ K; orange: C_p same as the blue curves in addition to a correction term computed from the average band energy of the electrons over the trajectories; green: C_p from the finite- T^{el} sampling. The error bars are computed from standard block analysis.

ities, $C_p = \left(\frac{\partial H}{\partial T}\right)_p$ that we obtain as finite differences of the enthalpy $H = \langle K \rangle + \langle A(T^{\text{el}}) \rangle + T^{\text{el}} \langle S^0(T^{\text{el}}) \rangle + p \langle V \rangle$, as described in the Supplemental Material [32]. Here, K is the kinetic energy of the ions, and the averages $\langle \dots \rangle$ are computed over finite- T^{el} NpT sampling. Figure 3 compares the heat capacity computed from $T^{\text{el}} = 0$ K simulations (blue) with that computed including the electronic contributions (green)—which amounts to almost 50% at the highest temperature considered. DFT and ML simulations agree with each other within their statistical uncertainty. The *a posteriori* incorporation of electronic excitation by adding $C_{\text{el}} \equiv \left(\frac{\partial \Delta E_{\text{el}}}{\partial T}\right)$ (orange) on top of the $T^{\text{el}} = 0$ K ionic contribution, as done in Ref. [22], cannot reproduce accurately the finite- T^{el} results.

These results demonstrate the accuracy of an ML model based on the ground-state DOS approximation in sampling the finite- T^{el} thermophysical properties of hydrogen in a challenging portion of its phase diagram. By treating explicitly the ionic and electronic degrees of freedom, our ML models eliminate one of the most glaring limitations of traditional MLIPs, that are restricted to perform simulations at a single (usually zero) electron temperature. We remark that no restriction occurs in applying our machinery to a two-temperature model where $T^{\text{el}} \neq T^i$ and the hot electrons are not in thermal equilibrium with the nuclei, even though electron-nuclei interactions should be included to allow for thermalization. Our approach can be easily extended to any electronic structure method based on the KS mapping, and can be naturally used also for multiple-species systems, opening the possibility of studying the complex phase diagram of metallic mixtures at high- pT conditions, which dictates the evolution of giant planets [57]. On a conceptual level, the idea of using a physical approximation in synergy with data-driven predictions indicates a promising research direction to further extend the scope of applicability of predictive atomic-scale simulations.

The data supporting the findings in this work are available on the Materials Cloud platform [58]. Archived versions of software used to run the DFT and ML simulations are available on Zenodo [59].

We thank G. Mazzola for discussing with us an early version of the manuscript, and G. Gil for useful comments. C.B.M. and M.C. acknowledge support by the Swiss

National Science Foundation (Project No. 200021-182057) and the NCCR MARVEL, a National Centre of Competence in Research, funded by the Swiss National Science Foundation (Grant No. 182892). F.G. acknowledges funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Action IF-EF-ST, Grant Agreement No. 101018557 (TRANQUIL).

-
- [1] A. P. Bartók, J. Kermode, N. Bernstein, and G. Csányi, *Phys. Rev. X* **8**, 041048 (2018).
- [2] L. Zhang, J. Han, H. Wang, R. Car, and W. E, *Phys. Rev. Lett.* **120**, 143001 (2018).
- [3] J. Behler, *J. Chem. Phys.* **134**, 074106 (2011).
- [4] D. P. Kovács, C. van der Oord, J. Kucera, A. E. A. Allen, D. J. Cole, C. Ortner, and G. Csányi, *J. Chem. Theory Comput.* **17**, 7696 (2021).
- [5] F. M. Paruzzo, A. Hofstetter, F. Musil, S. De, M. Ceriotti, and L. Emsley, *Nat. Commun.* **9**, 4501 (2018).
- [6] A. Y.-T. Wang, S. K. Kauwe, R. J. Muddock, and T. D. Sparks, *npj Comput. Mater.* **7**, 77 (2021).
- [7] T. Xie and J. C. Grossman, *Phys. Rev. Lett.* **120**, 145301 (2018).
- [8] P. Pegolo, S. Baroni, and F. Grasselli, *npj Comput. Mater.* **8**, 24 (2022).
- [9] A. Grisafi, A. Fabrizio, B. Meyer, D. M. Wilkins, C. Corminboeuf, and M. Ceriotti, *ACS Cent. Sci.* **5**, 57 (2019).
- [10] M. Veit, D. M. Wilkins, Y. Yang, R. A. DiStasio, Jr., and M. Ceriotti, *J. Chem. Phys.* **153**, 024113 (2020).
- [11] A. Chandrasekaran, D. Kamal, R. Batra, C. Kim, L. Chen, and R. Ramprasad, *npj Comput. Mater.* **5**, 22 (2019).
- [12] B. Cuevas-Zuñiría and L. F. Pacios, *J. Chem. Inf. Model.* **61**, 2658 (2021).
- [13] R. Scipioni, L. Stixrude, and M. P. Desjarlais, *Proc. Natl. Acad. Sci. USA* **114**, 9009 (2017).
- [14] D. I. Mihaylov, V. V. Karasiev, S. X. Hu, J. R. Rygg, V. N. Goncharov, and G. W. Collins, *Phys. Rev. B* **104**, 144104 (2021).
- [15] V. V. Karasiev, J. W. Dufty, and S. B. Trickey, *Phys. Rev. Lett.* **120**, 076401 (2018).
- [16] M. Bonitz, T. Dornheim, Z. A. Moldabekov, S. Zhang, P. Hamann, H. Kählert, A. Filinov, K. Ramakrishna, and J. Vorberger, *Phys. Plasmas* **27**, 042710 (2020).
- [17] J. M. McMahon, M. A. Morales, C. Pierleoni, and D. M. Ceperley, *Rev. Mod. Phys.* **84**, 1607 (2012).
- [18] B. Grabowski, L. Ismer, T. Hickel, and J. Neugebauer, *Phys. Rev. B* **79**, 134106 (2009).
- [19] D. Ma, B. Grabowski, F. Körmann, J. Neugebauer, and D. Raabe, *Acta Mater.* **100**, 90 (2015).
- [20] Y. Zhang, C. Gao, Q. Liu, L. Zhang, H. Wang, and M. Chen, *Phys. Plasmas* **27**, 122704 (2020).
- [21] C. Ben Mahmoud, A. Anelli, G. Csányi, and M. Ceriotti, *Phys. Rev. B* **102**, 235130 (2020).
- [22] N. Lopanitsyna, C. Ben Mahmoud, and M. Ceriotti, *Phys. Rev. Mater.* **5**, 043802 (2021).
- [23] J. A. Ellis, L. Fiedler, G. A. Popoola, N. A. Modine, J. A. Stephens, A. P. Thompson, A. Cangi, and S. Rajamanickam, *Phys. Rev. B* **104**, 035120 (2021).
- [24] N. D. Mermin, *Phys. Rev.* **137**, A1441 (1965).
- [25] A. Alavi, J. Kohanoff, M. Parrinello, and D. Frenkel, *Phys. Rev. Lett.* **73**, 2599 (1994).
- [26] N. Marzari and D. Vanderbilt, *Phys. Rev. B* **56**, 12847 (1997).
- [27] R. M. Wentzcovitch, J. L. Martins, and P. B. Allen, *Phys. Rev. B* **45**, 11372 (1992).
- [28] V. V. Karasiev, T. Sjöström, D. Chakraborty, J. W. Dufty, K. Runge, F. E. Harris, and S. B. Trickey, in *Frontiers and Challenges in Warm Dense Matter*, edited by F. Graziani, M. P. Desjarlais, R. Redmer, and S. B. Trickey, Lecture Notes in Computational Science and Engineering Vol. 96 (Springer, Berlin, 2014), pp. 61–85.
- [29] A. Pribram-Jones, S. Pittalis, E. K. U. Gross, and K. Burke, in *Frontiers and Challenges in Warm Dense Matter* (Ref. [28]), pp. 25–60.
- [30] M. Weinert and J. W. Davenport, *Phys. Rev. B* **45**, 13709 (1992).
- [31] S. Goedecker and K. Maschke, *Phys. Rev. B* **45**, 1597 (1992).
- [32] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevB.106.L121116> for details about the derivation of Eq. (5), the use of different DOS in Eq. (9), possible strategies to fit the forces from finite- T^{el} data, details about the DFT calculations and the ML models, and the calculation of the heat capacity, which includes Refs. [21,37,38,40,44–47,49–52,60–66].
- [33] X. Zhang, B. Grabowski, F. Körmann, C. Freysoldt, and J. Neugebauer, *Phys. Rev. B* **95**, 165126 (2017).
- [34] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, *Phys. Rev. Lett.* **104**, 136403 (2010).
- [35] N. Artrith, T. Morawietz, and J. Behler, *Phys. Rev. B* **83**, 153101 (2011).
- [36] K. T. Schütt, P. Kessel, M. Gastegger, K. A. Nicoli, A. Tkatchenko, and K.-R. Müller, *J. Chem. Theory Comput.* **15**, 448 (2019).
- [37] A. P. Bartók, R. Kondor, and G. Csányi, *Phys. Rev. B* **87**, 184115 (2013).
- [38] S. De, A. P. Bartók, G. Csányi, and M. Ceriotti, *Phys. Chem. Chem. Phys.* **18**, 13754 (2016).
- [39] M. J. Willatt, F. Musil, and M. Ceriotti, *Phys. Chem. Chem. Phys.* **20**, 29661 (2018).
- [40] F. Musil, M. Veit, A. Goscinski, G. Fraux, M. J. Willatt, M. Stricker, T. Junge, and M. Ceriotti, *J. Chem. Phys.* **154**, 114109 (2021).
- [41] V. Kapil, M. Rossi, O. Marsalek, R. Petraglia, Y. Litman, T. Spura, B. Cheng, A. Cuzzocrea, R. H. Meißner, D. M. Wilkins, B. A. Helfrecht, P. Juda, S. P. Bienvenue, W. Fang, J. Kessler, I. Poltavsky, S. Vandenbrande, J. Wieme, C. Corminboeuf, T. D. Kühne *et al.*, *Comput. Phys. Commun.* **236**, 214 (2019).

- [42] N. Nettelmann, A. Becker, B. Holst, and R. Redmer, *Astrophys. J.* **750**, 52 (2012).
- [43] B. Cheng, G. Mazzola, C. J. Pickard, and M. Ceriotti, *Nature (London)* **585**, 217 (2020).
- [44] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. Dal Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos *et al.*, *J. Phys.: Condens. Matter* **21**, 395502 (2009).
- [45] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. B. Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, N. Colonna, I. Carnimeo, A. D. Corso, S. de Gironcoli, P. Delugas, R. A. DiStasio, A. Ferretti, A. Floris, G. Fratesi, G. Fugallo *et al.*, *J. Phys.: Condens. Matter* **29**, 465901 (2017).
- [46] P. Giannozzi, O. Baseggio, P. Bonfà, D. Brunato, R. Car, I. Carnimeo, C. Cavazzoni, S. de Gironcoli, P. Delugas, F. F. Ruffino, A. Ferretti, N. Marzari, I. Timrov, A. Urru, and S. Baroni, *J. Chem. Phys.* **152**, 154105 (2020).
- [47] M. Schlipf and F. Gygi, *Comput. Phys. Commun.* **196**, 36 (2015).
- [48] J. Sun, B. K. Clark, S. Torquato, and R. Car, *Nat. Commun.* **6**, 8156 (2015).
- [49] K. Berland, V. R. Cooper, K. Lee, E. Schröder, T. Thonhauser, P. Hyldgaard, and B. I. Lundqvist, *Rep. Prog. Phys.* **78**, 066501 (2015).
- [50] T. Thonhauser, S. Zuluaga, C. A. Arter, K. Berland, E. Schröder, and P. Hyldgaard, *Phys. Rev. Lett.* **115**, 136402 (2015).
- [51] D. C. Langreth, B. I. Lundqvist, S. D. Chakarova-Käck, V. R. Cooper, M. Dion, P. Hyldgaard, A. Kelkkanen, J. Kleis, L. Kong, S. Li, P. G. Moses, E. Murray, A. Puzder, H. Rydberg, E. Schröder, and T. Thonhauser, *J. Phys.: Condens. Matter* **21**, 084203 (2009).
- [52] T. Thonhauser, V. R. Cooper, S. Li, A. Puzder, P. Hyldgaard, and D. C. Langreth, *Phys. Rev. B* **76**, 125112 (2007).
- [53] V. L. Deringer and G. Csányi, *Phys. Rev. B* **95**, 094203 (2017).
- [54] G. Bussi, D. Donadio, and M. Parrinello, *J. Chem. Phys.* **126**, 014101 (2007).
- [55] G. Bussi, T. Zykova-Timan, and M. Parrinello, *J. Chem. Phys.* **130**, 074101 (2009).
- [56] M. Ceriotti, G. Bussi, and M. Parrinello, *J. Chem. Theory Comput.* **6**, 1170 (2010).
- [57] R. Helled, G. Mazzola, and R. Redmer, *Nat. Rev. Phys.* **2**, 562 (2020).
- [58] L. Talirz, S. Kumbhar, E. Passaro, A. V. Yakutovich, V. Granata, F. Gargiulo, M. Borelli, M. Uhrin, S. P. Huber, S. Zoupanos *et al.*, *Sci. Data* **7**, 299 (2020); see doi: [10.24435/materialscloud:36-ff](https://doi.org/10.24435/materialscloud:36-ff) for access to the raw calculations of the training set of liquid hydrogen and code snippets to train the ML models and run ML MD simulations.
- [59] European Organization For Nuclear Research and OpenAIRE, Zenodo (2013); see doi: [10.5281/zenodo.7044489](https://doi.org/10.5281/zenodo.7044489) for access to archived versions of the software and libraries used in this Letter: LIBRASCAL, I-PI and QUANTUM ESPRESSO.
- [60] I. Lomonosov and V. Gryaznov, *Contrib. Plasma Phys.* **56**, 302 (2016).
- [61] R. W. Lee, S. J. Moon, H.-K. Chung, W. Rozmus, H. A. Baldis, G. Gregori, R. C. Cauble, O. L. Landen, J. S. Wark, A. Ng, S. J. Rose, C. L. Lewis, D. Riley, J.-C. Gauthier, and P. Audebert, *J. Opt. Soc. Am. B* **20**, 770 (2003).
- [62] N. Marzari, D. Vanderbilt, A. De Vita, and M. C. Payne, *Phys. Rev. Lett.* **82**, 3296 (1999).
- [63] J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
- [64] A. P. Thompson, H. M. Aktulga, R. Berger, D. S. Bolintineanu, W. M. Brown, P. S. Crozier, P. J. in 't Veld, A. Kohlmeyer, S. G. Moore, T. D. Nguyen, R. Shan, M. J. Stevens, J. Tranchida, C. Trott, and S. J. Plimpton, *Comput. Phys. Commun.* **271**, 108171 (2022).
- [65] M. J. Willatt, M. Ceriotti, and S. C. Althorpe, *J. Chem. Phys.* **148**, 102336 (2018).
- [66] G. Imbalzano, A. Anelli, D. Giofré, S. Klees, J. Behler, and M. Ceriotti, *J. Chem. Phys.* **148**, 241730 (2018).