

# Ab initio simulation of particle momentum distributions in high-pressure water

M. Ceriotti<sup>1</sup>

<sup>1</sup>Laboratory of Computational Science and Modelling, NQ Institute of Materials, École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland

E-mail: [michele.ceriotti@epfl.ch](mailto:michele.ceriotti@epfl.ch)

## Abstract.

Applying pressure to water reduces the average oxygen-oxygen distance, and facilitates the delocalisation of protons along the hydrogen bond. This pressure-induced delocalisation is further enhanced by the quantum nature of hydrogen nuclei, which is very significant even well above room temperature. Here we will evaluate the quantum kinetic energy and the particle momentum distribution of hydrogen and oxygen nuclei in water at extreme pressure, using ab initio path integral molecular dynamics. We will show that (transient) dissociation of water molecules induce measurable changes in the kinetic energy hydrogen atoms, although current deep inelastic scattering experiments are probably unable to capture the heterogeneity of the sample.

Treating the nuclei as classical particles is an approximation which is widely made in atomistic computer simulations, even in cases in which the electronic degrees of freedom are treated quantum mechanically by first-principles methods. Whenever light nuclei such as hydrogen are present, however, their quantum nature will have a sizable effect on physical observables even well above room temperature, that must be considered if one wants to capture quantitatively the properties of the system.

A typical example in which nuclear quantum effects (NQE) cannot be neglected is the evaluation of the particle momentum distribution of light nuclei, that deviates dramatically from the Maxwell-Boltzmann predictions and that can be measured in deep inelastic neutron scattering (DINS) experiments [1]. The standard method to describe the quantum nature of nuclei in simulations is based on the path integral formalism [2–5], which however involves a very significant computational overhead compared to a classical simulation, making it very challenging to include NQEs in *ab initio* simulations. Recently it has been shown that correlated-noise Langevin dynamics can be used to model NQEs inexpensively [6, 7], and that a combination of path integral molecular dynamics and colored noise (PIGLET) dramatically reduces the cost of attaining quantitative accuracy [8, 9].

Here, such a combination of path integral methods and generalized Langevin dynamics will be used to evaluate the particle momentum distribution (PMD) of oxygen and hydrogen nuclei in water at 10GPa and 750K. In this region of the phase diagram the mean O–O distance is considerably reduced relative to ambient temperature and pressure, increasing the probability of water dissociation to a point at which a significant fraction of auto-ionisation events can be observed within the limited time accessible to first-principles simulations [10]. Quantum effects facilitate dissociation even further [11], and it is interesting to evaluate the impact of



autoionisation on the hydrogen and oxygen mean kinetic energy. After having reported the estimates that can be obtained by computer simulations, it will be discussed whether and how it may be possible to measure these effects by DINS experiments.

## 1. Computational details

Simulations were performed using i-PI, a Python wrapper that makes it possible to do path integral and PIGLET simulations using any electronic structure code as a client to compute the Born-Oppenheimer energy and forces [11]. In this case, CP2K was used as the back-end [12, 13].

The supercell contained 64 water molecules, equilibrated at the desired thermodynamic conditions using an empirical forcefield [14]. Then, 25 ps of *NVT* equilibration with classical nuclei were performed with *ab initio* forces, using the BLYP exchange-correlation functional [15, 16] and GTH pseudopotentials [17]. Wave functions were expanded in the Gaussian DZVP basis set, while the electronic density was represented using an auxiliary plane wave basis, with a kinetic energy cutoff of 300 Ry.

Then, constant-pressure production simulations were performed using the barostatting procedure described in [11], using a more converged TZV2P basis set, a density cutoff of 800 Ry, and a plane wave grid consistent with a cubic reference cell with a side of 10.5 Å (the higher cutoff was necessary to converge the stress tensor). Empirical Van der Waals corrections, which are essential to obtain a reasonable description of water at constant pressure [18], were included based on the D3 scheme [19–21]. The quantum nature of nuclei was modelled using a 4-bead ring polymer, supplemented with PIGLET colored noise. The use of colored noise makes it possible to converge simulation of water at room temperature using as little as six path integral replicas [9], and so one can expect that at this much higher temperature a slightly reduced number of replicas will be sufficient for convergence of quantum properties. The cell momentum was thermostatted using an optimal-sampling, classical GLE thermostat. A time step of 0.5fs was used, and a total of 50ps of PIGLET dynamics were performed to accumulate statistical averages.

NQEs enhance greatly the delocalisation of protons along the hydrogen bond, so that even in room-temperature neutral water there is a sizeable probability of observing transient autoionization events where a hydrogen atom momentarily visits configurations in which it is closer to the acceptor oxygen atom than to the oxygen it is covalently bound to [22]. This makes it somewhat problematic to identify charged  $\text{H}_3\text{O}^+$  and  $\text{OH}^-$  species based only on structural information, so here we use a simple-minded approach to identify configurations that are significantly distorted relative to a neutral water molecule environment, without trying to distinguish between transient excursions and genuine autoionization.

Following Refs. [10, 11], we attributed to each oxygen atom the hydrogens that are within its Voronoi polyhedron (i.e. those that are closest to it than to any other oxygen), and labelled as neutral species the oxygen atoms with two hydrogen partners, positive ‘+’ species the oxygen atoms to which three H have been assigned, and ‘-’ species the oxygen atoms to which a single proton belongs. We also refer to  $\text{H}^+$  as the hydrogens that are assigned to a  $\text{O}^+$ ,  $\text{H}^-$  as the hydrogens that are assigned to a  $\text{O}^-$ , and to just H as the protons that are attached to a ‘neutral’ oxygen. According to this definition, simulations using PIGLET contain about 1% of  $\text{O}^+$  and  $\text{O}^-$  species, almost 20 times as much as simulations in the same conditions that ignore NQEs [11].

## 2. Quantum kinetic energies and particle momentum distributions

Atomistic computer simulations make it possible to interrogate with great detail the properties of the system being studied. For instance, one can compute conditional averages that evaluate physical observables for only a subset of the atoms. In particular, one can ask what is the kinetic energy of hydrogen and oxygen nuclei as a function of the formal charge that has been

assigned based on the procedure outlined above. What is more, one can evaluate for each of the species the transient anisotropic Gaussian (TAG) approximation of the particle momentum distribution [9], giving even further insight into the difference between environments.

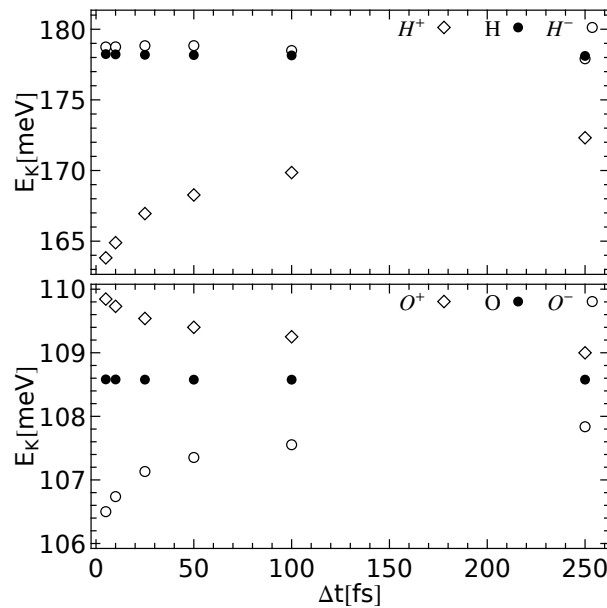
The TAG approximation evaluates an approximant of the kinetic energy tensor for a particle at a given time by performing a moving average of the centroid-virial kinetic tensor estimator  $T_{\alpha\beta}$  over a triangular window of width  $\Delta t$ :

$$T_{\alpha\beta} = \frac{\delta_{\alpha\beta}}{2\beta} + \frac{1}{4P} \sum_i \left[ (q_{i\alpha} - \bar{q}_\alpha) \frac{\partial V}{\partial q_{i\beta}} + (q_{i\beta} - \bar{q}_\beta) \frac{\partial V}{\partial q_{i\alpha}} \right];$$

$$T_{\alpha\beta}(t; \Delta t) = \frac{1}{\Delta t} \int_{t-\Delta t}^{t+\Delta t} T_{\alpha\beta}(t') \left( 1 - \frac{|t-t'|}{\Delta t} \right) dt'. \quad (1)$$

Here,  $P$  is the number of path integral replicas,  $q_{i\alpha}$  identifies the  $\alpha$  Cartesian coordinate of the tagged particle in the  $i$ -th replica,  $\bar{\mathbf{q}}$  is the position of the centroid of the particle,  $V$  is the physical potential and  $\beta$  the inverse temperature. One can compute the three principal components of  $\mathbf{T}(t; \Delta t)$ ,  $E_\alpha$ , and average them over the course of the trajectory. This procedure gives only an approximation of the three components of the mean kinetic energy tensor: a too short window means that the average picks up an instantaneous value of the centroid virial estimator (which is not physically meaningful), while a too large window would average over different orientations of the molecule and tend towards giving an isotropic tensor. By conditionally averaging over particles that are in a given formal charge state at time  $t$  one can evaluate a TAG approximation to the kinetic energy resolved between  $\text{O}^+$ ,  $\text{O}$ ,  $\text{O}^-$  and  $\text{H}^+$ ,  $\text{H}$ ,  $\text{H}^-$ .

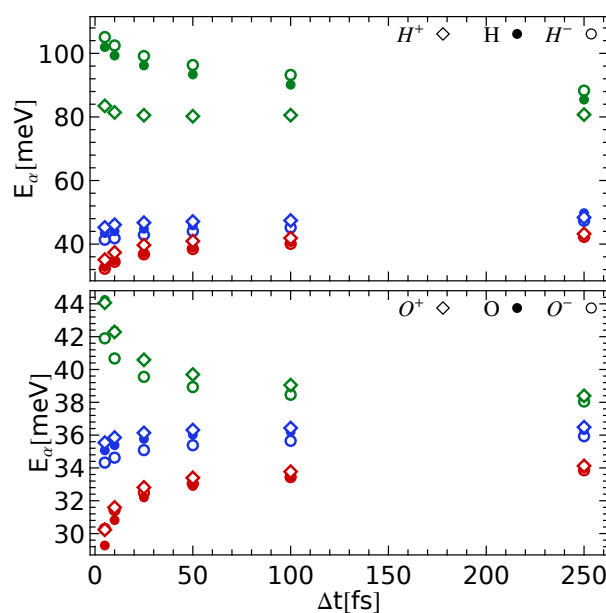
**Figure 1.** Total mean kinetic energy of hydrogen (top) and oxygen (bottom) atoms. Moving averages are performed with a time window of amplitude  $\Delta t$ , and conditionally based on the charge state of each atom at the centre of the moving window.  $\text{O}^+$ ,  $\text{O}$ ,  $\text{O}^-$  and  $\text{H}^+$ ,  $\text{H}$ ,  $\text{H}^-$  species are indicated by different symbols, as indicated in the legend.



One should stress that using the TAG approximation in this context is stretching considerably its range of applicability. For one thing, since the charged defects have a fluxional, rapidly changing nature, when averaging over larger  $\Delta t$  one often picks up configurations that are not

in the same formal charge state as the central frame that determines the conditional averaging. This means that as  $\Delta t$  increases the *total* mean kinetic energies (that are not affected by the TAG procedure if the average is performed over all atoms of the same element) of the different charged species become less and less different from each other (see Figure 1).

**Figure 2.** Principal components of the mean kinetic energy of hydrogen (top) and oxygen (bottom) atoms, as computed within the TAG approximation with a time window of amplitude  $\Delta t$ . Averages are performed conditionally based on the charge state of each atom at the centre of the moving window.  $O^+$ ,  $O$ ,  $O^-$  and  $H^+$ ,  $H$ ,  $H^-$  species are indicated by different symbols, as indicated in the legend, and the three components of the kinetic energy tensor are drawn in red, green and blue.



Furthermore, at 750K the orientational relaxation of water is much faster than at room temperature, and there is no obvious plateau between the  $\Delta t$  that is needed to converge the mean centroid virial estimator and the  $\Delta t$  at which the rotation of water molecules averages out the instantaneous environment felt by protons (see Figure 2).

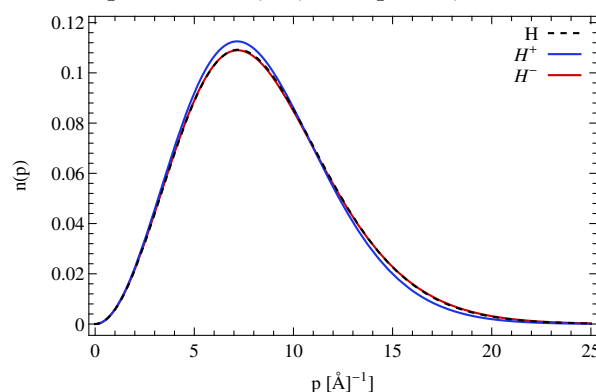
Nevertheless, use of the TAG approximation is necessary here, as the alternative approach of defining molecular axes along which to perform a kinetic energy decomposition is not viable for the charged species, and since an open path calculation would be exceedingly hard to converge, given that the approximation of opening one path per molecule [23] is hard to justify when there is a significant degree of proton exchange between molecules. Let us take rather arbitrarily  $\Delta t = 50$ fs as the averaging window for which we will discuss results. In the absence of a clear plateau, this seems a reasonable compromise between the conflicting goals of averaging the unphysical instantaneous values of the centroid virial estimator and the need to capture a rapidly changing instantaneous environment.

A first observation is that even at 750K atoms in water are far from a classical behaviour. At this temperature the classical mean kinetic energy per atom would have a value of 97 meV. Oxygen atoms have 10% higher  $E_K$ , and hydrogen atoms have nearly double  $E_K$  compared to the classical predictions. This is mostly due to the zero-point energy in the stretching mode. Differences between charged species is noticeable but small, amounting to a few meV for most components. The only large difference is observed for  $H^+$  species, that have a weakened stretching mode resulting in a drop of more than 15 meV in the largest principal component

of the kinetic energy tensor. As it is often the case in water [24–26], there is a degree of cancellation between the changes of the different components of the kinetic energy, so that the overall kinetic energy change between H and  $H^+$  is reduced to about 11 meV when evaluated with the same moving window of 50fs.  $H^-$  does not show large differences to “neutral” H atoms, with the components  $E_\alpha$  changing by less than 3 meV, and a near-perfect cancellation when the difference in the total  $E_K$  is considered.

Note that due to the low concentration of charged species, the kinetic energy components averaged over *all* the H or O atoms is almost indistinguishable from the  $E_\alpha$ s for the dominant neutral H and O species. This points at an inherent limitation of DINS experiments, that shall be discussed in more detail in the next section.

**Figure 3.** Particle momentum distribution approximated as the spherical average of a multivariate Gaussian, with the principal components obtained from the TAG approximation with  $\Delta t = 50$ fs. The three lines correspond to  $H^+$ , H,  $H^-$  species, as indicated in the legend.



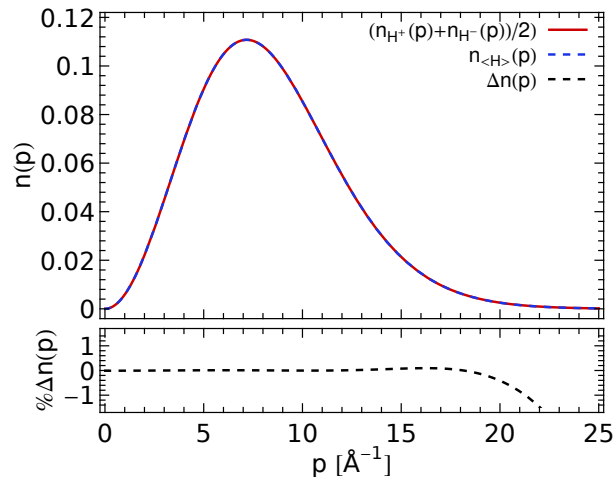
### 3. Detecting heterogeneity in DINS experiments

Neutron Compton scattering is perhaps the most direct probe of the quantum nature of nuclei in the condensed phase, fully accounting for anharmonicity [27], and having been used to probe the environment of protons in a variety of different materials [28–30]. In the present case it is clear that the system – albeit homogeneous on a macroscopic scale – contains a variety of different local environments, that in the specific case of the charged  $H^+$  species imply a sizeable change in the kinetic energy tensor compared to the H species.

Fig. 3 shows that the particle momentum distribution of  $H^+$ , taken separately, would differ significantly from that of the “neutral” hydrogen atoms, by an amount sufficient to be unambiguously detected by DINS. As we discussed before, the small concentration of charged species means that the signal would be probably too weak to be detected against that of the majority species. This lead us to ask ourselves, more in general, whether DINS can provide information on the heterogeneity of a material, a question that is relevant to the interpretation of experimental results and to the comparison with simulations.

Let us consider for example the most favourable case of a hypothetical sample containing 50% of  $H^-$  and 50% of  $H^+$  species. Let us assume that the overall PMD arising from the sample will then be the arithmetic mean of the multivariate-Gaussian TAG approximant of the two species as obtained in our simulations. Figure 4 shows that one can obtain a near-perfect fit (undistinguishable within the current experimental uncertainty) to this average distribution using a *single* multivariate Gaussian model. This observation highlights one problem of DINS experiments – namely that spherical averaging washes away much of the information content of

**Figure 4.** The top panel shows the mean between the multivariate-Gaussian particle momentum distributions of  $H^+$  and  $H^-$  species, and the distribution obtained by fitting a single multivariate Gaussian model to the former curve. The relative discrepancy between the two curves is also plotted in the lower panel.



the particle momentum distribution, so that it is almost impossible to single out contributions from different environments in the sample, and just a “mean field” description can be obtained.

#### 4. Outlook

The present work shows that the quantum nature of light nuclei is still very pronounced even much above room temperature, and that the quantum kinetic energy of particles varies significantly between nuclei that are experiencing different environments. At the same time, it highlights limitations of the current approaches to simulate and measure NQEs in the condensed phase. From the modelling point of view, it shows the limits of a recently-developed approximation to estimate a Gaussian model of the particle momentum distribution based on an “instantaneous” kinetic energy tensor. While this TAG approximation makes it possible to obtain the PMD inexpensively, it can only provide qualitative information, so an approach to make conventional open path simulations more affordable is much needed.

On the experimental side, this paper points at a significant limitation of current DINS techniques – namely the fact that it is practically impossible to extract information on the heterogeneity of the sample, as the PMD arising from a very inhomogeneous sample would be undistinguishable from that of a homogeneous system made up of an effective, mean environment. A possible approach that we can suggest to extract some site-specific information would be to exploit the different propensity of hydrogen isotopes to occupy different sites, and to perform measurements with different isotope tagging. In systems where proton exchange is minimal, this would allow one to precisely measure the PMD of specific sites. In systems such as water in which proton mixing happens much more rapidly than a measurement can be performed, it might still be possible to infer some insight on the heterogeneity of the material by comparing measurements performed at different isotope concentrations.

#### Acknowledgments

We thank Greg Voth for suggesting to analyse the output of these simulations to compute the particle momentum distribution, David Manolopoulos for suggestions and insightful discussion, and acknowledge generous allocations of computer time from the Swiss National Supercomputing

Centre (project ID s338 and s466).

## References

- [1] Andreani C, Colognesi D, Mayers J, Reiter G F and Senesi R 2005 *Adv. Phys.* **54** 377–469
- [2] Feynman R P and Hibbs A R 1964 *Quantum Mechanics and Path Integrals* (New York: McGraw-Hill)
- [3] Chandler D and Wolynes P G 1981 *J. Chem. Phys.* **74** 4078–4095
- [4] Parrinello M and Rahman A 1984 *J. Chem. Phys.* **80** 860
- [5] Ceperley D M 1995 *Rev. Mod. Phys.* **67** 279–355
- [6] Ceriotti M, Bussi G and Parrinello M 2009 *Phys. Rev. Lett.* **103** 30603
- [7] Dammak H, Chalopin Y, Laroche M, Hayoun M and Greffet J J 2009 *Phys. Rev. Lett.* **103** 190601
- [8] Ceriotti M, Manolopoulos D E and Parrinello M 2011 *J. Chem. Phys.* **134** 84104
- [9] Ceriotti M and Manolopoulos D E 2012 *Phys. Rev. Lett.* **109** 100604
- [10] Schwegler E, Galli G, Gygi F and Hood R 2001 *Phys. Rev. Lett.* **87** 265501
- [11] Ceriotti M, More J and Manolopoulos D E 2014 *Comput. Phys. Commun.* **185** 1019–1026
- [12] CP2K <http://www.cp2k.org>
- [13] VandeVondele J, Krack M, Mohamed F, Parrinello M, Chassaing T and Hutter J 2005 *Comput. Phys. Commun.* **167** 103–128
- [14] Habershon S, Markland T E and Manolopoulos D E 2009 *J. Chem. Phys.* **131** 24501
- [15] Becke A D 1988 *Phys. Rev. A* **38** 3098
- [16] Lee C, Yang W and Parr R G 1988 *Phys. Rev. B* **37** 785
- [17] Goedecker S, Teter M and Hutter J 1996 *Phys. Rev. B* **54** 1703–1710
- [18] Schmidt J, VandeVondele J, Kuo I F W, Sebastiani D, Siepmann J I, Hutter J and Mundy C J 2009 *J. Phys. Chem. B* **113** 11959–64
- [19] Grimme S, Antony J, Ehrlich S and Krieg H 2010 *J. Chem. Phys.* **132** 154104
- [20] Guidon M, Hutter J and VandeVondele J 2009 *J. Chem. Theory Comput.* **5** 3010–3021
- [21] Guidon M, Hutter J and VandeVondele J 2010 *J. Chem. Theory Comput.* **6** 2348–2364
- [22] Ceriotti M, Cuny J, Parrinello M and Manolopoulos D E 2013 *Proc. Natl. Acad. Sci. USA* **110** 15591–6
- [23] Morrone J A, Srinivasan V, Sebastiani D and Car R 2007 *J. Chem. Phys.* **126** 234504
- [24] Li X Z, Walker B and Michaelides A 2011 *Proc. Natl. Acad. Sci. USA* **108** 6369–6373
- [25] Markland T E and Berne B J 2012 *Proc. Natl. Acad. Sci. USA* **109** 7988–7991
- [26] Romanelli G, Ceriotti M, Manolopoulos D E, Pantalei C, Senesi R and Andreani C 2013 *J. Phys. Chem. Lett.* **4** 3251–3256
- [27] Mayers J, Reiter G F and Platzman P 2002 *J. Mol. Struct.* **615** 275–282
- [28] Andreani C, Degiorgi E, Senesi R, Cilloco F, Colognesi D, Mayers J, Nardone M and Pace E 2001 *J. Chem. Phys.* **114** 387–398
- [29] Reiter G F, Senesi R and Mayers J 2010 *Phys. Rev. Lett.* **105** 148101
- [30] Reiter G F, Kolesnikov I, Paddison S J, Platzman P M, Moravsky P, Adams M A and Mayers J 2012 *Phys. Rev. B* **85** 045403