**Title**
Family of Oxygen–Oxygen Radial Distribution Functions for Water

**Permalink**
https://escholarship.org/uc/item/0cg0r43p

**Journal**
The Journal of Physical Chemistry Letters, 6(15)

**ISSN**
1948-7185

**Authors**
Brookes, David H
Head-Gordon, Teresa

**Publication Date**
2015-08-06

**DOI**
10.1021/acs.jpclett.5b01066

Peer reviewed

# The Family of Oxygen-Oxygen Radial Distribution Functions for Water

**David H. Brookes[2] and Teresa Head-Gordon[1,2,3,4*]**

*[1]Department of Chemistry, [2]Department of Bioengineering, [3]Department of Chemical and Biomolecular Engineering, [4]Chemical Sciences Division, Lawrence Berkeley National Laboratory, University of California, Berkeley, CA 94720*

**Abstract**

In a typical X-ray diffraction experiment, the elastically scattered intensity, $I(Q)$, is the experimental observable. $I(Q)$ contains contributions from both intramolecular as well as intermolecular correlations embodied in the scattering factors $H_{OO}(Q)$ and $H_{OH}(Q)$, with negligible contributions from $H_{HH}(Q)$. Thus in order to accurately define the oxygen-oxygen radial distribution function $g_{OO}(r)$, a model of the electron density is required to accurately weight the $H_{OO}(Q)$ component relative to the intramolecular and oxygen-hydrogen correlations from the total intensity observable. In this work we carefully define the electron density model, and its underlying assumptions, and more explicitly utilize two restraints on the allowable $g_{OO}(r)$ functions which must conform to both very low experimental errors at high-$Q$ and the need to satisfy the isothermal compressibility at low-$Q$. Although highly restrained by these conditions, the underdetermined nature of the problem is such that we present a *family* of $g_{OO}(r)$'s that provide equally good agreement with the high-$Q$ intensity and compressibility restraints and with physically correct behavior at small $r$.

**\*Corresponding author**
**thg@berkeley.edu**

TOC
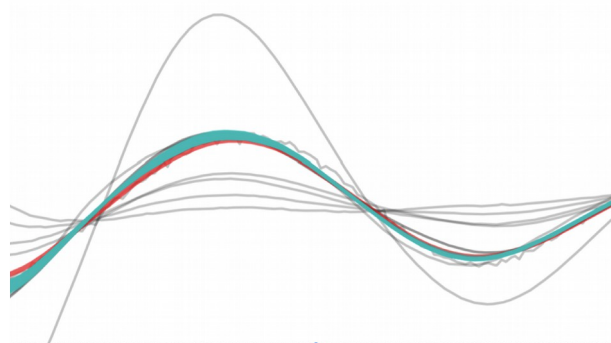
There has been significant progress in the acquisition of high quality X-ray intensity data for liquid water obtained at synchrotron sources over the last 15 years[1-9]. Over the period of 2000-2003, an X-ray scattering experiment was obtained at a synchrotron beam line at the Advanced Light Source (ALS) at Lawrence Berkeley National Laboratory[1, 10-12]. The advances made in the ALS experiment compared to the seminal work by Narten and co-workers[13-18] included data acquisition on a 3rd generation and highly monochromatic light source, a well characterized polarization correction, a Compton scattering correction that included electron correlation, and use of a more accurate Charge Coupled Device (CCD) detector, while also providing for an estimate of any remaining random or "shot noise" errors[1, 10]. However, one of the primary limitations of the original ALS experiment was that the intensity, $I(Q)$, where $Q$ is the momentum transfer, was carried out to a maximum value of $Q_{max}$ of only 11.0 Å$^{-1}$.
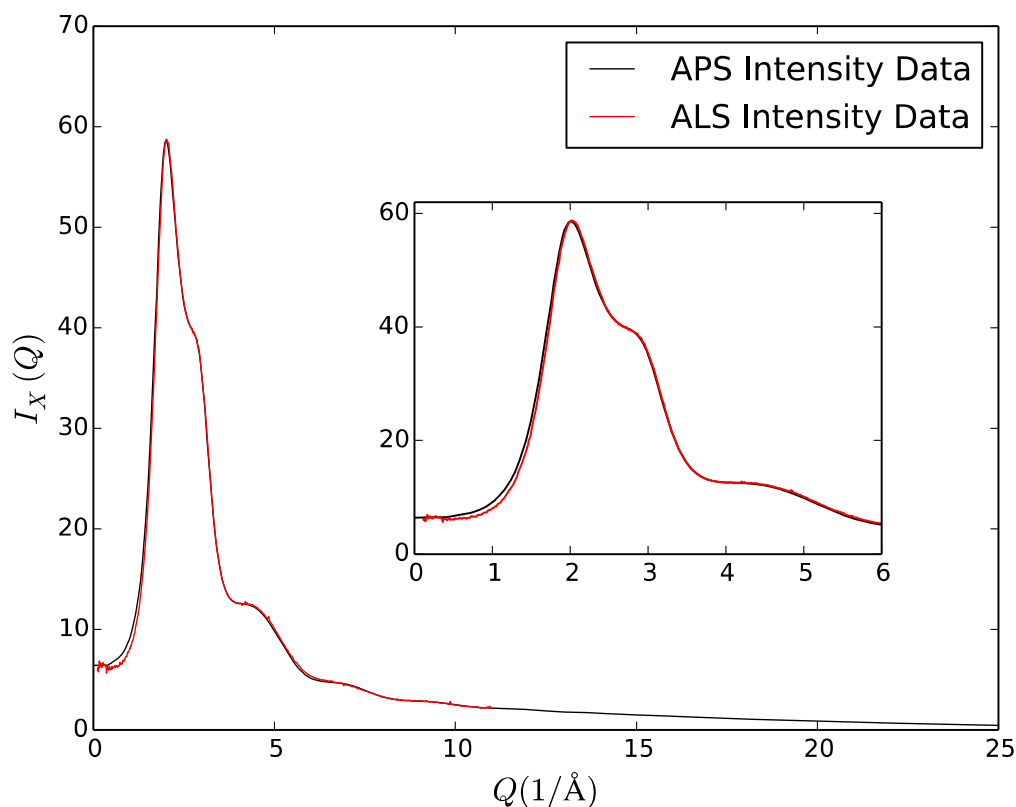


**Figure 1.** A comparison of the X-ray scattering intensity on liquid water at ambient conditions taken at the Advanced Light Source (ALS) in 2000 and Advanced Photon Source (APS) in 2013.

This limitation was overcome in 2013 and 2014 when Skinner and co-workers reported an X-ray measurement taken on liquid water at the Advanced Photon Source (APS)[2, 4]. The APS $I(Q)$ data also carefully included needed systematic corrections and estimates of uncertainty in the intensity observable from statistical errors, and greatly improved on the $Q$ range covered compared to the original ALS experiment, with $Q_{max} \sim 25.0$ Å$^{-1}$. Figure 1 shows a comparison of the $I(Q)$ data generated from the two

sets of measurements over their complete $Q$-range in which they show good agreement but for which there are clear differences as well. More importantly, however, the APS experimental intensity for $Q >$ 10.0 Å$^{-1}$ exhibits very small errors, and the amplitude, decay frequency and phase in this region is especially sensitive to the height and position of the first peak and low-$r$ behavior in $g_{OO}(r)$. Thus while the original ALS experiment provided an important milestone in data acquisition and error control for the generation of more reliable water intensity data, the APS data has superseded that work with its greater and more accurate $Q$-range, and hence we exclusively reanalyze that intensity data here.

Since the advent of high energy X-ray diffraction, it is possible to measure $I(Q)$ to a very high degree of accuracy, especially at high-$Q$, if a lot of care is taken, since it minimizes many of the difficult absorption and angle dependent corrections present for lower energy measurements, such that the collected intensity data yields experimental statistical errors of ~0.1% for $10\text{Å}^{-1} < Q < 25\text{Å}^{-1}$. However, this collected intensity is a combination of both elastic and inelastic (i.e. Compton) scattering, and after subtraction of the Compton scattering from the total intensity, the resulting elastic intensity is known to within ~1% uncertainty, most of which is concentrated in the $r < 2.5$ Å region. Hence while there is certainly a range of $g_{OO}(r)'s$ that agrees with the elastic intensity, they are nonetheless highly restrained by this ~1% error in the experimental intensity at high-$Q$.

The importance of the measured x-ray differential elastic cross-section, $I(Q)$ is that it can be further analyzed to understand the real space structural correlations under several assumptions. Formally $I(Q)$ is the square of the Fourier transform of the real space electron density, $\rho(r)$, where $r$ is the atomic separation.

$$I(Q) = \left\langle |F(Q)|^2 \right\rangle \qquad \int_0^\infty \rho(r)\exp(iQ\times r)dr$$

(1)

where the angle brackets denote a thermal average, and $\rho(r)$ is the electron density at point $r$. The real space integration can be approximated by a sum over atom positions $i,j$ in the liquid, leading to the following expression for $I(Q)$

$$I(Q) = \left\langle \sum_i \sum_j x_i x_j f_i(Q)\, f_j(Q)\exp\left(iQ\times r_{ij}\right) \right\rangle$$

(2)

where, $x_i$ is the atomic fraction of chemical species $i$, $f_i(Q)$ is the atomic scattering factor for atom $i$, and $r_{ij}$ is the distance between atoms $i$ and $j$. Specifically Eq. (2) states that the superposition of atom centered electron densities approximates the continuous real electron density distribution through the atomic scattering factor function. Assuming spherical symmetry, and the assumption that the liquid is composed of discrete atomic units, Eq. (2) can be reformulated as

$$I(Q) = I_{self}(Q) + I_{intra}(Q) + I_{inter}(Q)$$

(3a)

and more specifically in the case where the f(Q) functions have negligible imaginary component can be given as

$$I(Q) \cong n \sum_{\alpha} x_{\alpha} f_{\alpha}^2(Q) + \frac{1}{N} \sum_{m} \sum_{i} \sum_{j \neq i} f_i^{(m)}(Q) f_j^{(m)}(Q) \frac{sin(Qr_{ij}^{(m)})}{Qr_{ij}^{(m)}} + n^2 \sum_{\alpha} \sum_{\beta} x_{\alpha} x_{\beta} f_{\alpha}(Q) f_{\beta}(Q) H_{\alpha\beta}(Q)$$

(3b)

The indices $\alpha$, $\beta$ represent atom species (i.e. O and H), and the $i,j$ represent the three atomic sites in the molecule, which have separation $r_{ij}$; $n$ is the number of atoms per molecule (i.e. $n = 3$). The first term arises from self-scattering, the second term is due to scattering from intramolecular bonds with contributions from $N$ molecules, and the final term reports on the desired intermolecular correlations. The partial structure factors $H_{\alpha\beta}(Q)$ are related to partial radial distribution functions $g_{\alpha\beta}(r)$ through the back Fourier Transform:

$$H_{\alpha\beta}(Q) = 4\pi\rho \int_0^{\infty} r^2 \left( g_{\alpha\beta}(r) - 1 \right) \frac{sinQr}{Qr} dr$$

(4)

where $\rho$ is the number density. The important point that should be emphasized is that the experimental $I(Q)$ is a weighted sum of structure factor correlations $H_{OO}(Q)$, $H_{OH}(Q)$, and $H_{HH}(Q)$ for liquid water. The transformations of Eq. (3b) to the corresponding real space $g_{OO}(r)$, $g_{OH}(r)$, and $g_{HH}(r)$ (Eq. 4) requires a *model* of the electron density in the condensed phase (the $f_i(Q)$ functions) to weight and thus accurately isolate the individual radial distribution functions (rdfs) from the intensity data.

Often a free atom model of the atomic form factors, $f_i^{free}(Q)$, is used which in the limit as $Q \rightarrow 0$ corresponds to the atoms full electron number (Z), and the X-ray scattering experiment on liquid water is certainly dominated by the oxygen-oxygen correlations due to its larger Z number. However, it is now accepted that $f_i^{free}(Q)$ for describing the electron density of a molecular liquid is a poor description due to electron density redistributions that arise from chemical bonding and to molecular interactions in the condensed phase that shifts net electron density from hydrogen to the oxygen.[11] In 2000 the ALS group advanced a new theoretical model to account for these electron density shifts that gave rise to the use of a modified atomic form factor (MAFF), $f_i(Q)$:

$$f_i(Q) = f_i^{free}(Q) \left[ 1 + (\alpha_i - 1) exp\left( -\frac{Q}{2\delta^2} \right) \right]$$

(5)

where $\alpha_i$ and $\delta$ describe the electron density of the water molecule in a vacuum or polarized liquid environment. In particular $\alpha_i$ adjusts to conform to a condensed phase dipole moment that arises from polarizability, charge transfer, charge penetration, or even nuclear quantum effects, that can distort the

electron density away from the gas phase value of 1.85D. Estimates of the condensed phase dipole moment for water from *ab initio* simulations and x-ray scattering yield an average of ~2.95D but with large variation of up to ±0.6D[19-21], since the dipole moment is actually a distribution of dipole moments in the liquid. Therefore this introduces a level of model uncertainty that we use as part of an optimization variable for deriving the $g_{OO}(r)$ family, which is explained further below. Here we also increase water's gas phase covalent bond length $r_{OH}$ = 0.957 Å to its recently determined experimental ambient liquid value $r_{OH}$ = 0.990(5) Å, which was derived from high quality neutron diffraction data taken on a sample of $H_2^{18}O$ - $H_2^{16}O$, thereby avoiding problems with the incoherent and inelastic scattering differences of H and D.[22]

Thus in order to extract the $g_{OO}(r)$ from the experimental $I(Q)$ observable, we next require subtraction of $I_{self}(Q)$, $I_{intra}(Q)$ which account for self-scattering and the intensity that arises from the intramolecular bonds based on the MAFF electron density model. As is intuitive based on the MAFFs for hydrogen, and shown quantitatively in the previous ALS and APS studies, we can neglect $H_{HH}(Q)$ with very small error (~0.1%). Furthermore, the isolation of the dominant correlations in the X-ray measurement, the $H_{OO}(Q)$, must rely on a good estimate of $H_{OH}(Q)$, although not a perfect one. We have found that the resulting $H_{OO}(Q)$ is not particularly sensitive to the $H_{OH}(Q)$ used, as long as it is reasonable; in this work we have used the averaged $H_{OH}(Q)$ data generated by Skinner and co-workers[2, 4].

Based on the MAFF electron density model, the APS study used a straightforward numerical Fourier Transform procedure on the resulting structure factor $H_{OO}(Q)$ to determine the corresponding $g_{OO}(r)$. This (by construction) provides a $g_{OO}(r)$ that gives an excellent reproduction of the $I(Q)$ data when back-transformed to $Q$-space. However, systematic errors, statistical uncertainty and truncation artifacts for the limited Q-range give rise to spurious peaks and improper limiting behavior at small $r$ in the $g_{OO}(r)$ obtained by Fourier transform of the isolated structure factor $H_{OO}(Q)$. This can give erroneous area under the $r^2 g_{OO}(r)$ curve used to determine low-$Q$ behavior. In order to eliminate the unphysical oscillations in the Fourier Transform caused by high frequency noise and the finite $Q$ range of the APS measurement, a variable Lorch function was included in the integrand in Eq. (4) which suppresses the unphysical high frequencies in $r$-space. The 2013 APS oxygen-oxygen radial distribution function is shown in Figure 2; even with application of the variable Lorch modification function, the 2013 "benchmark" $g_{OO}(r)$ retains the unphysical oscillations visible at *r < 2.5 Å*. This is because the Lorch modification function was designed to be weak between 1.5-4.0 Å in the 2013 APS study so as to avoid artificial broadening of the 1st O-O peak. In 2014 the APS data was reanalyzed and removed these low-r correlations so that the $g_{OO}(r)$ beyond $r$ > 2.5Å was not significantly perturbed between the two APS real-space data sets (also shown in Figure 2 and compared to the ALS $g_{OO}(r)$ for completeness).
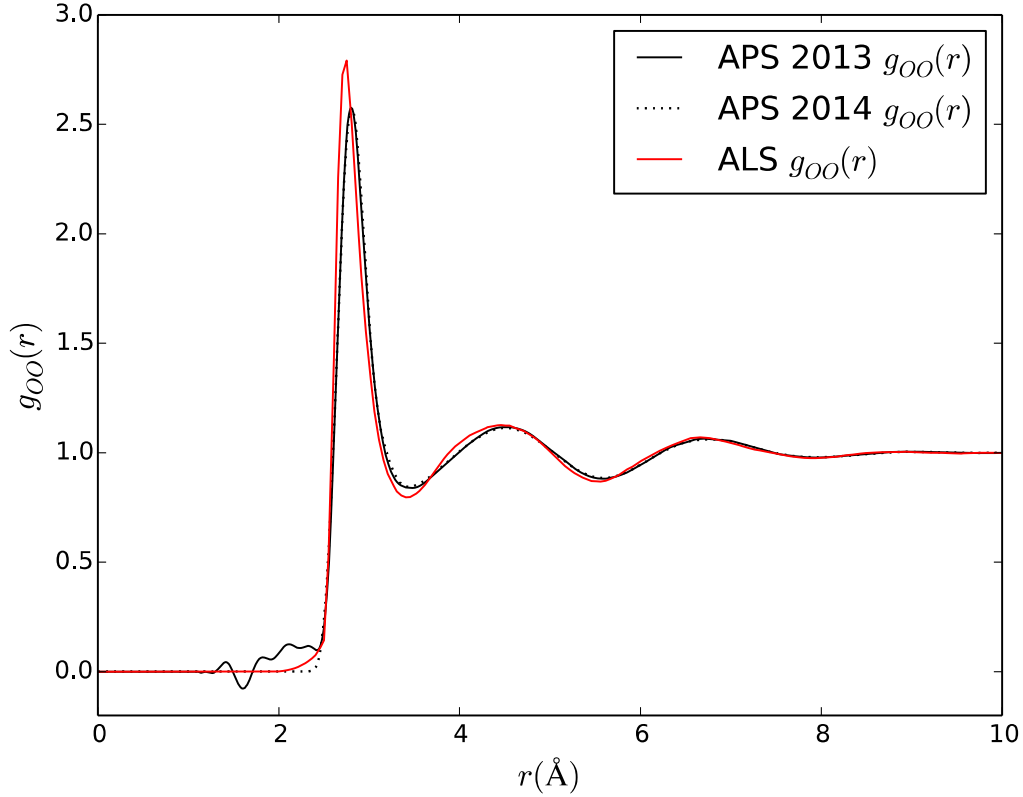
**Figure 2.** A comparison of the experimental models for $g_{OO}(r)$ for liquid water at ambient conditions derived from ALS 2000[11], APS 2103[2], and APS 2014[4] studies.

The importance of this low $r$ behavior becomes apparent when we consider the isothermal compressibility limit. The intermolecular scattering factor for molecular water in the limit as $Q \rightarrow 0$ can be written in terms of the intensity as

$$H_X(Q \rightarrow 0) = \frac{\left[I_{self}(Q) + I_{intra}(Q) + I_{inter}(Q)\right]_{Q \rightarrow 0}}{\left(n \sum_\alpha x_\alpha f_\alpha(Q)\right)^2_{Q \rightarrow 0}} = k_b T \rho \kappa_T \tag{6}$$

where $\rho$ is the mass density, $\kappa_T$ is the isothermal compressibility, and where $X$ designates that the structure factor corresponding to the water molecule scattering factor, in analogy to that of an atomic liquid, and the sum in the denominator corresponds to our normalization scheme. Eq. (6) can thus be expressed in terms of the partial structure factors using our MAFF weighting scheme which assumes an average molecular dipole moment of 2.8D such that

$$H_X(Q \rightarrow 0) = 0.81 H_{OO}(Q \rightarrow 0) + 0.18 H_{OH}(Q \rightarrow 0) = k_b T \rho \kappa_T \tag{7}$$

where again the hydrogen-hydrogen correlations are a negligible contribution even at very low $Q$. Thus to determine whether the MAFF weighted sum of $H_{OO}(Q)$ and $H_{OH}(Q)$ conform to the isothermal

compressibility limit requires the evaluation of Eq. (7) using quadrature, that in turn requires a finite value $R_{max}$ for the upper bound of the integration. We adopt the result by Salacuse et al. to correct for finite size errors[23] due to truncation of the upper bound in Eq. (4)

$$H_{\alpha\beta}(Q \to 0) \cong \int_{R=0}^{R_{MAXx}} r^2 \left( g_{\alpha\beta}(r) - 1 \right) dr \Big/ \left( 1 - 4\pi\rho R_{MAX}^3 / 3N \right)$$

(8)

which for our results corresponds to an $R_{max}$=19.2 Å, a value that exceeds any correlation length present in the system. We also adopt the procedure of Overduin and Patey to reduce any statistical noise by averaging Eq. (8) over 16 Å < $R_{max}$ < 19.2 Å to further minimize the truncation error[24]. Based on this procedure, we found that the 18% weighting of the $H_{OH}(Q)$ correlations term in Eq. (7) contributes at most 3-5% to the isothermal compressibility, and is thus comparable to the error in the quadrature itself. Therefore to a very good approximation we can assume that

$$H_{OO}(Q \to 0) \cong k_B T \rho \kappa_T$$

(9)

i.e. that the entirety of the water electron density is spherically symmetric and is largely centered on the oxygen[25]. This approximation has been used previously with good success[24-25], and thus Eq. (9) can provide a physically important and reasonable restraint on an allowable real space function. Normally, the measured $I(Q)$ tends approximately towards the correct isothermal compressibility limit by normalizing to the correct number of electrons at high $Q$, in which values of the intensity $I(Q \to 0)$ should reach a value of 6.2 electron units at room temperature and pressure. However, there is a gap in the experimental data between 0 < $Q$ < $Q_{min}$, (where $Q_{min}$ ~ 0.4 Å$^{-1}$), so that the $g_{OO}(r)$ derived from simple back-transform may not reproduce the very low-$Q$ behavior adequately. Thus when we evaluate the compressibility limit using Eq. (9) for the $g_{OO}(r)$ models derived from the ALS and APS 2013 and APS 2014 experiments, we find e.u. values of 1.4, 10.7, and 2.53, respectively, all well outside an acceptable range for the isothermal compressibility $I(Q=0)$ limit.

To reduce Fourier Transform problems we consider a procedure that uses a combination of real-space functions to find the optimal $g_{OO}(r)$ function(s) which best fit the experimental data[11] and which conform to a good approximation to the isothermal compressibility limit. More specifically, a "basis set" of radial distribution functions were culled from simulations, experiments, and theoretical predictions, such that all of these functions had the useful properties of displaying the proper small-$r$ behavior (no negative densities) and possessing a desirable degree of smoothness[11]; as part of the basis set we also include the ALS and 2013 APS and 2014 APS $g_{OO}(r)$ functions. The $g_{OO}(r)$ basis set all extend out to $r$=19.2 Å, thereby providing meaningful values of predicted scattering down to $Q$=~0.16 Å$^{-1}$. A

Langrangian constrained optimization procedure was chosen to find the optimal coefficients [$a_i$] in the following equation

(10)

$$g_{OO}(r) = \sum_i^m a_i g_{OO}^i(r)$$

where the $g_{OO}^i(r)$ are the "basis" functions taken from the sources described above to fit the experimental data to within the designated error of the experiment, as were the MAFF $\alpha_i$'s in Eq. (5) which were allowed to vary over a corresponding dipole moment range of 2.4D-3.1D (see Methods); this optimization procedure was performed under the restraint that the resulting $g_{OO}(r)$ conforms to the isothermal compressibility limit and fit $I(Q)$ well throughout the $Q$-range, but especially at high-$Q$ where the APS data is most accurate. We used Powell optimization[26] as implemented in the NLopt nonlinear-optimization package, http://ab-initio.mit.edu/nlopt to perform the restrained optimization. Due to the underdetermined nature of the problem, multiple optimal solutions can be found for the $g_{OO}(r)$ function. To explore the family of allowable solutions, we ran 1000's of optimization runs starting from independent initial conditions (i.e. different $\alpha_i$'s and different $a_i$'s). Our measure of convergence is when we could find no new $g_{OO}(r)$ functions given our basis set.

To better quantify the resulting agreement with the experiment, especially in the high-Q region, we have generated a filtered $H'_X(Q)$ function which removes the large background due to self and intramolecular scattering in the intensity, and in addition removes consideration of any Fourier components for $r < 1.6$Å and $r > 30$ Å. Furthermore, we weight $H'_X(Q)$ with $Q^3$ in order to endow the high-$Q$ region with far greater importance since error is smaller in this region compared to low-$Q$. The resulting formulated error function is

(11)

$$\chi^2 = \frac{\sum_{Q=0}^{Q_{max}} \left(Q^3 H'_{X,\text{model}}(Q) - Q^3 H'_{X,\text{expt}}(Q)\right)^2}{\sum_{Q=0}^{Q_{max}} \left(Q^3 H'_{X,\text{expt}}(Q)\right)^2}$$

where $Q_{max}$ = 20.0 Å$^{-1}$. Based on error estimates of the original APS data and uncertainty of the electron density model adopted to remove the background, an acceptable $\chi^2$ difference of 0.038 was suggested to us[27]; to provide greater even stringency for matching the high-$Q$ region, we reduced this acceptable $\chi^2$ difference further to 0.03. To place this in perspective, the corresponding difference with the simulated $I(Q)$ compared to the experimental value can range between a minimum of 0.07% to a maximum of 0.23% in the high-$Q$ region.

In Figure 3 we compare the resulting $Q^3*H'_x(Q)$ functions based on back-transformation of the family of optimized $g_{OO}(r)$ functions and the 2014 APS $g_{OO}(r)$ function against the experimental quantity. All of the back-transformed model $g_{OO}(r)$'s conform to the $\chi^2$ difference of 0.03 or less, and show good agreement with the amplitude and frequency of decay of the filtered experimental quantity, but in different areas compared to the APS 2014 real space function. It is interesting to note that the APS 2014 $g_{OO}(r)$ is flattening out beyond ~17 Å$^{-1}$, although the experiment still shows evidence of oscillations whose amplitude is well-fit by the family of $g_{OO}(r)$'s for Q > 17 Å$^{-1}$.
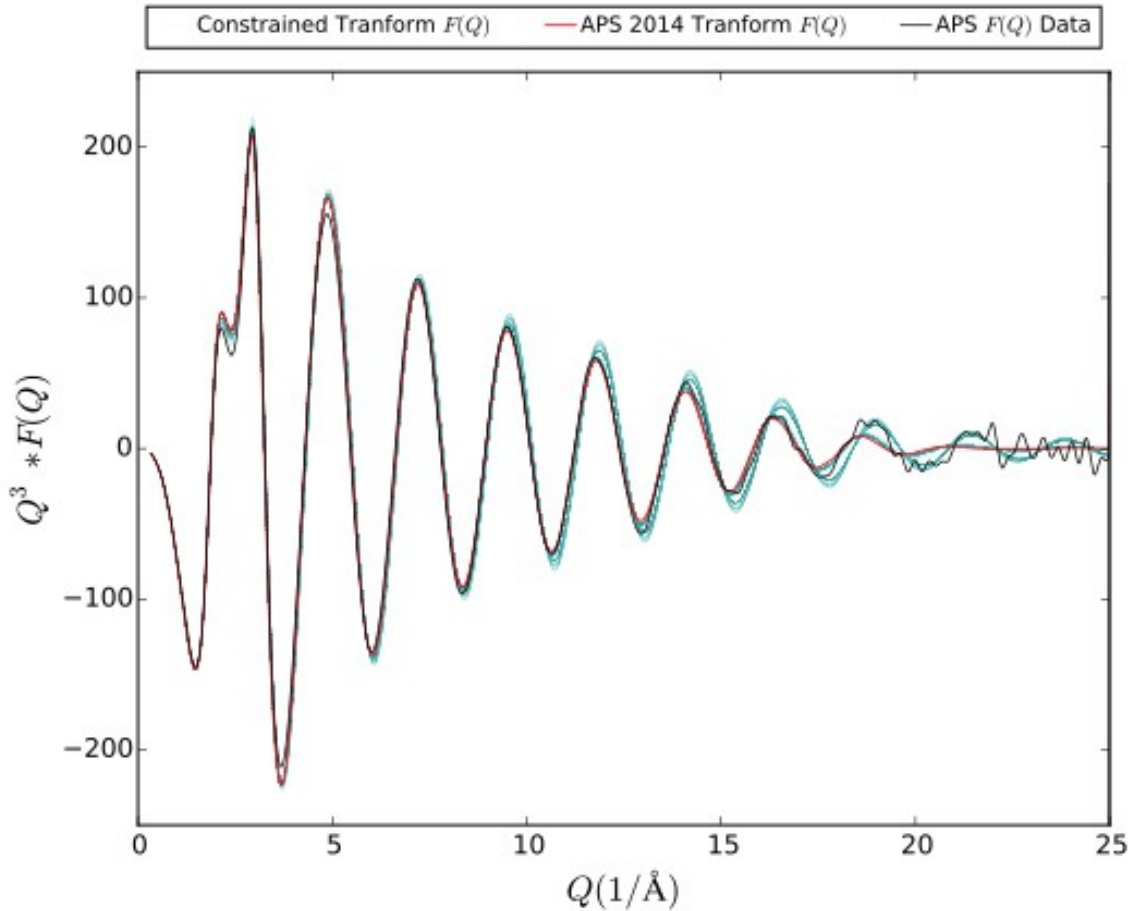


**Figure 3.** A comparison of the APS filtered $Q^3*H'_x(Q)$ function (black) and predicted from the corresponding quantity using the back-transformation of the APS 2014 data (red) and the optimized family of oxygen-oxygen radial distribution functions (blue).

The resulting family of $g_{OO}(r)$ functions are shown in Figure 4. The family of real space correlations emphasizes the variation that is possible given the underdetermined nature of the electron density model, the need to conform to the isothermal compressibility of 45.6 +/- 3.0 x 10$^{-11}$ Pa$^{-1}$ (6.2 +/- 0.4 in electron units) within the uncertainty pertaining to the quadrature error, and which satisfies the

entire intensity curve when back used in Eq. (3), *but* with agreement more highly weighted in the high-$Q$ region based on the fit to $Q^3 * H'_X(Q)$ with $\chi^2 < 0.03$.
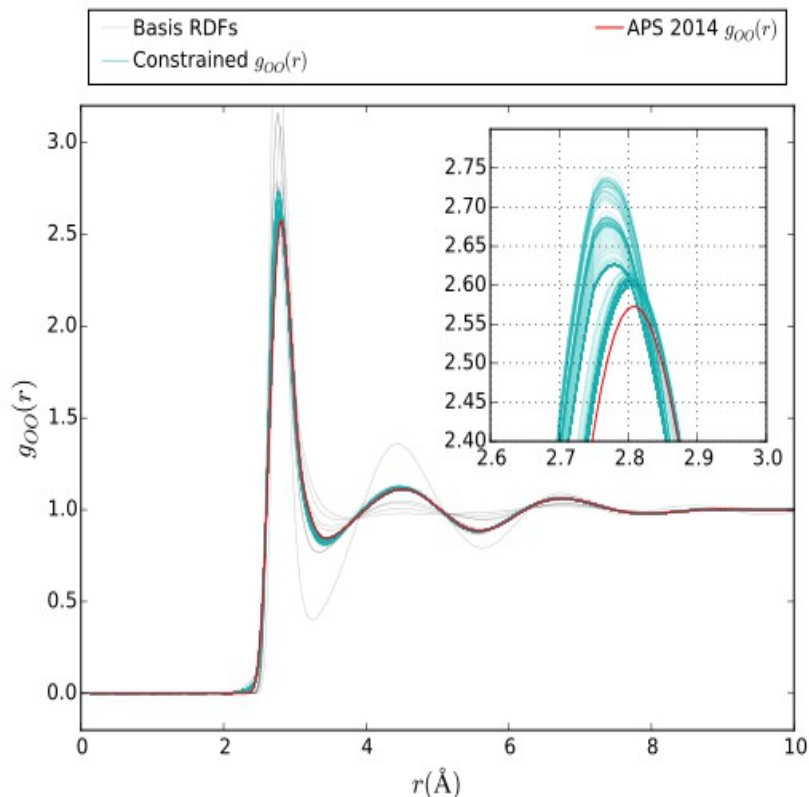


**Figure 4.** The 2014 APS $g_{OO}(r)$ (red) and the *family* of $g_{OO}(r)$'s (blue) that provide equally good agreement with the intensity restraints and with physically correct behavior at small $r$.

While the primary variation occurs in the first peak, some but smaller variations are observed in the first trough and second peak as well. When compared to the APS 2013 and 2014 $g_{OO}(r)$, we find that the first peak of the $g_{OO}(r)$ family is slightly shifted on average to smaller $r$, and the peak height on average is somewhat larger and varies over the range 2.53 to 2.73. Each member of the $g_{OO}(r)$ family is isolated from the intensity using a slightly different MAFF weighting since the $\alpha_i$'s were optimized as well. We find that the set of converged optimizations settle on $\alpha_i$'s that correspond to an average dipole moment of 2.7D with a spread of +/-0.2D, in good agreement with the single value of 2.8D derived in the original ALS study to define the MAFF parameters and which were subsequently adopted in the APS work (see Methods).

To demonstrate the general applicability of our approach, we analyze the recently reported $g_{OO}(r)$ derived from the APS X-ray data combined with neutron scattering using the empirical potential structure refinement (EPSR) method[3], as well as the $g_{OO}(r)$ functions derived from the AMOEBA14 force field when simulated under ambient conditions in the NPT ensemble[28] (both shown in Figure 5a). These were chosen since they fall outside of the range of 1$^{st}$ peak height values found in our optimization approach. When we evaluate the compressibility limit for the 2013 EPSR $g_{OO}(r)$ we get 4.6 e.u, which is well outside the bounds of the quadrature error. By contrast the 2014 AMOEBA model yields reasonable agreement with the correct compressibility limit via structure per Eq. (9), consistent with its known and reasonably accurate reported value derived from the volume fluctuation formulation[28]. Although these two $g_{OO}(r)$ functions are fixed in our investigation, we allow the MAFF model parameters to adjust to best optimize the agreement with the low-$Q$ and high-$Q$ intensity. We find that in both cases the agreement with $Q^3 * H'_X(Q)$ and the intensity is well outside experimental error compared to the 2014 APS and family of $g_{OO}(r)$ functions (Figure 5b) in which $\chi^2 = 0.035$ for 2013 EPSR and $\chi^2 = 0.105$ for 2014 AMOEBA. It is likely that the $Q$-scale errors from different experimental measurements that were averaged in creation of the 2013 EPSR function likely led to partial damping of the high-$Q$ structural oscillations. By contrast while the AMOEBA model conforms to the compressibility limit, it is nonetheless over-structured and leads to under-damping of the high-$Q$ structural oscillations.
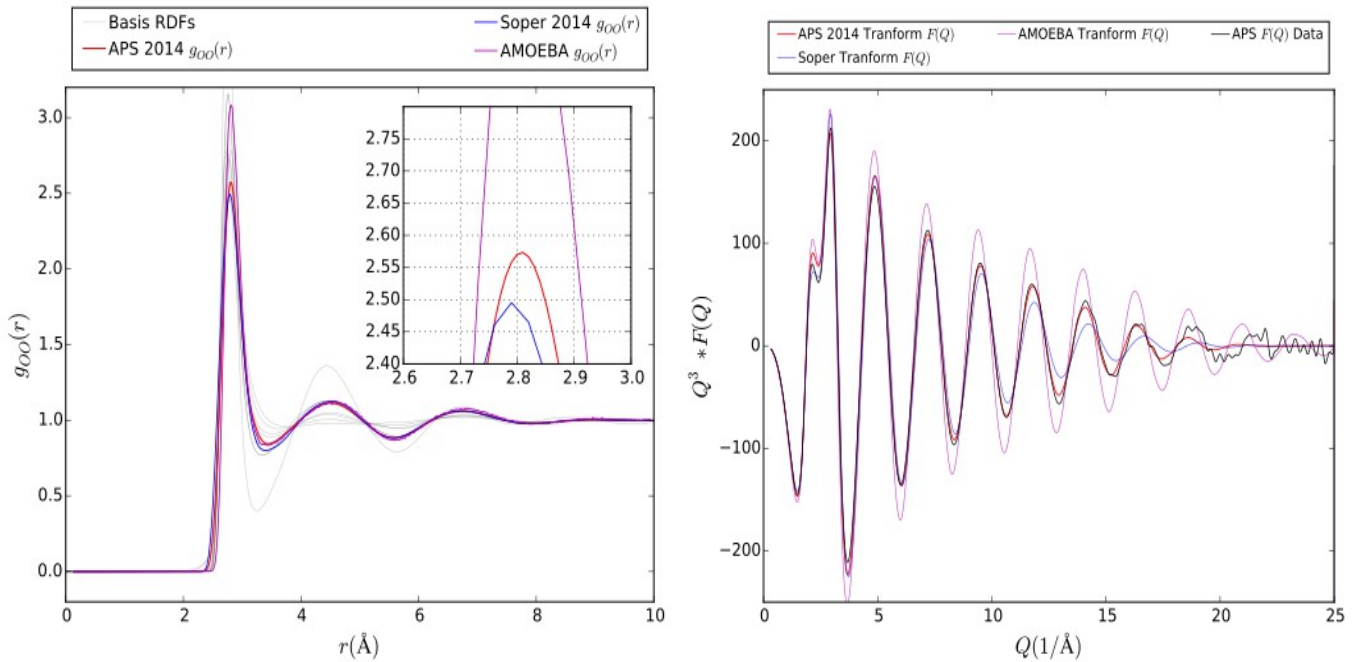
**Figure 5.** (a) A comparison of different $g_{OO}(r)$ functions: 2014 APS (red), 2013 EPSR (blue) and from the 2014 AMOEBA model (magenta). (b) The latter two functions are outside the acceptable error based on back-transformation and comparison to the the APS filtered $Q^3*H'_X(Q)$ function (black).

In summary, this work demonstrates that the extraction of the real-space conjugate $g_{OO}(r)$ from the X-ray intensity observable is an under-determined problem, but that we can determine an experimental procedure and data analysis that places bounds on the *family* of possible $g_{OO}(r)$ functions that equally well reproduce water's coherent X-ray scattering cross section, $I(Q)$. We have utilized the best known restraints on what is the most physically acceptable model for the oxygen-oxygen correlations. In particular our optimization procedure presents a $g_{OO}(r)$ family which all equally well reproduce the intensity data, especially at high $Q$ where the high energy APS intensity data is most accurate, do not exhibit unphysical Fourier oscillations at small $r$, and are bounded by a fairly precise estimate of the isothermal compressibility limit. This work attempts to quantify experimental and model uncertainty such that it provides a realistic benchmark *range* for real-space rdfs generated from theory and experiment. As such, we have made the intensity-rdf analysis utility available at [http://thglab.berkeley.edu](http://thglab.berkeley.edu) in which model $g_{OO}(r)$'s can be processed upon request.

## METHODS

*Electron density model.* In the original ALS study we chose to adjust $\alpha_i$ on each atom such that it reproduced the known gas phase molecular dipole moment of 1.85D, specifically by shifting $4/3e^-$ to oxygen ($9.333e^-$ on oxygen) and depleting the two hydrogens by $1/3e^-$ each ($0.333e^-$ on each hydrogen), such that this model conforms to the accepted water bond dipole moment of 1.5D when $\alpha_O$=1.167 and $\alpha_H$=0.333. These same $\alpha_i$ parameters were used in the APS study, but the two studies differed in their fitted ⎕ parameter, in which ⎕ = 2.2 Å$^{-1}$ in the ALS study but was adjusted in the most recent APS study to a value of ⎕ = 2.0 Å$^{-1}$. The differences in ⎕ values are completely negligible in their impact on the electron density model for water. In both the ALS[11] and recent APS studies[2, 4], the $\alpha$'s in Eq. (5) were readjusted from their gas phase values to reproduce an average condensed phase dipole estimate of 2.8D such that $1.5e^-$ are shifted to the oxygen and the two hydrogen's are depleted by $0.5e^-$ each. This gives rise to MAFF parameters $\alpha_O$=1.125 and $\alpha_H$=0.5 that were used in both the ALS and APS experimental studies to weight the relative importance of condensed phase intramolecular scattering as well as the weighting of the partial structure factors, $S_{ij}(Q)$, due to intermolecular correlations.

Here we provide an explicit example of how the MAFF parameters are calculated using these physical assumptions since this has caused some confusion in the past. First note that $\alpha_i$-1 corresponds to the fractional gain in electron density, when a charge $\Delta q$ is transferred off each H and onto the O atoms .

To conserve charge $\sum_i \Delta q_i = 0$ and since there are twice as many hydrogen as oxygen atoms

$$\alpha_O - 1 = 2\frac{\Delta q}{Z_O} \qquad \alpha_H - 1 = -\frac{\Delta q}{Z_H}$$

(12a)

It then follows that $\alpha_O -1 = \Delta q/4$ and $\alpha_H - 1 = \Delta q$-1. For a given dipole moment of the water molecule, $\mu$, the bond length $r_{OH}$, and bond angle $\theta$ (half the H-O-H angle), we can calculate the $\Delta q$ using via,:

$$\mu_{OH} = \frac{\mu}{2\cos\theta} \qquad \Delta q = \frac{\mu}{2r_{OH}\cos\theta}$$

(12b)

Using $r_{OH} = 0.990^{22}$, $\mu = 2.8$, and $\theta = 104.5°$, gives $\Delta q = 0.5$ which was used in the ALS and 2013 and 2014 APS studies.

**REFERENCES**

1.      Hura, G.; Russo, D.; Glaser, R.; Head-Gordon, T.; Krack, M.; Parrinello, M. Water Structure as a Function of Temperature from x-ray Scattering Experiments and Ab Initio Molecular dynamics. *Phys. Chem. Chem. Phys.* **2003,** *5*, 1981-1991.
2.      Skinner, L. B.; Huang, C.; Schlesinger, D.; Pettersson, L. G.; Nilsson, A.; Benmore, C. J. Benchmark oxygen-oxygen pair-distribution function of ambient water from x-ray diffraction measurements with a wide Q-range. *J. Chem. Phys.* **2013,** *138* (7), 074506.
3.      Soper, A. K. The Radial Distribution Functions of Water as Derived from Radiation Total Scattering Experiments: Is There Anything We Can Say for Sure? *ISRN Phys. Chem.* **2013,** *2013*, 279463.
4.      Skinner, L. B.; Benmore, C. J.; Neuefeind, J. C.; Parise, J. B. The structure of water around the compressibility minimum. *J. Chem. Phys.* **2014,** *141* (21), 214507.
5.      Soper, A. K. The radial distribution functions of water and ice from 220 to 673 K and at pressures up to 400 MPa. *Chem. Phys.* **2000,** *258* (1), 121-137.
6.      Huang, C.; Wikfeldt, K. T.; Nordlund, D.; Bergmann, U.; McQueen, T.; Sellberg, J.; Pettersson, L. G. M.; Nilsson, A. Wide-angle X-ray diffraction and molecular dynamics study of medium-range order in ambient and hot water. *Phys. Chem. Chem. Phys.* **2011,** *13* (44), 19997-20007.

7.      Weck, G.; Eggert, J.; Loubeyre, P.; Desbiens, N.; Bourasseau, E.; Maillet, J.-B.; Mezouar, M.; Hanfland, M. Phase diagrams and isotopic effects of normal and deuterated water studied via x-ray diffraction up to 4.5 GPa and 500 K. *Phys. Rev. B* **2009,** *80* (18), 180202.

8.      Soper, A. K.; Benmore, C. J. Quantum differences between light and heavy water. *Phys. Rev. Lett.* **2008,** *101*, 065502.

9.      Fu, L.; Bienenstock, A.; Brennan, S. X-ray study of the structure of liquid water. *J. Chem. Phys.* **2009,** *131* (23), 234702.

10.     Hura, G.; Sorenson, J. M.; Glaeser, R. M.; Head-Gordon, T. A high-quality x-ray scattering experiment on liquid water at ambient conditions. *J. Chem. Phys.* **2000,** *113* (20), 9140-9148.

11.     Sorenson, J. M.; Hura, G.; Glaeser, R. M.; Head-Gordon, T. What can x-ray scattering tell us about the radial distribution functions of water? *J. Chem. Phys.* **2000,** *113* (20), 9149-9161.

12.     Head-Gordon, T.; Hura, G. Water structure from scattering experiments and simulation. *Chem. Rev.* **2002,** *102* (8), 2651-70.

13.     Narten, A. H.; Danford, M. D.; Levy, H. A. X-Ray Diffraction Study of Liquid Water in Temperature Range 4-200 Degrees C. *Discuss Faraday Soc* **1967,** (43), 97-&.

14.     Narten, A. H.; Levy, H. A. Observed Diffraction Pattern and Proposed Models of Liquid Water. *Science* **1969,** *165* (3892), 447-&.

15.     Narten, A. H.; Levy, H. A. Liquid Water - Molecular Correlation Functions from X-Ray Diffraction. *J. Chem. Phys.* **1971,** *55* (5), 2263-&.

16.     Narten, A. H. Liquid Water - Atom Pair Correlation-Functions from Neutron and X-Ray-Diffraction. *J. Chem. Phys.* **1972,** *56* (11), 5681-&.

17.     Narten, A. H.; Thiessen, W. E.; Blum, L. Atom Pair Distribution-Functions of Liquid Water at 25-Degrees-C from Neutron-Diffraction. *Science* **1982,** *217* (4564), 1033-1034.

18.     Thiessen, W. E.; Narten, A. H. Neutron-Diffraction Study of Light and Heavy-Water Mixtures at 25-Degrees-C. *J. Chem. Phys.* **1982,** *77* (5), 2656-2662.

19.     Gubskaya, A. V.; Kusalik, P. G. The total molecular dipole moment for liquid water. *J. Chem. Phys.* **2002,** *117*, 5290.

20.     Badyal, Y. S.; Saboungi, M.-L.; Price, D. L.; Shastri, S. D.; Haeffner, D. R.; Soper, A. K. Electron distribution in water. *J. Chem. Phys.* **2000,** *112*, 9206.

21.     Silvestrelli, P. L.; Parrinello, M. Water dipole moment in the gas and liquid phase. *Phys. Rev. Lett.* **1999,** *82*, 3308.

22.     Zeidler, A.; Salmon, P. S.; Fischer, H. E.; Neuefeind, J. C.; Simonson, J. M.; Markland, T. E. Isotope effects in water as investigated by neutron diffraction and path integral molecular dynamics. *J Phys-Condens Mat* **2012,** *24* (28).

23.     Salacuse, J. J.; Denton, A. R.; Egelstaff, P. A. Finite-size effects in molecular dynamics simulations: Static structure factor and compressibility .1. Theoretical method. *Phys Rev E* **1996,** *53* (3), 2382-2389.

24.     Overduin, S. D.; Patey, G. N. Understanding the Structure Factor and Isothermal Compressibility of Ambient Water in Terms of Local Structural Environments. *J. Phys. Chem. B* **2012,** *116* (39), 12014-12020.

25.     Sedlmeier, F.; Horinek, D.; Netz, R. R. Spatial Correlations of Density and Structural Fluctuations in Liquid Water: A Comparative Simulation Study. *J. Amer. Chem. Soc.* **2011,** *133*, 1391-1398.

26.     Powell, M. J. D. A direct search optimization method that models the objective and constraint functions by linear interpolation. In *Advances in Optimization and Numerical Analysis*, Gomez, S.; Hennart, J.-P., Eds. Kluwer Academic: Dordrecht, 1994; pp 51-67.

27.     Skinner, L. B.; Benmore, C. J. private communication **2015**.

28.     Laury, M. L.; Wang, L. P.; Pande, V. S.; Head-Gordon, T.; Ponder, J. W. Revised Parameters for the AMOEBA Polarizable Atomic Multipole Water Model. *J. Phys. Chem. B* **2015**.