

Supporting Information: Nuclear Quantum Effects on the Electronic Structure of Water and Ice

Margaret Berrens,[†] Arpan Kundu,[‡] Marcos F. Calegari Andrade,[¶] Tuan Anh Pham,[¶] Giulia Galli,^{‡,§,||} and Davide Donadio^{*,†}

[†]*Department of Chemistry, University of California Davis, One Shields Ave. Davis, CA, 95616.*

[‡]*Pritzker School of Molecular Engineering, University of Chicago, Chicago, IL 60637*

[¶]*Quantum Simulations Group, Materials Science Division, Lawrence Livermore National Laboratory, Livermore, California, 94550-5507*

[§]*Department of Chemistry, University of Chicago, Chicago, Illinois 60637, United States*

^{||}*Materials Science Division and Center for Molecular Engineering, Argonne National Laboratory, Lemont, Illinois 60439, United States*

E-mail: ddonadio@ucdavis.edu

Computational Methods

We ran classical MD simulations of a 64-molecule bulk liquid water system at 300 K and a 96-molecule bulk hexagonal ice system at 230 K. The bulk liquid water system had a cubic simulation box (12.45 Å) and the bulk hexagonal ice system had an orthorhombic simulation box that measures $13.500 \times 15.5 \times 14.8$ Å³ (2x2x2 supercell). To generate our proton-disordered Ice I_h system we used the GenIce¹ which ensures the generation of completely

randomized hydrogen-disordered networks obeying the ice rules. Both systems had periodic boundary conditions (PBC) along all three dimensions. The equations of motion were integrated with the velocity Verlet algorithm with a time step of 0.5 fs and the temperature was controlled by stochastic velocity rescaling² with a relaxation time of 1 ps. Classical MD simulations using the DNNP were performed using the combination of the DEEPMD-KIT^{3,4} and the LAMMPS package.⁵ Classical MD simulations using the NEP were performed using the GPUMD package.⁶ The systems were equilibrated for 10 ps followed by a production run of 100 ps. The path-integral MD simulations were carried out on the same systems with the same time step and simulation length but instead were accelerated with a generalized Langevin equation thermostat (PIGLET),⁷ to allow for convergence on the number of beads, such that only 8 beads were used for the simulation. The convergence in the number of beads used can be seen in Figure S2. The PIGLET simulations with the DNNP used the i-PI driver⁸ with the DEEPMD-KIT with LAMMPS. The PIGLET simulations with the NEP used the i-PI driver⁸ with the NEP3 interface⁹ with LAMMPS. For the PIGLET simulations of the 192 molecule I_h system, the simulations were only carried out for 20 ps. The parameters of the colored noise thermostat were downloaded from the GLE4MD web interface and are included in the online repository uploaded to Materials Cloud.¹⁰

For the calculations of the electronic properties of the bulk water and bulk hexagonal ice at the DFT and G_0W_0 level. DFT calculations, including IPR calculations, were carried out using the Qbox code,¹¹ with a plane-wave energy cutoff of 85 Ry, and norm-conserving pseudopotentials.¹² G_0W_0 calculations were performed using the West code,¹³ starting from wavefunctions obtained with the PBE¹⁴ functionals. Quasiparticle energies were determined using the same exchange-correlation potential as the one employed in the self-consistent-field DFT calculations. All of our G_0W_0 calculations were carried out with 1000 eigenpotentials. DFT calculations were done for 100 equally spaced snapshots along each bulk trajectory. G_0W_0 calculations were performed for 50 snapshots and the convergence on the number of frames sampled can be seen in Figure S3. The sole Γ point is sampled throughout the

MD simulations. The CP2K 8.2 package,¹⁵ was used for electronic structure calculations employing the van der Waals corrected hybrid functional (revPBE0)^{16–18} for the water and ice systems. Valence Kohn–Sham orbitals are expanded on a triple- ζ localized basis set¹⁹ in real space, and core states are treated implicitly using Goedecker–Teter–Hutter pseudopotentials.²⁰ Plane waves up to a cutoff energy of 400 Ry are used as a basis set for the density in reciprocal space.

Comparison of MLPs

A noteworthy difference between the two functionals that the MLPs were trained on is that the use of revPBE0-D3 predicts a slightly higher frequency for the O-H stretching vibration of water,²¹ whereas SCAN is found to match the vibrational spectra quite well. This shift to higher frequency is also reproduced by the two MLPs used in this study (Figure S1). However, it was shown that including NQEs, for *ab initio* MD simulations with revPBE0-D3, leads to an excellent agreement in the vibrational spectra.²² It is also important to note that the SCAN functional has been found to overstructure liquid water while revPBE0-D3 does a satisfactory job of modeling the structure of liquid water.^{23,24}

Supporting Information Figures

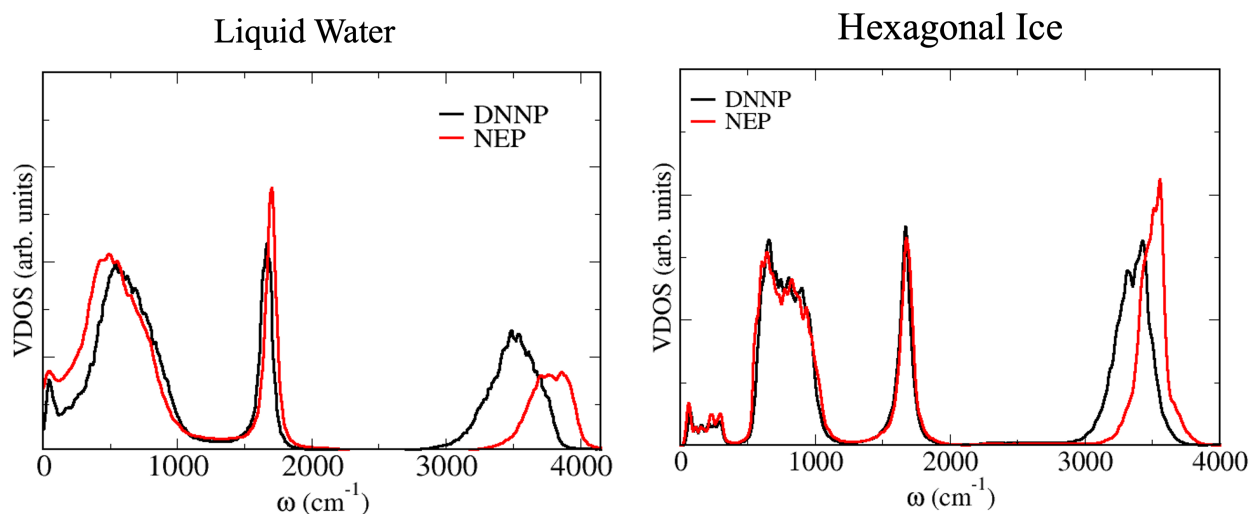


Figure S1: Vibrational density of states of water and ice using both sampled MLPs (DNNP and NEP) calculated over a 10 ps MD run, with velocities sampled every 0.5 fs.

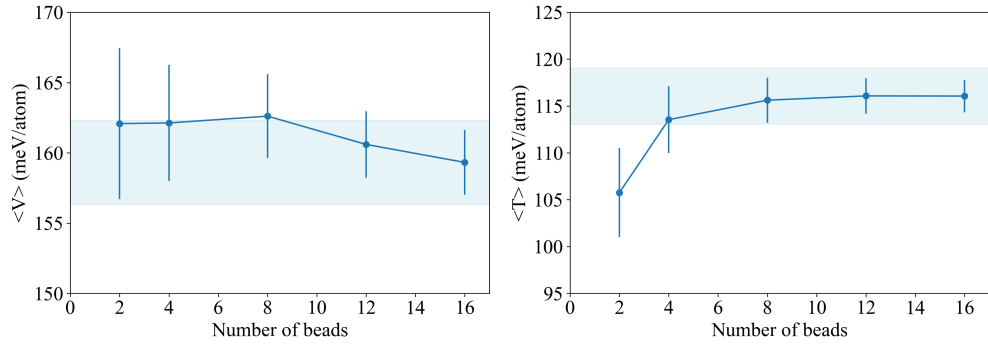


Figure S2: Convergence of the vibrational potential and kinetic energies (quantum estimates) of hexagonal ice with respect to the number of beads used in PIGLET simulations. It can be seen that 8 beads reproduce the results within 3 meV/atom (region highlighted in light blue) of the 16 bead results. Energies for each bead are averaged over a 1 ps run with a 0.1 fs time step.

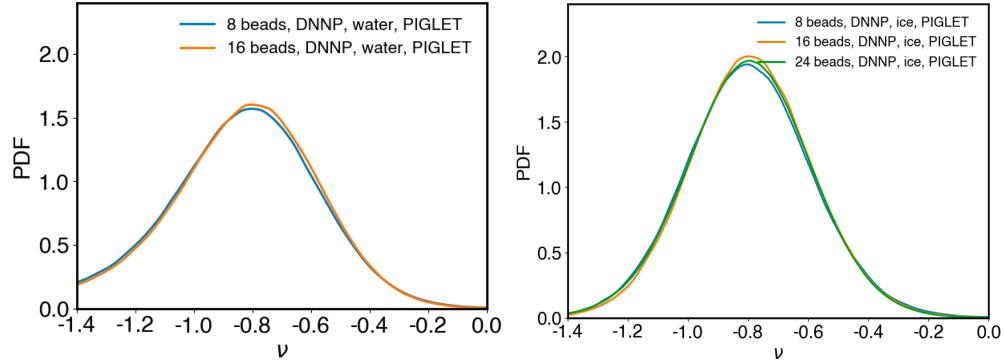


Figure S3: Convergence of the proton transfer coordinate probability distribution for both water and Ice I_h for number of beads. Probability distributions are calculated over 100 ps simulations for each quantum simulation with the respective number of beads. The distributions with 8 beads are well converged with the 16 and 24 bead distributions

Table S1: Convergence check of VBM, CBM, and gap values with corresponding standard deviations calculated for liquid water and Ice I_h calculated using the DNNP@SCAN

MLP used	Electronic Structure Method	Type of MD simulation	VBM energy value (eV)	CBM energy value (eV)	Gap energy value (eV)
DNNP (Liquid Water)	DFT-revPBE0-aug-TZV2P	PIGLET (8 beads)	-4.09 ± 0.21	2.82 ± 0.010	6.91 ± 0.24
DNNP (Liquid Water)	DFT-revPBE0-aug-TZV2P	PIGLET (16 beads)	-4.07 ± 0.22	2.80 ± 0.11	6.87 ± 0.24
DNNP (Ice I_h)	DFT-revPBE0-aug-TZV2P	PIGLET (8 beads)	-4.32 ± 0.22	2.72 ± 0.060	7.04 ± 0.24
DNNP (Ice I_h)	DFT-revPBE0-aug-TZV2P	PIGLET (16 beads)	-4.35 ± 0.19	2.75 ± 0.044	7.09 ± 0.20
DNNP (Ice I_h)	DFT-revPBE0-aug-TZV2P	PIGLET (24 beads)	-4.34 ± 0.21	2.74 ± 0.044	7.08 ± 0.22

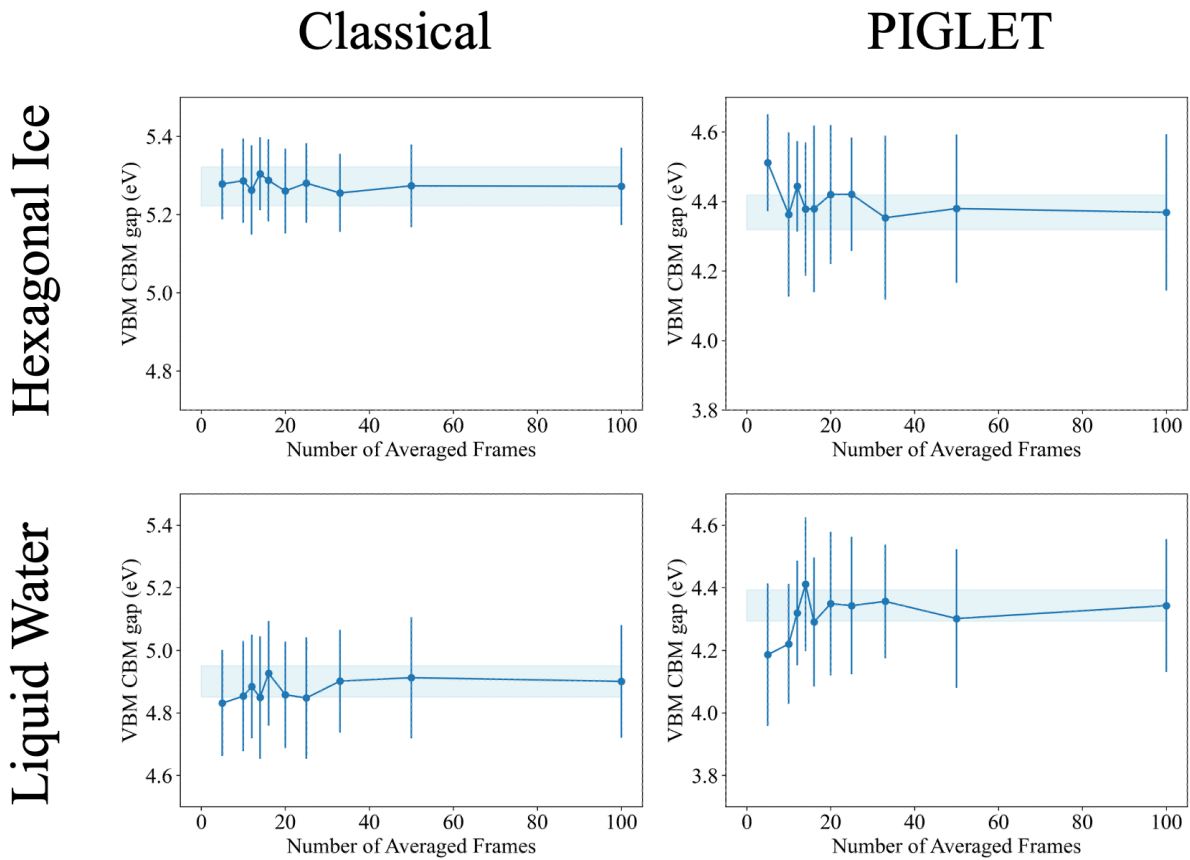


Figure S4: Convergence of the average VBM CBM gap as a function of the number of frames. It can be seen that 50 frames practically reproduce the 100-frame results within 50 meV (region highlighted in light blue). Thus, 50 frames were chosen to be averaged over for MBPT calculations

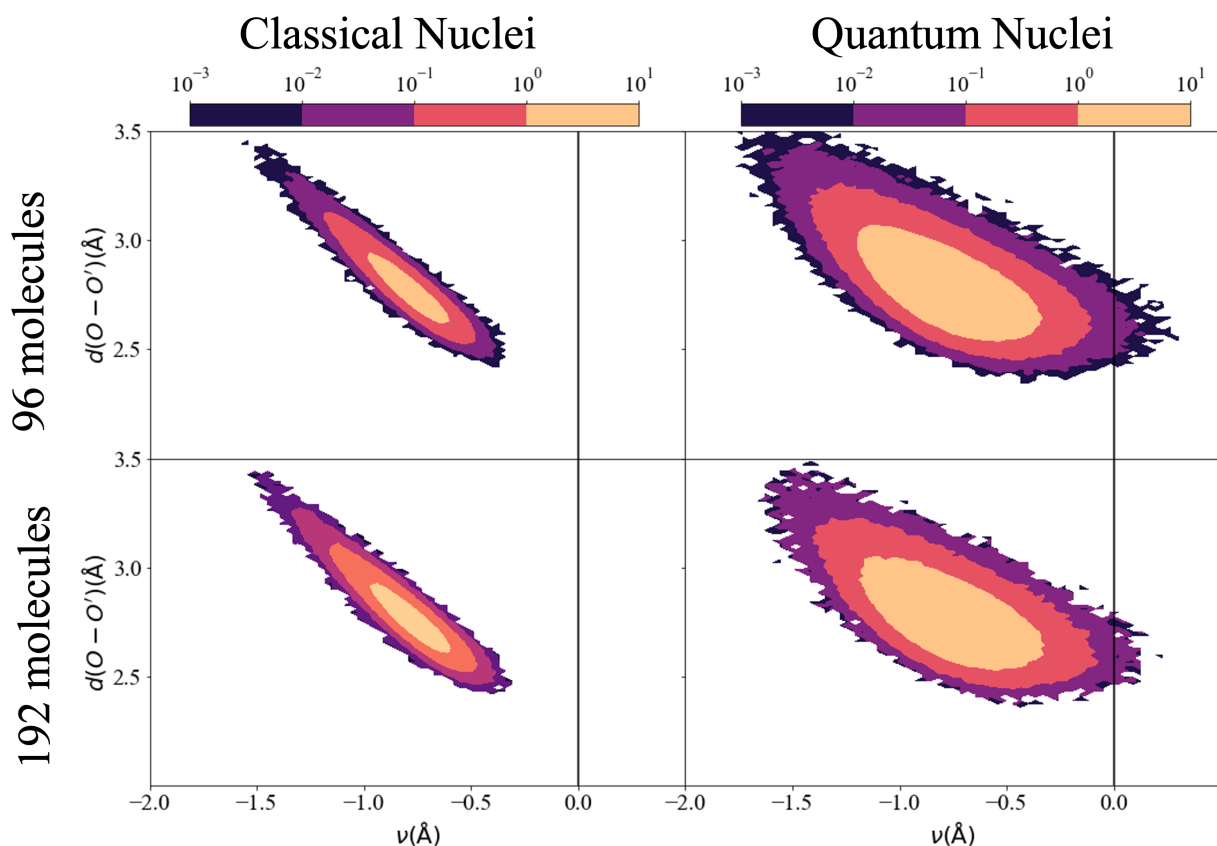


Figure S5: Joint probability distribution of the proton-transfer coordinate ν and the distance $d(\text{O}-\text{O}')$ between the covalently bound and acceptor oxygen atoms for 96 H_2O molecules of hexagonal ice (top) and 192 H_2O molecules hexagonal ice (bottom), with frames extracted from 100 equally spaced frames from 100 ps DNNP classical and quantum MD runs. MD runs for the 192-molecule system utilized the same parameters as described in the computational methods. For both sized systems of bulk ice, the classical and quantum joint probability distributions are in agreement.

Table S2: VBM, CBM, and gap values with corresponding standard deviations calculated for 96 and 192 water molecules for hexagonal ice for classical and quantum simulations. MD runs for the 192-molecule system utilized the same parameters as described in the computational methods. Between the two supercells, VB and CB energies are within 0.1 eV of each other for both classical and quantum systems.

Size of Structure	Electronic Structure Method	Type of MD simulation	VBM energy value (eV)	CBM energy value (eV)	Gap energy value (eV)
96 H ₂ O molecules	DFT/SCAN/85R _y	Classical (100 frames)	-10.45 ± 0.10	-5.18 ± 0.013	5.28 ± 0.10
96 H ₂ O molecules	DFT/SCAN/85R _y	PIGLET (100 frames)	-9.79 ± 0.21 (NQE = +0.66)	-5.42 ± 0.046 (NQE = -0.24)	4.37 ± 0.23 (NQE = -0.90 ± 0.25)
192 H ₂ O molecules	DFT/SCAN/60R _y	Classical (30 frames)	-10.40 ± 0.095	-5.20 ± 0.020	5.18 ± 0.095
192 H ₂ O molecules	DFT/SCAN/60R _y	PIGLET (30 frames)	-9.67 ± 0.15 (NQE = +0.73)	-5.47 ± 0.043 (NQE = -0.27)	4.20 ± 0.16 (NQE = -0.98 ± 0.19)
96 H ₂ O molecules	DNNP/SCAN/85R _y	One energy minimized frame	-10.95	-5.06	5.90
192 H ₂ O molecules	DNNP/SCAN/60R _y	One energy minimized frame	-10.98	-5.07	5.91
436 H ₂ O molecules	DNNP/SCAN/60R _y	One energy minimized frame	-10.94	-5.05	5.90

Table S3: VBM, CBM, and gap values with corresponding standard deviations calculated for liquid water for classical and quantum simulations for both MLPs

MLP used	Electronic Structure Method	Type of MD simulation	VBM energy value (eV)	CBM energy value (eV)	Gap energy value (eV)
DNNP	DFT-SCAN	Classical	-10.46 ± 0.17	-5.56 ± 0.043	4.90 ± 0.18
DNNP	DFT-SCAN	PIGLET	-10.05 ± 0.21 (NQE = +0.40)	-5.71 ± 0.058 (NQE = -0.15)	4.34 ± 0.21 (NQE = -0.56±0.28)
DNNP	G ₀ W ₀ -PBE	Classical	-5.12 ± 0.19	3.48 ± 0.025	8.60 ± 0.19
DNNP	G ₀ W ₀ -PBE	PIGLET	-4.49 ± 0.25 (NQE = +0.63)	3.30 ± 0.05 (NQE = -0.18)	7.79 ± 0.27 (NQE = -0.81± 0.33)
NEP	DFT-SCAN	Classical	-10.39 ± 0.16	-5.72 ± 0.037	4.66 ± 0.17
NEP	DFT-SCAN	PIGLET	-10.11 ± 0.25 (NQE = +0.28)	-5.87 ± 0.064 (NQE = -0.14)	4.24 ± 0.27 (NQE = -0.45±0.32)
NEP	G ₀ W ₀ -PBE	Classical	-5.13 ± 0.18	3.39 ± 0.028	8.52 ± 0.18
NEP	G ₀ W ₀ -PBE	PIGLET	-4.60 ± 0.38 (NQE = +0.53)	3.24 ± 0.057 (NQE = -0.15)	7.84 ± 0.38 (NQE = -0.69±0.42)

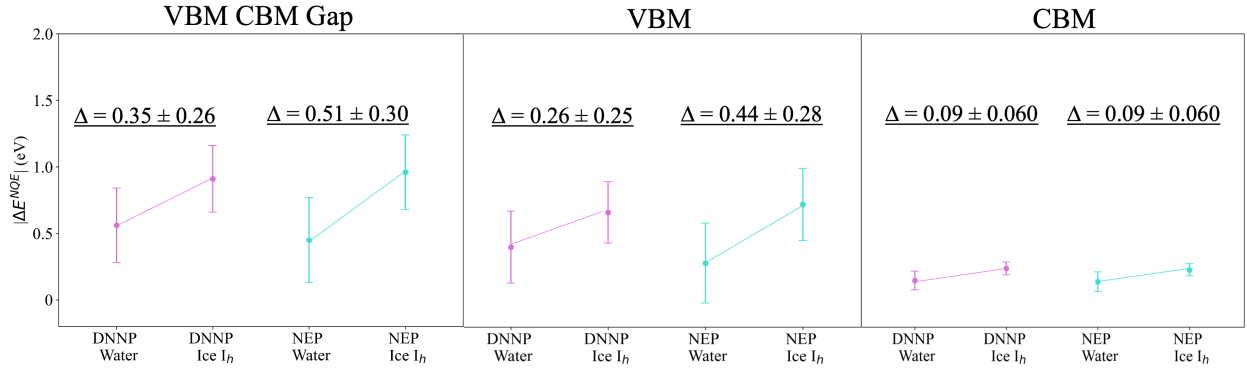


Figure S6: The ΔE_g^{NQE} of the VBM-CBM gap, VBM, and CBM for liquid water and hexagonal ice at the DFT level using the SCAN functional for 100 frames sampled from the classical MD and quantum path-integral simulations using two different machine learning models DNNP and NEP.

Table S4: VBM, CBM, and gap values calculated using DFT (SCAN) and MBPT with corresponding standard deviations calculated for hexagonal ice for classical and quantum simulations for both MLPs.

MLP used	Electronic Structure Method	Type of MD simulation	VBM energy value (eV)	CBM energy value (eV)	Gap energy value (eV)
DNNP	DFT-SCAN	Classical	-10.45 ± 0.10	-5.18 ± 0.013	5.28 ± 0.10
DNNP	DFT-SCAN	PIGLET	-9.79 ± 0.21 (NQE = +0.66)	-5.42 ± 0.046 (NQE = -0.24)	4.37 ± 0.23 (NQE = -0.91 ± 0.25)
DNNP	G ₀ W ₀ -PBE	Classical	-5.68 ± 0.091	3.43 ± 0.011	9.11 ± 0.18
DNNP	G ₀ W ₀ -PBE	PIGLET	-4.76 ± 0.25 (NQE = +0.93)	3.22 ± 0.040 (NQE = -0.21)	7.98 ± 0.27 (NQE = -1.13 ± 0.32)
NEP	DFT-SCAN	Classical	-10.51 ± 0.086	-5.16 ± 0.011	5.35 ± 0.090
NEP	DFT-SCAN	PIGLET	-9.79 ± 0.26 (NQE = +0.72)	-5.39 ± 0.044 (NQE = -0.23)	4.39 ± 0.27 (NQE = -0.96 ± 0.28)
NEP	G ₀ W ₀ -PBE	Classical	-5.82 ± 0.10	3.43 ± 0.0076	9.24 ± 0.10
NEP	G ₀ W ₀ -PBE	PIGLET	-4.78 ± 0.35 (NQE = +1.04)	3.25 ± 0.031 (NQE = -0.18)	8.04 ± 0.36 (NQE = -1.20 ± 0.37)
DNNP optimized	DNNP-SCAN	One energy minimized frame	-10.95	-5.06	5.90
DNNP optimized	G ₀ W ₀ -PBE	One energy minimized frame	-6.19	3.51	9.70

Table S5: VBM, CBM, and gap values with corresponding standard deviations calculated for liquid water and I_h for classical and quantum simulations for comparing the DNNP trained on MB-Pol simulation data to the other two MLPs used.

MLP used	Electronic Structure Method	Type of MD simulation	VBM energy value (eV)	CBM energy value (eV)	Gap energy value (eV)
DNNP -MB-Pol (Liquid Water)	DFT-revPBE0-aug-TZV2P	Classical	-4.53 ± 0.19	2.85 ± 0.061	7.37 ± 0.20
DNNP -MB-pol (Liquid Water)	DFT-revPBE0-aug-TZV2P	PIGLET	-4.25 ± 0.21 (NQE = +0.28)	2.55 ± 0.091 (NQE = -0.30)	6.79 ± 0.25 (NQE = -0.58±0.32)
DNNP -SCAN (Liquid Water)	DFT-revPBE0-aug-TZV2P	Classical	-4.54 ± 0.17	3.05 ± 0.061	7.59 ± 0.19
DNNP -SCAN (Liquid Water)	DFT-revPBE0-aug-TZV2P	PIGLET	-4.090 ± 0.21 (NQE = +0.45)	2.82 ± 0.10 (NQE = -0.23)	6.91 ± 0.24 (NQE = -0.68±0.32)
NEP (Liquid Water)	DFT-revPBE0-aug-TZV2P	Classical	-4.49 ± 0.17	2.80 ± 0.052	7.29 ± 0.18
NEP (Liquid Water)	DFT-revPBE0-aug-TZV2P	PIGLET	-4.20 ± 0.23 (NQE = +0.29)	2.64 ± 0.080 (NQE = -0.16)	6.85 ± 0.26 (NQE = -0.44±0.32)
DNN -MB-Pol (I_h)	DFT-revPBE0-aug-TZV2P	Classical	-5.04 ± 0.10	3.01 ± 0.019	8.05 ± 0.11
DNNP -MB-Pol (I_h)	DFT-revPBE0-aug-TZV2P	PIGLET	-4.34 ± 0.26 (NQE = +0.70)	2.71 ± 0.060 (NQE = -0.30)	7.06 ± 0.29 (NQE = -0.99 ± 0.31)
DNN -SCAN (I_h)	DFT-revPBE0-aug-TZV2P	Classical	-5.02 ± 0.10	3.02 ± 0.016	8.04 ± 0.11
DNNP -SCAN (I_h)	DFT-revPBE0-aug-TZV2P	PIGLET	-4.32 ± 0.22 (NQE = +0.70)	2.72 ± 0.06 (NQE = -0.30)	7.04 ± 0.24 (NQE = -1.00± 0.26)
NEP (I_h)	DFT-revPBE0-aug-TZV2P	Classical	-5.08 ± 0.095	3.04 ± 0.029	8.12 ± 0.11
NEP (I_h)	DFT-revPBE0-aug-TZV2P	PIGLET	-4.31 ± 0.29 (NQE = +0.77)	2.75 ± 0.056 (NQE = -0.29)	7.06 ± 0.30 (NQE = -1.06±0.32)
DNNP optimized	DFT-revPBE0-aug-TZV2P	One energy minimized frame	-5.57	3.14	8.71

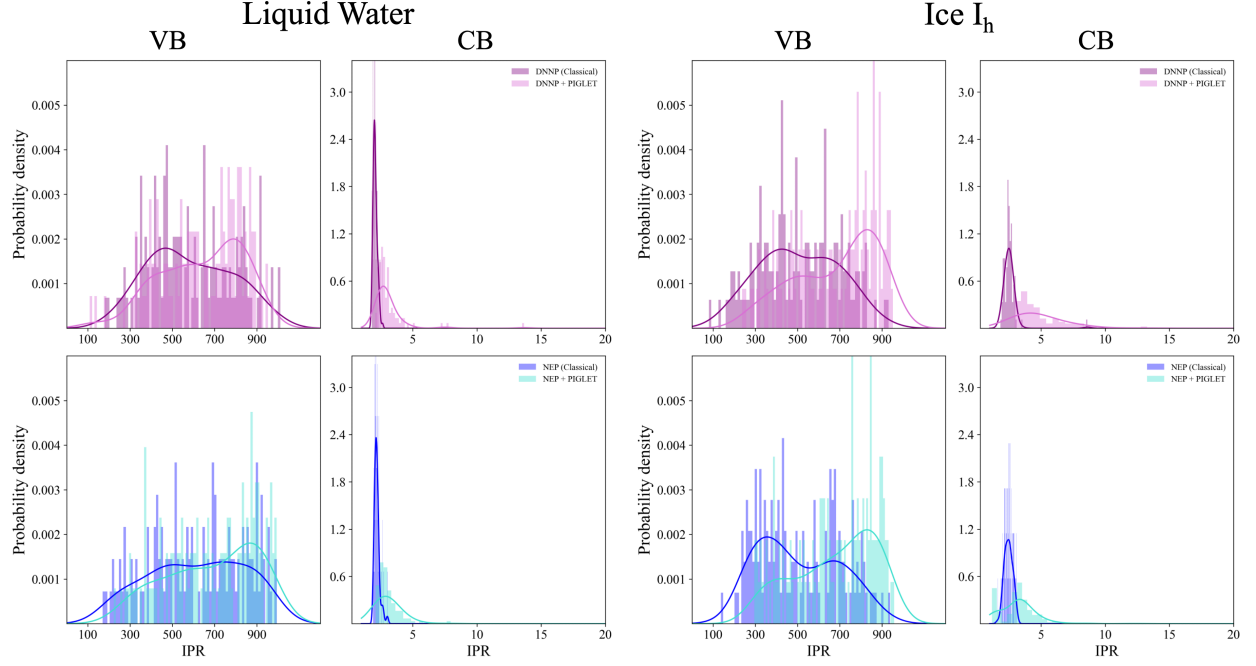


Figure S7: Histogram of inverse participation ratio (IPR) values calculated for 100 extracted frames for Kohn-Sham orbitals with energy values within 50 meV of the VB (valence band) and CB (conduction band) respectively, using 75 bins. The Gaussian kernel density estimation (KDE) for each histogram is shown as solid lines. Histograms are generated for water and ice using classical (MD) and quantum (PIGLET) simulations employing NEP and DNNP. The IPR values are scaled by 10000.

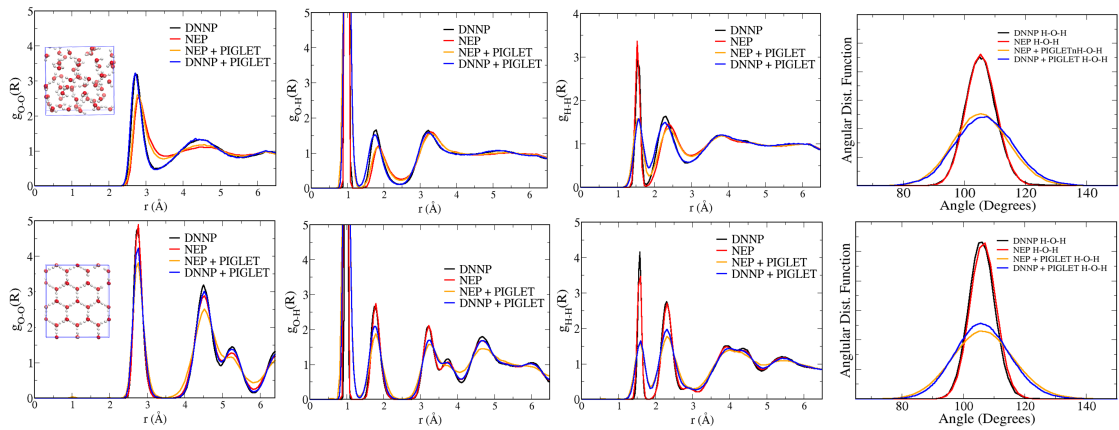


Figure S8: Radial distribution functions (g_{OO} , g_{OH} , g_{HH}) and angular distribution function ($H-O-H$) of water (top row) and ice (bottom row) for classical and quantum (PIGLET) MD simulations for both sampled MLPs.

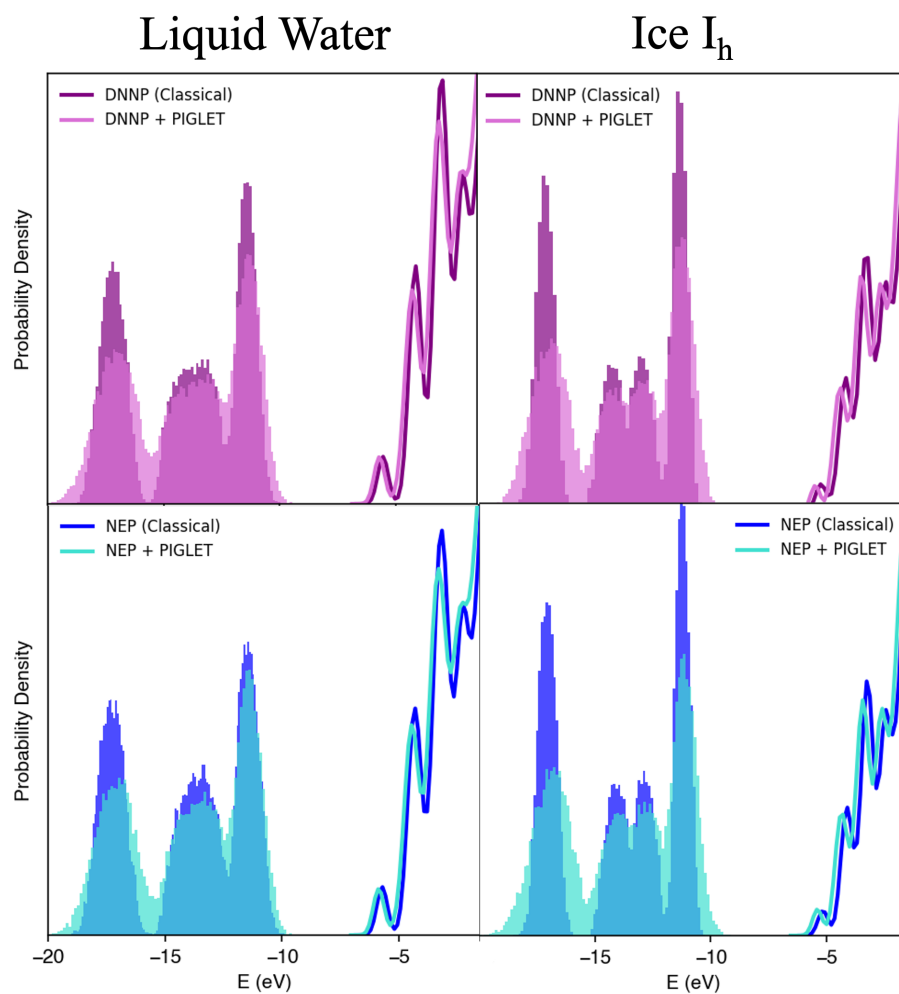


Figure S9: Electronic density of states (EDOS) of water and ice calculated employing DFT (SCAN) calculations using snapshots extracted from DNNP (NEP) trajectories obtained with classical and quantum (PIGLET) MD simulations.

References

- (1) Matsumoto, M.; Yagasaki, T.; Tanaka, H. GenIce: Hydrogen-Disordered Ice Generator. *J. Comp. Chem.* **2018**, *39*, 61–64.
- (2) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.
- (3) Wang, H.; Zhang, L.; Han, J.; E, W. DeePMD-kit: A deep learning package for many-body potential energy representation and molecular dynamics. *Comput. Phys. Commun.* **2018**, *228*, 178–184.
- (4) Zeng, J. et al. DeePMD-kit v2: A software package for deep potential models. *J Chem Phys* **2023**, *159*, 054801.
- (5) Thompson, A. P.; Aktulga, H. M.; Berger, R.; Bolintineanu, D. S.; Brown, W. M.; Crozier, P. S.; in 't Veld, P. J.; Kohlmeyer, A.; Moore, S. G.; Nguyen, T. D.; Shan, R.; Stevens, M. J.; Tranchida, J.; Trott, C.; Plimpton, S. J. LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Comput. Phys. Commun.* **2022**, *271*, 108171.
- (6) Fan, Z. Improving the accuracy of the neuroevolution machine learning potential for multi-component systems. *J. Phys.: Condens. Matter* **2022**, *34*, 125902.
- (7) Ceriotti, M.; Manolopoulos, D. E. Efficient First-Principles Calculation of the Quantum Kinetic Energy and Momentum Distribution of Nuclei. *Phys. Rev. Lett.* **2012**, *109*, 100604.
- (8) Kapil, V. et al. i-PI 2.0: A universal force engine for advanced molecular simulations. *Comput. Phys. Commun.* **2019**, *236*, 214–223.
- (9) Fan, Z. et al. GPUMD: A package for constructing accurate machine-learned poten-

- tials and performing highly efficient atomistic simulations. *J. Chem. Phys.* **2022**, *157*, 114801.
- (10) Talirz, L. et al. Materials Cloud, a platform for open computational science. *Sci Data* **2020**, *7*, 299, Publisher: Nature Publishing Group.
 - (11) Gygi, F. Architecture of Qbox: A scalable first-principles molecular dynamics code. *IBM J. Res. & Dev.* **2008**, *52*, 137–144.
 - (12) Schlipf, M.; Gygi, F. Optimization algorithm for the generation of ONCV pseudopotentials. *Comput. Phys. Commun.* **2015**, *196*, 36–44.
 - (13) Govoni, M.; Galli, G. Large Scale GW Calculations. *J. Chem. Theory Comput.* **2015**, *11*, 2680–2696.
 - (14) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
 - (15) VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach. *Comp. Phys. Commun.* **2005**, *167*, 103–128.
 - (16) Perdew, J. P.; Ernzerhof, M.; Burke, K. Rationale for mixing exact exchange with density functional approximations. *J. Chem. Phys.* **1996**, *105*, 9982–9985.
 - (17) Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110*, 6158–6170.
 - (18) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J Chem Phys* **2010**, *132*, 154104.
 - (19) VandeVondele, J.; Hutter, J. Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases. *J. Chem. Phys.* **2007**, *127*, 114105.

- (20) Goedecker, S.; Teter, M.; Hutter, J. Separable dual-space Gaussian pseudopotentials. *Phys. Rev. B* **1996**, *54*, 1703.
- (21) Zhong, K.; Yu, C.-C.; Dodia, M.; Bonn, M.; Nagata, Y.; Ohto, T. Vibrational mode frequency correction of liquid water in density functional theory molecular dynamics simulations with van der Waals correction. *Phys. Chem. Chem. Phys.* **2020**, *22*, 12785–12793.
- (22) Marsalek, O.; Markland, T. E. Quantum Dynamics and Spectroscopy of Ab Initio Liquid Water: The Interplay of Nuclear and Electronic Quantum Effects. *J. Phys. Chem. Lett.* **2017**, *8*, 1545–1551.
- (23) Yao, Y.; Kanai, Y. Temperature dependence of nuclear quantum effects on liquid water via artificial neural network model based on SCAN meta-GGA functional. *J. Chem. Phys.* **2020**, *153*, 044114.
- (24) Li, C.; Paesani, F.; Voth, G. A. Static and Dynamic Correlations in Water: Comparison of Classical Ab Initio Molecular Dynamics at Elevated Temperature With Path Integral Simulations at Ambient Temperature. *J Chem Theory Comput* **2022**, *18*, 2124–2131.