Letter

High-pressure hydrogen by machine learning and quantum Monte Carlo

Andrea Tirelli [1,*] Giacomo Tenti [1] Kousuke Nakano [1,2] and Sandro Sorella [1,3,†]

[1] *International School for Advanced Studies (SISSA), Via Bonomea 265, 34136 Trieste, Italy*

[2] *School of Information Science, JAIST, Asahidai 1-1, Nomi, Ishikawa 923-1292, Japan*

[3] *Computational Materials Science Research Team, RIKEN Center for Computational Science (R-CCS), Kobe, Hyogo 650-0047, Japan*

(Received 22 December 2021; accepted 13 June 2022; published 8 July 2022)

We have developed a technique combining the accuracy of quantum Monte Carlo in describing the electron correlation with the efficiency of a machine learning potential (MLP). We use kernel regression in combination with the smooth overlap of atomic position (SOAP) features, implemented here in a very efficient way. The key ingredients are as follows: (i) a sparsification technique, based on farthest point sampling, ensuring generality and transferability of our MLPs, and (ii) the so-called Δ–learning, allowing a small training data set, a fundamental property for highly accurate but computationally demanding calculations, such as the ones based on quantum Monte Carlo. As an application we present a benchmark study of the liquid-liquid transition of high-pressure hydrogen and show the quality of our MLP, by emphasizing the importance of high accuracy for this very debated subject, where experiments are difficult in the laboratory, and theory is still far from being conclusive.

DOI: 10.1103/PhysRevB.106.L041105

Introduction. The determination of the phase diagram of hydrogen at high pressure has been a particularly long-standing problem in the physics community since the pioneering work of Wigner and Huntington [1]. Recently, it has received renewed interest, e.g., for the still controversial properties of the newly discovered H-based superconductors such as H_3S [2] and LaH_{10} [3], for establishing the polymorphs of the solid phase of hydrogen [4,5], and, last but not least, for characterizing the nature of the liquid-liquid transition (LLT) between molecular and atomic phases.

In this Letter we are mainly addressing the last aspect, that is of paramount importance for determining the composition of giant planets, and, by consequence, of our universe. Unfortunately, many contradictory experimental [6–11] and computational [12–19] results have been reported so far, about whether the LLT is a first-order or a smooth transition, and the location of the LLT on the p-T phase diagram. Indeed, for studying phase transition phenomena, molecular dynamics simulations with long times and several atoms are required, in order to exclude equilibration and finite-size effect artifacts, respectively. In this respect, the construction of machine learning potentials (MLPs), aimed to reproduce accurate electronic energies and force fields, is one of the most promising approaches to overcome the large computational cost of present *ab initio* techniques for electronic simulation [20]. In fact, very recently, Cheng *et al.* [19] constructed MLPs for hydrogen, using the Behler-Parrinello artificial neural network framework (NNPs) [21,22] based on density functional theory (DFT) references. Here, we show that this approach can be extended to a state-of-the-art wave function method, quantum Monte Carlo (QMC) [23], that is well known to describe very accurately the electronic correlation but has a computational cost at least three orders of magnitude larger than DFT. Indeed, in hydrogen, it is important to have a very reliable technique, because experiments are difficult at extreme conditions and DFT, on the other hand, depends too much on the functional used [24,25]. In this Letter we aim to produce a MLP that is efficient enough to allow long and large size simulations, but on the other hand, retains the same accuracy of QMC, and, importantly, the training can be achieved by a relatively small number of simulations. We will show that this is possible, by employing the Δ-learning method combined with the Gaussian kernel regression (GKR) framework [26].

Model and methods. We outline the theoretical and technical details of the machine learning (ML) models that can be trained using our algorithmic approach [27].

Smooth overlap of atomic position (SOAP) kernel method. One of the key features of modern MLPs is their applicability to systems with an arbitrary number of atoms N, allowing simulations with sizes much larger than the ones used during the training. One possibility is to write the total energy of the system as a sum of contributions depending on the local environment around each atom placed at the position $\mathbf{r}_i$ [21], $E = \sum_{i=1}^{N} e(R_i)$, where $R_i = (\mathbf{r}_1 - \mathbf{r}_i, \ldots, \mathbf{r}_N - \mathbf{r}_i) \equiv (\mathbf{r}_{1i}, \ldots, \mathbf{r}_{Ni})$. The function $e(R_i)$ depends only on the neighbors of atom i inside a sphere with radius r_c, namely such that $r_{ji} = |\mathbf{r}_{ji}| \leqslant r_c$ $(j = 1, \cdots, N)$. Several approaches are possible to parametrize $e(R_i)$. In this work we used a kernel regression model,

$$e(R_i) = \sum_{j=1}^{N_{\text{train}}} c_j \mathcal{K}(R_i, R_j), \tag{1}$$

where $\mathcal{K}(R_i, R_j)$ is the normalized kernel between two local environments R_i (the target one) and R_j (the ones used for the training, that may or may not contain R_i), and c_j are

*atirelli@sissa.it

†sorella@sissa.it

 ©2022 American Physical Society

coefficients to be determined. Here, N_{train} is the total number of local environments used for the construction of the MLP.

The local environment around each atom is characterized using the smooth overlap of atomic position (SOAP) [28]: In such an approach, a density function $\rho(R_i, \mathbf{r})$ is associated to each local environment R_i. Hence

$$\rho(R_i, \mathbf{r}) = \sum_{j|r_{ij}\leqslant r_c} f_c(r_{ij}) \exp\left(-\frac{(\mathbf{r}_{ij} - \mathbf{r})^2}{2\sigma^2}\right),$$

where f_c is a cutoff function and σ is a hyperparameter of the model. In this work a piecewise defined polynomial, vanishing smoothly at $r_{ij} = r_c$, was used as $f_c(r_{ij})$ [29]. The *similarity kernel* between two local environments R_i and R_j is defined as the average over all rotations $\hat{U}$ of the overlap between the densities raised to some exponent n:

$$\tilde{k}(R_i, R_j) = \int d\hat{U}\left(\int d\mathbf{r}\rho(R_i, \mathbf{r})\rho(\hat{U}R_j, \mathbf{r})\right)^n. \quad (2)$$

With this definition the resulting kernel is explicitly invariant under permutation of the particles and rotations in three-dimensional (3D) space. In order to compute the integral in Eq. (2), the usual approach is to expand ρ in radial and angular components and evaluate the power spectrum [28]. Here, instead, in analogy with the nowadays standard implementation of nonlocal pseudopotentials within QMC [30] (where a similar angular integration appears), we have used an approximate version of Eq. (2) which takes the average of the overlap only over a finite subgroup (e.g., all the symmetries of the cube) of SO(3), that is spanned by n_{sym} 3×3 real orthogonal matrices U_k:

$$\tilde{k}(R_i, R_j) = \frac{1}{n_{\text{sym}}}\sum_{k=1}^{n_{\text{sym}}}\left(\int d\mathbf{r}\rho(R_i, \mathbf{r})\rho(\hat{U}_k R_j, \mathbf{r})\right)^n. \quad (3)$$

The resulting kernel is explicitly invariant under the set of transformations considered, that approaches the full continuous symmetry for large n_{sym}. Just as for QMC pseudopotential calculations, we have experienced that a relatively small n_{sym} is sufficient to reach the desired accuracy. Moreover, Eq. (3) can be computed in a very efficient way since each term in the sum has a simple expression in terms of Gaussian functions. Finally, in order to obtain $\mathcal{K}$ in Eq. (1), we normalize $\tilde{k}$ and raise it to some power η, so that $0 \leqslant \mathcal{K} \leqslant 1$.

Local sparsification. A key aspect in the implementation and training of a ML model is the prevention of overfitting, i.e., the statistical phenomenon for which a model can fit almost exactly the training data but fails to predict unknown data reliably.

Overfitting problems can arise when the number of variational parameters of a given model is too large. Moreover, since the number of variational parameters in a GKR model depends linearly on the size of the training set, particular care should be taken when selecting the corresponding reference samples. This becomes even more important in the present setting: Indeed, adding one configuration with N atoms to the training data set implicitly defines N new local environments corresponding to further N variational degrees of freedom of the model in Eq. (1). Additionally, the computational effort needed in the training of a GKR model is polynomial (at most cubic) in the size of the training set—while instead stochastic optimization algorithms employed in neural networks allow the complexity to be independent of the size of the number of samples.

Therefore, within GKR, the careful choice of the local environments used in the training phase is of paramount importance. This is usually obtained by means of the so-called *sparsification*. We implement a variant of it that we name *local sparsification*, since we discard from each configuration some or even a large fraction of the corresponding N local environments. This prevents overfitting and allows an increased predictive power of the model on unseen data.

The local sparsification preprocessing can be performed with multiple algorithms, and we employ the *farthest point sampling scheme*, which is a greedy algorithm first adopted as an approximate solution of the so-called center selection problem (see Sec. 11.2 in Ref. [31]). A detailed explanation of our algorithm is contained in the Supplemental Material (SM) [27].

Δ-machine learning. In this Letter we also follow the Δ-machine learning (Δ-ML) approach, first introduced in Ref. [32] as a composite strategy reproducing a target *ab initio* method, by adding ML corrections to a much cheaper baseline reference potential. Here, we take as the baseline *ab-initio* reference a DFT calculation with a Perdew-Burke-Ernzerhof (PBE) functional and variational Monte Carlo (VMC) as the target method.

There is a significant gain in exploiting the Δ-ML approach: By training on the difference between two *ab initio* methods, the machine learning task becomes easier. This is crucial for the case at hand: If we were to train a GKR model to predict VMC observables, the amount of training configurations to achieve the desired accuracy would require a computational effort which is at present not feasible. In particular, we are able to obtain a VMC potential with an accuracy comparable with the error bar of the VMC predictions, by training the model on a relatively small data set. Moreover, the computational cost of the ML inference (also owing to the local sparsification preprocessing) is negligible compared to the baseline DFT calculation. Hence, we are able to compute observables with VMC accuracy at the cost of a DFT simulation.

Results. We discuss now a successful application of our method to the study of the LLT in hydrogen at high pressures. Our data set is composed of 684 configurations of 128 hydrogen atoms in cubic simulation cells of different edge sizes, from 10.56 to 11.53 bohrs. For each of these configurations we computed the ground state energy, the force components acting on each atom, and the pressure of the system using two different *ab initio* methods: DFT with the PBE and Becke-Lee-Yang-Parr (BLYP) functionals, with the QUANTUM ESPRESSO software package [33], and VMC, with the TURBORVB software package [34]. Concerning the VMC simulations, the predicted energies, force components and pressure have average error bars of 2×10^{-3} hartree, 1.8×10^{-3} hartree/bohrs, and 0.1 GPa, respectively. Note that such averages can be taken as theoretical lower bounds on the achievable performances of any MLP trained on such noisy data. Furthermore, we point out that performing a molecular

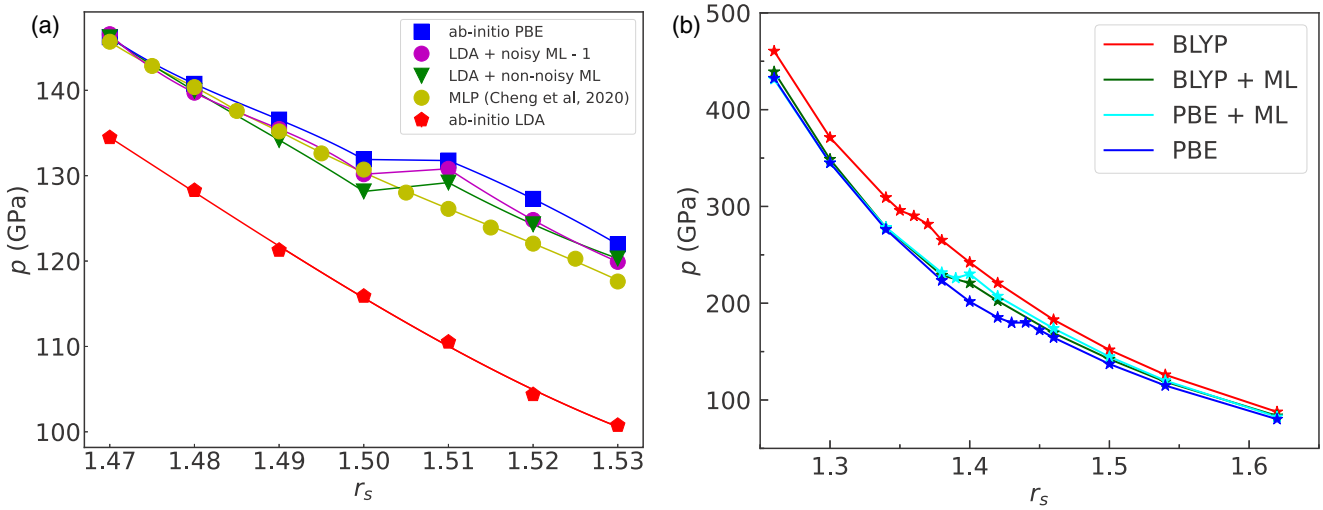


FIG. 1. (a) Comparison between molecular dynamics simulations via MLPs (cyan and green lines for our GKR, yellow line for the NN-MLP of Cheng *et al.* [19]), and the AIMD PBE and LDA references (blue and red lines, respectively) for 128 H at 1400 K. The values of the hyperparameters used for all the MLPs are reported in the SM [27]. (b) EOS of 128 H at 1000 K using DFT with PBE and BLYP functionals and the appropriate ML corrections.

dynamics (MD) simulation using VMC with such an accuracy at each step would be computationally prohibitive.

Performances of the Δ-ML correction. Although computing the root-mean-square error (RMSE) on target quantities is a standard approach to access the predictive power of a ML model, it seems necessary in our case to have more tailored qualitative and quantitative performance metrics. Since QMC provides training data with intrinsic noise, the first step in our study is to determine how this affects the ML approach. To this purpose, since it is not possible to have a noiseless VMC reference, we have performed the following test: We trained our GKR on two data sets, such that one was simply the difference between a DFT baseline reference [calculated with the local density approximation (LDA) functional] and a DFT-PBE target, while the other was a noisy version of the first, where we introduced Gaussian white noises on the observables in order to emulate QMC inputs. As shown in Fig. 1(a), the effect of a noise $\simeq 10^{-3}$ a.u. on forces and total energies comparable with the intrinsic noise affecting the VMC calculations performed for our study [35] is almost negligible, yielding a pressure plateau almost indistinguishable from the reference PBE one and the MLP trained with the noiseless data set. The performances of the proposed MLP are also much better than the ones of the previous attempt [19], unable to reproduce the pressure plateau in this case. This is remarkable because in this region the LDA baseline is very far from the reference and also featureless. Note also that our finding is compatible with Fig. S6 of Ref. [19], since in that graph the r_s grid employed for the *ab initio* molecular dynamics (AIMD) PBE simulations is very coarse and away from the plateau.

Given this remarkable stability property of our GKR against noise, we further show the independence of the method [see Fig. 1(b)] when trained (1) with true QMC noisy data and (2) starting from two references, namely PBE and BLYP. As shown in Fig. 1(b), we have obtained an excellent agreement between the two MLP equations of state (EOS), even though the two DFT-based isotherms differ significantly from each other. In particular, the transition pressure changes by almost 100 GPa with the two chosen DFT functionals. Nevertheless, the two models corrected with the MLP are much closer, with a difference in the predicted p not larger than 3–4 GPa (or anyway within the error bars) in the range of densities explored. For the rest of this work, we used the PBE-DFT method as our baseline potential, since it appears to provide results closer to the VMC ones.

LLT for high-pressure hydrogen. We have considered two systems of 128 and 256 hydrogen atoms, respectively, and performed a series of classical NVT simulations for several values of the density and temperature. At each step of MD we first compute DFT-PBE forces for the given configuration and then sum the correction given by our MLP. We have used a time step of 10 a.u. (i.e., 0.242 fs) and controlled the temperature using either second-order Langevin dynamics [36] or a Bussi-Parrinello thermostat [37]. A cubic supercell has been considered for all the simulations. For each point we have chosen a time simulation length between 2 and 10 ps, from which we have calculated the average of the thermodynamic quantities of interest. We remark that, even though our Δ-ML approach is still expensive (it has the same cost of a standard DFT Born-Oppenheimer MD), it is about 10^3 times faster than computing directly VMC forces at each step, with the chosen accuracy.

For all cases considered we have found a narrow but clear plateau, suggesting a first-order nature of the LLT transition in this system [Fig. 2(a)]. Both the transition pressures and corresponding densities are quantitatively much different from the DFT ones, suggesting the importance of the VMC correction for the accurate evaluation of those quantities.

We have also computed the average of the molecular fraction along the trajectory, defined as the percentage of atoms

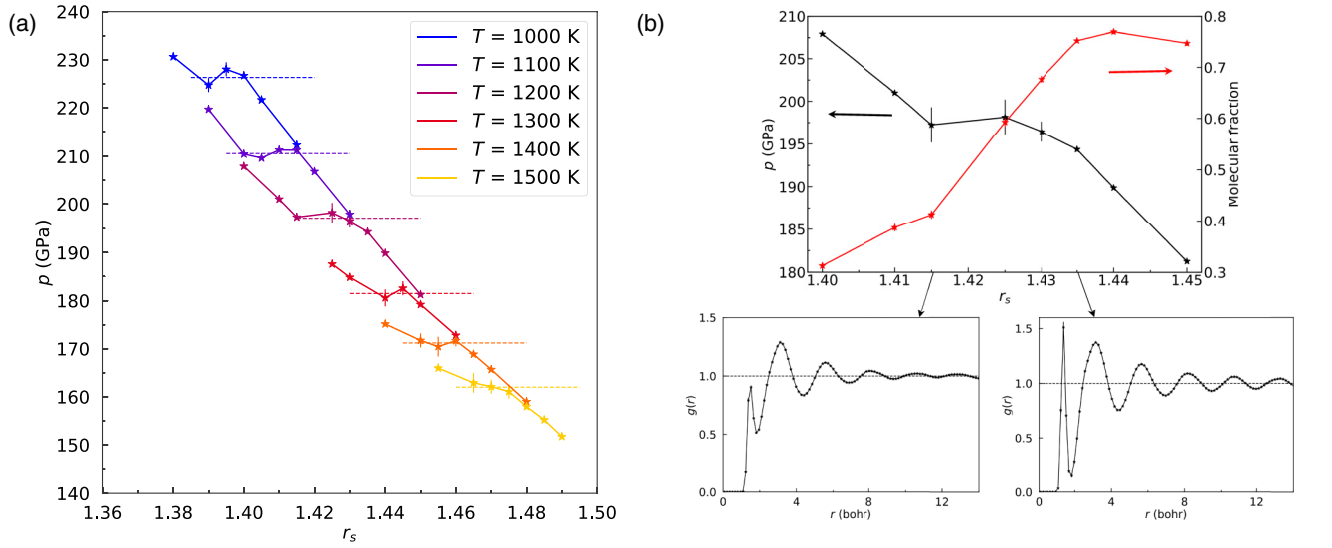


FIG. 2. (a) Isotherm curves at different values of temperature for a system of 256 hydrogen atoms. Dashed lines represent the transition pressure obtained when p becomes constant (within the error bars) in some range of densities. (b) Isotherm for $T = 1200$ K (black line). The molecular fraction (defined in the text) as a function of the Wigner-Seitz radius r_s (red line) and the radial distribution functions for $r_s = 1.415$ and $r_s = 1.435$ (boxes in the bottom) are also shown.

with at least one neighbor inside a cutoff of 1.4 a.u., i.e.,

$$m = \frac{1}{N}\sum_i g_c\left(\min_{j,j\neq i}|\mathbf{r}_i - \mathbf{r}_j|\right),$$

where $g_c(r)$ is a cutoff function that is equal to 1 for $r < 1.4$ a.u. and goes smoothly to 0 for $r > 1.6$ a.u.

As evident in Fig. 2(b), at the location of the pressure plateau the molecular fraction jumps from small values, corresponding to an atomic liquid, to values closer to 1, typical for a molecular system. This can be also noticed by computing the radial distribution function for two values of the density at opposite sides of the transition.

In Fig. 3 we report the computed transition line, alongside previous numerical and experimental results. As expected, we have found good agreement with previous VMC-based molecular dynamics simulations that used the same wave function [17] but the statistical errors on energy, forces, and pressure were an order of magnitude larger, thus explaining the small differences found. Our calculations give transition pressures that are 30 GPa below the ones predicted by the coupled electron-ion Monte Carlo (CEIMC) method [38]. The reason for this discrepancy is not clear at present; in principle the calculations done in Ref. [38] could be more accurate due to the presence of backflow correlation and diffusion Monte Carlo (DMC) projection, that could be further included for the training of our MLP. However, from our benchmark tests DMC gives only a small correction to the VMC forces, that is compatible with the error bars as shown in the Supplemental Material [27] (in future work, we plan to study the system using a MLP constructed on DMC data, where forces are computed using Reynold's approximation, [42]) Our results are also in excellent agreement with the diamond anvil cell (DAC) experiments of Refs. [8,10], especially if we take into account that quantum corrections tend to further lower the transition pressure.

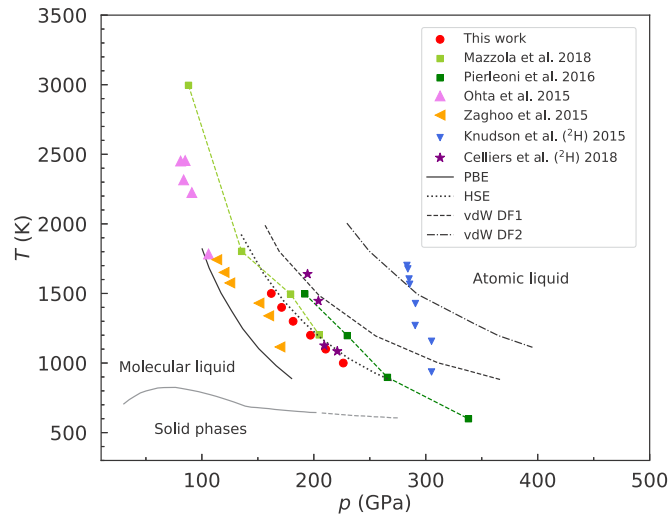


FIG. 3. Proposed transition lines of the LLT in hydrogen. Our results (red circles). Previous (classical) QMC calculations taken from Refs. [17,38] (light green and dark green squares). DFT results for several functionals as reported in Ref. [38] (black lines). Recent classical PBE results calculated in Ref. [39] (black solid lines) and the deorbitalized version of the strongly constrained and appropriately normed (SCAN-L) functional from Ref. [40] (dashed-dotted black line) are also shown. DAC experimental data taken from Refs. [8,10] (pink and orange triangles). Dynamical compression experiments taken from Refs. [7] and [6] (blue triangles and violet stars, respectively). The literature data were digitalized using WEBPLOTDIGITIZER [41].

We conclude this discussion by pointing out that our method gives almost the same line as the hybrid Heyd-Scuseria-Ernzerhof (HSE) functional. This could be useful for future studies based on DFT for high-pressure hydrogen.

Conclusion. In this Letter we have introduced a successful approach, combining the ML and QMC techniques: In particular, through the Δ-ML framework, we are able to obtain energies, forces, and pressure predictions with QMC accuracy, at a computational cost comparable to DFT, allowing simulations that were previously impossible (see Fig. S7, where MD simulations with $N > 1000$, at the QMC accuracy, are shown [27]). Our technique is perfectly suited for being employed in a context where high accuracy is needed, but small training data sets are available. We benchmarked our approach by studying the LLT in hydrogen at high pressures, whose phase diagram has been under a very lively debate, especially recently. In this context, our method is capable of reproducing the pressure plateau computed directly with PBE at 1400 K, in contrast to the previously proposed NN MLP. This outcome appears in agreement with Ref. [39] but the subtle question of whether a genuine first-order transition or a sharp crossover occurs is outside the scope of this Letter, and will be dealt with in future works with much larger N and simulation times. Moreover, we believe that further substantial improvements will be possible by using baseline potentials (e.g., the recent ultrafast MLP [43]) that are much more efficient than DFT ones. Indeed, we expect that our approach will enable the study of systems that have been so far impossible to simulate, due to the almost intractable computational cost required by advanced many-body techniques, such as QMC.

Acknowledgments. We thank Matthias Rupp for useful discussions and careful reading of the manuscript. The computations in this work have been mainly performed using the supercomputer Fugaku provided by RIKEN through the HPCI System Research Projects (Project IDs: hp210038 and hp220060) and Marconi100 provided by CINECA through the PRACE project No. 2019204934 and ISCRA project AI-H-QMC, HP10BGJH1X. K.N. is also grateful for computational resources from the facilities of Research Center for Advanced Computing Infrastructure at Japan Advanced Institute of Science and Technology (JAIST). A.T. and S.S. acknowledge financial support from the MIUR Progetti di Ricerca di Rilevante Interesse Nazionale (PRIN) Bando 2017 Grant No. 2017BZPKSZ. K.N. acknowledges support from the JSPS Overseas Research Fellowships, from the Grant-in-Aid for Early-Career Scientists Grant No. JP21K17752, and from the Grant-in-Aid for Scientific Research(C) Grant No. JP21K03400. This work is supported by the European Centre of Excellence in Exascale Computing TREX - Targeting Real Chemical Accuracy at the Exascale. This project has received funding from the European Union's Horizon 2020 Research and Innovation program under Grant Agreement No. 952165.

[1] E. Wigner and H. Huntington, J. Chem. Phys. **3**, 764 (1935).
[2] A. Drozdov, M. Eremets, I. Troyan, V. Ksenofontov, and S. I. Shylin, Nature (London) **525**, 73 (2015).
[3] M. Somayazulu, M. Ahart, A. K. Mishra, Z. M. Geballe, M. Baldini, Y. Meng, V. V. Struzhkin, and R. J. Hemley, Phys. Rev. Lett. **122**, 027001 (2019).
[4] N. D. Drummond, B. Monserrat, J. H. Lloyd-Williams, P. L. Ríos, C. J. Pickard, and R. J. Needs, Nat. Commun. **6**, 7794 (2015).
[5] A. Drozdov, P. Kong, V. Minkov, S. Besedin, M. Kuzovnikov, S. Mozaffari, L. Balicas, F. Balakirev, D. Graf, V. Prakapenka *et al.*, Nature (London) **569**, 528 (2019).
[6] P. M. Celliers, M. Millot, S. Brygoo, R. S. McWilliams, D. E. Fratanduono, J. R. Rygg, A. F. Goncharov, P. Loubeyre, J. H. Eggert, J. L. Peterson *et al.*, Science **361**, 677 (2018).
[7] M. D. Knudson, M. P. Desjarlais, A. Becker, R. W. Lemke, K. Cochrane, M. E. Savage, D. E. Bliss, T. Mattsson, and R. Redmer, Science **348**, 1455 (2015).
[8] K. Ohta, K. Ichimaru, M. Einaga, S. Kawaguchi, K. Shimizu, T. Matsuoka, N. Hirao, and Y. Ohishi, Sci. Rep. **5**, 16560 (2015).
[9] R. S. McWilliams, D. A. Dalton, M. F. Mahmood, and A. F. Goncharov, Phys. Rev. Lett. **116**, 255501 (2016).
[10] M. Zaghoo, A. Salamat, and I. F. Silvera, Phys. Rev. B **93**, 155128 (2016).
[11] M. Zaghoo and I. F. Silvera, Proc. Natl. Acad. Sci. USA **114**, 11873 (2017).
[12] S. Scandolo, Proc. Natl. Acad. Sci. USA **100**, 3051 (2003).
[13] M. A. Morales, C. Pierleoni, E. Schwegler, and D. M. Ceperley, Proc. Natl. Acad. Sci. USA **107**, 12799 (2010).
[14] W. Lorenzen, B. Holst, and R. Redmer, Phys. Rev. B **82**, 195107 (2010).
[15] K. T. Delaney, C. Pierleoni, and D. M. Ceperley, Phys. Rev. Lett. **97**, 235702 (2006).
[16] J. Vorberger, I. Tamblyn, B. Militzer, and S. A. Bonev, Phys. Rev. B **75**, 024206 (2007).
[17] G. Mazzola, R. Helled, and S. Sorella, Phys. Rev. Lett. **120**, 025701 (2018).
[18] H. Y. Geng, Q. Wu, M. Marqués, and G. J. Ackland, Phys. Rev. B **100**, 134109 (2019).
[19] B. Cheng, G. Mazzola, C. J. Pickard, and M. Ceriotti, Nature (London) **585**, 217 (2020).
[20] W. Jia, H. Wang, M. Chen, D. Lu, L. Lin, R. Car, W. E, and L. Zhang, Pushing the limit of molecular dynamics with ab initio accuracy to 100 million atoms with machine learning, in *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis* (IEEE, New York, 2020).
[21] J. Behler and M. Parrinello, Phys. Rev. Lett. **98**, 146401 (2007).
[22] J. Behler, J. Chem. Phys. **134**, 074106 (2011).
[23] W. Foulkes, L. Mitas, R. Needs, and G. Rajagopal, Rev. Mod. Phys. **73**, 33 (2001).
[24] S. Azadi and W. M. C. Foulkes, Phys. Rev. B **88**, 014115 (2013).
[25] R. C. Clay, J. Mcminis, J. M. McMahon, C. Pierleoni, D. M. Ceperley, and M. A. Morales, Phys. Rev. B **89**, 184106 (2014).
[26] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, Phys. Rev. Lett. **104**, 136403 (2010).
[27] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevB.106.L041105 for technical details of the methods, benchmark, and validation tests.

[28] A. P. Bartók, R. Kondor, and G. Csányi, Phys. Rev. B **87**, 184115 (2013).

[29] $f(r_{ij}) = 1$ if $0 < r_{ij} \leqslant r_c/2$ and $f(r_{ij}) = 4(r_{ij}/r_c - 1)^2(4r_{ij}/r_c - 1)$ for $r_c/2 \leqslant r_{ij} \leqslant r_c$. We also assume that $f(0) = 0$, in order to exclude the trivial $i = j$ contribution.

[30] L. Mitáš, E. L. Shirley, and D. M. Ceperley, J. Chem. Phys. **95**, 3467 (1991).

[31] J. Kleinberg and E. Tardos, *Algorithm Design* (Pearson Education India, Delhi, 2006).

[32] R. Ramakrishnan, P. O. Dral, M. Rupp, and O. A. von Lilienfeld, J. Chem. Theory Comput. **11**, 2087 (2015).

[33] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo *et al.*, J. Phys.: Condens. Matter **21**, 395502 (2009).

[34] K. Nakano, C. Attaccalite, M. Barborini, L. Capriotti, M. Casula, E. Coccia, M. Dagrada, C. Genovese, Y. Luo, G. Mazzola, A. Zen, and S. Sorella, J. Chem. Phys. **152**, 204121 (2020).

[35] This error bar can be achieved, for instance, by a single QMC simulation on about 1000 processors for about a couple of hours within the TURBORVB package [34]. Clearly this error bar cannot be sustained for an MD run with several thousand steps, as the ones employed here.

[36] N. V. Kampen, in *Stochastic Processes in Physics and Chemistry*, edited by N. Van Kampen, North-Holland Personal Library, 3rd ed. (Elsevier, Amsterdam, 2007), pp. 219–243.

[37] G. Bussi, D. Donadio, and M. Parrinello, J. Chem. Phys. **126**, 014101 (2007).

[38] C. Pierleoni, M. A. Morales, G. Rillo, M. Holzmann, and D. M. Ceperley, Proc. Natl. Acad. Sci. USA **113**, 4953 (2016).

[39] V. V. Karasiev, J. Hinz, S. X. Hu, and S. B. Trickey, Nature (London) **600**, E12 (2021).

[40] J. Hinz, V. V. Karasiev, S. X. Hu, M. Zaghoo, D. Mejía-Rodríguez, S. B. Trickey, and L. Calderín, Phys. Rev. Research **2**, 032065(R) (2020).

[41] A. Rohatgi, Webplotdigitizer: Version 4.5, https://automeris.io/WebPlotDigitizer (2021).

[42] K. Nakano, A. Raghav, and S. Sorella, J. Chem. Phys. **156**, 034101 (2022).

[43] S. R. Xie, M. Rupp, and R. G. Hennig, arXiv:2110.00624.