

# Extracting the electronic structure signal from X-ray and electron scattering in the gas phase

Thomas Northey,<sup>a</sup> Adam Kirrander<sup>b\*</sup> and Peter M. Weber<sup>a\*</sup>

<sup>a</sup>Department of Chemistry, Brown University, Providence, RI 02912, USA, and <sup>b</sup>Physical and Theoretical Chemistry Laboratory, Department of Chemistry, University of Oxford, South Parks Road, Oxford OX1 3QZ, United Kingdom.

\*Correspondence e-mail: adam.kirrander@chem.ox.ac.uk, peter\_weber@brown.edu

Received 1 November 2023

Accepted 3 January 2024

Edited by K. Kvashnina, ESRF – The European Synchrotron, France

This article is part of a collection of articles from the IUCr 2023 Congress in Melbourne, Australia, and commemorates the 75th anniversary of the IUCr.

**Keywords:** X-ray scattering; electron scattering; gas phase; *ab initio*.

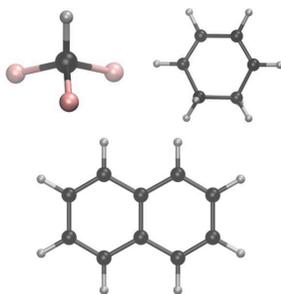
X-ray and electron scattering from free gas-phase molecules is examined using the independent atom model (IAM) and *ab initio* electronic structure calculations. The IAM describes the effect of the molecular geometry on the scattering, but does not account for the redistribution of valence electrons due to, for instance, chemical bonding. By examining the total, *i.e.* energy-integrated, scattering from three molecules, fluoroform (CHF<sub>3</sub>), 1,3-cyclohexadiene (C<sub>6</sub>H<sub>8</sub>) and naphthalene (C<sub>10</sub>H<sub>8</sub>), the effect of electron redistribution is found to predominantly reside at small-to-medium values of the momentum transfer ( $q \leq 8 \text{ \AA}^{-1}$ ) in the scattering signal, with a maximum percent difference contribution at  $2 \leq q \leq 3 \text{ \AA}^{-1}$ . A procedure to determine the molecular geometry from the large- $q$  scattering is demonstrated, making it possible to more clearly identify the deviation of the scattering from the IAM approximation at small and intermediate  $q$  and to provide a measure of the effect of valence electronic structure on the scattering signal.

## 1. Introduction

### 1.1. Background on gas-phase X-ray and electron scattering

Scattering has provided an indispensable tool in advancing our understanding of the structure of matter (von Laue, 1915; Bragg & Bragg, 1915; Watson & Crick, 1953; Perutz *et al.*, 1960). Gas-phase scattering from molecules was a key component in early advances (Debye, 1915; Debye *et al.*, 1929; Debye, 1930; Pirene, 1939, 1946) and the invention of X-ray free-electron lasers (XFELs) has sparked a renewed interest in gas-phase X-ray scattering (Küpper *et al.*, 2014; Minitti *et al.*, 2014), not the least in the domain of ultrafast X-ray scattering (Minitti *et al.*, 2015; Glowina *et al.*, 2016; Kirrander *et al.*, 2016; Stankus *et al.*, 2019; Ruddock *et al.*, 2019; Yong *et al.*, 2020, 2021a; Gabalski *et al.*, 2022; Odate *et al.*, 2023), alongside advances in ultrafast electron diffraction (UED) experimental capabilities (King *et al.*, 2005; Sciaini & Miller, 2011; Weathersby *et al.*, 2015; Zandi *et al.*, 2017). We note that in the context of scattering, X-ray and electron scattering are close analogues (Stefanou *et al.*, 2017; Ma *et al.*, 2020) and in this paper we consider both.

A chemical bonding effect has been observed in the electron scattering signal of molecules in the gas phase, observed mainly at small scattering angles (Iijima *et al.*, 1965; Fink & Kessler, 1967; Duguet & Jaegle, 1975; Hirota *et al.*, 1981; An-Ding & Xiao-Lei, 1995). This has been theoretically studied to quantify the effect in a number of molecules (Bonham & Iijima, 1963; Wang *et al.*, 1994; Hoffmeyer *et al.*, 1998; Shibata *et al.*, 1999, 2002). We discuss this effect in the total X-ray and electron scattering of gas-phase molecules, further quantifying the error that occurs in the structure determination, and



proposing an approach to separate out the molecular structure contribution in the signal from the contribution from the electronic structure.

## 1.2. Aim of the paper

The aim of this paper is to establish a procedure to identify the valence electronic structure contribution to the molecular scattering signal in the gas-phase and to determine the molecular structure without resorting to full *ab initio* calculations of the scattering signal. The valence contribution is characteristic of the redistribution of electrons away from spherical atom-centred distributions, predominantly due to electrons in the valence shells forming molecular orbitals as in chemical bonding and it may be comparatively localized in  $q$  (Bredtmann *et al.*, 2014). Although scattering is commonly viewed as a method to probe molecular geometry, X-rays scatter from all the electrons in the target and thus the scattering relates to the electron density (Ben-Nun *et al.*, 1996; Kirrander, 2012; Suominen & Kirrander, 2014; Northey *et al.*, 2014, 2016; Northey & Kirrander, 2019), and even the pairwise correlation between electrons (Moreno Carrascosa *et al.*, 2019, 2022; Zotev *et al.*, 2020). As a consequence, effects such as the redistribution of electrons due to chemical bonding, the delocalization of electrons in aromatic rings, and the localization of electrons in valence molecular orbitals appear in the scattering signal. Time-resolved experiments can thus detect the rearrangement of electrons due to photoexcitation (Yong *et al.*, 2020), changes in inelastic scattering due to changes in electronic state populations (Yang *et al.*, 2020), dynamic charge transfer (Yong *et al.*, 2021b), the breaking of chemical bonds (Ruddock *et al.*, 2019), or carry out tomography on rotational wave packets (Zhang *et al.*, 2021). Interesting future directions for scattering experiments are also actively explored via simulations (Mu *et al.*, 2023; Bertram *et al.*, 2023). Given sufficient time-resolution, it should become possible to track the dynamics of electrons in a molecular system (Simmermacher *et al.*, 2019a,b; Ziemis *et al.*, 2023). These opportunities apply to both X-ray and electron scattering, however with additional terms in the electron scattering due to the interactions between the incoming electrons and the nuclei (see Section 2).

For gas-phase scattering, the absence of a crystalline lattice means that energy-integrating detectors pick up the total, rather than just the elastic scattering. In this paper, we use the independent atom model (IAM) approximation (Debye *et al.*, 1929; Debye, 1930) of the scattering signal as a baseline that does not include any redistribution of electrons due to bonding, as the IAM assumes a non-interacting spherical distribution of electrons around each individual atom. These results are compared with accurate *ab initio* calculations of the total X-ray scattering that fully account for the redistribution of electrons (Moreno Carrascosa *et al.*, 2019; Zotev *et al.*, 2020; Carrascosa *et al.*, 2022). The difference between the IAM and the *ab initio* signal is identified as the valence electronic structure component. However, it is important to note that the exact molecular geometry is not necessarily known *a priori*.

We therefore require a procedure to determine the molecular geometry as accurately as possible before the electronic component can be calculated. We show in this paper that the molecular geometry can be reliably determined from the large values of the momentum transfer  $q$ , while small and intermediate  $q$  values are most affected by electronic effects. In doing this, we use a recently developed simulated annealing algorithm to fit the molecular geometry to a target X-ray signal (Northey *et al.*, 2024) for various ranges of the momentum transfer vector  $q$ .

## 2. Methods

### 2.1. X-ray and electron scattering

**2.1.1. *Ab initio* calculation.** The *ab initio* X-ray and electron scattering calculations were carried out using an in-house code from the Kirrander group (Northey *et al.*, 2014; Moreno Carrascosa *et al.*, 2019; Zotev *et al.*, 2020) that calculates the scattering signal directly from the molecular wavefunction expressed in a Gaussian basis and obtained via *ab initio* electronic structure methods, such as Hartree–Fock (HF) or multiconfigurational wavefunction methods (CASSCF, MRCI, MCCI *etc.*). The code calculates the total scattering, *i.e.* both the elastic and inelastic components of the signal. In this paper, HF theory with the 6-31G\* Pople basis set is used.

**2.1.2. Independent atom model.** According to the IAM, for X-ray scattering from an isotropic ensemble of  $N_{\text{at}}$ -atom molecules the total scattering intensity is

$$I_{\text{IAM}}(q) = \sum_{i=1}^{N_{\text{at}}} f_i(q)^2 + \sum_{i \neq j}^{N_{\text{at}}} f_i(q) f_j(q) \frac{\sin(qR_{ij})}{qR_{ij}} + \sum_{i=1}^{N_{\text{at}}} S_i(q), \quad (1)$$

where the first sum constitutes the atomic scattering,  $I_{\text{at}}(q)$ , and runs over all atoms  $i$  with tabulated atomic X-ray scattering factors  $f_i(q)$  (Prince, 2006). The first sum contains no structural information about the molecule; instead, structure is contained in the second molecular scattering term,  $I_{\text{mol}}(q)$ , a double sum which runs over all pairs of atoms  $i$  and  $j$  (excluding  $i = j$ ). This term involves the distance between atoms,  $R_{ij} = |\mathbf{R}_i - \mathbf{R}_j|$ , where  $\mathbf{R}_i$  and  $\mathbf{R}_j$  are the positions of atoms  $i$  and  $j$ , respectively. The final term accounts for inelastic scattering and is a sum of tabulated inelastic scattering factors,  $S_i(q)$ . The amplitude of the scattering vector is  $q = |\mathbf{q}|$ , defined as  $\mathbf{q} = \mathbf{k}_1 - \mathbf{k}_0$ , with  $\mathbf{k}_1$  and  $\mathbf{k}_0$  the wavevectors of the scattered and incident X-ray photons. Finally, we note that equation (1) is appropriate for scattering from rotationally averaged free gas-phase molecules in their electronic ground state, as considered in this paper.

For electron scattering, the Mott–Bethe formula (Mott, 1930) can be used to transform the X-ray atomic factors to electron factors,

$$f_i^{\text{electron}}(s) \propto \frac{Z_i - f_i^{\text{xray}}(s)}{s^2}, \quad (2)$$

with proportionality constant  $2m_e c^2 / \hbar$ . This means that the IAM electron scattering is very similar to the X-ray scattering equation aside from the  $1/s^2$  and  $Z_i$  terms, where  $Z_i$  is the

atomic number of atom  $i$ , due to the additional scattering of the electrons by the positive charge of the nuclei. By convention, in electron scattering the scattering vector is denoted as  $s$  instead of  $q$ .

In this paper, comparison between *ab initio* and IAM scattering is in terms of the percent difference, defined as

$$\% \Delta I(q) = 100 \times \frac{I_{\text{IAM}}(q) - I_{\text{abinitio}}(q)}{I_{\text{abinitio}}(q)}. \quad (3)$$

Note that this is a percentage, and that the reference signal  $I_{\text{abinitio}}(q)$  is subtracted from  $I_{\text{IAM}}(q)$  and also appears in the denominator.

## 2.2. Fitting procedure

A recently developed simulated annealing (SA) approach is used to fit the predicted IAM signal to the target data. This approach is described in detail by Northey *et al.* (2024). It minimizes the target function,

$$\zeta_{\text{targ}} = \zeta_{\text{signal}} + \zeta_{\text{aux}}, \quad (4)$$

by changing the molecular geometry  $\mathbf{R}'$  and recalculating  $\zeta_{\text{targ}}$  iteratively, where

$$\zeta_{\text{signal}} = \frac{[I_{\text{IAM}}(q; \mathbf{R}') - I_{\text{abinitio}}(q; \mathbf{R}_{\text{targ}})]^2}{I_{\text{abinitio}}(q; \mathbf{R}_{\text{targ}})}, \quad (5)$$

for the predicted X-ray (or electron) IAM scattering signal which depends on  $\mathbf{R}'$ , and the target signal is calculated using the *ab initio* scattering code, where  $\mathbf{R}_{\text{targ}}$  is the target geometry. Auxiliary harmonic oscillator terms  $\zeta_{\text{aux}} = \sum_{ij} A_{ij} (d'_{ij} - d_{ij}^{\text{start}})^2$  are included in  $\zeta_{\text{targ}}$ , with interatomic distances  $d'_{ij}$  and  $d_{ij}^{\text{start}}$  between atoms  $i$  and atom  $j$  corresponding to distances from  $\mathbf{R}'$  and the starting geometry  $\mathbf{R}_{\text{start}}$ , respectively. The index  $i = 1, 2, \dots, N_{\text{at}}$  runs over each atom in the molecule, whereas index  $j \neq i$  runs over each nearest-neighbour and second-nearest-neighbour atom (counting from atom  $i$ ). This restrains the molecular geometry, ensuring that the simulated annealing algorithm does not waste effort exploring unphysical regions of the conformational space. The auxiliary terms have weighting factors  $A_{ij}$  which are tuned such that the scattering term  $\zeta_{\text{signal}}$  is the predominant driving force in the minimization, *i.e.*  $\zeta_{\text{signal}}/\zeta_{\text{aux}} \simeq 10$ .

The fitting procedure minimizes the squared-difference functions contained in equation (4) iteratively by randomly perturbing the molecular geometry along normal modes and accepting perturbations if the fit improves. The method is robust and can escape local minima by accepting non-favourable (uphill) steps with non-zero probability, corresponding to an effective temperature. Simulated annealing works quite well here, but other optimization methods should be capable of reproducing the same results. Notably, the focus of the current paper is not the overall optimal structure determination but rather to establish the information content in various  $q$ -ranges of the scattering signal. Additional data, for instance from other experiments or *ab initio* calculations, which are exploited in structural retrieval methods such as the SARACEN method developed by Mitzel & Rankin (2003),

would be counterproductive in this context as they would distort this analysis. In this work, the target geometry is the Hartree-Fock(HF)/6-31G\* ground state optimized geometry,  $\mathbf{R}_{\text{targ}} = \mathbf{R}_0$ , calculated using *MOLPRO* (Werner *et al.*, 2012). Notably, in an experiment, the target geometry would not be known *a priori*; the goal is to find it by sampling around a reasonable initial guess.

The starting geometry for the fitting procedure is initialized by small random perturbations along each mode, away from  $\mathbf{R}_0$ . Thus, the starting geometry is close to the target geometry,  $\mathbf{R}_{\text{start}} \simeq \mathbf{R}_{\text{targ}}$ , and the fitting procedure predominantly depends on the difference between the IAM and *ab initio* signals rather than an extensive conformational search. A frequency calculation on the ground state  $\mathbf{R}_0$  geometry is performed to obtain the normal mode unit vectors, which are used to sample all dimensions of the nuclear coordinate space when minimizing the target function. The final molecular geometry  $\mathbf{R}_{\text{best}}$  is determined from the minimum of  $\zeta_{\text{targ}}$ . Due to the stochastic nature of the SA algorithm, it was run 20 times for each fitting and the outcome with the lowest  $\zeta_{\text{targ}}$  is selected. This avoids getting stuck in higher-lying minima and increases the probability that a structure close to the global minimum is obtained.

A metric used in this work to assess how close a given molecular geometry is to the reference geometry is the mean absolute percent deviation (MAPD) (Yong *et al.*, 2021c), defined as

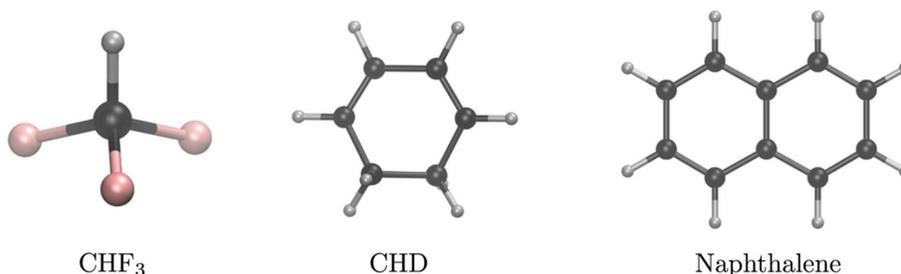
$$\text{MAPD} = 100 \times \frac{2}{N_{\text{at}}^{\circ}(N_{\text{at}}^{\circ} - 1)} \sum_{i=1}^{N_{\text{at}}^{\circ}} \sum_{j>i}^{N_{\text{at}}^{\circ}} \frac{|d_{ij} - d'_{ij}|}{d'_{ij}}, \quad (6)$$

for the total number of atoms of interest  $N_{\text{at}}^{\circ}$ , where  $N_{\text{at}}^{\circ} \leq N_{\text{at}}$ . In this work, the non-hydrogen atoms are chosen in the calculation of the MAPD, *i.e.*  $N_{\text{at}}^{\circ}$  equals the number of non-hydrogen atoms. The distances between atom  $i$  and atom  $j$  are  $d_{ij}$  and  $d'_{ij}$ , where the prime (') denotes the reference structure, which is the ground state optimized structure  $\mathbf{R}_0$  unless otherwise specified. In the following, we proceed to consider three molecules: fluoroform ( $\text{CHF}_3$ ), 1,3-cyclohexadiene ( $\text{C}_6\text{H}_{10}$ , CHD) and naphthalene ( $\text{C}_{10}\text{H}_8$ ), which are shown in Fig. 1.

## 3. Results

### 3.1. IAM and *ab initio* scattering calculations

The top panels in Fig. 2 show a comparison between IAM and *ab initio* X-ray (top) and electron (bottom) scattering for the three molecules in Fig. 1. The  $I(q)$  signal is multiplied by  $q$  to enhance the visibility of the signal at large  $q$  [see the unscaled scattering intensity  $I(q)$  plotted in Fig. S1 of the supporting information]. The  $I(s)$  signal is multiplied by  $s^4$  for the same reason. Often in electron scattering experiments,  $M(s) = I_{\text{mol}}(s)/I_{\text{at}}(s)$  is plotted, which cancels out the  $s^{-4}$  Rutherford scattering term in  $I(s)$ ; however, the *ab initio*  $I(s)$  signal in this work cannot be decomposed into separate  $I_{\text{mol}}(s)$  and  $I_{\text{at}}(s)$  terms, so  $s^4 I(s)$  is shown instead, cancelling the Rutherford scaling. The bottom panels in Fig. 2 show the


**Figure 1**

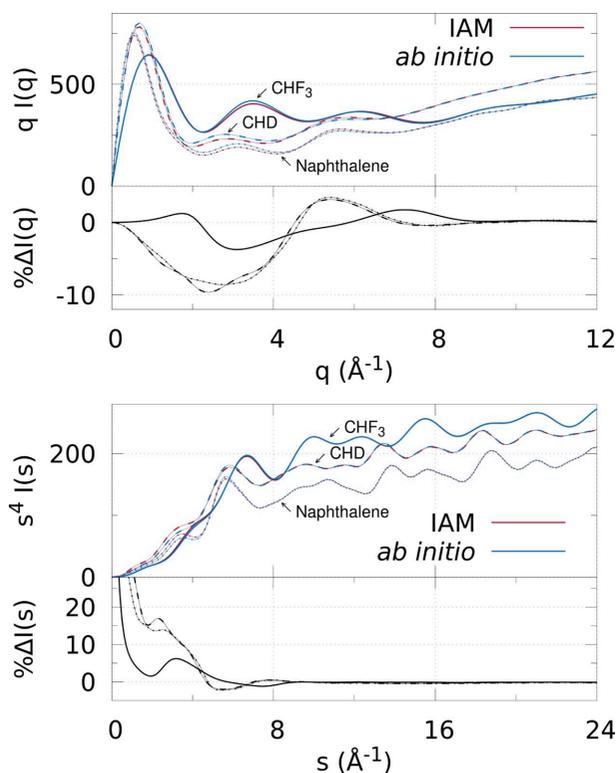
The three molecules fluoroform (CHF<sub>3</sub>), 1,3-cyclohexadiene (C<sub>6</sub>H<sub>8</sub>, CHD) and naphthalene (C<sub>10</sub>H<sub>8</sub>).

percent difference [as in equation (3)] between IAM and *ab initio* scattering.

For X-ray scattering, all three molecules are similar in that the IAM underestimates signal in the approximate range  $0 \leq q \leq 4.6 \text{ \AA}^{-1}$  ( $0 \leq q \leq 5.9 \text{ \AA}^{-1}$  for CHF<sub>3</sub>), albeit in CHF<sub>3</sub> IAM slightly overestimates for  $q < 2.1 \text{ \AA}^{-1}$ , and in all three IAM overestimates the scattering at larger  $q$  (approximately  $4.6 \leq q \leq 8.4 \text{ \AA}^{-1}$  for CHD and naphthalene, and approximately  $5.9 \leq q \leq 8.8 \text{ \AA}^{-1}$  for CHF<sub>3</sub>) up until  $q \gtrsim 8 \text{ \AA}^{-1}$  when IAM and *ab initio* become very similar; excellent agreement

is seen here with  $|\% \Delta I(q > 8)| < 0.3\%$ . The peak percent differences for CHD and naphthalene are relatively large, with  $\% \Delta I(q = 2.4) = -9.6\%$  and  $\% \Delta I(q = 2.7) = -8.6\%$ , respectively. They are pure hydrocarbons which have been reported to contain larger chemical bonding effects compared with molecules such as CCl<sub>4</sub>, N<sub>2</sub>, O<sub>2</sub> and CS<sub>2</sub> (Shibata *et al.*, 2002). This is due to delocalization of electrons by chemical bonding, which the IAM does not take into account, including double and triple bonds, aromatic rings, and hydrogen atoms bonded to heavier atoms. Conversely, CHF<sub>3</sub> has a smaller peak percentage difference of  $\% \Delta I(q = 3.0) = -3.7\%$  due to the three polarized C–F single bonds and an absence of double bonds or delocalized electrons, which means that the molecule is quite well described by IAM.

Similar to X-ray scattering, the electron scattering results show a substantial difference between IAM and *ab initio*, with the IAM performing the best for CHF<sub>3</sub>, whereas CHD and naphthalene have larger peaks in percentage difference  $\% \Delta I(s)$ . The maximum peak differences for each molecule are: CHF<sub>3</sub> has  $\% \Delta I(s = 3.2) = 6.2\%$ , CHD has  $\% \Delta I(s = 2.3) = 17.0\%$ , and naphthalene has  $\% \Delta I(s = 2.5) = 13.9\%$ . These peak percent difference values are similar to the X-ray scattering result in terms of magnitude and molecule order, showing that the redistribution of electrons away from atom-centred spherical distributions is similarly visible. Finally, at  $8 < s < 24 \text{ \AA}^{-1}$  the mean absolute percent difference is  $\langle |\% \Delta I(s)| \rangle < 0.3\%$  for each molecule, comparable with the X-ray scattering results at  $8 < q < 12 \text{ \AA}^{-1}$ .

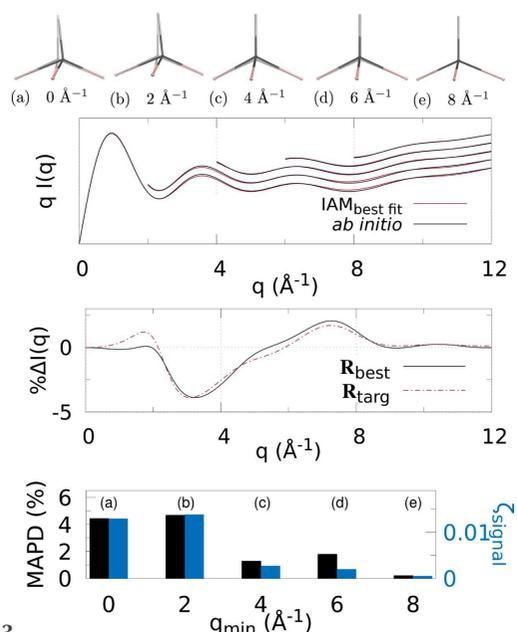

**Figure 2**

Comparison between IAM and *ab initio* scattering (top figure: X-ray scattering; bottom figure: electron scattering) for the three molecules CHF<sub>3</sub>, CHD and naphthalene in their optimized ground-state geometries  $\mathbf{R}_0$ . The scattering signals are shown in the top panels (IAM by red lines and *ab initio* by blue), with the signals scaled by  $q$  and  $s^4$ , respectively, *i.e.*  $qI(q)$  and  $s^4I(s)$ . In both top and bottom figures the naphthalene signal has been multiplied by 0.5 for visualization purposes. The bottom panels show percent differences between *ab initio* and IAM scattering, as defined in equation (3). Note that the electron scattering  $\% \Delta I(s)$  becomes very large at small  $s$  due to division by small numbers, so the y-axis is truncated.

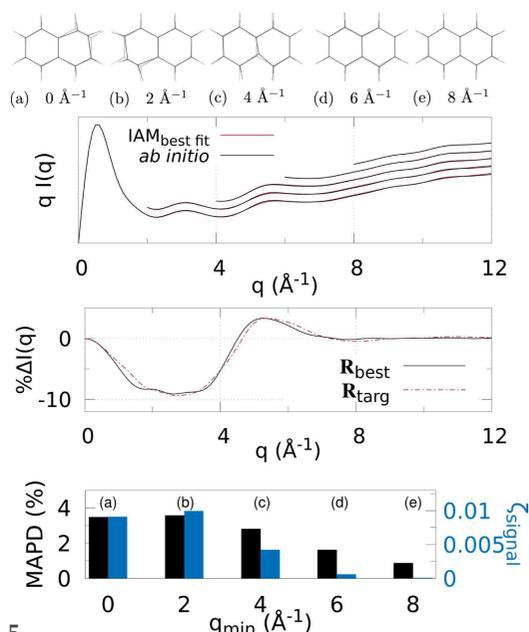
### 3.2. Fitting to the target signal

Figs. 3–5 and 6–8 show the results of fitting the IAM signals to *ab initio* X-ray and electron scattering signals, respectively. The predicted data are  $I_{\text{IAM}}(q)$  for the X-ray scattering, as in equation (1) [not  $qI(q)$  as shown in the figures], and  $s^4I_{\text{IAM}}(s)$  for electron scattering. The corresponding target data is calculated by *ab initio* scattering theory at  $\mathbf{R}_{\text{targ}} = \mathbf{R}_0$  for both X-ray and electron scattering, *i.e.*  $I_{\text{abinitio}}(q)$  or  $s^4I_{\text{abinitio}}(s)$ , respectively.

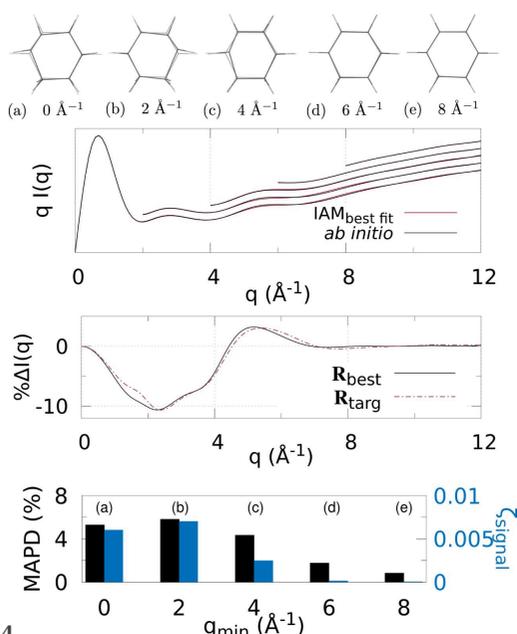
Various  $q$  (or  $s$ ) ranges were fitted to assess the accuracy of IAM as a function of the values of the momentum transfer, to find the region most suitable to determine the molecular geometry with minimal non-IAM contamination. This, correspondingly, is the region where the valence electronic struc-


**Figure 3**

X-ray scattering for  $\text{CHF}_3$  using different  $q$ -ranges for structure determination: (a)–(e) The best fit geometry for each  $q_{\min}$  (solid) versus the target structure  $\mathbf{R}_{\text{targ}}$  (translucent), where  $q \in [q_{\min}, q_{\max}]$ , and  $q_{\min} = [0, 2, 4, 6, 8] \text{ \AA}^{-1}$  and  $q_{\max} = 12 \text{ \AA}^{-1}$ . (Top) The IAM best fits to the *ab initio* X-ray scattering calculated at  $\mathbf{R}_{\text{targ}}$  as a function of  $q$ -range. The  $q_{\min} \geq 2 \text{ \AA}^{-1}$  curves are shifted vertically for visibility. (Middle) The percent difference  $\% \Delta I(q)$  [equation (3)] (solid black line) for the lowest  $\zeta_{\text{targ}}$  structure. For comparison,  $\% \Delta I(q; \mathbf{R}_{\text{targ}})$  is also shown, using  $I_{\text{IAM}}(q; \mathbf{R}_{\text{targ}})$  and  $I_{\text{abinitio}}(q; \mathbf{R}_{\text{targ}})$ . (Bottom) Bar chart showing the MAPD [equation (6)] and  $\zeta_{\text{signal}}$  for each value of  $q_{\min}$ .


**Figure 5**

X-ray scattering of naphthalene using different  $q$ -ranges for structure determination: (a)–(e) The best fit geometry for each  $q_{\min}$  (solid) versus the target structure  $\mathbf{R}_{\text{targ}}$  (translucent), where  $q \in [q_{\min}, q_{\max}]$ , and  $q_{\min} = [0, 2, 4, 6, 8] \text{ \AA}^{-1}$  and  $q_{\max} = 12 \text{ \AA}^{-1}$ . (Top) The IAM best fits to the *ab initio* X-ray scattering calculated at  $\mathbf{R}_{\text{targ}}$  as a function of  $q$ -range. The  $q_{\min} \geq 2 \text{ \AA}^{-1}$  curves are shifted vertically for visibility. (Middle) The percent difference  $\% \Delta I(q)$  [equation (3)] (solid black line) for the lowest  $\zeta_{\text{targ}}$  structure. For comparison,  $\% \Delta I(q; \mathbf{R}_{\text{targ}})$  is also shown, using  $I_{\text{IAM}}(q; \mathbf{R}_{\text{targ}})$  and  $I_{\text{abinitio}}(q; \mathbf{R}_{\text{targ}})$ . (Bottom) Bar chart showing the MAPD [equation (6)] and  $\zeta_{\text{signal}}$  for each value of  $q_{\min}$ .

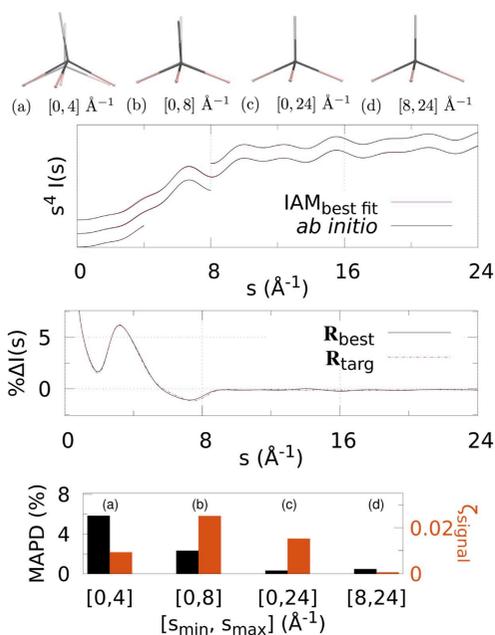

**Figure 4**

X-ray scattering of CHD using different  $q$ -ranges for structure determination: (a)–(e) The best fit geometry for each  $q_{\min}$  (solid) versus the target structure  $\mathbf{R}_{\text{targ}}$  (translucent), where  $q \in [q_{\min}, q_{\max}]$ , and  $q_{\min} = [0, 2, 4, 6, 8] \text{ \AA}^{-1}$  and  $q_{\max} = 12 \text{ \AA}^{-1}$ . (Top) The IAM best fits to the *ab initio* X-ray scattering calculated at  $\mathbf{R}_{\text{targ}}$  as a function of  $q$ -range. The  $q_{\min} \geq 2 \text{ \AA}^{-1}$  curves are shifted vertically for visibility. (Middle) The percent difference  $\% \Delta I(q)$  [equation (3)] (solid black line) for the lowest  $\zeta_{\text{targ}}$  structure. For comparison,  $\% \Delta I(q; \mathbf{R}_{\text{targ}})$  is also shown, using  $I_{\text{IAM}}(q; \mathbf{R}_{\text{targ}})$  and  $I_{\text{abinitio}}(q; \mathbf{R}_{\text{targ}})$ . (Bottom) Bar chart showing the MAPD [equation (6)] and  $\zeta_{\text{signal}}$  for each value of  $q_{\min}$ .

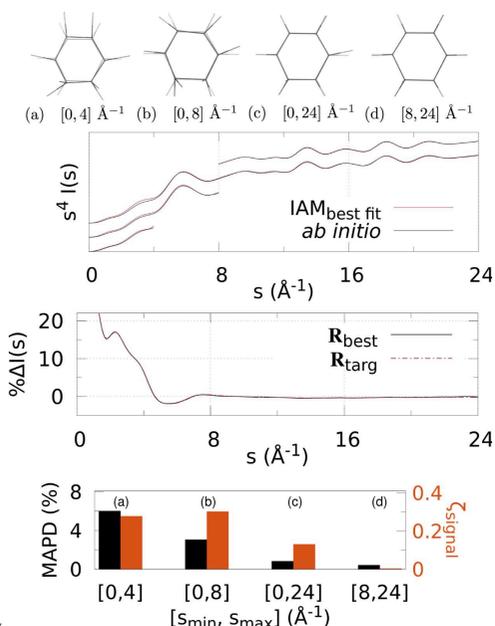
ture component of the signal is strongest (that is, where the IAM and *ab initio* signals are significantly different).

At the top of each figure the best-fit molecular geometries  $\mathbf{R}_{\text{best}}$  (in solid) are shown together with the reference geometries  $\mathbf{R}_{\text{targ}}$  (translucent) for the various  $q$  (or  $s$ ) ranges. For X-ray scattering, the top graph shows the best-fit signal scaled by  $q$ ,  $qI_{\text{IAM}}(q; \mathbf{R}_{\text{best}})$ , compared with  $qI_{\text{abinitio}}(q; \mathbf{R}_{\text{targ}})$  for each signal range  $q_{\min} \leq q \leq q_{\max}$ . For electron scattering, similarly the best fits for each range are shown  $s^4 I_{\text{IAM}}(s; \mathbf{R}_{\text{best}})$  compared with  $s^4 I_{\text{abinitio}}(s; \mathbf{R}_{\text{targ}})$  for  $s_{\min} \leq s \leq s_{\max}$ . The curves for each range of  $q$  or  $s$  used in the fitting are shifted vertically for visualization purposes.

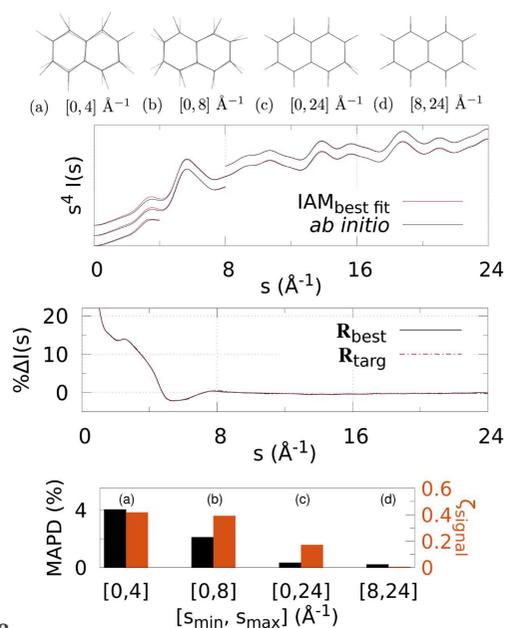
The bar charts at the bottom summarize the fitting for each signal range in terms of  $\zeta_{\text{signal}}$  (notably, not  $\zeta_{\text{targ}}$ ) as in equation (4), and MAPD [equation (6)]. Tables 1 and 2 show the values from these charts. It is clear that  $8 \leq q \leq 12 \text{ \AA}^{-1}$  finds the molecular geometry closest to  $\mathbf{R}_{\text{targ}}$ ; in fact, it is exceptionally close to the target, with mean absolute atom–atom distance deviations  $\leq 0.9\%$  for all three molecules. Similarly, for electron scattering, the range  $8 \leq s \leq 24 \text{ \AA}^{-1}$  gives the best geometry with  $\text{MAPD} \leq 0.5\%$ . This can be seen clearly from the overlap of the solid and translucent structures in (e) [or (d) for Figs. 6–8], whereas (a), (b), etc. have quite large visible deviations from  $\mathbf{R}_{\text{targ}}$ , such as stretching of C–F bonds, aromatic C–C bonds compacted/stretched, and different hydrogen positions (and C–H distances). Importantly, MAPD and  $\zeta_{\text{signal}}$  are generally correlated (aside from the smallest electron scattering range  $0 \leq s \leq 4 \text{ \AA}^{-1}$  outliers), and



**Figure 6**  
 Electron scattering of CHF<sub>3</sub> using different  $s$ -ranges for structure determination: (a)–(d) The best fit geometry for each  $s$ -range (solid) versus the reference structure  $\mathbf{R}_{\text{targ}}$  (translucent), where  $s \in [s_{\text{min}}, s_{\text{max}}] = [0, 4], [0, 8], [0, 24], [8, 24] \text{ \AA}^{-1}$ . (Top) The IAM best fits to the *ab initio* electron scattering calculated at  $\mathbf{R}_{\text{targ}}$  as a function of  $s$ -range. The  $s_{\text{max}} > 4 \text{ \AA}^{-1}$  curves are shifted vertically for visibility. (Middle) The percent difference  $\% \Delta I(s)$  [equation (3)] (solid black line) for the lowest  $\zeta_{\text{targ}}$  structure. For comparison,  $\% \Delta I(s; \mathbf{R}_{\text{targ}})$  is also shown, using  $I_{\text{IAM}}(s; \mathbf{R}_{\text{targ}})$  and  $I_{\text{abinitio}}(s; \mathbf{R}_{\text{targ}})$ . (Bottom) Bar chart showing the MAPD [equation (6)] and  $\zeta_{\text{signal}}$  for each value of  $[s_{\text{min}}, s_{\text{max}}]$ .



**Figure 7**  
 Electron scattering of CHD using different  $s$ -ranges for structure determination: (a)–(d) The best fit geometry for each  $s$ -range (solid) versus the reference structure  $\mathbf{R}_{\text{targ}}$  (translucent), where  $s \in [s_{\text{min}}, s_{\text{max}}] = [0, 4], [0, 8], [0, 24], [8, 24] \text{ \AA}^{-1}$ . (Top) The IAM best fits to the *ab initio* electron scattering calculated at  $\mathbf{R}_{\text{targ}}$  as a function of  $s$ -range. The  $s_{\text{max}} > 4 \text{ \AA}^{-1}$  curves are shifted vertically for visibility. (Middle) The percent difference  $\% \Delta I(s)$  [equation (3)] (solid black line) for the lowest  $\zeta_{\text{targ}}$  structure. For comparison,  $\% \Delta I(s; \mathbf{R}_{\text{targ}})$  is also shown, using  $I_{\text{IAM}}(s; \mathbf{R}_{\text{targ}})$  and  $I_{\text{abinitio}}(s; \mathbf{R}_{\text{targ}})$ . (Bottom) Bar chart showing the MAPD [equation (6)] and  $\zeta_{\text{signal}}$  for each value of  $[s_{\text{min}}, s_{\text{max}}]$ .



**Figure 8**  
 Electron scattering of naphthalene using different  $s$ -ranges for structure determination: (a)–(d) The best fit geometry for each  $s$ -range (solid) versus the reference structure  $\mathbf{R}_{\text{targ}}$  (translucent), where  $s \in [s_{\text{min}}, s_{\text{max}}] = [0, 4], [0, 8], [0, 24], [8, 24] \text{ \AA}^{-1}$ . (Top) The IAM best fits to the *ab initio* electron scattering calculated at  $\mathbf{R}_{\text{targ}}$  as a function of  $s$ -range. The  $s_{\text{max}} > 4 \text{ \AA}^{-1}$  curves are shifted vertically for visibility. (Middle) The percent difference  $\% \Delta I(s)$  [equation (3)] (solid black line) for the lowest  $\zeta_{\text{targ}}$  structure. For comparison,  $\% \Delta I(s; \mathbf{R}_{\text{targ}})$  is also shown, using  $I_{\text{IAM}}(s; \mathbf{R}_{\text{targ}})$  and  $I_{\text{abinitio}}(s; \mathbf{R}_{\text{targ}})$ . (Bottom) Bar chart showing the MAPD [equation (6)] and  $\zeta_{\text{signal}}$  for each value of  $[s_{\text{min}}, s_{\text{max}}]$ .

the high- $q$  (or  $s$ ) ranges  $8 \leq q \leq 12 \text{ \AA}^{-1}$  and  $8 \leq s \leq 24 \text{ \AA}^{-1}$  have the lowest MAPD and  $\zeta_{\text{signal}}$  in all cases except for one outlier in the CHF<sub>3</sub> electron scattering where the  $0 \leq s \leq 24 \text{ \AA}^{-1}$  range has slightly lower (albeit very close) MAPD, despite much higher  $\zeta_{\text{signal}}$ . The correlation between MAPD and  $\zeta_{\text{signal}}$  is still promising however, as experimentally we do not know the MAPD (because we do not inherently know the molecular geometry) but do know the value of  $\zeta_{\text{signal}}$  from comparison with theory. Therefore, fitting the X-ray (or electron) scattering curve via minimization of  $\zeta_{\text{signal}}$  (or equivalent) should give a structure close to the true structure.

Finally, the middle graphs shows the valence electronic structure component as a percentage,  $\% \Delta I(q \text{ or } s)$ , for the best-fit geometry,  $\mathbf{R}_{\text{best}}$ , *i.e.* the structure with lowest  $\zeta_{\text{signal}}$  for the ranges  $8 \leq q \leq 12 \text{ \AA}^{-1}$  and  $8 \leq s \leq 24 \text{ \AA}^{-1}$  for X-ray and electron scattering, respectively, using the *ab initio* signal calculated at  $\mathbf{R}_{\text{targ}}$  as the reference. It is compared with the percent differences between IAM and *ab initio* both at  $\mathbf{R}_{\text{targ}}$  (dashed red line). The results show good agreement between  $\% \Delta I(q \text{ or } s; \mathbf{R}_{\text{best}})$  and  $\% \Delta I(q \text{ or } s; \mathbf{R}_{\text{targ}})$  for both X-ray and electron scattering. Notably, the electron results find near perfect agreement in this regard, revealing that  $\mathbf{R}_{\text{best}}$  is closer to  $\mathbf{R}_{\text{targ}}$  compared with the X-ray data fitting, hinting that electron scattering could be a better tool for molecular structure determination at larger scattering vector amplitudes compared with X-ray scattering. This is likely due to the dominance of elastic electron scattering at high values of  $s$ , whereas high- $q$  X-ray scattering is dominated by unmodulated

**Table 1**

Fitting results between IAM X-ray scattering and *ab initio* calculated at the target geometry  $\mathbf{R}_{\text{targ}}$  at different  $q$ -ranges.

$q_{\text{min}}$  is shown in the first column and  $q_{\text{max}} = 12 \text{ \AA}^{-1}$ , and the data correspond to the bar charts in Figs. 3–5. The target function,  $\zeta_{\text{signal}}$ , is described in equation (4), and the mean absolute percent deviation (MAPD) is given by equation (6)

$q_{\text{min}} (\text{\AA}^{-1})$	$\zeta_{\text{signal}}$	MAPD (%)
CHF <sub>3</sub>		
0	$1.0 \times 10^{-2}$	4.5
2	$9.5 \times 10^{-3}$	4.2
4	$2.7 \times 10^{-3}$	1.3
6	$1.9 \times 10^{-3}$	1.9
8	$4.7 \times 10^{-4}$	0.2
CHD		
0	$6.0 \times 10^{-3}$	5.3
2	$7.0 \times 10^{-3}$	5.8
4	$2.5 \times 10^{-3}$	4.3
6	$1.3 \times 10^{-4}$	1.7
8	$1.6 \times 10^{-5}$	0.8
Naphthalene		
0	$9.1 \times 10^{-3}$	3.5
2	$9.9 \times 10^{-3}$	3.6
4	$4.2 \times 10^{-3}$	2.8
6	$5.8 \times 10^{-4}$	1.6
8	$2.4 \times 10^{-5}$	0.9

**Table 2**

Fitting results between IAM electron scattering and *ab initio* calculated at the target geometry  $\mathbf{R}_{\text{targ}}$  at different  $s$ -ranges.

$s_{\text{min}}$  is shown in the first column and  $s_{\text{max}}$  in the second, and the data correspond to the bar charts in Figs. 6–8. The target function,  $\zeta_{\text{signal}}$ , is described in equation (4), and the mean absolute percent deviation (MAPD) is given by equation (6).

$s_{\text{min}} (\text{\AA}^{-1})$	$s_{\text{max}} (\text{\AA}^{-1})$	$\zeta_{\text{signal}}$	MAPD (%)
CHF <sub>3</sub>			
0	4	$9.3 \times 10^{-3}$	5.8
0	8	$2.5 \times 10^{-2}$	2.3
0	24	$1.5 \times 10^{-2}$	0.3
8	24	$5.1 \times 10^{-4}$	0.5
CHD			
0	4	$2.8 \times 10^{-1}$	6.0
0	8	$3.0 \times 10^{-1}$	3.0
0	24	$1.3 \times 10^{-1}$	0.8
8	24	$1.9 \times 10^{-3}$	0.4
Naphthalene			
0	4	$4.2 \times 10^{-1}$	4.0
0	8	$3.9 \times 10^{-1}$	2.1
0	24	$1.7 \times 10^{-1}$	0.3
8	24	$3.2 \times 10^{-3}$	0.2

inelastic scattering. Despite this, X-ray scattering still performed well in this region, finding low MAPD structures for each molecule. It is striking that the elastic scattering structural information persists here (see Fig. S2 of the supporting information).

## 4. Conclusion

The molecular geometry can be determined from large momentum transfer or equivalently large-angle scattering, with  $q > 8 \text{ \AA}^{-1}$  (in the following,  $q$  also encompasses  $s$ ), with

good results achieved already with the simple IAM approximation if only the large- $q$  data are used. Although the elastic scattering component containing structural information drops off more quickly for X-ray than electron scattering, we find that sufficient elastic scattering persists to retrieve the correct molecular geometry using IAM theory in the range  $8 < q < 12 \text{ \AA}^{-1}$  for both modalities of scattering. An important aspect to note is that in this high  $q$  regime the contribution to the scattering due to chemical bonding is negligible, allowing the structure to be determined reliably using IAM theory. Conversely, if using IAM across the full available range of  $q$ , the resulting structure may be distorted from the correct  $\mathbf{R}_0$  geometry. Another point to note is that the concept of structure may become ill-defined in excited-state dynamics, where the dispersion of the nuclear wave packet leads to the co-existence of a range of structures that manifest as an effective damping of the high- $q$  signal (Kirrander & Weber, 2017).

Nevertheless, for states of reasonably well defined geometry, such as molecules in their ground electronic state, we can determine the molecular geometry using the large- $q$  scattering, allowing us in the next step to extract the contribution to the scattering from the bonding valence electrons, at small and intermediate  $q$ . In this  $q$ -range, we find the deviation from IAM to be significant:  $\sim 10\%$  for CHD and naphthalene, and  $\sim 5\%$  for CHF<sub>3</sub>, for both X-ray and electron scattering.

We note that any robust inversion algorithm that can transform the one-dimensional scattering signal into reasonable molecular geometries would work (Yang *et al.*, 2014; Ishikawa *et al.*, 2015; Acheson & Kirrander, 2023). A practical challenge is that the large- $q$  signal is small and is detected on a background of featureless inelastic scattering, which is demanding in terms of experimental signal-to-noise. The reliable detection of large- $q$  signals might therefore be most appropriate at facilities such as the upgraded LCLS-II where high repetition rates and photon energies upwards of 18 keV can help overcome such shortcomings. Nevertheless, given data of sufficient quality, the procedure outlined in this paper demonstrates that it should be possible to isolate the electronic contributions to the scattering signal, potentially opening the door for exciting new insights into electronic structure (Carrascosa *et al.*, 2022).

## 5. Related literature

The following reference, not cited in the main body of the paper, has been cited in the supporting information: Mai *et al.* (2014).

## Funding information

The following funding is acknowledged: National Science Foundation (award No. CHE-1953839 to Peter M. Weber); Engineering and Physical Sciences Research Council (grants Nos. EP/V049240/2, EP/X026698/1 and EP/X026973/1 to Adam Kirrander); Leverhulme Trust (grant No. RPG-2020-208).

References

Acheson, K. & Kirrander, A. (2023). *J. Chem. Theory Comput.* **19**, 2721–2734.

An-Ding, J. & Xiao-Lei, Z. (1995). *Acta Physico-Chim. Sin.* **11**, 663–666.

Ben-Nun, M., Martínez, T. J., Weber, P. M. & Wilson, K. R. (1996). *Chem. Phys. Lett.* **262**, 405–414.

Bertram, L., Weber, P. M. & Kirrander, A. (2023). *Faraday Discuss.* **244**, 269–293.

Bonham, R. & Iijima, T. (1963). *J. Phys. Chem.* **67**, 2266–2272.

Bragg, W. H. & Bragg, W. L. (1915). Nobel Lecture.

Bredtmann, T., Ivanov, M. & Dixit, G. (2014). *Nat. Commun.* **5**, 5589.

Debye, P. (1915). *Annal. Phys.* **351**, 809–823.

Debye, P. (1930). *Phys. Z.* **31**, 419–428.

Debye, P., Bewilogua, L. & Ehrhardt, F. (1929). *Phys. Z.* **30**, 84.

Duguet, A. & Jaegle, A. (1975). *Chem. Phys. Lett.* **33**, 436–440.

Fink, M. & Kessler, J. (1967). *J. Chem. Phys.* **47**, 1780–1782.

Gabalski, I., Sere, M., Acheson, K., Allum, F., Boutet, S., Dixit, G., Forbes, R., Glownia, J. M., Goff, N., Hegazy, K., Howard, A. J., Liang, M., Minitti, M. P., Minns, R. S., Natan, A., Peard, N., Rasmus, W. O., Sension, R. J., Ware, M. R., Weber, P. M., Werby, N., Wolf, T. J. A., Kirrander, A. & Bucksbaum, P. H. (2022). *J. Chem. Phys.* **157**, 164305.

Glownia, J. M., Natan, A., Cryan, J. P., Hartsock, R., Kozina, M., Minitti, M. P., Nelson, S., Robinson, J., Sato, T., van Driel, T., Welch, G., Weninger, C., Zhu, D. & Bucksbaum, P. H. (2016). *Phys. Rev. Lett.* **117**, 153003.

Hirota, F., Kakuta, N. & Shibata, S. (1981). *J. Phys. B: At. Mol. Phys.* **14**, 3299–3304.

Hoffmeyer, R. E., Bündgen, P. & Thakkar, A. J. (1998). *J. Phys. B At. Mol. Opt. Phys.* **31**, 3675–3692.

Iijima, T., Bonham, R. A., Tavaré, C., Roux, M. & Cornille, M. (1965). *Bull. Chem. Soc. Jpn.* **38**, 1757–1760.

Ishikawa, T., Hayes, S. A., Keskin, S., Corthey, G., Hada, M., Pichugin, K., Marx, A., Hirscht, J., Shionuma, K., Onda, K., Okimoto, Y., Koshihara, S., Yamamoto, T., Cui, H., Nomura, M., Oshima, Y., Abdel-Jawad, M., Kato, R. & Miller, R. J. D. (2015). *Science*, **350**, 1501–1505.

King, W. E., Campbell, G. H., Frank, A., Reed, B., Schmerge, J. F., Siwick, B. J., Stuart, B. C. & Weber, P. M. (2005). *J. Appl. Phys.* **97**, 111101.

Kirrander, A. (2012). *J. Chem. Phys.* **137**, 154310.

Kirrander, A., Saita, K. & Shalashilin, D. V. (2016). *J. Chem. Theory Comput.* **12**, 957–967.

Kirrander, A. & Weber, P. M. (2017). *Appl. Sci.* **7**, 534.

Küpper, J., Stern, S., Holmegaard, L., Filsinger, F., Rouzée, A., Rudenko, A., Johnsson, P., Martin, A. V., Adolph, M., Aquila, A., Bajt, S., Barty, A., Bostedt, C., Bozek, J., Caleman, C., Coffee, R., Coppola, N., Delmas, T., Epp, S., Erk, B., Foucar, L., Gorkhover, T., Gumprecht, L., Hartmann, A., Hartmann, R., Hauser, G., Holl, P., Hömke, A., Kimmel, N., Krasniqi, F., Kühnel, K.-U., Maurer, J., Messerschmidt, M., Moshhammer, R., Reich, C., Rudek, B., Santra, R., Schlichting, I., Schmidt, C., Schorb, S., Schulz, J., Soltau, H., Spence, J. C. H., Starodub, D., Strüder, L., Thøgersen, J., Vrakking, M. J. J., Weidenspointner, G., White, T. A., Wunderer, C., Meijer, G., Ullrich, J., Stapelfeldt, H., Rolles, D. & Chapman, H. N. (2014). *Phys. Rev. Lett.* **112**, 083002.

Laue, M. von (1915). *Concerning the detection of X-ray interferences*. Nobel Lecture.

Ma, L., Yong, H., Geiser, J. D., Moreno Carrascosa, A., Goff, N. & Weber, P. M. (2020). *Struct. Dyn.* **7**, 034102.

Mai, S., Richter, M., Ruckebauer, M., Oppel, M., Marquetand, P. & González, L. (2014). *SHARC: Surface Hopping Including Arbitrary Couplings - Program Package for Non-Adiabatic Dynamics*, <https://sharc-md.org>.

Minitti, M. P., Budarz, J. M., Kirrander, A., Robinson, J., Lane, T. J., Ratner, D., Saita, K., Northey, T., Stankus, B., Cofer-Shabica, V., Hastings, J. & Weber, P. M. (2014). *Faraday Discuss.* **171**, 81–91.

Minitti, M. P., Budarz, J. M., Kirrander, A., Robinson, J. S., Ratner, D., Lane, T. J., Zhu, D., Glownia, J. M., Kozina, M., Lemke, H. T., Sikorski, M., Feng, Y., Nelson, S., Saita, K., Stankus, B., Northey, T., Hastings, J. B. & Weber, P. M. (2015). *Phys. Rev. Lett.* **114**, 255501.

Mitzel, N. W. & Rankin, D. W. (2003). *Dalton Trans.* pp. 3650–3662.

Moreno Carrascosa, A., Coe, J. P., Simmermacher, M., Paterson, M. J. & Kirrander, A. (2022). *Phys. Chem. Chem. Phys.* **24**, 24542–24552.

Moreno Carrascosa, A., Yong, H., Crittenden, D. L., Weber, P. M. & Kirrander, A. (2019). *J. Chem. Theory Comput.* **15**, 2836–2846.

Mott, N. F. (1930). *Proc. R. Soc. London. Ser. A*, **127**, 658–665.

Mu, X., Zhang, M., Feng, J., Yang, H., Medvedev, N., Liu, X., Yang, L., Wu, Z., Xu, H. & Li, Z. (2023). *Ultrafast Sci.* **3**, 0015.

Northey, T. & Kirrander, A. (2019). *J. Phys. Chem. A*, **123**, 3395–3406.

Northey, T., Moreno Carrascosa, A., Schäfer, S. & Kirrander, A. (2016). *J. Chem. Phys.* **145**, 154304.

Northey, T., Weber, P. M. & Kirrander, A. (2024). In preparation.

Northey, T., Zotev, N. & Kirrander, A. (2014). *J. Chem. Theory Comput.* **10**, 4911–4920.

Odate, A., Kirrander, A., Weber, P. M. & Minitti, M. P. (2023). *Adv. Phys. X*, **8**, 2126796.

Perutz, M. F., Rossmann, M. G., Cullis, A. F., Muirhead, H., Will, G. & North, A. (1960). *Nature*, **185**, 416–422.

Pirenne, M. H. (1939). *J. Chem. Phys.* **7**, 144–155.

Pirenne, M. H. (1946). *The Diffraction of X-rays and Electrons by Free Molecules*. Cambridge University Press.

Prince, E. (2006). Editor. *International Tables for Crystallography*, Vol. C. International Union of Crystallography.

Ruddock, J. M., Zotev, N., Stankus, B., Yong, H.-W., Bellshaw, D., Boutet, S., Lane, T. J., Liang, M., Carbajo, S., Du, W., Kirrander, A., Minitti, M. P. & Weber, P. M. (2019). *Angew. Chem. Int. Ed.* **58**, 6371–6375.

Sciaini, G. & Miller, R. D. (2011). *Rep. Prog. Phys.* **74**, 096101.

Shibata, S., Hirota, F. & Shioda, T. (1999). *J. Mol. Struct.* **485–486**, 1–11.

Shibata, S., Sekiyama, H., Tachikawa, K. & Moribe, M. (2002). *J. Mol. Struct.* **641**, 1–6.

Simmermacher, M., Henriksen, N. E., Møller, K. B., Moreno Carrascosa, A. & Kirrander, A. (2019a). *Phys. Rev. Lett.* **122**, 073003.

Simmermacher, M., Moreno Carrascosa, A. E., Henriksen, N. B., Møller, K. & Kirrander, A. (2019b). *J. Chem. Phys.* **151**, 174302.

Stankus, B., Yong, H., Zotev, N., Ruddock, J. M., Bellshaw, D., Lane, T. J., Liang, M., Boutet, S., Carbajo, S., Robinson, J. S., Du, W., Goff, N., Chang, Y., Koglin, J. E., Minitti, M. P., Kirrander, A. & Weber, P. M. (2019). *Nat. Chem.* **11**, 716–721.

Stefanou, M., Saita, K., Shalashilin, D. V. & Kirrander, A. (2017). *Chem. Phys. Lett.* **683**, 300–305.

Suominen, H. J. & Kirrander, A. (2014). *Phys. Rev. Lett.* **112**, 043002.

Wang, J., Tripathi, A. N. & Smith, V. H. Jr (1994). *J. Chem. Phys.* **101**, 4842–4854.

Watson, J. D. & Crick, F. H. (1953). *Nature*, **171**, 737–738.

Weathersby, S. P., Brown, G., Centurion, M., Chase, T. F., Coffee, R., Corbett, J., Eichner, J. P., Frisch, J. C., Fry, A. R., Gühr, M., Hartmann, N., Hast, C., Hettel, R., Jobe, R. K., Jongewaard, E. N., Lewandowski, J. R., Li, R. K., Lindenbergh, A. M., Makasyuk, I., May, J. E., McCormick, D., Nguyen, M. N., Reid, A. H., Shen, X., Sokolowski-Tinten, K., Vecchione, T., Vetter, S. L., Wu, J., Yang, J., Dürr, H. A. & Wang, X. J. (2015). *Rev. Sci. Instrum.* **86**, 073702.

Werner, H.-J., Knowles, P. J., Knizia, G., Manby, F. R. & Schütz, M. (2012). *WIRES Comput. Mol. Sci.* **2**, 242–253.

Yang, J., Makhija, V., Kumarappan, V. & Centurion, M. (2014). *Struct. Dyn.* **1**, 044101.

- Yang, J., Zhu, X. F., Nunes, J. P., Yu, J. K., Parrish, R. M., Wolf, T. J., Centurion, M., Gühr, M., Li, R., Liu, Y., Moore, B., Niebuhr, M., Park, S., Shen, X., Weathersby, S., Weinacht, T., Martinez, T. J. & Wang, X. (2020). *Science*, **368**, 885–889.
- Yong, H., Cavaletto, S. M. & Mukamel, S. (2021a). *J. Phys. Chem. Lett.* **12**, 9800–9806.
- Yong, H., Moreno Carrascosa, A., Ma, L., Stankus, B., Minitti, M. P., Kirrander, A. & Weber, P. M. (2021c). *Faraday Discuss.* **228**, 104–122.
- Yong, H., Xu, X., Ruddock, J. M., Stankus, B., Carrascosa, A. M., Zotev, N., Bellshaw, D., Du, W., Goff, N., Chang, Y., Boutet, S., Carbajo, S., Koglin, J. E., Liang, M., Robinson, J. S., Kirrander, A., Minitti, M. P. & Weber, P. M. (2021b). *Proc. Natl Acad. Sci. USA*, **118**, e2021714118.
- Yong, H., Zotev, N., Ruddock, J. M., Stankus, B., Simmermacher, M., Carrascosa, A. M., Du, W., Goff, N., Chang, Y., Bellshaw, D., Liang, M., Carbajo, S., Koglin, J. E., Robinson, J. S., Boutet, S., Minitti, M. P., Kirrander, A. & Weber, P. M. (2020). *Nat. Commun.* **11**, 2157.
- Zandi, O., Wilkin, K. J., Xiong, Y. & Centurion, M. (2017). *Struct. Dyn.* **4**, 044022.
- Zhang, M., Zhang, S., Xiong, Y., Zhang, H., Ischenko, A. A., Vendrell, O., Dong, X., Mu, X., Centurion, M., Xu, H., Miller, R. J. D. & Li, Z. (2021). *Nat. Commun.* **12**, 5441.
- Ziems, K. M., Simmermacher, M., Gräfe, S. & Kirrander, A. (2023). *J. Chem. Phys.* **159**, 044108.
- Zotev, N., Moreno Carrascosa, A., Simmermacher, M. & Kirrander, A. (2020). *J. Chem. Theory Comput.* **16**, 2594–2605.