Quantum Dynamics of Water from Møller-Plesset Perturbation Theory via a Neural Network Potential

Jinggang Lan
Department of Chemistry, University of Zurich, Winterthurerstrasse 190, Zurich 8057, Switzerland and
Current address: Chaire de Simulation à l'Echelle Atomique (CSEA),
Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland;

David M. Wilkins
Atomistic Simulation Centre, School of Mathematics and Physics,
Queen's University Belfast, Belfast BT7 1NN, Northern Ireland, United Kingdom

Vladimir V. Rybkin
Department of Chemistry, University of Zurich, Winterthurerstrasse 190, Zurich 8057, Switzerland and
Current address: HQS Quantum Simulations GmbH, Haid-und-Neu-Straße 7, D-76131 Karlsruhe, Germany

Marcella Iannuzzi, Jürg Hutter
Department of Chemistry, University of Zurich, Winterthurerstrasse 190, Zurich 8057, Switzerland
(*V.R: vladimir.rybkin@quantumsimulations.de)
(*J.L:jinggang.lan@epfl.ch)
(Dated: December 17, 2021)

We report the static and dynamical properties of liquid water at the level of second-order Møller-Plesset perturbation theory (MP2) with classical and quantum nuclear dynamics using a neural network potential. We examined the temperature-dependent radial distribution functions, diffusion, and vibrational dynamics. MP2 theory predicts over-structured liquid water as well as a lower diffusion coefficient at ambient conditions compared to experiments, which may be attributed to the incomplete basis set. A better agreement with experimental structural properties and the diffusion constant are observed at an elevated temperature of 340 K from our simulations. Although the high-level electronic structure calculations are expensive, training a neural network potential requires only a few thousand frames. The approach is promising as it involves modest human effort and is straightforwardly extensible to other simple liquids.

I. INTRODUCTION

Water is both an extremely common and an extremely unusual liquid, with anomalous properties such as high surface tension, high viscosity, and temperature of maximum density.[1, 2] Although these macroscopic properties can be well characterized experimentally, resolving the atomistic picture remains challenging. In fact, even the structure of the water is still subject to intense debate.[2, 3] It is also important to understand how water behaves as a solvent or reactant, whether in the bulk or near an interface as these properties are of central importance in many fields, such as from atmospheric science, photochemistry, and biochemistry.[4–7] Hence, water has received great attention in both experimental and theoretical studies across diverse scientific disciplines.

Molecular dynamics (MD) simulations can provide useful insights into the atomic details of water. There is a long standing effort to simulate water using MD methods.[3, 8–10] The driving forces for MD can be obtained from either empirical force-field or *ab initio* calculations. Many water models have been proposed such as the TIP*n*P family[11–13], SPC[14, 15], and MB-Pol[16]. It is also common to use density functional theory (DFT) for *ab initio* molecular dynamics. However, despite the fact that the DFT formalism is exact for ground-state properties in principle, in practice approximations must be made in the exchange-correlation energy. In principle, the accuracy increases according to the "Jacob's ladder" of electronic structure theory, with the local density approximation (LDA), generalized gradient approximations (GGAs), meta-GGAs, hybrid functionals, and beyond e.g. double hybrids, random phase approximation (RPA), and correlated wave function methods, such as second-order Møller-Plesset perturbation theory (MP2). A properly chosen exchange-correlation functional has been found to predict reasonable binding energy for the water dimer, trimer, larger clusters, and ice as well as the structure of liquid water when combined with MD. LDA is unacceptable as it overbinds the water dimer by nearly a factor of 2.[17] Unsurprisingly, LDA predicts overstructured liquid water and a very small diffusion constant.[18] GGA functionals perform only slightly better. DFT-based MD at the GGA level, without dispersion corrections, often requires a much higher temperature than 300 K to maintain water in the liquid state. For example, one of the most popular functionals, PBE, can reproduce experimental radial distribution functions (RDFs) and density at T=440 K and P = 0.3 GPa[19]. Even when a dispersion correction is added a higher temperature is still required to reproduce these properties.[20, 21] Grimme's dispersion correction to PBE does not significantly improve the structure of water.[21] BLYP with this correction requires simulation at 360 K to reproduce the correct RDF.[22] Among the GGA functionals, revPBE-D3 tends to perform best for RDF, density, and diffusion coefficient at ambient conditions without increasing the temperature.[17, 21] However, the results are very sen-

sitive to the particular choice of dispersion functional. The revPBE+DRSLL functional tends to overestimate the volume of ice VIII by 20%. [17]

Recent developments in computational power and methodology have made it possible to climb higher rungs of the ladder beyond GGA, namely meta-GGA, hybrid functionals, and many-body correlated methods.[23] Recent works show that significant improvement can be achieved by the non-empirical meta-GGA functional SCAN, as developed by Perdew and co-workers.[24] SCAN yields the correct ordering of densities between liquid water and ice, at the same time predicting quantitative agreement with experiments at an elevated temperature of 330 K.[25] However, the bare SCAN functional systematically overbinds water[26]. Moreover, adding rVV10 to SCAN makes overbinding even stronger, and predicts even higher water density.[27] Fortunately, a very recent study by Paesani and co-workers found that density-corrected SCAN improved the accuracy to the same degree as the "gold standard" coupled-cluster theory. [26] At the same time, some empirically parametrized meta-GGA functionals such as B97M with rVV10 correction appear to perform quite well. [28]

At hybrid functional level, The revPBE0-D3 functional is able to predict the correct experimental RDF and density at room temperature.[29] However, revPBE0-D3 still underestimates the temperature-dependent density of water by about 5%[30], which may be attributed to the choice of van der Waals interactions. When the hybrid functional is combined with rVV10 van der Waals interactions, the equilibrium density of water agrees better with the experiment when the parameter from the rVV10 van der Waals interaction is fine-tuned.[31] Fortunately, it is possible to go beyond hybrid functionals; including virtual orbitals allows for long-range van der Waals dispersion interactions from parameter-free ab-initio calculations. With tremendous computational cost, one can reach the fifth rung of the ladder with methods such as RPA and MP2. The MP2 theory gives excellent predictions of the density at room temperature, while the calculated radial distribution functions are in fair agreement with experimental data[32] that tend to be attributed to the stronger dispersion interaction. Previous studies suggest that "MP2 water" is denser and cooler at ambient conditions compared both with experiment and with "DFT water". [33]

In addition to the underlying electronic structure theory, another issue is the accurate account of the quantum nature of the nuclei.[34, 35] Neglect of nuclear quantum effects (NQEs) may become one of the largest sources of error.[35] For instance, there are subtle differences in RDFs upon the inclusion of NQEs, especially those that involve hydrogen atoms.[34] In a classical simulation, to produce an equivalent change in the RDF, one has to increase the temperature by about 30 K.[9, 25] Although different quantum effects on the O-O RDF can be canceled, quantum effects on individual H-bond coordinates can be significant. In comparison with a classical simulation, proton fluctuations along a covalent bond direction increase almost 10-fold. In addition, recent work from Voth and co-workers found that higher temperatures do not accurately replicate NQEs at room temperature, which is evident in different three-body correlations as well as dynamics.[36] Furthermore, classical molecular dynamics overestimates the vibrational frequencies of the O-H stretch region and the H-O-H bend region compared to results that take the NQEs into account.[29] Regarding the reactivity of water, "classical" water is more basic than "quantum" water (i.e., water in nature) with a pH of 8.6, which is about a 30-fold change in the ionization constant.[34] Similar effects have also been found at metal-water interfaces[37] and for electron transfer in aqueous solutions[38]. Hence, the inclusion of nuclear quantum effects is often necessary to achieve high accuracy in molecular dynamics simulations.

Path integral molecular dynamics (PIMD) is a method to incorporate quantum mechanics into molecular dynamics by mapping each quantum particle onto a classical representation[39]. Combining electronic structure theory with PI methods is exceedingly computationally expensive as each particle is represented by several classical replicas, each requiring a separate on-the-fly electronic structure calculation. For this reason, high-level electronic structure calculations coupled with quantum nuclear dynamics have only recently been made possible.[29] Thanks to the recent development of neural network potentials (NNPs),[40, 41] an extensive study using a high-quality functional beyond local DFT is feasible even when combined with path-integral molecular dynamics. For instance, the densities of water as predicted by NNP trained on revPBE0-D3 data agree with the experiment to within 3% for both liquid water and ice Ih and Ic.[30] Accurate and efficient quantum vibrational spectra of water can also be obtained for bulk and interfacial systems.[42, 43] Recently, quantum dynamics simulations with forces from a fifth-rung electronic structure level have provided an accurate determination of the structure, diffusion, and vibrational features of water and aqueous solvated electron .[44, 45]

In this work, the structural and dynamical properties of bulk water have been studied by means of MD and PIMD simulations via an NNP. Radial distribution functions, diffusion, and vibrational dynamics were investigated at different temperatures. At ambient conditions, MP2 theory predicts overstructured liquid water, which might be due to an incomplete basis set. It is found that experimental structural properties and the diffusion coefficients are in better agreement with simulations at 340 K.

II. COMPUTATIONAL DETAILS

All MD simulations are performed using the i-PI code[46, 47] interfaced with LAMMPS [48], which implements the NNP potential using N2P2 [49]. We used 128 water molecules at experimental density (0.997 g/mL) to perform long-timescale molecular dynamics simulations. To achieve efficient canonical sampling while minimally perturbing the dynamics, the temperature of classical MD simulations was controlled using a stochastic velocity rescaling (SVR) thermostat [50], with a time constant of 1000 fs. A timestep of 0.5 fs was used.

Path integral molecular dynamics (PIMD) [39] was used to incorporate nuclear quantum effects into the calculation of

static properties. 64 ring-polymer beads were used, which is sufficient to converge the properties of interest even at low temperatures.[30, 51] For dynamical properties, thermostatted ring polymer molecular dynamics (TRPMD) [52], which has been shown to predict accurate vibrational dynamics of liquid water [29, 53], and partially adiabatic centroid molecular dynamics (PACMD) [54]. A timestep of 0.25 fs was used for TRPMD calculations, and 0.05 fs for PACMD. For PACMD simulations, the Parrinello-Rahman mass matrix was adjusted to shift all normal mode frequencies to to ω_{NM} = 14000 cm^{-1}. The temperature in PIMD calculations was controlled using the global path integral Langevin equation (PILE-G) thermostat attached to all the non-centroid normal modes,[55] with a time constant of 1000 fs. Both methods predict similar diffusion constants but either may have deficiencies when computing vibrational spectra especially in stretching regions.[29, 56] To obtain excellent quantitative agreement with experimental vibrational spectra, it is necessary to combine both high-precision electronic structure theory and quantum dynamics. Both of these questions are still open and under development. In particular, neither TRPMD or PACMD is recognized as the "gold standard" for uncompromising precision for most systems and properties.[53] In fact, the transition frequency for Morse Oscillator as parametrized to model an O-H bond can be solved analytically. Compared to the analytical solution, the TRPMD simulation tends to overestimate the O-H vibrational frequency, while PACMD systematically improves the description of the stretching region of the O-H bond.[43] Therefore, additional vibrational density of states based on PACMD have been performed.

Static and dynamical properties are calculated based on a 300 ps trajectory at each temperature, other than PACMD, which used only 20 ps for additional vibrational density of states (VDOS) calculations. The dynamical properties of liquid water can be measured by the diffusion coefficient (D) which can be estimated by Einstein's relation:

$$D = \frac{1}{6t}\langle|\boldsymbol{r}(t) - \boldsymbol{r}(0)|^2\rangle. \tag{1}$$

Where $r(t)$ are the atomic positions at time t, and $r(0)$ is the atomic positions at 0. The finite size correction based on the experimental viscosity.[57] (Details see Supporting Information)

The VDOS from classical MD or PIMD is calculated using the Fourier transform of the velocity autocorrelation (C_{vv}),

$$C_{vv}(\omega) = \int \langle \boldsymbol{v}(\tau)\cdot\boldsymbol{v}(t+\tau)\rangle_\tau e^{-i\omega t}dt, \tag{2}$$

with the quantum autocorrelation function computed using the centroid velocity $\boldsymbol{v}$.

The starting point for our NNP is that of Ref. 45, at the MP2 level. The reference data obtained at room temperature may be highly correlated, leading to incorrect dynamics at temperatures other than 300 K. The potential was retrained by generating a large set of configurations of 64 molecules at temperatures from 240 K to 350 K using replica exchange molecular dynamics in the NPT ensemble. 1000 additional structures were selected using CUR selection based on their atomic fingerprints [58] and added to the training set. MP2 calculations [59, 60] with triple-zeta quality correlation-consistent basis sets [61] were carried out using CP2K [62, 63].NNP training was implemented in N2P2 [49].

III. RESULTS

A. Structure of MP2 water at different temperatures

To shed light on the structure of water at different temperatures, the radial distribution functions (RDFs) of oxygen-oxygen, oxygen-hydrogen, and hydrogen-hydrogen pairs have been calculated, as shown in Fig.1 The first point to note here is that the RDFs as obtained from an NNP at 300 K are in good agreement with the previous results from Ref. 32: g_{OO}(r) tends to agree with experimental observations regardless of a much higher first maximum. (As shown in Fig. 2 in Supporting Information) In general, MP2 is more likely to overbind noncovalent complexes.

The imperfect agreement between MP2 simulations and the experimental structure may be attributed to an insufficient basis set. Calculations are carried out using a triple-ζ basis set. MP2 is known to overestimate dispersion when a relatively small basis set is applied.[66–69] Previous studies indicate that as the basis set increases from double-ζ to triple-ζ to quadruple-ζ, the mean absolute error in the binding energies relative to the complete basis set (CBS) values decreases from 2.18 to 1.74 to 0.99 kcal/mol for the water clusters (n=2-10). The relatively small basis set overestimates the binding energies compared to the CBS.[69] The global minima of water clusters (n=2-6) at the MP2 level agree with the CCSD(T) level of theory.[70] However, MP2 systematically contracts the nearest-neighbor O-O separation in water clusters by 0.005 - 0.022 Å with the average of 0.016 Å. In this respect, MP2 tends to overestimate binding energies compared to CCSD(T) results at the same quality of basis set (aug-cc-pVDZ). [70] When the binding energies are extrapolated to the CBS limit, both MP2 and CCSD(T) predict very similar values. The differences between the two methods are generally lower than 0.5 kcal/mol.[70]

Given the fact MP2 with triple-ζ overbinds and overestimates the dispersion, it is not surprising that MP2 water at room temperature over-structures and behaves like ice. The temperature dependence of the RDFs as shown in Fig.1 from 240 to 350 K was also examined. The RDFs are strongly affected by the temperature with a negative correlation at each peak. Temperature effects are more pronounced at the first peaks of $g_{\mathrm{OO}}(r)$ which concerns the structure of the first solvation shell. As the temperature increases, the water becomes less ordered with a relatively lower intensity of $g_{\mathrm{OO}}(r)$ at the first peak. NQEs broaden and weaken the intensity of the first peaks in $g_{\mathrm{OO}}(r)$ and these effects are more pronounced at a lower temperature. A temperature difference of $\sim$ 30 K in classical simulations is required to give the same height of $g_{\mathrm{OO}}(r)$ as in quantum simulations, indicating that the quantum simulation is "hotter".[9, 25] Indeed, classical MD simu-

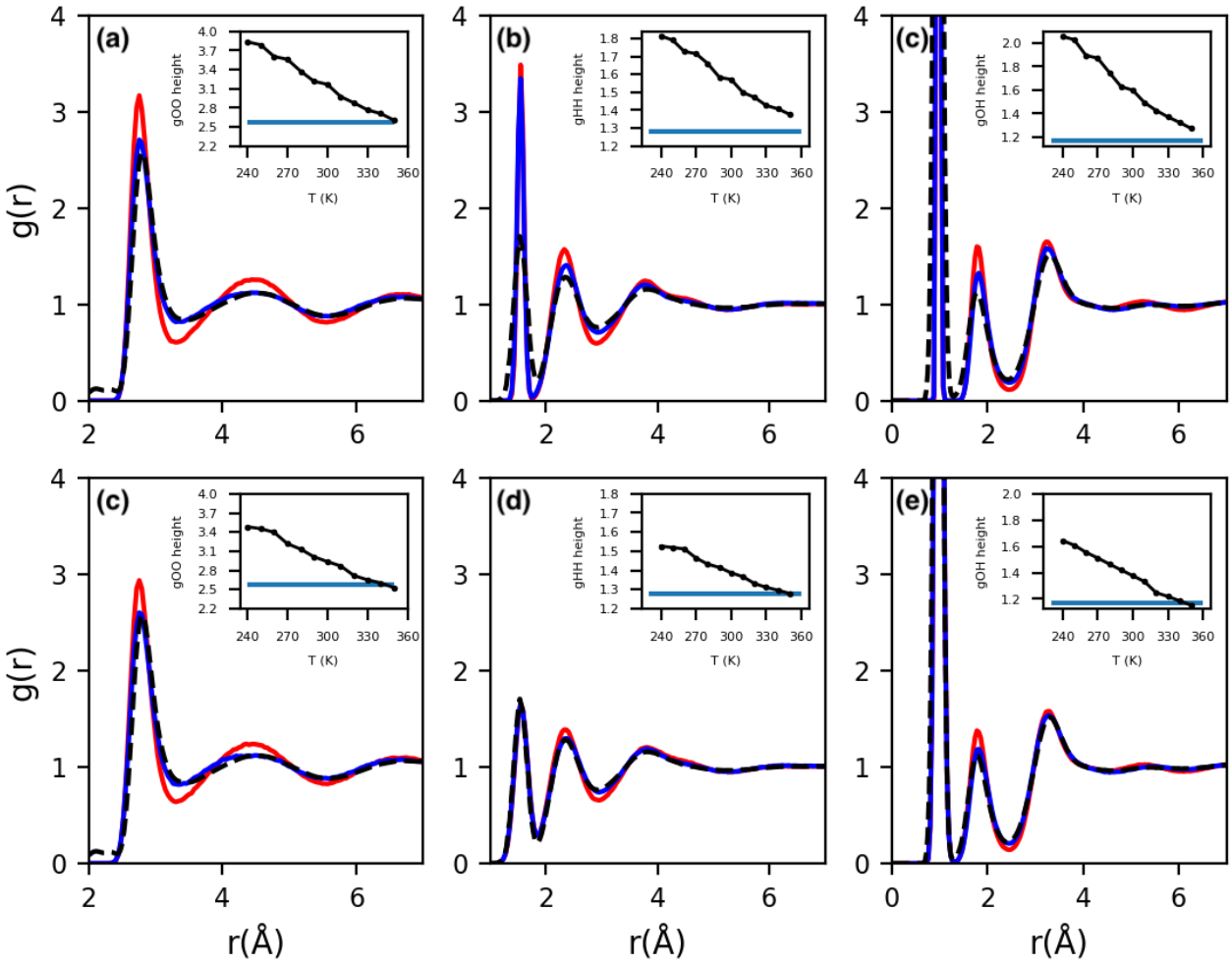

FIG. 1. Radial distribution function gOO (a,c), gHH(b,d) and gOH(c,e) as obtained from classical (a-c) and quantum (c-e) simulations at 300 K (red) and 340 K (blue) from 128 water models, in comparison to the experimental values at room temperature (black dashed line).[64, 65] Subplots show the heights at different temperatures of gOO at the first peak, gHH and gOH at second peak. The light blue lines in subplots represent RDF heights of the experimental value at room temperature.

lations of water are often carried out at an elevated temperature to mimic the quantum effect of oxygen and yield an improved description of the local structure of water.[9] In fact, NQEs are more significant when light atoms are involved. This simulation protocol with an artificial elevation of 30 K may fail dramatically in describing the $g_{\mathrm{HH}}(r)$ and $g_{\mathrm{OH}}(r)$.[9] As shown in Fig.1(b,d), NQEs significantly broaden the first and second peaks of $g_{\mathrm{HH}}(r)$ and $g_{\mathrm{OH}}(r)$. Similarly, we also plot the temperature dependence of the height of the second peaks of $g_{\mathrm{HH}}(r)$ and $g_{\mathrm{OH}}(r)$. It is evident that the temperature difference may be more than 70 K for $g_{\mathrm{HH}}(r)$ (See subplots in Fig.1(b,d)) and 50 K for g_{OH} (See subplots in Fig.1(c,e)) at the second peak of the RDFs. More importantly, the broadening of the first peaks of $g_{\mathrm{HH}}(r)$ and $g_{\mathrm{OH}}(r)$ cannot be reproduced with an elevated temperature even up to 70 K. Classical MD simulations only sample within the range of the thermal energy kT, which is far below the zero-point energy of an O–H stretch. Accurate RDFs can be achieved from MP2 theory with a higher temperature of 340 K, as shown by our classical and quantum dynamics. As in Fig.1, the experimental RDFs at room temperature are given in the dashed black lines. The calculated $g_{\mathrm{OO}}(r)$, $g_{\mathrm{HH}}(r)$ and $g_{\mathrm{OH}}(r)$ at 340 K match with the experimental value at room temperature. To further understand the local structure of the water molecules, we calculate the distribution oxygen-oxygen-oxygen triplet angles within the first solvation shell and the tetrahedral order parameter q, defined as

$$q = 1 - \frac{3}{8}\sum_{i=1}^{3}\sum_{j=i+1}^{4}(cos(\theta_{ij} + \frac{1}{3})^2, \tag{3}$$

where θ_{ij} is the angle formed by the central water molecules and its two neighboring water molecules i and j. To calculate the distribution of oxygen-oxygen-oxygen triplet angles within the first solvation shell, we take the cutoff distance of two oxygen atoms of 3.35 Å, which yields an average coordination number of 4. An order parameter of $q = 1$ defines a

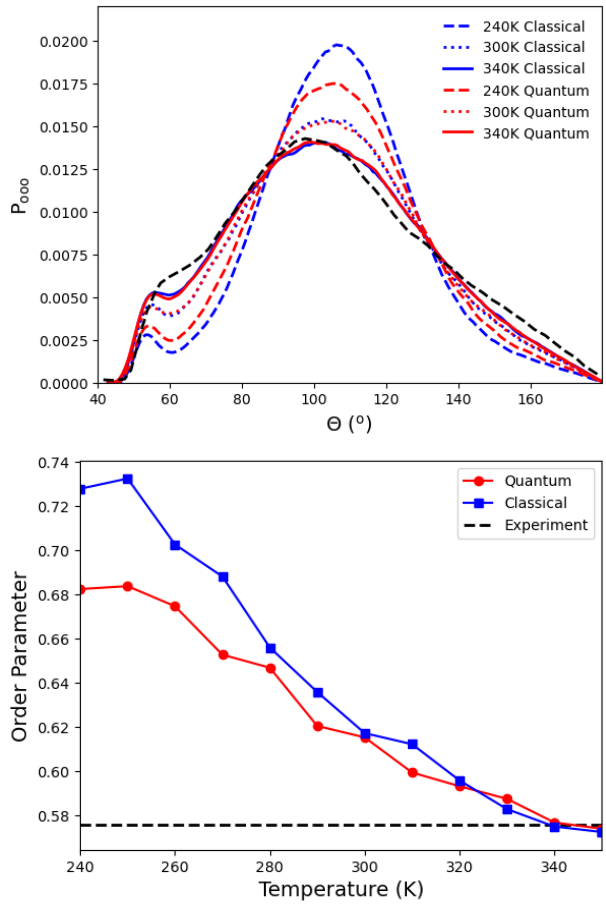

FIG. 2. The oxygen–oxygen–oxygen triplet angular distribution as obtained from classical and quantum dynamics at multiple temperatures. The resulting tetrahedral order parameter of liquid water is shown as a function of temperature. The experimental data at 298 K are marked in black dashed line as comparison.[65]

perfect tetrahedral local environment, and q decreases as the structure of water becomes less ordered and tetrahedral. As shown in Fig.2, the experimental triplet angles show a weak shoulder at around 60° and a broad strong peak at around 100°. We present the calculated P_{ooo} at three different temperatures – 240 K, 300 K, and 340 K. At the lower temperature of 240 K, NQEs strongly adjust the local environment of water, resulting in a less ordered water structure. The order parameter as estimated from quantum simulations is about 0.68 which is lower than that of 0.72 from classical simulations. The difference of local parameters becomes less with increasing the temperature. At room temperature, MP2 predicts a local parameter of 6.15 for quantum simulations (6.17 for classical), which is higher than the prediction from the fragment-based MP2/aug-cc-pVDZ simulation.[71] The inconsistency may come from the choice of basis sets and the employment of fragmentation methods. Beyond room temperature, the local parameter from classical and quantum simulations is almost identical (see Fig.2). The local order parameter at 340 K ($q_{quantum}$=0.577, $q_{classical}$=0.575) predict almost exact number compared to the experimental value of 0.570. The calculated triplet angles distributions predict an accurate line shape at around 100°, but a weaker shoulder at 56° instead of 60°.

B. Dynamical properties

It is difficult to converge the diffusion coefficient using *ab initio* molecular dynamics due to its high computational cost, especially when combined with high-level electronic structure theory beyond standard GGA. Diffusion coefficients from high-level electronic structure theory have been only made available by using a multiple time stepping (MTS) method. For instance, Marsalek and Markland carried out revPBE0-D3 calculations using ring-polymer contraction with MTS methods.[29] The calculated diffusion coefficient of liquid water at 300 K is determined to be 2.67 and 2.29 $\times$ 10^{-5} cm^2/s from classical and quantum simulations respectively. The inclusion of NQEs decreases the diffusion coefficient by 30%. [29] Del Ben and co-workers combined a hybrid functional (PBEW1- D3) with MP2, using fast (0.25 fs) and slow (2.5 fs) time steps to calculate the dynamical property of MP2 water.[72] Their analysis is based on two trajectories of each roughly 10 ps as obtained in the NVE ensemble, and the self-diffusion constant values obtained are 0.67 and 0.77 $\times$ 10^{-5} cm^2/s at 300 and 307 K, respectively. At room temperature, our simulations predict a diffusion constant of 0.629 (0.693) and 0.996 (1.060) $\times$ 10^{-5} cm^2/s from classical and quantum dynamics (the quantities in brackets are after finite-size correction[73]), which is in fair agreement with previous simulations[72]. In fact, a classical MD simulation with a limited timescale of 10-20 ps may yield statistical error bars of more than 20% on dynamical quantities such as the diffusion coefficient. That may explain the minor inconsistency between our values and previous results.[29, 72] Nevertheless, these values are below the experimental diffusion at room temperature, and more similar to a lower-temperature diffusion. In Fig.3, we plot the temperature-dependent diffusion coefficient as obtained from classical and quantum dynamics. The diffusion constants are systematically underestimated by MP2 theory for both classical and quantum simulations. When the experimental diffusion constant is shifted to a higher temperature by 40 K, the diffusion constant as calculated from our simulations matches with experimental values. Previous studies show that NQEs contribute a 30% decrease to the self-diffusion of liquid water using revPBE0-D3 functional and a similar slowdown has also been observed using revPBE-D3 functional.[29] Unlike the revPBE0-D3 functional, the diffusion of water upon including NQEs are enhanced by 10-70% from 240 to 320 K, while a slow down of diffusion by 3-7% have been observed beyond 330 K.

To understand the vibrational dynamics of MP2 water, we calculate the density of states (VDOS) from the Fourier Transform of the velocity-velocity auto-correlation function at different temperatures. As shown in Fig. 4, classical MD simulations overestimate the OH-stretch mode and bend mode by around 300 cm^{-1} and 40 $^{-1}$ compared to experiment at room temperature (with the experimental OH stretch at 3380 cm). In quantum simulations, the O–H stretching frequency is red-

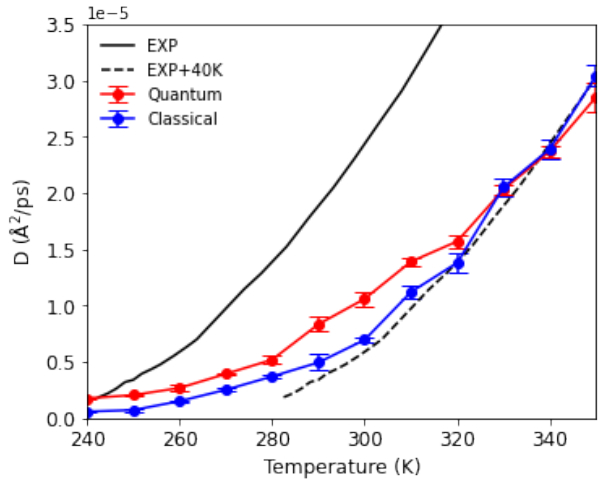

FIG. 3. Temperature dependence of the diffusion coefficient after finte size correction as obtained from classical and quantum simulations, in comparison to the experimental values (black line)[74] and its shifted value by 40 K.

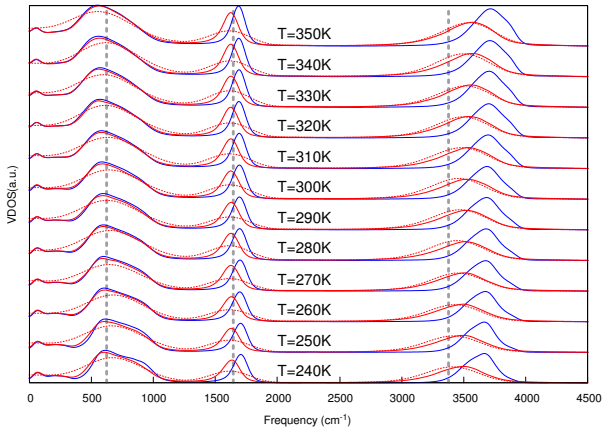

FIG. 4. Temperature dependent vibrational density of states from 240 K to 350 K. The blue lines are obtained from classical simulations, while the red lines are from TRPMD, the red dashed-lines are from PACMD simulations. Vertical lines show experimental vibrational frequencies at 300K.

shifted from 3685 cm^{-1} to 3524 cm^{-1} (from TRPMD) and 3442 cm^{-1} (PACMD), with the latter agreeing better with the experiment. [75]

NQEs also contribute to the improvement of the bending mode (experimentally 1637 cm^{-1} in the IR measurement) from 1693 (classical) to 1624 (quantum). While the agreement of MP2 calculations with experiments is very good, that of revPBE0-D3 water is somewhat better, particularly for the O–H stretching peak [35]. The PACMD results for the librational and bending modes in water are reasonably good across the range of temperatures studied, but the agreement between the experimental and PACMD O–H stretching modes improves as the temperature is *decreased*, with quantum results at 240 K giving the best agreement with experiments.[75]

IV. CONCLUSIONS

In this work, we simulated MP2 water using an NNP at different temperatures using classical and quantum dynamics. MP2 theory, though belonging to the fifth rung of the electronic structure "Jacob's ladder", is unable to predict accurately the static and dynamical properties of liquid water. This may be because an insufficiently large basis set causes MP2 theory to overestimate the Van der Waals interaction. This may be improved upon by increasing the basis set to quadruple-ζ and larger, or by estimating the difference between the results of small and large basis sets using cluster models and decomposing the error into two- and three-body interaction terms, which can be used to further cancel the basis-set error. Alternatively, double-hybrid density functionals based on MP2 and random-phase approximation may be used to train the NNP: they exhibit faster convergence with basis set size at similar computational cost as the parent correlated methods [76].

Although nuclear quantum effects reduce the height of the first peak of the oxygen-oxygen radial distribution function at room temperature, the match with the experimental RDF is still not perfect.[72] However, a better agreement with experimental RDFs is observed at an elevated temperature of 340 K, at which MP2 predicts stratifying static properties of liquid water compared to the experimental data, as characterized by the RDFs, triplet angles distribution, and tetrahedral local order parameters. It is possible that the elevated temperature cancels the error due to an insufficient basis set, higher kinetic energy compensating for the overestimation of dispersion by MP2 theory.

Although our NNP accurately captures a range of static and dynamical properties, many properties of water such as the dielectric constant and vibrational spectra are not yet available. Future work will focus on the modelling of these properties using symmetry-adapted machine learning methods [77]. Other future prospects include clarifying the basis set effect on the computed properties, probing the performance of double-hybrid density functionals, and extending simulations to other important simple liquids such as ammonia and hydrogen sulfide.

ACKOWLEDGEMENTS

This work is funded by the Swiss National Science Foundation (SNSF) Sinergia grant and the University Research Priority Program (URPP) for solar light to chemical energy conversion (LightChEC) of the University of Zurich. This work was also supported by a grant from the Swiss National Supercomputing Centre (CSCS) under Project ID uzh1, s1043. VRR has been supported by the SNSF in the form of Ambizione grant No. PZ00P2_174227.

DATA AVAILABILITY

Data generated and analyzed for this study that are not included in this article are available at https://www.materialscloud.org.

[1] L. G. M. Pettersson, R. H. Henchman, and A. Nilsson, Chemical reviews **116**, 7459 (2016).

[2] P. Gallo, K. Amann-Winkel, C. A. Angell, M. A. Anisimov, F. Caupin, C. Chakravarty, E. Lascaris, T. Loerting, A. Z. Panagiotopoulos, J. Russo, *et al.*, Chemical reviews **116**, 7463 (2016).

[3] G. A. Cisneros, K. T. Wikfeldt, L. Ojamäe, J. Lu, Y. Xu, H. Torabifard, A. P. Bartók, G. Csányi, V. Molinero, and F. Paesani, Chemical reviews **116**, 7501 (2016).

[4] B. Cheng, M. Bethkenhagen, C. J. Pickard, and S. Hamel, arXiv preprint arXiv:2103.09035 (2021).

[5] C. J. Mundy and I.-F. W. Kuo, Chemical reviews **106**, 1282 (2006).

[6] M. Karplus and J. A. McCammon, Nature structural biology **9**, 646 (2002).

[7] W. Yang, R. R. Prabhakar, J. Tan, S. D. Tilley, and J. Moon, Chemical Society Reviews **48**, 4979 (2019).

[8] A. Rahman and F. H. Stillinger, The Journal of Chemical Physics **55**, 3336 (1971).

[9] J. A. Morrone and R. Car, Physical review letters **101**, 017801 (2008).

[10] K. Laasonen, M. Sprik, M. Parrinello, and R. Car, The Journal of chemical physics **99**, 9080 (1993).

[11] R. W. Hockney and J. W. Eastwood, *Computer simulation using particles* (crc Press, 2021).

[12] M. W. Mahoney and W. L. Jorgensen, The Journal of Chemical Physics **114**, 363 (2001).

[13] H. W. Horn, W. C. Swope, J. W. Pitera, J. D. Madura, T. J. Dick, G. L. Hura, and T. Head-Gordon, The Journal of chemical physics **120**, 9665 (2004).

[14] H. Berendsen, J. Grigera, and T. Straatsma, Journal of Physical Chemistry **91**, 6269 (1987).

[15] H. J. Berendsen, J. P. Postma, W. F. van Gunsteren, and J. Hermans, in *Intermolecular forces* (Springer, 1981) pp. 331–342.

[16] V. Babin, G. R. Medders, and F. Paesani, Journal of chemical theory and computation **10**, 1599 (2014).

[17] M. J. Gillan, D. Alfe, and A. Michaelides, The Journal of chemical physics **144**, 130901 (2016).

[18] E. Liberatore, R. Meli, and U. Rothlisberger, Journal of chemical theory and computation **14**, 2834 (2018).

[19] S. Yoo, X. C. Zeng, and S. S. Xantheas, The Journal of chemical physics **130**, 221102 (2009).

[20] A. Bankura, A. Karmakar, V. Carnevale, A. Chandra, and M. L. Klein, The Journal of Physical Chemistry C **118**, 29401 (2014).

[21] K. Forster-Tonigold and A. Groß, The Journal of chemical physics **141**, 064501 (2014).

[22] S. Yoo and S. S. Xantheas, The Journal of chemical physics **134**, 121105 (2011).

[23] V. V. Rybkin, Chemistry – A European Journal **26**, 362 (2019).

[24] J. Sun, A. Ruzsinszky, and J. P. Perdew, Physical review letters **115**, 036402 (2015).

[25] M. Chen, H.-Y. Ko, R. C. Remsing, M. F. C. Andrade, B. Santra, Z. Sun, A. Selloni, R. Car, M. L. Klein, J. P. Perdew, *et al.*, Proceedings of the National Academy of Sciences **114**, 10846 (2017).

[26] S. Dasgupta, E. Lambros, J. P. Perdew, and F. Paesani, Nature communications **12**, 1 (2021).

[27] J. Wiktor, F. Ambrosio, and A. Pasquarello, The Journal of chemical physics **147**, 216101 (2017).

[28] L. R. Pestana, N. Mardirossian, M. Head-Gordon, and T. Head-Gordon, Chemical science **8**, 3554 (2017).

[29] O. Marsalek and T. E. Markland, The journal of physical chemistry letters **8**, 1545 (2017).

[30] B. Cheng, E. A. Engel, J. Behler, C. Dellago, and M. Ceriotti, Proceedings of the National Academy of Sciences **116**, 1110 (2019).

[31] F. Ambrosio, G. Miceli, and A. Pasquarello, The Journal of Physical Chemistry B **120**, 7456 (2016).

[32] M. Del Ben, M. Schönherr, J. Hutter, and J. VandeVondele, The journal of physical chemistry letters **4**, 3753 (2013).

[33] S. Y. Willow, X. C. Zeng, S. S. Xantheas, K. S. Kim, and S. Hirata, The journal of physical chemistry letters **7**, 680 (2016).

[34] M. Ceriotti, W. Fang, P. G. Kusalik, R. H. McKenzie, A. Michaelides, M. A. Morales, and T. E. Markland, Chemical reviews **116**, 7529 (2016).

[35] T. E. Markland and M. Ceriotti, Nature Reviews Chemistry **2**, 1 (2018).

[36] C. Li, F. Paesani, and G. A. Voth, ChemRxiv (2021), 10.33774/chemrxiv-2021-c603x.

[37] J. Lan, V. V. Rybkin, and M. Iannuzzi, The journal of physical chemistry letters **11**, 3724 (2020).

[38] V. V. Rybkin and J. VandeVondele, The Journal of Physical Chemistry Letters **8**, 1424 (2017), pMID: 28296416, https://doi.org/10.1021/acs.jpclett.7b00386.

[39] M. Parrinello and A. Rahman, The Journal of chemical physics **80**, 860 (1984).

[40] J. Behler and M. Parrinello, Physical Review Letters **98**, 146401 (2007).

[41] L. Zhang, J. Han, H. Wang, R. Car, and E. Weinan, Physical review letters **120**, 143001 (2018).

[42] S. Shepherd, J. Lan, D. M. Wilkins, and V. Kapil, arXiv preprint arXiv:2108.03056 (2021).

[43] V. Kapil, D. M. Wilkins, J. Lan, and M. Ceriotti, The Journal of chemical physics **152**, 124104 (2020).

[44] Y. Yao and Y. Kanai, The journal of physical chemistry letters **12**, 6354 (2021).

[45] J. Lan, V. Kapil, P. Gasparotto, M. Ceriotti, M. Iannuzzi, and V. V. Rybkin, Nature communications **12**, 1 (2021).

[46] V. Kapil, M. Rossi, O. Marsalek, R. Petraglia, Y. Litman, T. Spura, B. Cheng, A. Cuzzocrea, R. H. Meißner, D. M. Wilkins, *et al.*, Computer Physics Communications **236**, 214 (2019).

[47] M. Ceriotti, J. More, and D. E. Manolopoulos, Computer Physics Communications **185**, 1019 (2014).

[48] S. Plimpton, Journal of computational physics **117**, 1 (1995).

[49] A. Singraber, J. Behler, and C. Dellago, Journal of Chemical Theory and Computation **15**, 1827 (2019).

[50] G. Bussi, D. Donadio, and M. Parrinello, The Journal of Chemical Physics **126**, 014101 (2007).

[51] V. Kapil, A. Cuzzocrea, and M. Ceriotti, The Journal of Physical Chemistry B **122**, 6048 (2018).

[52] M. Rossi, M. Ceriotti, and D. E. Manolopoulos, The Journal of Chemical Physics **140**, 234116 (2014).

[53] R. L. Benson, G. Trenins, and S. C. Althorpe, Faraday Discussions **221**, 350 (2019).

[54] T. D. Hone, P. J. Rossky, and G. A. Voth, J. Chem. Phys. **124**, 154103 (2006).

[55] M. Ceriotti, M. Parrinello, T. E. Markland, and D. E. Manolopoulos, The Journal of chemical physics **133**, 124104 (2010).

[56] S. Habershon, G. S. Fanourgakis, and D. E. Manolopoulos, The Journal of chemical physics **129**, 074501 (2008).

[57] A. Dehaoui, B. Issenmann, and F. Caupin, Proceedings of the National Academy of Sciences **112**, 12020 (2015).

[58] G. Imbalzano, A. Anelli, D. Giofré, S. Klees, J. Behler, and M. Ceriotti, The Journal of Chemical Physics **148**, 241730 (2018).

[59] M. Del Ben, J. Hutter, and J. VandeVondele, The Journal of Chemical Physics **143**, 102803 (2015).

[60] V. V. Rybkin and J. VandeVondele, Journal of Chemical Theory and Computation **12**, 2214 (2016).

[61] M. Del Ben, J. Hutter, and J. VandeVondele, Journal of Chemical Theory and Computation **9**, 2654 (2013).

[62] J. Hutter, M. Iannuzzi, F. Schiffmann, and J. VandeVondele, Wiley Interdisciplinary Reviews: Computational Molecular Science **4**, 15 (2014).

[63] T. D. Kühne, M. Iannuzzi, M. Del Ben, V. V. Rybkin, P. Seewald, F. Stein, T. Laino, R. Z. Khaliullin, O. Schütt, F. Schiffmann, D. Golze, J. Wilhelm, S. Chulkov, M. H. Bani-Hashemian, V. Weber, U. Borštnik, M. Taillefumier, A. S. Jakobovits, A. Lazzaro, H. Pabst, T. Müller, R. Schade, M. Guidon, S. Andermatt, N. Holmberg, G. K. Schenter, A. Hehn, A. Bussy, F. Belleflamme, G. Tabacchi, A. Glöß, M. Lass, I. Bethune, C. J. Mundy, C. Plessl, M. Watkins, J. VandeVondele, M. Krack, and J. Hutter, The Journal of Chemical Physics **152**, 194103 (2020).

[64] L. B. Skinner, C. Huang, D. Schlesinger, L. G. Pettersson, A. Nilsson, and C. J. Benmore, The Journal of chemical physics **138**, 074506 (2013).

[65] A. Soper and C. Benmore, Physical review letters **101**, 065502 (2008).

[66] J. Kim and K. S. Kim, The Journal of chemical physics **109**, 5886 (1998).

[67] S. S. Xantheas, C. J. Burnham, and R. J. Harrison, The Journal of chemical physics **116**, 1493 (2002).

[68] S. S. Xantheas and E. Aprà, The Journal of chemical physics **120**, 823 (2004).

[69] B. Temelso, K. A. Archer, and G. C. Shields, The Journal of Physical Chemistry A **115**, 12034 (2011).

[70] E. Miliordos, E. Aprà, and S. S. Xantheas, The Journal of chemical physics **139**, 114302 (2013).

[71] J. Liu, X. He, J. Z. Zhang, and L.-W. Qi, Chemical science **9**, 2065 (2018).

[72] M. Del Ben, J. Hutter, and J. VandeVondele, The Journal of chemical physics **143**, 054506 (2015).

[73] I.-C. Yeh and G. Hummer, The Journal of Physical Chemistry B **108**, 15873 (2004).

[74] P. Tofts, D. Lloyd, C. Clark, G. Barker, G. Parker, P. McConville, C. Baldock, and J. Pope, Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine **43**, 368 (2000).

[75] M. W. Chase Jr, J. Phys. Chem. Ref. Data, Monograph **9** (1998).

[76] F. Stein, J. Hutter, and V. V. Rybkin, Molecules **25** (2020), 10.3390/molecules25215174.

[77] A. Grisafi, D. M. Wilkins, G. Csányi, and M. Ceriotti, Phys. Rev. Lett. **120**, 036002 (2018).

[78] C. Herrero, G. Tocci, S. Merabia, and L. Joly, Nanoscale **12**, 20396 (2020).

[79] L. Joly, R. H. Meißner, M. Iannuzzi, and G. Tocci, ACS Nano (2021), 10.1021/acsnano.1c05931, pMID: 34491721, https://doi.org/10.1021/acsnano.1c05931.

[80] J. VandeVondele and J. Hutter, The Journal of Chemical Physics **127**, 114105 (2007).

[81] I. R. Craig and D. E. Manolopoulos, The Journal of Chemical Physics **121**, 3368 (2004).

[82] W. Wagner and A. Pruß, Journal of physical and chemical reference data **31**, 387 (2002).

[83] J. Schmidt, J. VandeVondele, I.-F. W. Kuo, D. Sebastiani, J. I. Siepmann, J. Hutter, and C. J. Mundy, The Journal of Physical Chemistry B **113**, 11959 (2009).

[84] V. Daggett, Chemical reviews **106**, 1898 (2006).

[85] M. G. Walter, E. L. Warren, J. R. McKone, S. W. Boettcher, Q. Mi, E. A. Santori, and N. S. Lewis, Chemical reviews **110**, 6446 (2010).