PCCP

Structure retrieval in liquid-phase electron scattering

Journal:	*Physical Chemistry Chemical Physics*
Manuscript ID	CP-ART-11-2020-006045
Article Type:	Paper
Date Submitted by the Author:	21-Nov-2020
Complete List of Authors:	Yang, Jie; SLAC National Accelerator Laboratory; SLAC National Accelerator Laboratory, PULSE institute F. Nunes, J. Pedro; University of Nebraska-Lincoln Ledbetter, Kathryn; SLAC National Accelerator Laboratory, PULSE institute; Stanford University, Physics Biasin, Elisa; SLAC, Stanford PULSE Institute; SLAC National Accelerator Laboratory, PULSE institute Centurion, Martin; University of Nebraska-Lincoln Chen, Zhijiang; SLAC National Accelerator Laboratory cordones , Amy; SLAC National Accelerator Laboratory, PULSE institute Crissman, Christopher; SLAC National Accelerator Laboratory DePonte, Daniel; SLAC National Accelerator Laboratory, LCLS Glenzer, Siegfried; SLAC National Accelerator Laboratory Lin, Ming-Fu; SLAC National Accelerator Laboratory Mo, Mianzhen; SLAC National Accelerator Laboratory Rankine, Conor; Newcastle University Shen, Xiaozhe; SLAC National Accelerator Laboratory Wolf, Thomas; SLAC National Accelerator Laboratory, PULSE institute Wang, Xijie; SLAC National Accelerator Laboratory

ARTICLE

Structure retrieval in liquid-phase electron scattering†

Jie Yang*‡[ab], J. Pedro F. Nunes‡[c], Kathryn Ledbetter[bd], Elisa Biasin[ab], Martin Centurion[c], Zhijiang Chen[a], Amy A. Cordones[b], Christ Crissman[a], Daniel P. Deponte[a], Siegfried H. Glenzer[a], Ming-Fu Lin[a], Mianzhen Mo[a], Conor D. Rankine[e], Xiaozhe Shen[a], Thomas J. A. Wolf[b], Xijie Wang*[a]

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Electron scattering on liquid samples has been enabled recently by the development of ultrathin liquid sheet technologies. The data treatment of liquid-phase electron scattering has been mostly reliant on methodologies developed for gas electron diffraction, in which theoretical inputs and empirical fittings are often needed to account for the atomic form factor and remove the inelastic scattering background. In this work, we present an alternative data treatment method that is able to retrieve the radial distribution of all the charged particle pairs without the need of either theoretical inputs or empirical fittings. The merits of this new method are illustrated through the retrieval of real-space molecular structure from experimental electron scattering patterns of liquid water, carbon tetrachloride, chloroform, and dichloromethane.

1. Introduction

Scattering provides a direct glimpse into the molecular structure of matter. For samples in the liquid phase, X-ray and neutron scattering have been the two methods of choice,[1-4] with the use of electron scattering being largely hampered by the shallow penetration depth of electrons (<1 μm) compared to hard X-rays (>100 μm) and neutrons (>1 cm). Nevertheless, the use of electron scattering in liquid samples remains an enticing proposition, as in principle these experiments could be carried out using miniaturized and inexpensive table-top instruments in university labs.[5-8] To avoid the loss of information due to multiple scattering, sample thicknesses on the order of a few hundred nm are required for high energy (~100 keV and above) electrons.[9] Since the 1970s, liquid electron scattering (LES) experiments using evaporating films,[10] vapor deposition,[11, 12] and nanofluidic cells[13, 14] have been developed with different levels of success. More recently, the use of free-flowing ultrathin liquid sheet jet[15, 16] in LES has overcome the limitation posed by the shallow penetration depth of electrons and enabled time-resolved LES. These advances in instrumentation had, however, remained unmatched by the development of LES specific data treatment methodologies. In fact, existing LES data analysis methods have largely been adopted from gas electron diffraction (GED). These methods often rely on theoretical inputs and empirical fittings inherited from GED and not benchmarked for liquid samples.

In this work, we introduce an alternative data analysis method, charge-pair distribution function (*CPDF*), which is able to deliver real-space structural information directly from scattering patterns without relying on theoretical inputs or empirical fittings. We present high-quality experimental LES data for four common solvents: H_2O, CCl_4, $CHCl_3$, and CH_2Cl_2.

2. Experimental Setup

The experimental data were recorded at the SLAC MeV-UED facility using gas-accelerated and converging liquid jets capable of producing free-flowing ultrathin liquid sheets.[15-17] A schematic drawing of the experimental setup is shown in Fig.1(a) and a sample scattering pattern of CCl_4 is shown in Fig.1(b). The detailed design and characterization of the experimental setup is reported in an earlier publication.[16] In brief terms, the electron beam is generated and accelerated by an rf-type photoinjector to a kinetic energy of 3.7 MeV. The pulse contains roughly 60,000 electrons per pulse at birth, and roughly 15,000 electrons per pulse at sample location. The electron beam size is 88 (horizontal) × 37 μm (vertical) in full-width-at-half-maximum (FWHM), measured by a knife-edge scan. The electron pulse has a repetition rate of 360 Hz.

The two types of ultrathin liquid jet used in this work was described in detail in previous publications[15-17]. The first type is a gas-accelerated sheet jet[15], which is based on a microfluidic chip with two gas channels and a liquid channel. High-pressure helium is used to flatten the liquid sheet at interaction region. The operating condition for datasets acquired in this work were as follows: H_2O—72 psi helium pressure, 0.20 mL/min flow rate; CCl_4—75 psi helium pressure, 0.20 mL/min flow rate; $CHCl_3$—80 psi helium pressure, 0.25 mL/min flow rate; CH_2Cl_2—77 psi helium pressure, 0.18 mL/min flow rate. The second type is a converging jet which is also based on microfluidic chip technology[17]. It includes two liquid channels that are angled towards each other to produce a flatten liquid sheet. In this work, only water was studied with the converging jet. The sample flow rate was set to 2.4 mL/min. The gas-accelerated jet and the converging jet produces a liquid sheet with a thickness of ~100 nm and ~650 nm, respectively, at the interaction point.

[a] SLAC National Accelerator Laboratory, Menlo Park, California, 94025, USA. Email: jieyang@slac.stanford.edu; wangxj@slac.stanford.edu
[b] Stanford PULSE Institute, SLAC National Accelerator Laboratory, Menlo Park, California, 94025, USA.
[c] Department of Physics and Astronomy, University of Nebraska–Lincoln, Lincoln, Nebraska, 68588, USA.
[d] Physics Department, Stanford University, Stanford, California, 94305, USA.
[e] School of Natural and Environmental Sciences, Newcastle University, Newcastle upon Tyne, NE1 7RU, UK.

† Electronic Supplementary Information (ESI) available: (1) Simulating *CPDF* contributions under IAM, (2) Influence of maximum *Q* range on *CPDF*, (3) Simulated $CPDF_{inter}$ by atom pairs in four liquids. (4) Experimental background-removed *CPDF* in four liquids. See DOI: 10.1039/x0xx00000x

‡ These authors contributed equally to this work.

The electron scattering patterns are recorded with a phosphor screen-based detector that is located approximately 3 meters downstream of the sample location[6]. The 40×40 mm phosphor screen contains a 2.9 mm diameter hole at the center position for transmitting the unscattered electron beam. The 1-D scattering curves $I(Q)$ are generated by radially average the recorded 2-D scattering patterns.

3. Theory of Data Treatment

In this section, we will first review the theory of electron scattering and the conventional pair distribution function (*PDF*) data treatment method commonplace in GED and then present CPDF method developed for LES.

Under the 1st Born approximation, the total electron scattering intensity (including both elastic and inelastic components) for a molecule can be written as:[18]

$$I(\vec{Q})=\frac{1}{Q^4}\int\sum_u\sum_v Z_u Z_v P_{uv}(\vec{r})\exp(i\vec{Q}\cdot\vec{r})d\vec{r} \quad (1)$$

Here $\vec{Q}$ is the momentum transfer, indices u and v run over all charged particles (nuclei and electrons), and Z_u is the charge of the u^{th} particle in units of elementary charge (−1 for electrons). $P_{uv}(\vec{r})$ represents the probability distribution of finding uv pairs at position $\vec{r}$. When $u = v$, $P_{uv}(\vec{r}) = \delta(\vec{r})$, the Dirac delta function. To keep the format concise, a constant $4\gamma^2/a_0^2$ is omitted from Eq. (1), where γ is the Lorentz factor and a_0 is the Bohr radius. Eq. (1) was initially derived for gas-phase molecules, where the space between molecules are typically much larger than the coherence length of the probe electrons and thus can be considered as isolated molecules in GED experiments. For liquid samples, the double sum in Eq. (1) is no longer limited to within a single molecule, but needs to be performed over all charged particles within the coherent volume of the probe electron beam.

The total scattering intensity, $I(\vec{Q})$, can be separated into elastic and inelastic components,[18-20]

$$I(\vec{Q}) = I_{elastic}(\vec{Q}) + I_{inelastic}(\vec{Q}) \quad (2)$$

where,

$$I_{elastic}(\vec{Q}) = Q^{-4}\,|\sum_u Z_u e^{i\vec{Q}\cdot\vec{R}_u} - F_1(\vec{Q})\,|^2 \quad (3)$$

$$I_{inelastic}(\vec{Q}) = Q^{-4}[n + F_2(\vec{Q}) - |\,F_1(\vec{Q})\,|^2] \quad (4)$$

in which $\vec{Q}$ is the momentum transfer, $\vec{R}_u$ is the position of the u^{th} nucleus, n is the number of electrons in the target, F_1 and F_2 are the Fourier transform of the one-electron and two-electron density, defined as

$$F_1(\vec{Q}) = \int e^{i\vec{Q}\cdot\vec{r}}\rho(\vec{r})d\vec{r} \quad (5)$$

$$F_2(\vec{Q}) = \int e^{i\vec{Q}\cdot(\vec{r}-\vec{r}')}\rho^{(2)}(\vec{r},\vec{r}')d\vec{r}d\vec{r}' \quad (6)$$

where $\rho(\vec{r})$ and $\rho^{(2)}(\vec{r},\vec{r}')$ are the one-electron and two-electron reduced density operator, respectively.

Directly applying Eq. (3-6) for scattering pattern prediction requires accurate one- and two-electron densities and is therefore computationally very expensive. A convenient approximation that is often used in GED data processing is the independent atom model (IAM), in which electrons are assumed to be distributed around their parent nuclei as if these were isolated atoms. Therefore, the IAM fails to account for electron redistribution due to the formation of chemical bonds. Under this approximation, the elastic component can be separated into atomic and molecular parts,[21, 22]

$$I_{elastic}(\vec{Q}) = I_{at}(Q) + I_{mol}(\vec{Q}) \quad (7)$$

where

$$I_{at}(Q) = \sum_u |f_u(Q)|^2 \quad (8)$$

$$I_{mol}(\vec{Q}) = \sum_u\sum_{v\neq u} f_u^*(Q)f_v(Q)\exp(i\vec{Q}\cdot\vec{r}_{uv}) \quad (9)$$

where f_u is the atomic form factor (AFF) of elastic electron scattering for the u^{th} atom and $\vec{r}_{uv}$ is a vector between the u^{th} and v^{th} nuclei. This treatment parallels the IAM method used for X-ray scattering: the relation between AFF for elastic electron scattering and AFF for elastic X-ray scattering is given by the Mott-Bethe formula,

$$f_u(Q) = Q^{-2}\,|\,Z_u - f_u^{X-ray}(Q)\,| \quad (10)$$

For liquid or gas samples, as molecules are randomly oriented, I_{mol} becomes isotropic and Eq. (9) can be reduced to:

$$I_{mol}(Q) = \sum_u\sum_{v\neq u} f_u^*(Q)f_v(Q)\frac{\sin(Qr_{uv})}{Qr_{uv}} \quad (11)$$

Under the IAM, the inelastic component only depends on the type of atom, which can be written as

$$I_{inelastic}(Q) = Q^{-4}\sum_u S_u(Q) \quad (12)$$

where S_u is the inelastic form factor (IFF) for the u^{th} atom. Note S_u is identical to the IFF for X-ray scattering, so that the inelastic component for electron and X-ray scattering are only differ by the Q^{-4} factor.[18]

Under the IAM, typically only I_{mol} contains structural information of the target system, as I_{at} and $I_{inelastic}$ are properties of the constituent atoms used to construct the model. The rapid drop in I_{mol} amplitude, imparted a Q^{-5} dependency is typically overcome through the use of modified scattering intensity curves, $sM(Q)$, in which each atom pair distance appears sinusoidal without damping across the Q range.

$$sM(Q) = \frac{I_{mol}(Q)}{I_{at}(Q)}Q \quad (13)$$

The *PDF* can then be calculated from $sM(Q)$ using a sine transform

$$PDF(r)=\int sM(Q)\sin(Qr)dQ \quad (14)$$

In principle, *PDF* analysis must always be preceded by the removal of inelastic and atomic scattering contribution from the experimental data, a process which often requires theoretical inputs and empirical fittings, even in the well-established GED data analysis.[22] In X-ray scattering, for example, it has been found that even high-level quantum chemical calculations are insufficient to remove the inelastic scattering background adequately in liquid water.[23] Moreover, even after getting $I_{mol}(Q)$, theoretical AFFs are still needed [Eq. (8) and (13)] in order to calculate the $sM(Q)$. For liquid, AFFs under IAM often need to be modified, for example, to incorporate the dipole moment in the case of a polar molecule such as water.[4]

Now we present an alternative method to retrieve real-space information directly from LES patterns with neither background removal nor AFFs. The scattering intensity from a randomly-oriented molecular ensemble can be obtained by performing an angular integral on Eq.(1),

$$I(Q) = \frac{1}{Q^4}\int\sum_u\sum_v Z_u Z_v P_{uv}(r)\frac{\sin(Qr)}{Qr}dr \quad (15)$$

We define a charge-pair distribution function, *CPDF*(*r*), as:

$$CPDF(r) = \sum_u \sum_v Z_u Z_v P_{uv}(r) \quad (16)$$

It includes all nucleus-nucleus, nucleus-electron, and electron-electron pairs, with their amplitude and sign being determined by the product of the charge of the two particles.

Substituting Eq. (16) into Eq. (15), we obtain:

$$Q^5 I(Q) = \int_0^\infty CPDF(r) \frac{\sin(Qr)}{r} dr \quad (17)$$

the sine transform of which yields

$$CPDF(r) = r \int_0^\infty Q^5 I(Q) \sin(Qr) dQ \quad (18)$$

Since only a finite range of Q is measured in any experiment, I is multiplied by a damping term, $e^{-\alpha Q^2}$, to avoid edge effects during the transform.

$$CPDF(r) = r \int_0^{Q_{MAX}} Q^5 I(Q) e^{-\alpha Q^2} \sin(Qr) dQ \quad (19)$$

This edge effect mitigation has been widely used in X-ray and electron scattering experiments; see for example Ref. 24. The damping term is equivalent to a Gaussian smoothing for *CPDF*(*r*) in the real space.

Note that the formula of the *CPDF*(*r*) in Eq. (19) is a mathematically-rigorous derivation from Eq. (1) without invoking approximations such as the IAM. In addition, it requires neither empirical fitting nor theoretical inputs.

Dividing by AFFs during the transformation (Eq.13) is equivalent to a real-space deconvolution that replaces the "form" of atom with a point-like nucleus. For X-ray scattering, the "form" of an atom is a fuzzy electron cloud (Thomson scattering). For electron scattering, the "form" of an atom is a point-like positive nucleus embedded in a fuzzy negative electron cloud (Rutherford scattering). For this reason, the *CPDF* is still able to return sharp nucleus-nucleus pairs without the AFFs. However, this introduces negative shoulders that accompany the positive nucleus-nucleus pair peak, which arise from the fuzzy nature of electron clouds. Although this could be seen as an inconvenience in understanding liquid structure, it also could enhance the contrast of the nuclei, as per demonstrated in Fig. 6 in this work. In addition, the embedded electron-nucleus and electron-electron pairs may lead to a slight distortion of the nucleus-nucleus peaks. In this case, the peak positions will no longer directly reflect the internuclear distances. For liquid water, this distortion leads to a difference up to 4% in peak position between *CPDF* and *PDF*, as shown in Fig. 6 below. A detailed analysis of contributions from nucleus-nucleus, nucleus-electron and electron-electron pairs is given in Section 5.1.

4. Experimental Data

4.1 *CPDF* of liquid water

The experimental and simulated scattering patterns of liquid water, along with the key steps in calculating the *CPDF* [I, Q^5I, $Q^5 Ie^{-\alpha Q^2}$ and *CPDF*(*r*)] are shown in Fig. 2(a)-(d). The blue curve was measured from a gas-accelerated liquid jet with a thickness of ~100 nm and an electron transmission of 88%. The black curve was measured from a converging liquid jet with a thickness of ~650 nm and an electron transmission of only 40%. The jet thickness was measured, in each case, using an optical interferometer.[16] A classical molecular dynamics (MD) simulation for liquid water was performed using the GROMACS software suite with the TIP4P-Ew force-field.[25] The scattering simulation was performed under the IAM, as outlined in the Supporting Information (SI). A single scaling factor is used between experimental and simulated curves, which is obtained by matching the amplitude of 1st shell O-O peak around 2.9 Å in the *CPDF* (Fig. 2d). A discrepancy between simulated and experimental Q^5I signal is observed [Fig. 2(b)-(c)], with the experimental signal being significantly higher than the simulated signal at high Q. This discrepancy is likely due to some small camera- or sample-related residual background being amplified by the Q^5 factor. Nevertheless, this smooth background only contributes at small distances, and very good agreement was found between the experimental and simulated *CPDF* for r >2 Å [Fig. 2(d). An advantage of this method is its robustness to the absence of low Q scattering. It is notoriously challenging to acquire electron diffraction patterns down to Q = 0. In this work the minimum Q value measured is between 0.3 and 0.5 Å^{-1}. Moreover, low-Q scattering tends to be overestimate under the IAM, as it fails to account for all of the "binding" and the majority of the "correlation" effects in inelastic electron scattering.[26] Both the IAM inaccuracy and the missing data in low Q have negligible contributions to the *CPDF*, since the Q^5 factor significantly reduces the weight of low-Q scattering in the sine transform.

The large difference in electron transmission between the two datasets suggests very different jet conditions and background levels. Nevertheless, the *CPDF* of the two datasets are still very similar, with the majority of differences concentrated at the small distances. This comparison shows that our method is relatively insensitive to the thickness of the jet – in the range between 100 and 650 nm, at least.

To further understand the simulated *CPDF* of water, we separate it into three components—elastic intraatomic contributions $CPDF_{intra}$, elastic interatomic contributions $CPDF_{inter}$, and inelastic contributions $CPDF_{ine}$. The IAM simulations of three components are shown in Fig. 3(a). $CPDF_{intra}$ covers the interference of charge pairs within each atom, $CPDF_{inter}$ covers the interference of charge pairs across different atoms, and $CPDF_{ine}$ covers the correction from electron correlation effect. The definition of each component is given in the SI. For r >~2 Å the $CPDF_{inter}$ contribution dominates over the other two terms. This feature arises from the fact that both intraatomic and inelastic component are short-ranged, and explains how the *CPDF* shown in Fig. 2(d) captures all nucleus pairs for r >~2 Å without removing any background. A more detailed analysis of each component in *CPDF* is given in Section 5.1. Nevertheless, background removal is still required in order to retrieve nuclear pairs with r <2 Å. In water, the two shortest nuclear pair distances, the bonded O-H (~1.0 Å) and the hydrogen-bonded O···H (~1.9 Å), are obscured by $CPDF_{intra}$ and $CPDF_{ine}$. Figure 3(b) shows the scattering pattern for each term after being multiplied by Q^5, with the molecular component, Q^5I_{inter}, oscillating around zero while Q^5I_{intra} and Q^5I_{ine} increase monotonically and smoothly. Therefore, a smooth background fitting to the Q^5I curve could potentially remove Q^5I_{intra} and Q^5I_{ine}, isolating the Q^5I_{inter} contribution. Here we use a simple method to remove this smooth background: a low-order polynomial (3rd order here) is fitted to the full range (0.3 < Q < 11.8 Å^{-1}) of experimental Q^5I, and Q^2 is multiplied to the range 0 < Q < 1 Å^{-1} to make the background smoothly go to zero at Q = 0. The fitted background is shown in Fig. 3(c), and the background-removed experimental data is plotted together with the simulated Q^5I_{inter} in Fig. 3(d). The polynomial fitting is a simple yet rudimentary background removal method that does not require any prior knowledge about the target system. The *CPDF*s, as retrieved from

experimental data with background removal, are shown in Fig. 3(e) together with simulation. The peaks in *CPDF* appear sharper in Fig. 3(e) than Fig. 2(d); this is because after background removal, Q^5I is no longer increasing rapidly with Q, and the damping factor α can be much relaxed. We used α = 0.03 for Fig. 3(e) and α = 0.06 for Fig. 2(d). While the background removal helps to reveal the two peaks under 2 Å, it also introduces certain artifacts, including, noticeably, a negative shift at 4-6 $Å^{-1}$ region and a positive shift at 7-9 $Å^{-1}$ region in Fig. 3d, leading to an underestimation of the length of the hydrogen bond in Fig. 3e. Therefore, for all the r > 2 Å peaks, we give higher credence to the *CPDF* without background removal, and only use background removal to retrieve peaks under 2 Å. More advanced background removal methods, such as a χ^2 background fitting based on zero-crossings of I_{inter},[5] smooth background method,[27] and earlier empirical methods,[28] among others, have been utilized in GED data analysis. Nevertheless, these methods typically require a certain amount of *a priori* knowledge or assumptions to be made about the system under study. The applicability of these methods to LES are subject to future studies. This simple background removal method is able to retrieve the r < 2 Å atom pair distances for all four liquid samples we studied in this work, as shown in Fig. S3 in the SI. In the rest of the main text, we will focus on the *CPDF* without background removal, and only focus on the r > 2 Å pair distances.

4.2 Experimental data for four liquids

In order to test the generality and robustness of the data treatment method, we have performed LES experiments and *CPDF* analyses on four liquid solvents: H_2O, CCl_4, $CHCl_3$, and CH_2Cl_2, delivered using a gas-accelerated liquid jet. The electron transmission for the four datasets are 88%, 57%, 35%, and 68%, respectively, comprising a wide range of liquid sheet conditions. The thickness of the liquid sheet for CCl_4, $CHCl_3$, CH_2Cl_2 was not directly measured, but has been estimated to be between 100 and 200 nm based on scattering cross-section, solvent density, and measured electron transmission. Fig. 4 shows the radially averaged raw scattering pattern $I(Q)$, the *CPDF* without background removal, and a graphic representation of the four liquid solvents, along with theoretical $I(Q)$ an *CPDF(r)* calculated from classical MD simulations. The experimental *CPDF* is calculated using Eq. (19). The structure of the four liquid solvents was simulated *via* classical MD with 50 × 50 × 50 Å boxes using the GROMACS software suite.[29] Water was modelled using the TIP4P-Ew force field,[25] and CCl_4, $CHCl_3$, CH_2Cl_2 were modeled using the OPLS-AA force field.[30] The classical MD trajectories were transformed into time-averaged PDFs using the VMD package,[31] and the scattering pattern simulation was carried out using the method of Dohn *et al.*[32] under the IAM. For each sample, six temperatures (270 K, 285 K, 300 K, 315 K, 330 K, and 345 K) were simulated, and the best fits (270 K for CCl_4, $CHCl_3$, CH_2Cl_2 and 285 K for H_2O) are shown in Fig. 4. It is known that OPLS-AA force field is inadequate for treating the electrostatic interaction for halogen atoms.[33] The reader should note, therefore, that the classical MD simulations carried out for CCl_4, $CHCl_3$, and CH_2Cl_2 are not intended to provide any rigorous comparison with the experimental data but, rather, to serve as references for the general appearance of the corresponding *CPDF*.

5. Discussion

5.1 *Understanding the CPDF*

To better understand the information content of the *CPDF* and the interplay between nucleus-nucleus, electron-electron and nucleus-electron pair contribution, we simulated the *CPDF* of isolated, randomly-oriented CCl_4 molecules under the IAM. The elastic and inelastic atomic form factors are taken from Ref. 34. The charge pairs in the *CPDF* can be separated into three parts: elastic intraatomic contributions $CPDF_{intra}$, elastic interatomic contributions $CPDF_{inter}$, and inelastic contributions $CPDF_{ine}$, as shown in Fig. 5(a).

The elastic intraatomic contributions come from charge pairs within the same atom, and contains a large positive peak at r = 0.4 Å, a large negative peak at r = 1 Å and a small positive peak at r = 2 Å. The two positive peaks come from electron-electron pair and the negative peak represents electron-nucleus pair. The $CPDF_{intra}$ contribution dominates r <2 Å, and becomes negligible at r >3 Å, because intraatomic pairs are constrained to short distances. The physical origin of inelastic electron/X-ray scattering was shown by Bartell *et al.* to be electron correlation, *i.e.* electrons avoid each other spatially through Pauli exclusion and Coulomb repulsion.[18] Therefore, the contribution $CPDF_{ine}$ encodes a "correction" to the electron spatial distribution due to electron correlation. Note that most commonly used AFFs and IFFs are based on Hartree-Fock models[34, 35], in which only the Pauli correlation for electrons within individual atoms are included. P_{ine} becomes negligible at r >2 Å, since electron correlation is also a short-ranged interaction. Therefore, longer-range (r >~2 Å) interactions are dominated by the P_{inter} term, in which structural information about the molecule is encoded. The current separation point (r >~2 Å) is related to the maximum Q range of this experiment. A detailed analysis of the influence of the Q range on the *CPDF* analysis is given in the SI (section 2).

The $CPDF_{inter}$ contribution is plotted in Fig. 5(b) along with the nucleus-nucleus, electron-electron and nucleus-electron contributions. Despite the diverse nature of the various contributions, peaks in $CPDF_{inter}$ are found at r = 1.72 and r = 2.92 Å, demonstrating that $CPDF_{inter}$ encodes the C-Cl and Cl···Cl internuclear separation (1.77 and 2.89 Å, respectively) within 0.05 Å accuracy.[36] This net effect arises from the negative contributions from nucleus-electron pairs and the broad and weak nature of electron-electron pair contributions. Note that the electron-electron pair contribution (the yellow curve in Fig. 5(b)) is equivalent to the information content obtained from X-ray scattering but without applying AFF in the sine transformation.

5.2 *Comparison to PDF*

In our previous work, we performed the *PDF* analysis on the liquid water dataset[16]. Fig. 6 presents a direct comparison of the *CPDF* and the *PDF* analysis. The *PDF* analysis requires an empirical power fit for background removal and theoretical AFFs to calculate *sM* [Fig. 6 (d-f)], the *CPDF* analysis only requires a Q^5 scaling [Fig. 6(a-c)]. A comparison between Fig. 6(c) and (f) shows that both methods returned peaks corresponding to the 1st water shell (*CPDF*: 2.93 Å, *PDF*: 2.91 Å), 2nd water shell (*CPDF*: 4.59 Å, *PDF*: 4.41 Å), 3rd water shell (*CPDF*: 6.71 Å, *PDF*: 6.94 Å). In addition, the *CPDF* also revealed a weak 4th water shell at ~8.85 Å, which can be attributed to the contrast enhancement provided by two negative shoulders beside each positive peak in *CPDF*. This unique feature could be very useful for exploring the long-range weak structural correlations in other liquid phase samples.

6. Conclusion

In summary, we propose the *CPDF* as an alternative real-space analysis tool to the conventional *PDF* method for structure retrieval in liquid electron scattering experiments. The *CPDF* is a mathematically rigorous transformation that enable access to

structural information from LES patterns directly without invoking any theoretical model or empirical fitting. The generality and robustness of the *CPDF* method is demonstrated through the retrieval of major internuclear pairs between ~2 and ~10 Å in experimental LES data for H_2O, CCl_4, $CHCl_3$, and CH_2Cl_2.

CPDF analysis returns the real-space distribution of all charged particle pairs and is a unique feature of electron scattering. This is because that electron scattering probes both electrons and nuclei through Coulombic interaction, while X-ray scattering solely probes the electron distribution and neutron scattering solely probes the nuclei distribution. In addition, electron scattering with *CPDF* analysis shows strong structural resolving capability that, in certain cases, goes beyond X-ray and neutron scattering. For example, the hydrogen bond in liquid water (Fig. 3e) is challenging to resolve in X-ray scattering experiments and the 4th O-O shell (Fig. 6c) in liquid water is challenging to resolve in neutron scattering experiments.

Although easily calculated, the interpretation of the *CPDF* requires careful consideration as it overlays contributions from nucleus-nucleus pairs as sharp peaks, nucleus-electron pairs as negative shoulders on either side of each nucleus-nucleus pairs, and electron-electron pair contribution as broad and smooth features. The *CPDF* is especially useful when it is not straightforward to apply conventional *PDF* analyses. For example, the AFFs and inelastic backgrounds are difficult to predict for a mixed solvent with unknown mixing ratios, or for nanofluidic liquid cells where the scattering from the cell walls are not negligible. In future, *CPDF* analysis might be used to quantify detailed electron distributions in liquids, provided adequate signal-to-noise ratio is achieved. We anticipate *CPDF* analysis will become a valuable tool in future LES studies.

Conflicts of Interest

There are no conflicts to declare.

Acknowledgements

We gratefully acknowledge Dr. Xiaolei Zhu for the discussion that motivated this experiment. We thank Gregory M. Stewart from SLAC National Accelerator Laboratory for making Figure 1A. The experiment was performed at SLAC MeV-UED, which is supported in part by the DOE BES SUF Division Accelerator & Detector R&D program, the LCLS Facility, and SLAC under contract Nos. DE-AC02-05-CH11231 and DE-AC02-76SF00515. J. P. F. N., and M. C. are supported by the U.S. Department of Energy Office of Science, Basic Energy Sciences under Award No. DE-SC0014170. K. L. is supported by a Melvin and Joan Lane Stanford Graduate Fellowship and a Stanford Physics Department fellowship. C. D. R. is supported by Newcastle University and an EPSRC Doctoral Prize Fellowship (EP/R51309X/1). A. C. and T. J. A. W. are supported by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, Chemical Sciences, Geosciences, and Biosciences Division.

Notes and References

1. H. E. Fischer, A. C. Barnes and P. S. Salmon, *Rep. Prog. Phys.*, 2006, **69**, 233-299.
2. K. Amann-Winkel, M. C. Bellissent-Funel, L. E. Bove, T. Loerting, A. Nilsson, A. Paciaroni, D. Schlesinger and L. Skinner, *Chem. Rev.*, 2016, **116**, 7570-7589.
3. J. G. Powles, *Adv. Phys.*, 1973, **22**, 1-56.
4. J. M. Sorenson, G. Hura, R. M. Glaeser and T. Head-Gordon, *J. Chem. Phys.*, 2000, **113**, 9149-9161.
5. R. Srinivasan, V. A. Lobastov, C. Y. Ruan and A. H. Zewail, *Helv. Chim. Acta*, 2003, **86**, 1761-1838.
6. S. P. Weathersby, G. Brown, M. Centurion, T. F. Chase, R. Coffee, J. Corbett, J. P. Eichner, J. C. Frisch, A. R. Fry, M. Guhr, N. Hartmann, C. Hast, R. Hettel, R. K. Jobe, E. N. Jongewaard, J. R. Lewandowski, R. K. Li, A. M. Lindenberg, I. Makasyuk, J. E. May, D. McCormick, M. N. Nguyen, A. H. Reid, X. Shen, K. Sokolowski-Tinten, T. Vecchione, S. L. Vetter, J. Wu, J. Yang, H. A. Durr and X. J. Wang, *Rev. Sci. Instrum.*, 2015, **86**, 073702.
7. R. J. Miller, *Annu. Rev. Phys. Chem.*, 2014, **65**, 583-604.
8. O. Zandi, K. J. Wilkin, Y. W. Xiong and M. Centurion, *Struct. Dyn.*, 2017, **4**.
9. G. R. Anstis, Z. Liu and M. Lake, *Ultramicroscopy*, 1988, **26**, 65-69.
10. S. Lengyel and E. Kalman, *Nature*, 1974, **248**, 405-406.
11. M. Hada, Y. Shigeeda, S. Koshihara, T. Nishikawa, Y. Yamashita and Y. Hayashi, *J. Phys. Chem. A*, 2018, **122**, 9579-9584.
12. C. Y. Ruan, V. A. Lobastov, F. Vigliotti, S. Y. Chen and A. H. Zewail, *Science*, 2004, **304**, 80-84.
13. A. A. Petruk, C. Allen, N. Rivas, K. Pichugin and G. Sciaini, *Nanotechnology*, 2019, **30**.
14. C. Mueller, M. Harb, J. R. Dwyer and R. J. D. Miller, *J Phys Chem Lett*, 2013, **4**, 2339-2347.
15. J. D. Koralek, J. B. Kim, P. Bruza, C. B. Curry, Z. J. Chen, H. A. Bechtel, A. A. Cordones, P. Sperling, S. Toleikis, J. F. Kern, S. P. Moeller, S. H. Glenzer and D. P. DePonte, *Nature Communications*, 2018, **9**, 1353.
16. J. P. F. Nunes, K. Ledbetter, M. Lin, M. Kozina, D. P. DePonte, E. Biasin, M. Centurion, C. J. Crissman, M. Dunning, S. Guillet, K. Jobe, Y. Liu, M. Mo, X. Shen, R. Sublett, S. Weathersby, C. Yoneda, T. J. A. Wolf, J. Yang, A. A. Cordones and X. J. Wang, *Struct. Dyn.*, 2020, **7**, 024301.
17. B. Ha, D. P. DePonte and J. G. Santiago, *Physical Review Fluids*, 2018, **3**, 114202.
18. L. S. Bartell and R. M. Gavin, *J. Am. Chem. Soc.*, 1964, **86**, 3493-3498.
19. T. Iijima, R. A. Bonham and T. Ando, *J. Phys. Chem.*, 1963, **67**, 1472-1474.
20. J. Yang, X. Zhu, J. P. F. Nunes, J. K. Yu, R. M. Parrish, T. J. A. Wolf, M. Centurion, M. Gühr, R. Li, Y. Liu, B. Moore, M. Niebuhr, S. Park, X. Shen, S. Weathersby, T. Weinacht, T. J. Martinez and X. Wang, *Science*, 2020, **368**, 885-889.
21. L. O. Brockway, *Rev. Mod. Phys.*, 1936, **8**, 0231-0266.
22. L. Schafer, *Appl. Spectrosc.*, 1976, **30**, 123-149.
23. C. Huang, K. T. Wikfeldt, D. Nordlund, U. Bergmann, T. McQueen, J. Sellberg, L. G. M. Pettersson and A. Nilsson, *Phys. Chem. Chem. Phys.*, 2011, **13**, 19997-20007.
24. H. Ihee, *Acc. Chem. Res.*, 2009, **42**, 356-366.
25. H. W. Horn, W. C. Swope, J. W. Pitera, J. D. Madura, T. J. Dick, G. L. Hura and T. Head-Gordon, *J. Chem. Phys.*, 2004, **120**, 9665-9678.
26. J. H. Wang, A. N. Tripathi and V. H. Smith, *J. Chem. Phys.*, 1994, **101**, 4842-4854.
27. S. Shibata and L. S. Bartell, *J. Mol. Struct.*, 1971, **9**, 1-&.

28. R. A. Bonham and L. S. Bartell, *J. Am. Chem. Soc.*, 1959, **81**, 3491-3496.
29. D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark and H. J. C. Berendsen, *J. Comput. Chem.*, 2005, **26**, 1701-1718.
30. W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives, *J. Am. Chem. Soc.*, 1996, **118**, 11225-11236.
31. W. Humphrey, A. Dalke and K. Schulten, *Journal of Molecular Graphics*, 1996, **14**, 33-38.
32. A. O. Dohn, E. Biasin, K. Haldrup, M. M. Nielsen, N. E. Henriksen and K. B. Moller, *Journal of Physics B-Atomic Molecular and Optical Physics*, 2015, **48**.
33. W. L. Jorgensen and P. Schyman, *J. Chem. Theory Comput.*, 2012, **8**, 3895-3901.
34. J. H. Hubbell, W. J. Veigele, E. A. Briggs, R. T. Brown, D. T. Cromer and R. J. Howerton, *J. Phys. Chem. Ref. Data*, 1975, **4**, 471-538.
35. E. Prince, *International Tables for Crystallography, Volume C: Mathematical, Physical and Chemical Tables*, Springer, Dordrecht, The Netherlands, 3rd edn., 2006.
36. L. S. Bartell, L. O. Brockway and R. H. Schwendeman, *J. Chem. Phys.*, 1955, **23**, 1854-1859.

ARTICLE

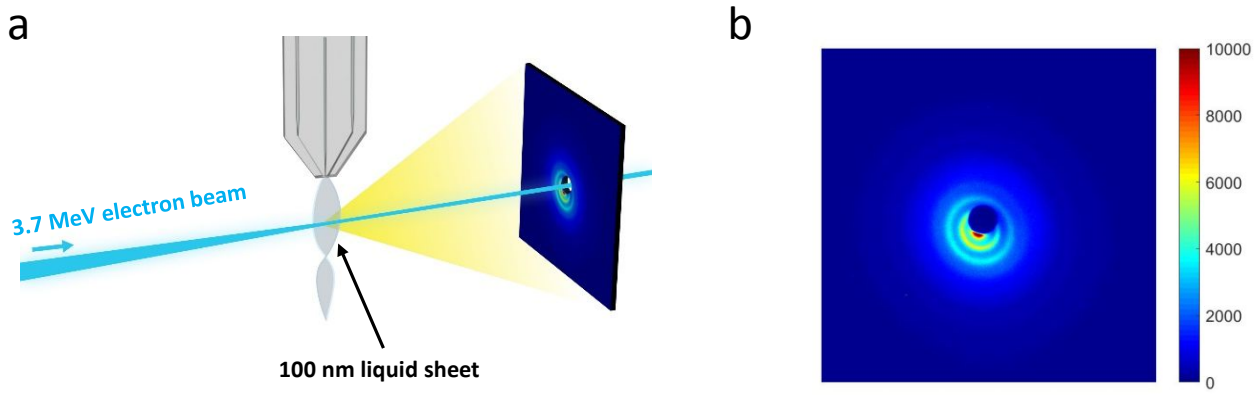

Figure 1. Experiment illustration. (a) Schematic illustration of the experimental setup. (b) Scattering pattern of liquid CCl_4. The pattern is deliberately off-center in order to access the small-angle scattering. The color bar represents the detector counts for a 1-second exposure.

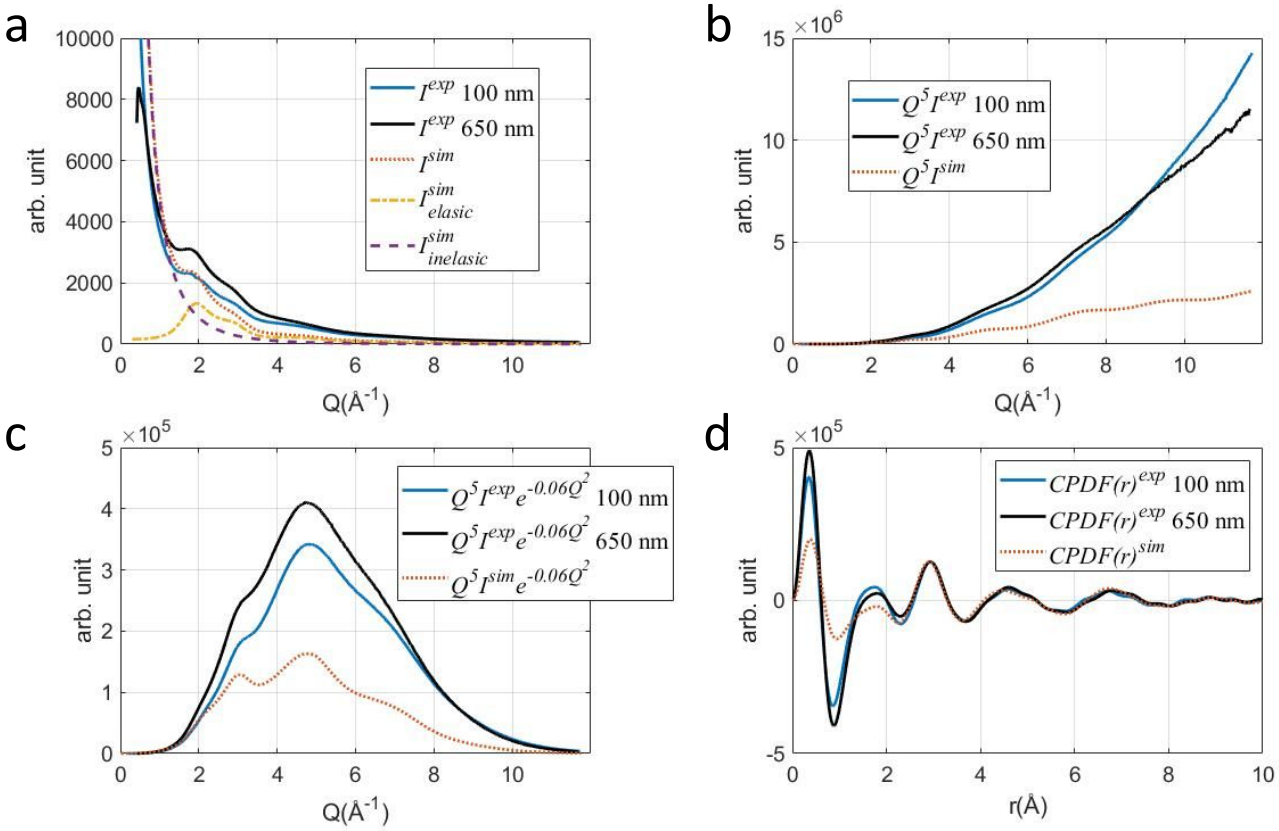

Figure 2. Generating *CPDF* from LES patterns. Experimental 100 nm sheet jet (blue solid), 650 nm sheet jet (black solid) and simulated (red dotted) (a) *I*, (b) Q^5I, (c) $Q^5 Ie^{-\alpha Q^2}$, (d) *CPDF*(*r*) for water. Elastic (yellow dash-dotted) and inelastic (purple dashed) components from simulation are also shown in part (a). Simulations were performed under the IAM, and the *CPDF* was calculated using Eq. (19) with a damping factor α = 0.06. Experimental and simulation curves are scaled by the first OO peak (~2.9 Å) in part (d).

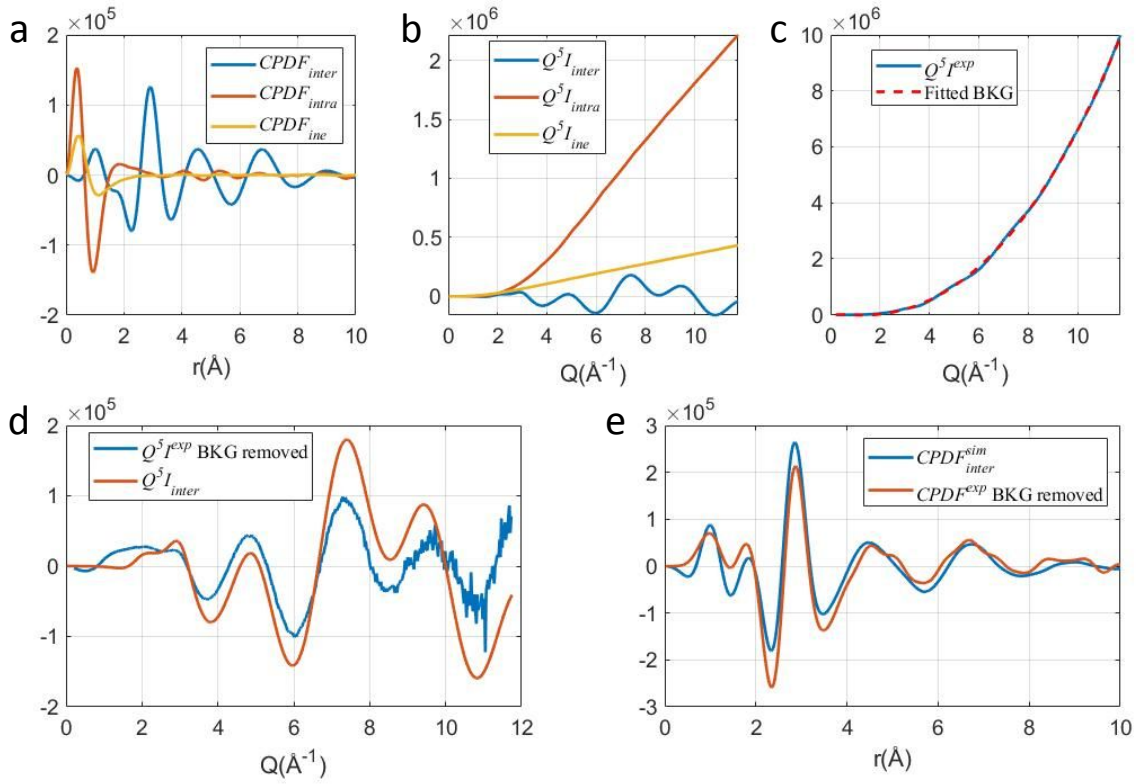

Figure 3. Detailed analysis on liquid water. (a) The interatomic elastic $CPDF_{inter}$, intraatomic elastic $CPDF_{intra}$ and inelastic $CPDF_{ine}$ term of *CPDF* for water simulated under the IAM with a damping factor α = 0.06. (b) Scattering pattern Q^5I_{inter} , Q^5I_{intra}, and Q^5I_{ine} corresponds to the three terms shown in part (a). (c) A 3rd order polynomial fitting for background removal in experimental Q^5I. (d) The background removed experimental data compare with the simulated Q^5I_{inter}. (e) The $CPDF_{inter}$ (interatomic contribution) from simulation (blue) and experimental *CPDF* with background removal (red), calculated using a damping factor α = 0.03. The scaling factor between experiment and simulation is the same as the one used in Fig. 2. All plots use Q_{max} =11.8 Å^{-1}. Experimental data is acquired using 100 nm water sheet jet.

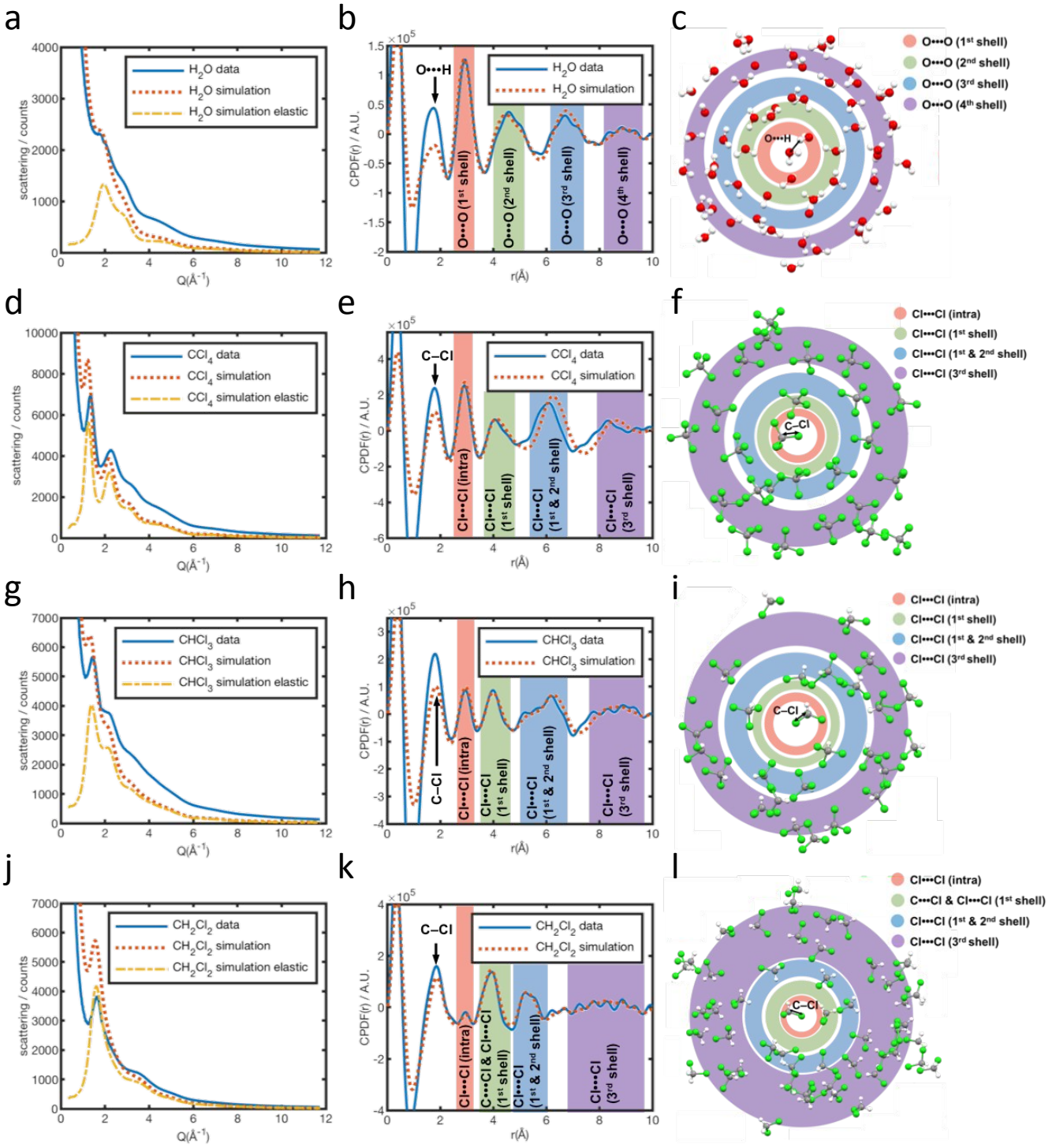

Figure 4. Scattering pattern and *CPDF* for four different liquids. Radially averaged electron scattering patterns (a, d, g, j), *CPDF* (b, e, h, k), and a graphic representation of atom pairs (c, f, i, l) of 4 liquids from experimental data and simulation. Q_{max} =11.8 Å^{-1} is used in both experimental and simulated patterns. A damping factor of α = 0.06 is used for panel (b), (e), (h), α = 0.05 is used for panel (k). All experimental data are taken with a gas-accelerated liquid jet.

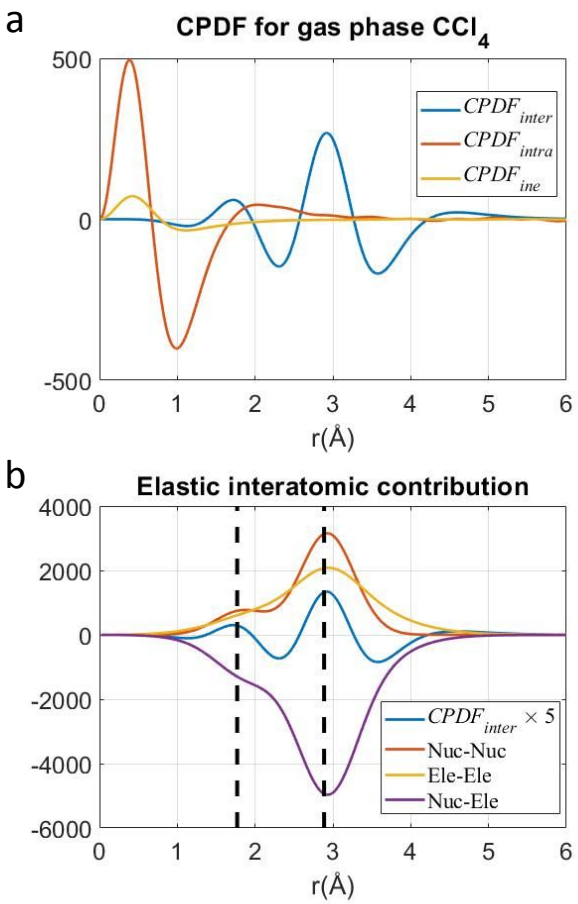

Figure 5. Simulated *CPDF* for gas phase CCl_4 molecule. (a) The elastic interatomic ($CPDF_{inter}$), elastic intraatomic ($CPDF_{intra}$), and inelastic ($CPDF_{ine}$) contributions to *CPDF*. (b) $CPDF_{inter}$ with nucleus-nucleus, nucleus-electron, and electron-electron contributions**.** Scattering patterns are simulated under IAM, and *CPDF* are calculated using Eq. (19) in the main text with damping factor α = 0.06 and Q_{max} =11.8 $Å^{-1}$. The black dashed line shows the internuclear separation for C-Cl (1.77 Å) and Cl···Cl (2.89 Å), taken from ref.[36].

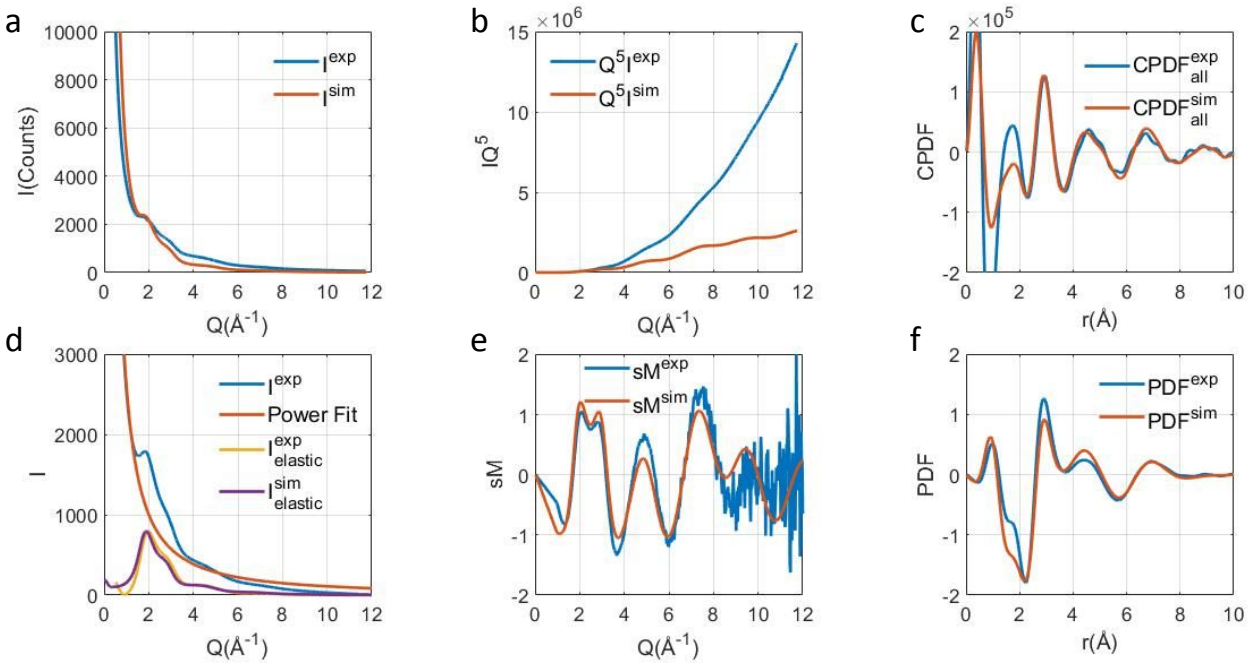

Figure 6. Comparison of *CPDF* (a-c) and *PDF* (d-f) analysis on water. (a) Raw scattering intensity *I*, (b) Q^5I, (c) *CPDF*, (d) experimental raw scattering intensity, empirical power fit, elastic components obtained by removing the power fit, and simulated elastic component, (e) *sM*, (f) *PDF*. Exp, experimental; sim, simulated. Parts (d-f) are adapted from Nunes et al.[16].